國 立 交 通 大 學

生物醫學研究所
碩士論文

利用飽和定點突變方法針對酵母菌氧化鯊烯環化酵素內 Ile705 和豌豆氧化鯊烯-β-麥胚固醇環化酵素假設活性區內的 Leu734 兩者的結構影響在環化及重組過程中的功能性分析

Site-Saturated Mutagenesis on Isoleucine 705 from

***Saccharomyces cerevisiae* Oxidosqualene-Lanosterol Cyclase**

and Leucine 734 from *Pisum sativum* β-amyrin Synthase

Generate Diverse Truncated Cyclization/Rearrangement

Products with Different Stereochemistry

研 究 生：張亦諄

指導教授：吳東昆 博士

中華民國 九十八年七月

Site-Saturated Mutagenesis on Isoleucine 705 from *Saccharomyces cerevisiae* Oxidosqualene-Lanosterol Cyclase and Leucine 734 from *Pisum sativum* β-amyrin Synthase Generate Diverse Truncated Cyclization/Rearrangement Products with Different Stereochemistry

研究生：張亦諄   Student: Yi-Chun Chang

指導教授：吳東昆 博士   Advisor: Dr. Tung-Kung Wu

國 立 交 通 大 學

生物醫學研究所

碩士論文

A Manuscript of Thesis

Submitted to Department of Biological Science and Technology

College of Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Biomedical Science

July, 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年七月

利用飽和定點突變方法針對酵母菌氧化鯊烯環化酵素內 **Ile705** 和豌豆氧化鯊烯-**β**-麥胚固醇環化酵素假設活性區內的 **Leu734** 兩者的結構影響在環化及重組過程中的功能性分析

學生: 張亦諄　　　　　　　　　　　　　　　指導教授: 吳東昆　博士

國立交通大學　生物醫學研究所碩士班

摘要

　　在近半個世紀以來,讓許多有機生物化學家為之著迷的酵素-氧化鯊烯環化酵素,以氧化鯊烯作為反應起始物,在不同生物體中經由各式各樣的氧化鯊烯環化酵素會形成特定的環化產物。催化過程中包含氧化鯊烯上環氧基開環起始反應,經由複雜的環化/重組反應以及最後高度專一性的去質子化步驟形成環化產物。為了比較不同類型的氧化鯊烯環化酵素,我們選用了酵母菌氧化鯊烯-羊毛硬脂醇環化酵素和豌豆-β-麥胚固醇環化酵素兩種作分析,前者在受質的摺疊會經由椅形-船形-椅形 (chair-boat-chair) 形成原脂醇碳陽離子中間物 (Protosteryl cation intermediate) 而後者則會經由椅形-椅形-椅形 (chair-chair-chair) 形成達瑪烯碳陽離子中間物 (Dammarenyl cation intermediate),兩種環化酵素分別會形成羊毛硬脂醇 (Lanosterol) 和 β-麥胚固醇 (β-Amyrin)。利用飽和定點突變的方式,分析存在於酵素假設活性區中的相對應胺基酸,Ile705 及 Leu734。由於在細菌 SHC 酵素活性區相對位置 L607 的研究中,L607 的重要性不容被忽視,所以希望能從實驗分析找出它們的功能以及重要性。

　　實驗結果方面,在Ile705位置突變過後的酵母菌氧化鯊烯-羊毛硬脂醇環化酵素,產生了七種產物,除了原先就會產生的羊毛硬脂醇之外,產生五種已知的環化中間物和一個先前未曾被發現過的未知物。四環的新產物較為被關注的焦點是它在 17 號碳的型態擁有向上的氫,也就是長碳側鏈是向下的,此特殊構型我們將其定義為17α 構

型，同時在 ERG7^{I705F} 突變株可發現兩種擁有 17α 構型的產物，所以可以得知 ERG7^{I705} 在決定 17 號碳向上或向下型態佔有一定的重要性。此外，除了新產物以外，其他環化中間物皆和 ERG7^{F699X} 的分析相同，加上與離受質較遠的 I705 相比，F699 為活性區第一層胺基酸且影響力大，可以推測 I705 和第一層胺基酸的穩定性息息相關。

在豌豆-β-麥胚固醇環化酵素中 Leu734 的功能性分析結果方面，和 Ile705 的結果大相逕庭，L734 的突變株沒有產生任何環化產物。而造成此結果的可能性是 L734 扮演著穩定受質的腳色，因此在 L734 突變過後，由於立障、酸鹼度、極性等環境的改變，從電腦模擬軟體的判斷推測鄰近胺基酸最有可能被影響是 F728，穩定受質環境被破壞以致於沒有任何環化產物產生。另一方面，由於豌豆-β-麥胚固醇環化酵素活性區胺基酸的分析非常稀少，對於 L734 的功能性判斷也無法深入，所以對於其鄰近胺基酸的分析也是未來所必須面臨的。

Site-Saturated Mutagenesis on Isoleucine 705 from *Saccharomyces*

cerevisiae Oxidosqualene-Lanosterol Cyclase and Leucine 734 from *Pisum sativum* β-amyrin Synthase Generate Diverse Truncated Cyclization/Rearrangement Products with Different Stereochemistry

Student: Yi-Chun Chang   Advisor: Dr. Tung-Kung Wu

Institute of Biomedical Science
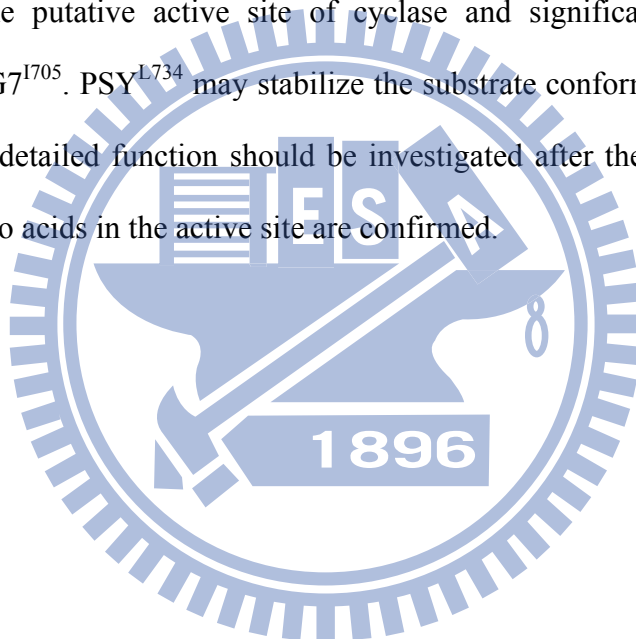
Naitonal Chiao Tung Unversity

Abstract

Oxidosqualene-lanosterol cyclase (*S. cerevisiae* ERG7) catalyzes the biotransformation of the linear form substrate, oxidosqualene, into tetracyclic lanosterol in yeast and mammals. Different species of organisms including *S. cerevisiae* OSC (ERG7) and *P. sativum* βAS (PSY) operate through different conformational intermediates within the oxidosqualene cyclization process. According to previous reports, by utilizing the diverse structural and stereochemical control in various catalytically important amino acid residue mutants, oxidosqualene cyclase produced diverse product profiles ranging from mono- to polycyclic triterpene alcohols. These data implied that the direction for the plastic enzyme was redesigned to obtain a novel reactivity from this complex enzyme, but with the characteristic of well-known high product specificity. Moreover, in order to further illustrate other critical amino acids involved in the catalytic significance and/or enzymatic plasticity of OSC and PSY, we describe herein a series of site-saturated mutations of the Ile705 residue of ERG7 and Leu734 of PSY.

In the mutations of I705, seven products including three known truncated cyclization tricyclic structures, three known tetracyclic structures, as well as one novel compound that

contains a tetracyclic scaffold with a 17α side chain and a △20/22 double bond, were identified from various ERG7[I705X] mutants. From the product distribution of ERG7[I705X] mutants, we deduce that the Ile705 residue may affect the first-tired residues and the stereochemistry of exocyclic long side chain during the final step of cyclization, to produce either 17α or 17β side chain derivatives in different mutants. The relationship of the structure-function-mechanisms of Ile705 on the catalysis activity of OSC will be discussed.

However, mutation of L734 causing disruption of catalytic cyclization, β-amyrin synthase did not work in PSY[L734X] mutants. This result revealed that the L734 residue is crucial within the putative active site of cyclase and significantly different with the mutations of ERG7[I705]. PSY[L734] may stabilize the substrate conformation with its neighbor residues, but the detailed function should be investigated after the functional roles of the neighboring amino acids in the active site are confirmed.

誌謝(Acknowledgement)

Table of Contents

List of Figures

List of Tables

Chapter 1   Introduction

1.1 Triterpenoids and its biosynthetic pathway

Terpenoids, sometimes called isopenoids, are a large class of organic chemicals similar to terpenes, which are the combination of several isoprene units. Some chemists use the term "terpene" extensively more than terpenoids. The classification of terpenoids could accord the the number of isoprene units or the number of cyclic structures they contain. Triterpenoids were derived from six isoprene units, usually have tetracyclic or pentacyclic structure and $C_{30}H_{50}O$ formula. (Fig. 1.1) They could be found in many classes of living things, and also are the largest group in the nature products. Moreover, the interest in various aspects of the biological activities on triterpenoids and triterpenoids saponins was increasing.[1] Belongs to triterpenoids, sterols have the same general ring structure and known as steroid alcohols with a hydroxyl group at the C-3 position of the A-ring. Thus, the hydroxyl group on the A-ring is polar, whereas the rest of aliphatic chain is non-polar. Sterols were vital constituents of cell membranes. In animals, cholesterol (Fig. 1.2) was found in the cell membranes and transported in the blood plasma. It was required to build and maintain cell membranes and regulate the membrane fluidity on the range of physiological temperatures. Also, cholesterol was the precursor molecule in many biochemical pathways. The phytosterols include campesterol, stigmasterol and β-sitosterol, and act as a sterol component in the membrane. Furthermore, food additive, medicines and cosmetics were all applications of phytosterols. Although sterols were usually absent in bacteria, some bacteria and protozoan could produce triterpenes such as hopene which was regarded as the sterol substitute.[2] Because these compounds play the important role in many ways, the complete understanding of the specific enzyme during their biosynthetic pathway was necessary.
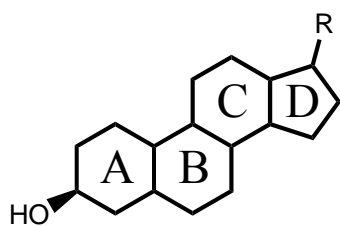
Figure 1.1 Triterpenoid backbone   **Figure 1.2 Cholesterol**
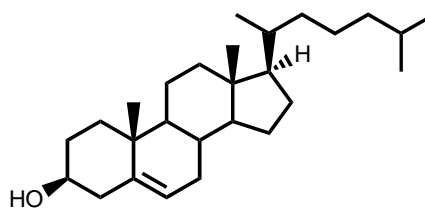
The sterol biosynthesis was generally through the mevalonic acid (MVA) pathway. Acetyl coenzyme A is to be the starting material, it is converted into acetoacetyl-CoA by acetoacetyl-CoA thiolase. After the catalysis of 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase) and 3-hydroxy-3-methygulutary-CoA reductase (HMG-CoA reductase), acetoacetyl-CoA is converted into mevalonate. Mevalonate is phosphorylated by 2 sequential Pi transfers from ATP, yielding the pyrophosphate derivative. ATP-dependent decarboxylation, with dehydration, yields isopentenyl pyrophosphate. Isopentenyl pyrophosphate in the pathway is referred to as isoprenoid, by a series of chemical reactions including isomerization and condensation, a linear molecule with 30 carbons, squalene, was produced. Oxidation of squalene by squalene synthase yields an acyclic polyolefin substrate, (3*S*)-2,3-oxidosqualene (OS). The common substrate oxidosqualene solely proceeded the biotransformation for the production of tetracyclic lanosterol in animals and fungi, whereas a variety of polycyclic triterpene alcohols including cycloartenol, lupeol, α-amyrin and β-amyrin are simultaneously generated in higher plants. (Fig. 1.3)

3 Acetyl CoA C2

↓

Acetoacetyl CoA

↓ HMG-CoA synthase

3-Hydroxy-3-methylglutaryl-CoA

↓ HMG-CoA reductase

Mevalonate

↓

Isopentenyl tRNA ←   → Dolichol   C5

Ubiquinone ←   Isoprenoid intermediate   → Heme A

geranyl pyrophosphate

farnesyl pyrophosphate

Hopene ← **SHC**   Squalene C30

Bacteria

(3*S*)2,3-Oxidosqualene   **βAS**

CAS   **High plant**   **High plant**

OSC

Cycloartenol   **Lanosterol**   **β-Amyrin**

Fungi   **Animal**

Phytosterol   **Ergosterol**   **Cholesterol**   **Oleanane-type triterpene saponins**

Figure 1.3 Sterol biosynthetic pathway

One of the triterpenoids that is crucial in our life is cholesterol. In the study of the treatment on hypercholesterolemia was targeting at the step catalyzed by HMG-CoA reductase, which is the rate-limiting step in the early part of the pathway. But the inhibition of HMG-CoA reductase not only affects the production of cholesterol, but also have the side effect that influence the isoprenoid intermediates and its derivatives. It may cause the adverse clinical results. On the other hand, in the downstream of the biosynthetic pathway, inhibitors of oxidosqualene-lanosterol cyclase (OSC) as anticholesteraemic drugs might provide insight of the development of a safer, more effective treatment. Furthermore, the highly-complex and specific reaction of oxidosqualene-lanosterol cyclase include protonation, cyclization, rearrangement of carbon skeleton, and deprotonation, have bring to the interest of chemists. Therefore, current research for OSC is increasing, and the structure-function relationship of this enzyme is to be worth to discuss in depth.
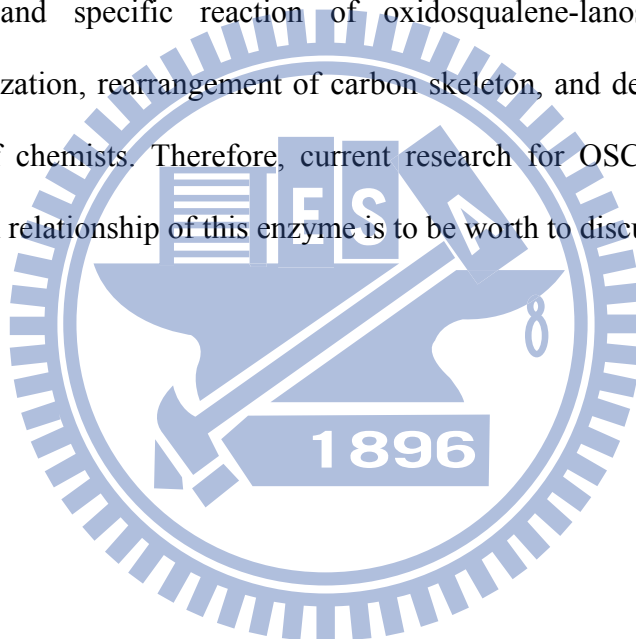
1.2 Triterpene cyclases

1.2.1 Product diversity

The triterpenoids are a large diverse group of natural products derived from squalene, oxidosqualene or related acyclic 30-carbon precursors. Nearly 200 tertripene compounds are known from nature or enzymatic reactions which are cyclization products of olefins such as squalene, oxidosqualene, or bis-oxidosqualene.[3] Several triterpene cyclase enzymes transform these olefins into complex and biologically important polycyclic products such as hopene (6-6-6-6-5 pentacycles), hopanol (6-6-6-6-5 pentacycles), lanosterol (6-6-6-5 tetracycles), cycloartenol (6-6-6-5 tetracycles), lupenol (6-6-6-6-5 pentacycles), and β-amyrin (6-6-6-6-6 pentacycles). (Fig. 1.4) The reason for triterpene diversity is that triterpene cyclases are groups of plastic enzymes and have novel catalytic properties from mutational alteration. The triterpenes serve as the precursors of all hopanoids and steroids, including cholesterol, glucocorticoid, estrogens, and hormones. Moreover, the mechanism of action of triterpene cyclases stimulated the interest of many biologists and chemists nearly a half century ago. Early studies on triterpene cyclases had used substrate analogs to provide important mechanistic information, the range of structures which cyclases can transform, as well as the nature of any resulting products.[4] However, in-depth and detailed understanding the structure-function-reaction mechanism relationships of the plastic triterpene cyclases, such as the role of cyclases in substrate prefolding and cation stabilization is required to develop new methods for use in protein expression and crystallization research.
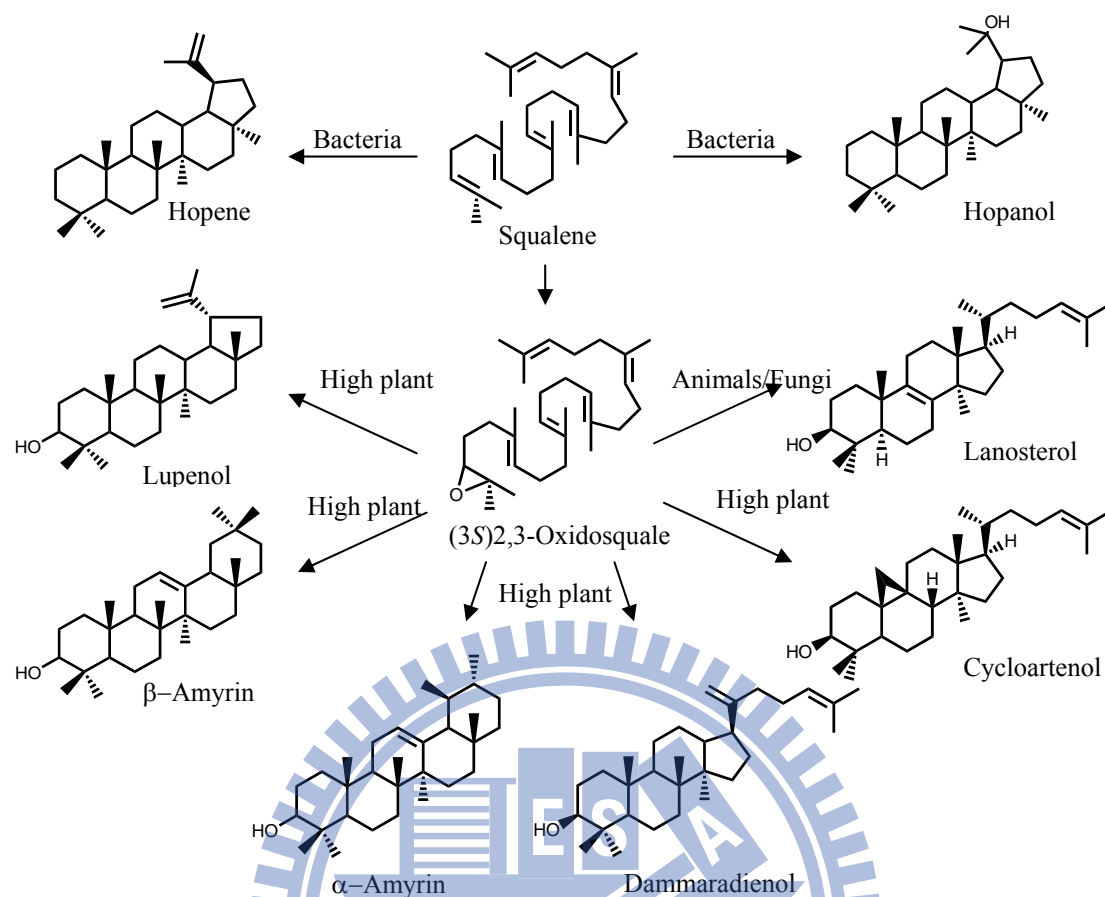
Figure 1.4 The product diversity of triterpene synthases

1.2.2 Mechanism

The general mechanism of triterpene cyclases involves: (1) binding on the olefin substrate in a folded conformation; (2) initiated by protonation of double bond (squalene) or one oxirane ring (oxidosqualene); (3) four or five-membered stereochemically controlled ring formations; (4) nucleophilic hydride and methyl group rearrangement; and, (5) termination either by deprotonation or water addition. Thus, the cyclized process involved sixteen bonds broken, and sixteen new bonds formed, and through the cation-olefin cyclization, the unstable cationic intermediates are formed. The transient cationic intermediates are stabilized by the putative active site of cyclases, hence, it could not be isolated by scientists. Due to different mechanisms, the diversity of triterpenes could be explained by different hydride / methyl shifts, and the early deprotonated truncations.

In contrast to oxidosqualene cyclase, squalene-hopene-cyclase (SHC) allows all-chair conformation of squalene, and there are two pre-folded conformational types in the oxidosqualene cyclization. One is so-called "chair-boat-chair" conformation of oxidosqualene, as a substrate. The "chair-boat-chair" conformation would form lanosterol in animal/ fungi or cycloartenol in higher plants. On the other way, oxidosqualene would pre-fold into "chair-chair-chair" conformation. This conformation is usually found in the plant species such as *Olea europaea* (lupenol synthase) and *Pisum sativum* (β-amyrin synthase).[5] These two different conformations via two different cation intermediates, cyclization of "chair-boat-chair" conformation would give the protosterol cation which had a positive charge at C-20, whereas the "chair-chair-chair" conformation would give the dammarenyl C-20 cation. In the protosteryl cationic pathway, B-ring boat products were produced, while B-ring chair products were formed in the dammarenyl cationic pathway. (Fig. 1.5)
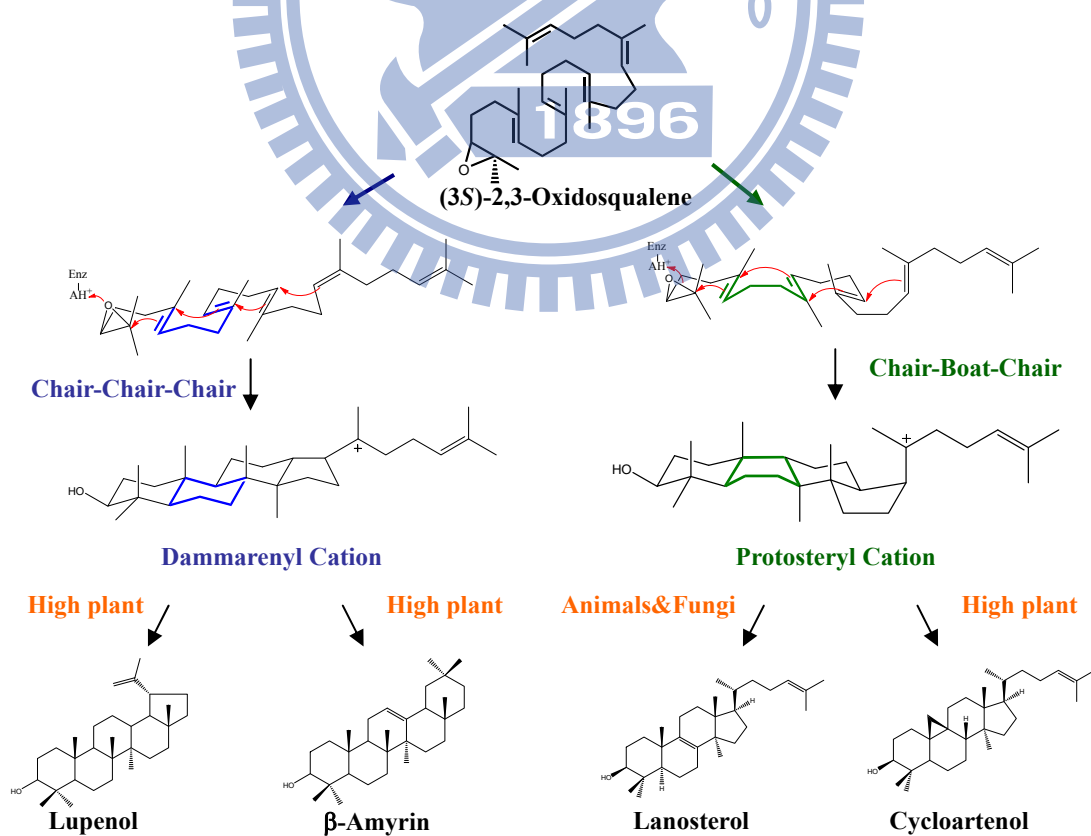
Figure 1.5 Cyclization of oxidosqualene to protosteryl and dammarenyl cation.

7

1.3 Oxidosqualene-lanosterol cyclase (OSC)

1.3.1 The hypothesis of oxidosqualene cyclases

The family of oxidosqualene cyclases extensively exist in the organism, and the common cyclases such as oxidosqualene-lanosterol cyclases (OSC) in animals and fungi, plant oxidosqualene cyclases including cycloartenol synthase (CAS), lupenol synthase (LUS), and β-amyrin synthase (βAS). These different enzyme systems convert the oxidosqualene, as a substrate, into various polycyclic triterpenes. And the relationship between the enzyme structure and cyclization mechanism is extremely interesting.

Research on oxidosqualene cyclases was remarkable for nearly half a century. Woodward and Bloch first proposed that in the cyclization of squalene in cholesterol synthesis followed by rearrangement to lanosterol in 1953.[6] In 1958 and 1965, Maudgal and Cornforth groups evidenced for 1.2-methyl and hydride shifts during lanosterol formation by incorporation experiments.[7,8] Corey and van Tamelen showed that 2,3-oxidosqualene is more efficiently incorporated in sterol synthesis than squalene, demonstrated intermediacy of 2,3-oxidosqualene in lanosterol biosynthesis.[9,10] Also, their laboratories had many substrate analogue compounds for synthesis and test. In addition, van Tamelen also showed the nonenzymatic cyclization of 2,3-oxidosqualene resulted in truncated cyclization to produce a tricyclic product, suggesting that direct enzymatic control is necessary for the prevention of the chemical tendency in the formation of five-membered C-ring, and for emergence of the biologically required six-membered C-ring.[11] In 1975, Barton confirmed that eukaryotic oxidosqualene cyclases accepted only (3S)- but not (3R)-enantiomer of 2,3-oxidosqualene as a substrate in the formation of lanosterol, demonstrating the highly substrate-specific property of oxidosqualene cyclases rather than squalene cyclases.[12] Guy Ourisson and his co-workers proposed the possible molecular evolution from the primitive squalene cyclases to oxidosqualene cyclases in higher organisms.[13] Furthermore, Corey and Matsuda showed that oxirane cleavage and

cyclization of the A-ring are concerted and it is essential for electrophilic activation of the oxirane function.[14,15]

In previous research, oxidosqualene would be stable in the neutral condition at room temperature for the whole day. And the more powerful acid, such as trichloroacetic acid, is required for strong epoxide activation to promote the cyclization step.[16] Then, it would form many unstable cation intermediates in the cyclization process, where there are two hypotheses that illustrated how the enzyme functions to stabilize the high-energy cation intermediate. In 1987, Johnson proposed the "cation-stabilizing auxiliary" model, which proposed that the Lewis acid residues on the active site provide a proton for initiation of the epoxide group, and a number of anionic sites in the cyclase enzyme would lead the cation generation and the formation of the proper ring system. The axial negative charge residues would face toward the transition states or the intermediates for cation stabilization, and it would demonstrate the ring formation of B-boat/chair ring. Therefore the B-boat ring could be promoted by delivery of a point charge to the α-face at pro-C-8, and lowering the activation energy of the boat form rather than the chair skeleton.[17,18] (Fig. 1.6) In addition, Griffin and co-workers proposed the "aromatic hypothesis" model which the electron-rich aromatic side chain such as Trp and Tyr, might stabilize the positively charged transition states or high-energy intermediates during the process of cyclization and rearrangement steps.[19] (Fig. 1.7) The hypothesis is that aromatic residues play the role of the anions group, just like that proposed in the Johnson model mechanism. Cation-π interactions are common features in enzyme-substrate complexes.

Dammarenyl Cation   **Protosteryl Cation**

⇧ **denotes site for delivery of negative point charge**
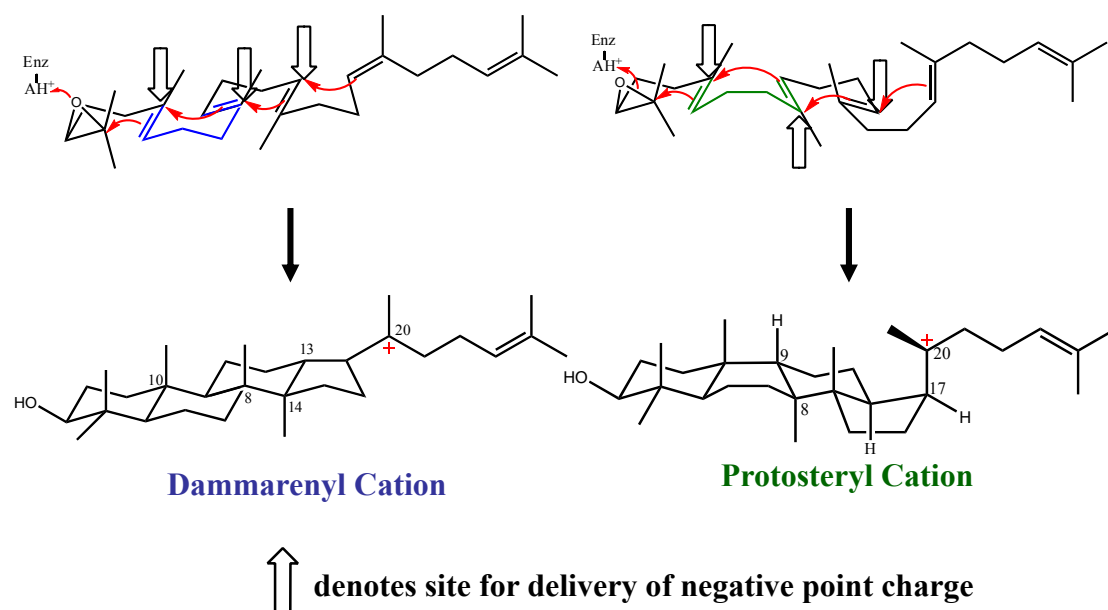
Figure 1.6 The proposed enzyme models by Johnson.
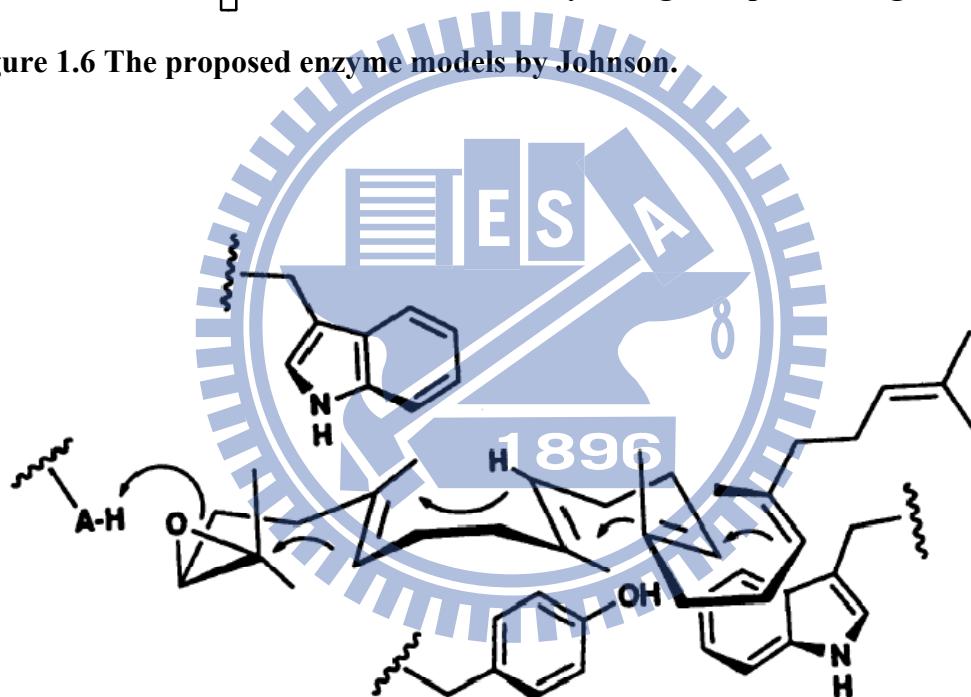
Figure 1.7 Griffin's hypothesis model for involvement of electron-rich aromatic side chains from Trp and Tyr residues in the cyclization of oxidosqualene to the protosteryl cation.[19]

1.3.2 The studies of mechanism and site-directed mutagenesis

Oxidosqualene-lanosterol cyclases (OSC), also called lanosterol synthases, are widely found in animals and fungi, and usually convert oxidosqualene into the tetracyclic

product, lanosterol. In *Saccharomyces cerevisiae*, the cyclases were encoded from *ERG7* gene, as a membrane protein. The protein contains 731 amino acids and the molecular mass is approximately 83 kDa. Because of the difficulty in purifying membrane protein, the crystal structure of *S. cerevisiae* oxidosqualene-lanosterol cyclases (*Sce*ERG7) enzyme has not been attained until now. In order to understand the structure-function relationships of oxidosqualene cyclase-catalyzed reactions in depth, site-directed saturated mutagenesis coupled with genetic selection and products analysis for illustrating the functional importance of these mutated residues, and describing the carboncationic intermediates in the reaction cascade.

Briefly, the mechanism of oxidosqualene-lanosterol cyclases include one oxirane ring opening, four stereochemically controlled ring formations, four nucleophilic hydride and methyl group migrations, and the final deprotonation reaction at C-8 or C-9. (Fig. 1.8)
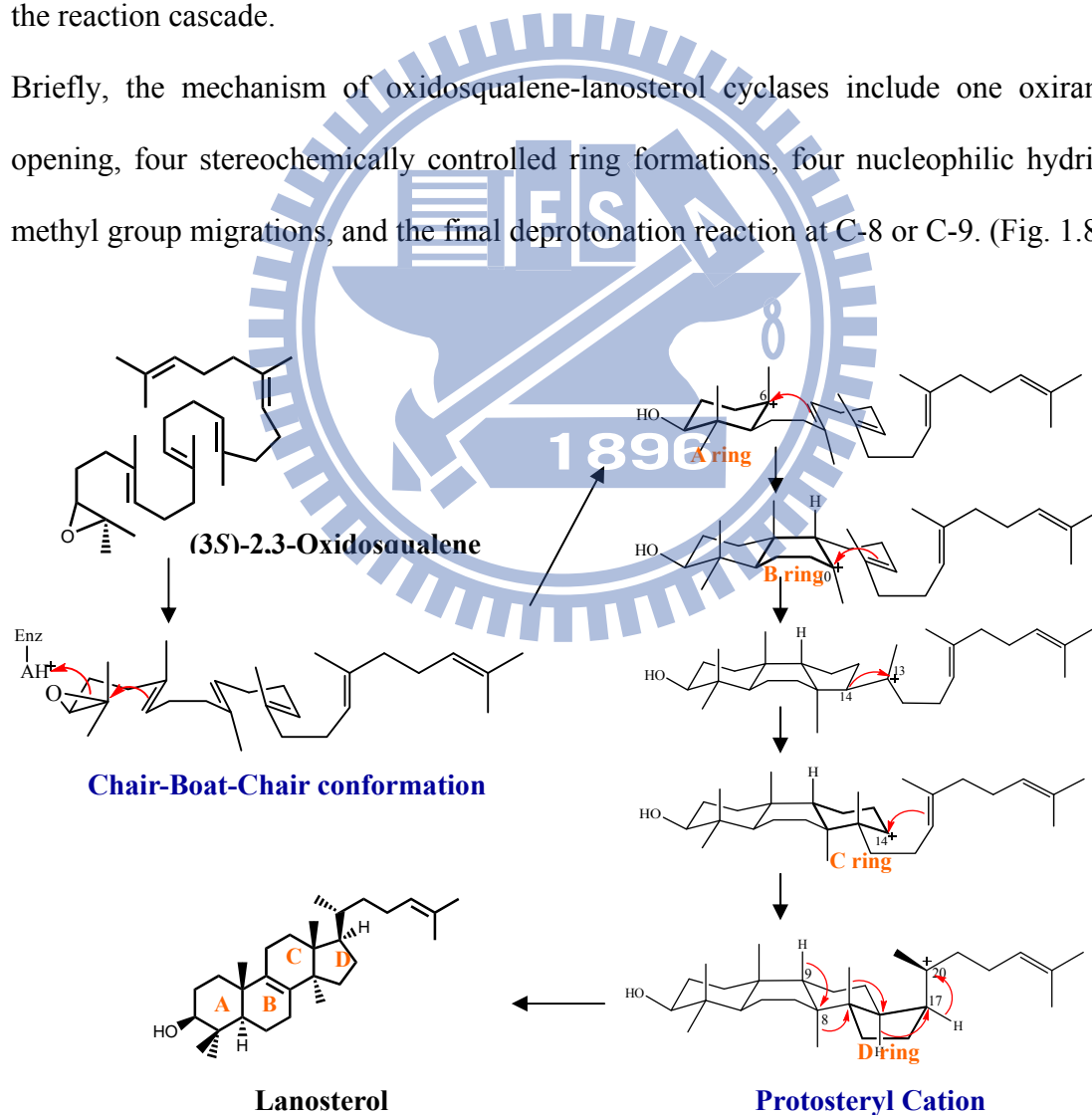
Figure 1.8 The detailed mechanism of conversion of 2,3-oxidosqualene into lanosterol.

In the detailed aspect, the enzyme catalyzed conversion of (3*S*)-2,3-oxidosqualene into a pre-folded chair-boat-chair conformation, a single aspartic acid residue (D456) in *Sce*ERG7 provided a proton to induce the epoxide opening. Before the structure of human OSC was crystallized, scientists used an affinity-labeling strategy of the substrate analogs, trying to identify the important residues within oxidosqualene cyclases.[20] In 1997, Corey and Matsuda used an "alanine scanning site-directed mutagenesis" method to identify the highly-conserved residues in the putative active site within *Sce*ERG7, revealing that His146, His234 and Asp456 are essential for catalytic activity.[11,14] Furthermore, these researchers also showed that His146 assists in forming hydrogen bonds with Asp456, and it would enhance the acidity of Asp456, to induce the oxirane ring opening.[21] (Fig. 1.9) Similarly, in human OSC research, Asp455 (which corresponds to Asp456 within *Sce*ERG7) was confirmed as the general acid that protonates the epoxide group of oxidosqualene. By hydrogen bonding association with Cys456 and Cys533, the acidity of Asp455 was increased and then activated the initial catalyzed reaction. After completing the entire reaction cycle, Asp455 would be reprotonated either from the bulk solvent through the carboxylate group of Glu459 and the water molecules, or by shifting the proton from the final deprotonation step back to the Asp455 residue.[22] (Fig. 1.10) In addition, some research has shown evidence to support concerted epoxide opening and A-ring closure by oxidosqualene-lanosterol cyclases, and these studies usually used oxidosqualene analogs to demonstrate the concerted C-O cleavage and A-ring formation.[14] In addition, some computational results agreed with these experimental findings.[23] Therefore, these two steps have been considered to be concerted up to the present day.
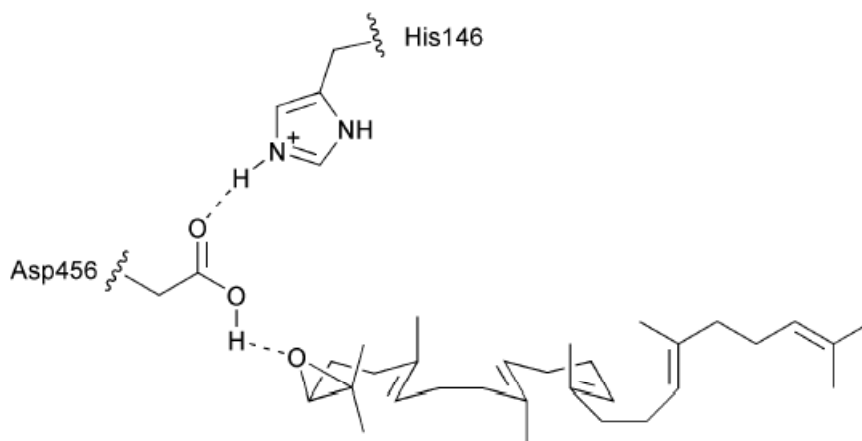
Figure 1.9 The proposed model for oxirane ring opening and cyclization initiation.[21]
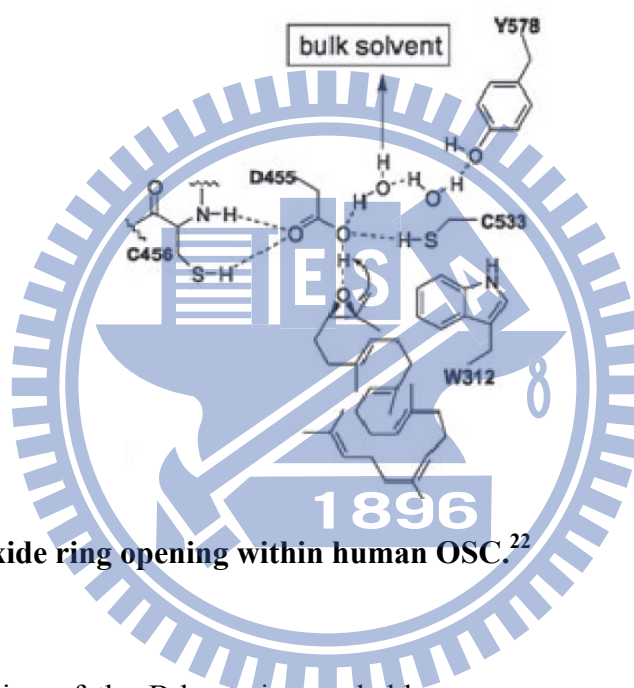
Figure 1.10 Epoxide ring opening within human OSC.[22]

The formation of the B-boat ring probably occurs very rapidly when the positive charge appears on the C-6 of the A-ring. The closure of the C-ring is puzzling because of the direct formation of the six-membered C-ring would represent the anti-Markovnikov form. In 1995, E. J. Corey and his co-workers first proposed the five-membered C-ring Markovnikov cyclization against the direct formation of the six-membered C-ring. By oxidosqualene analogs, 20-oxaoxidosqualene had been reacted with oxidosqualene cyclase, and the resulting products suggested that there was a five-membered C-ring intermediate appearing through an intramolecular reaction.[24] (Fig. 1.11) Moreover, incubations of additional substrate analogs within the oxidosqualene cyclases produced 6-6-5 cyclization

products, providing more evidence for a five-membered C-ring closure followed by a ring expansion.[4] In the molecular biological aspects, the highly-conserved residue Val454 within *Sce*ERG7 had been proven to have a role in the B-ring formation by its steric side chain in Matsuda's lab.[25] They used a series of *Sce*ERG7 mutants such as glycine, alanine, phenylalanine, leucine, and isoleucine to substitute for the Val454 residue. The Val454Phe mutant was inactive, while Val454Lue and Val454Ile mutations produced only lanosterol as the sole product. However, the Val454Ala and Val454Gly produced not only lanosterol by also achilleol A as a minor product. Decreasing the amino acid side chain allows formation of monocyclic product, suggesting that the valine steric bulk side chain may participate in pre-folding the 2,3-oxidosqualene to facilitate B-ring formation. Furthermore, Thoma *et al* determined the structure of human oxidosqualene cyclase in 2004, and they indicated the conserved residues of Phe444, Trp387 and Trp581 (which correspond to Phe445, Trp390 and Trp587 within *Sce*ERG7) stabilize the intermediate cations at C-6 and C-10 through the cation-π interactions, after closure of the A-ring and B-ring. It was considered that Tyr98 may affect the B-ring conformation to become the unfavored boat form, rather than the favorable chair conformation by its bulky side chain in their study.[26] (Fig. 1.12)
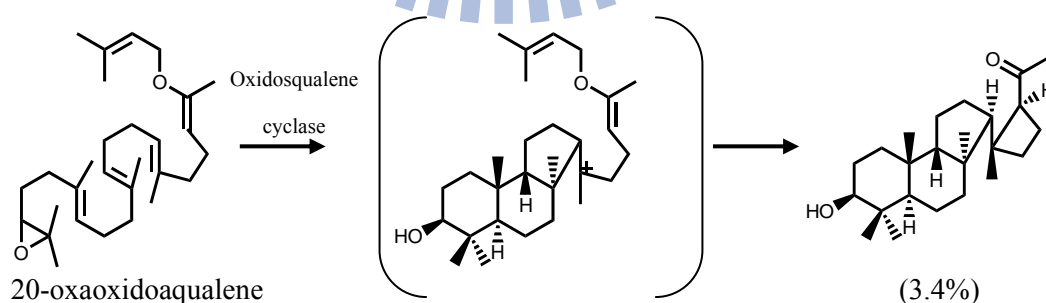
Figure 1.11 Substrate analog and product of oxidosqualene cyclase which are suggestive of a five-membered C-ring intermediate.[24]
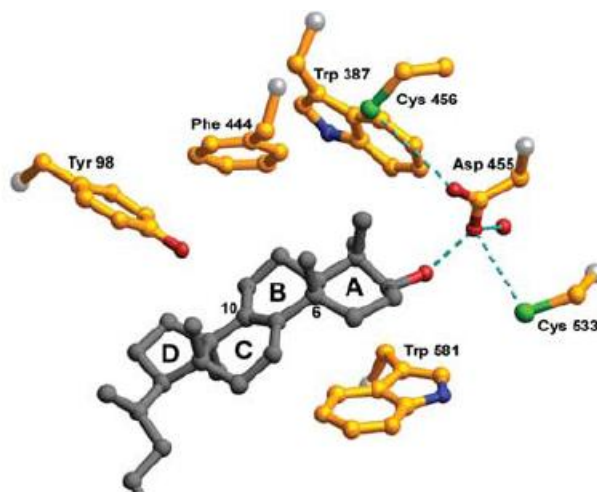
Figure 1.12 Cyclization mechanisms in human OSC. Phe444, Trp387 and Trp581 stabilize the intermediate cations at C-6 and C-10. In addition, the Tyr98 side chain contributes the energetically unfavorable boat conformation of the B-ring.[26]

The final step of the mechanism in lanosterol synthesis is protosterol cation formation and its rearrangement. Early studies of sterol biosynthesis assumed that the conversion of oxidosqualene into lanosterol proceeded via a protosterol cation where its cationic side chain at C-17 is α-oriented.[27] After hydride migration of the 17α- protosterol cation from C-17 to C-20, it does not form the natural sterol 20*R* skeleton, but the unnatural 20*S* configuration. Another explanation for the mechanism was the occurrence of a 120° rotation of C17-C20 bond via the hydride migration, in order to produce the 20*R* skeleton of lanosterol, whereas only a 60° rotation produces the unnatural 20*S* configuration.[28] (Fig. 1.13) In addition, it had been proposed that the 17α- protosterol cation may not be the intermediate, but a covalent adduct formed by the reaction of the 17α-protosterol cation with the nucleophilic group within the cyclase. In 1992, Corey proposed the evidence that the stereochemistry at C-17 of protosterol cation prefers the β rather than α orientation by the enzyme-catalyzed cyclization of 20-oxaoxidosqualene. (Fig. 1.14) This result overthrows the large rotation around the C17-C20 bond prior to rearrangement, and the covalent attachment of the short-lived enzyme nucleophile.[4,28] Moreover, the

17β-protosterol cation was also demonstrated to occur by the catalysis of 20,21-dehydrooxidosqualene.[29]
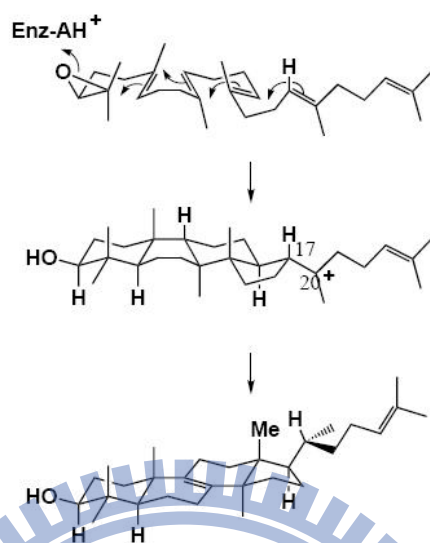
Figure 1.13 The previous study proposed an incorrect C17 stereochemistry of the protosteryl cation. The intermediate required a large side-chain rotation prior to rearrangement to account for the observed stereochemistry at C20.[4,28]
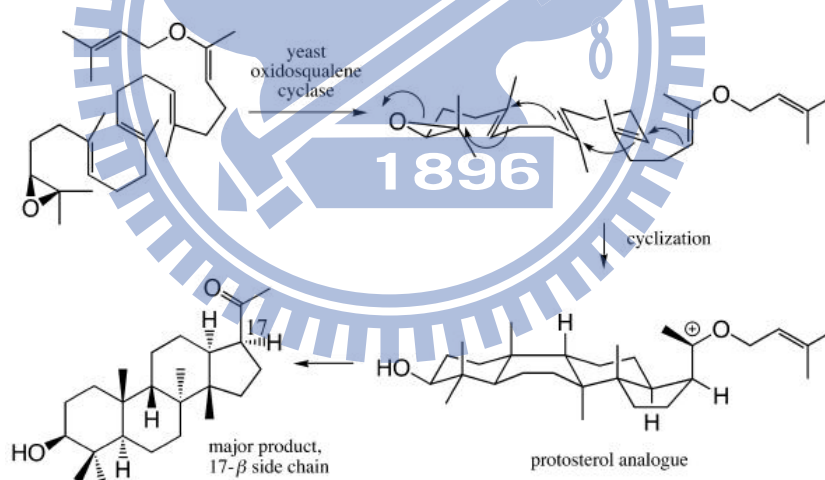
Figure 1.14 The evidence for the stereochemistry of the protosteryl cation intermediate at C17 prefers the β rather than α orientation by incubation of substrate analog.[4,28]

In theoretical calculation research, Hess performed the 6-6-5-tricyclic cation, which was the first intermediate during the formation of the protosterol cation, and suggested that five-membered C-ring expansion and formation of a D-ring were concerted.[30] (Fig. 1.15) In

the SHC study, Gao showed both none, and fewer intermediates of tricyclic and tetracyclic cations, suggesting that the formation of the C, D and E-ring were concerted.[31]
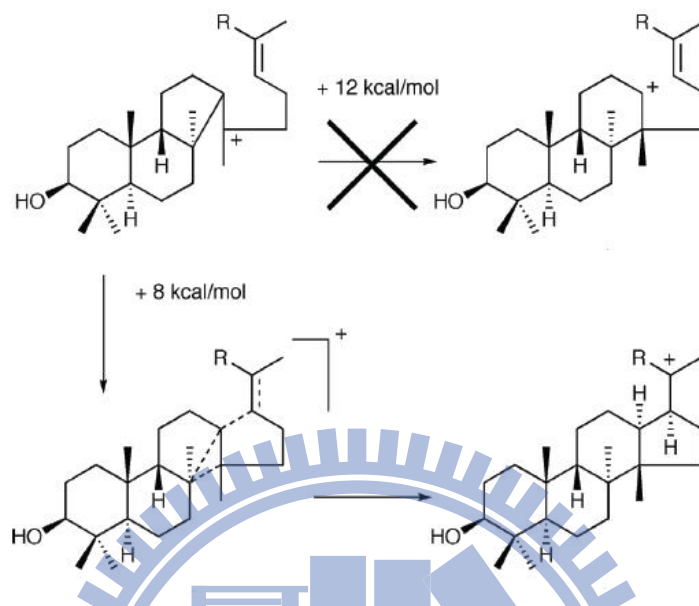
Figure 1.15 Proposed mechanisms for C-ring expansion and D-ring formation were concerted by Hess.[30]

The human OSC structure showed that the Phe696 and His232 (which correspond to Phe699 and His234 in *Sce*ERG7) could stabilize the positive charge at C-20 of the protosteryl cation through cation-π interactions, and then the highly conserved aromatic residues such as Trp192, Trp230, His232, Tyr237, Tyr503, Phe521 and Phe696 could stabilize the methyl/ hydride rearrangement process. Ultimately, His232 is the nearly basic residue that could accept a proton in the deprotonation of the C-8 or C-9 lanosterol cation, and complete the final step of catalysis. It is also possible that the Tyr503 residue which is hydrogen-bonded with His232 could be in the appropriate position for the final cyclization step. On the other hand, in the mutation of Phe699 within *Sce*ERG7, we suggested that it could stabilize the anti-Markovnikov 6-6-5-tricyclic C-14 cation and lanosteryl C-17 cation based on its product profiles. Unlike SHC, oxidosqualene cyclase lacks particular specific electron-rich residues that could promote the D-ring expansion and E-ring formation.

1.3.3 Human oxidosqualene cyclase

In 2004, Thoma and his co-workers solved the X-ray structure of human OSC which is in complex with lanosterol, published in *Nature*. Human OSC is a monomer which consists of two barrel domains that are connected by loops and three smaller β-structures, and the large active-site cavity is located in the center of the molecule between domains 1 and 2. It is a monotopic membrane protein that attached to the membrane from one side, and the membrane-inserted surface consists of a plateau 25 Å in diameter and a channel that leads to the active-site cavity. The channel is considered to admit oxidosqualene, as a substrate, to enter the hydrophobic active site and have the separated role from the putative active-site cavity. Achievement of the substrate passage would either by a change in the side chain of the residue such as Tyr237, Cys233 and Ile524 or by rearrangement of the strained loops 516 to 524 and 697 to 699.[26] (Fig. 1.16) Because of the human OSC crystal structure, more relationship between the functional residues could be understood in depth.
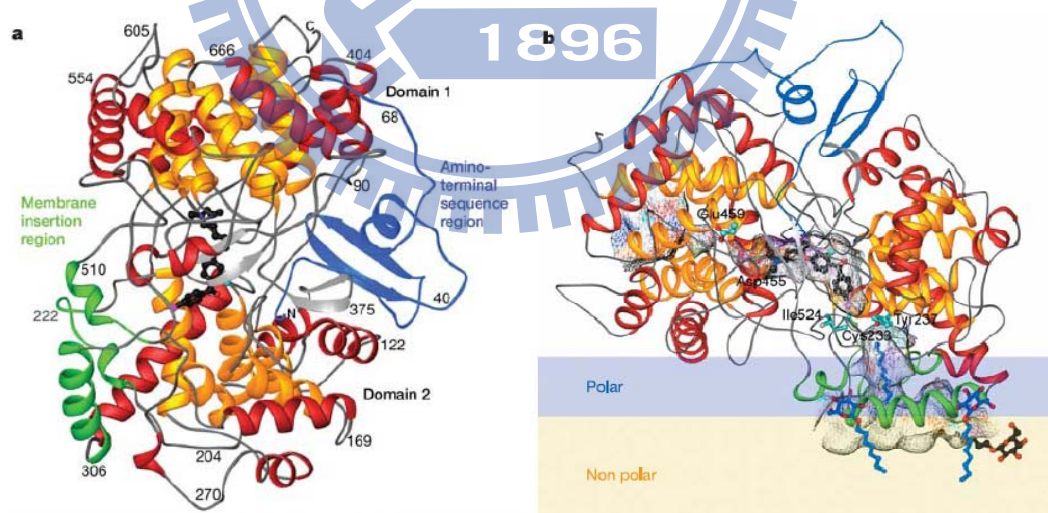
Figure 1.16 Human OSC structure. (a) The ribbon diagram of human OSC. The C and N termini and several sequence positions are labeled. The inner barrel helices are colored yellow. The bound inhibitor, Ro48-8071 (black), indicates the location of the active site. (b) The orientation of OSC relative to one leaflet of the membrane, Ro 48-8071 bind in the certral active-site cavity.[26]

Before the human OSC was solved, the homologue structure of SHC was the model for understanding the OSC catalysis. But SHC and OSC have different substrate, different B-ring conformation, further expandsion of D-ring, and cyclization of E-ring. To gain more information for the highly stereo selective cyclization, the highly sequence identity of human OSC with *Sce*ERG7 would be more useful than the SHC structure.

1.4 Oxidosqualene cyclase in plants

1.4.1 Cycloartenol synthase (CAS)

Cycloartenol is the sterol precursor in higher plants and its skeleton is similar to lanosterol found in animals and fungi. Cycloartenol synthase converts oxidosqualene into cycloartenol via protosteryl cation intermediate, following a series of methyl and hydride shifts to form the C-9 cation, and then terminating the reaction by cyclopropyl ring formation and deprotonation at C-19 to form cycloartenol. Whereas lanosterol synthesis via the extremely similar reaction intermediate that deprotonated from either C-8 or C-9 carbocationic intermediate after methyl and hydride shifts of protosteryl cation, and then yields lanosterol. (Fig. 1.17) Both cycloartenol synthase and lanosterol synthase have more than 700 amino acids, but about 400 amino acids are different in the sequence alignment, and therefore suggested the catalytic distinction for specific deprotonation or cyclopropyl ring formation of these two enzymes. Cycloartenol synthase was first cloned from *Arabidopsis thaliana* (*Ath*CAS1) and expressed and characterized in the yeast lanosterol synthase system by the Matsuda group.[32] The system used the mutated lanosterol synthase as an expression strain, incubated it by exogenous sterol for growth, and then analyzed the details of cycloartenol biosynthesis.

Figure 1.17 The difference between cyclization mechanisms of oxidosqualene-lanosterol cyclase and cycloartenol synthase.
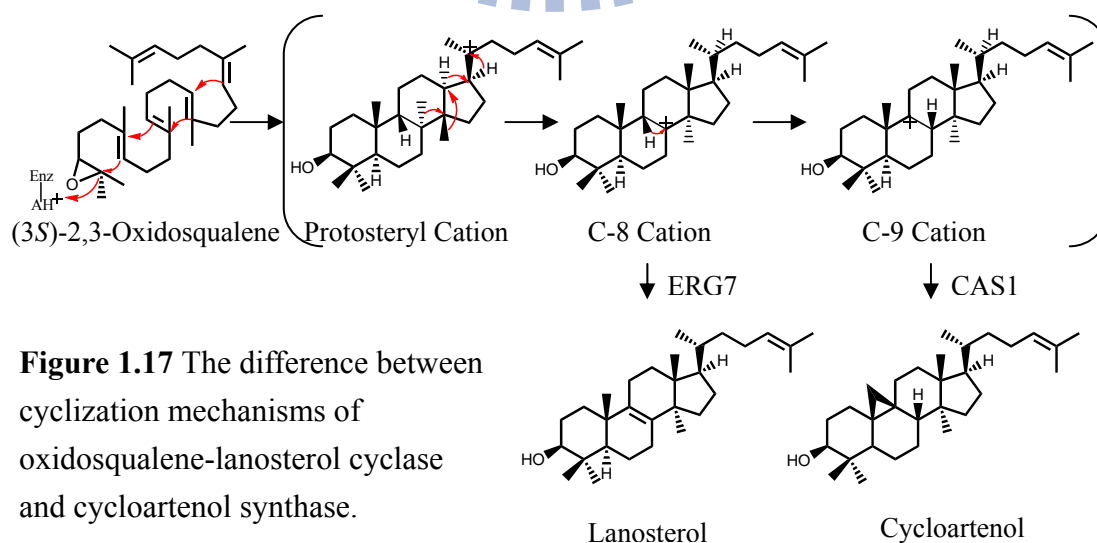
The study of cycloartenol synthase (*Ath*CAS1) by site-directed mutagenesis showed

some important residues such as Tyr410, His477 and Ile481. These residues are highly-conserved within cycloartenol synthase in many plant species, whereas these residues are totally different within ERG7 just like Thr, Cys, Gln and Val amino acids.[33] (Fig. 1.18) Speculation on this phenomenon, these residues may promote the formation of cyclopropyl ring within *Ath*CAS1. Corresponding to Ile481 that residue in the *Sce*ERG7 active site is Val454. Mutation on Ile481 into Val would have a subtle steric change by subsyituting the isopropyl side chain instead the *sec*-butyl side chain. The little steric change cause 25% lanosterol, 21% parleol and 54% cycloartenol formation, indicate that Ile481 is not catalytically necessary but may influence the substrate and active site base by its steric effect. Moreover, in order to understand the steric effect at Ile481, the Matsuda group constructed *Ath*CAS1 Ile481Phe, Ile481Leu, Ile481Ala and Ile481Gly mutations. The Ile481Phe was inactive, and suggested that Phe residue may occludes the active site to cause this result, whereas the similar size but different shape of Ile481Leu produces cycloartenol, parkeol and a little lanosterol. In addition, Ile481Ala and Ile481Gly mutations yielded lanosterol and parkeol, also produce monocyclic compounds achilleol A and camelliol C. These experiments indicated the steric bulky residue at this position is important in controlling cyclization and deprotonation. Also, the native cyclases may have less sterically bulky group at this position in order to form the monocyclic products.[21] (Table 1.1)

Figure 1.18 Conservation pattern between CAS1 and ERG7.[33]

In order to identify the catalytically important residues, Herrera and co-workers

performed a search for similar conservation patterns between CASs and ERG7s. Tyr410, Gly488, Phe717, and Met731 were strictly conserved within CASs, but only Tyr410Thr mutation altered catalysis, and produced lanosterol, parkeol, and 9β-lanosta-7,24-dien-3β-ol at the ratio of 65:2:33 instead of cycloartenol. Tyr410Thr abolishes cycloartenol formation that implicates Tyr410 is essential for cyclopropyl ring formation. Because Tyr410 mutant produced not only lanosterol but also the other by-products, thereby additional mutations are necessary to abolish parkeol and 9β-lanosta-7,24-dien-3β-ol formation for converting cycloartenol synthase into accurate oxidosqualene-lanosterol synthase. The previous study showed Ile481Val also induce the formation of lanosterol, therefore Tyr410Thr Ile481Val was constructed and the products profile was lanosterol, parkeol, and 9β-lanosta-7,24-dien-3β-ol at the ratio of 75: 0.6: 24. The double mutants yielded lanosterol more efficiently than single mutant. However, the Tyr410Thr and Ile481Val mutation have synergistic effects.[21] (Table 1.1)

The other research indicated that the second-tired residue His477 in the putative active site of *Ath*CAS1 which hydrogen bonding with Tyr410. His477 is considered to pull Tyr410 out of the active site and allow reorientation of the intermediate cation to form other products.[21,33] Corresponds to His477 within *Ath*CAS1, the residues within ERG7s is Cys or Gln. The His477Gln mutant is the most accurate parkeol synthase because the ratio of parkeol is 73%, whereas His477Asn is currently the most accurate lanosterol synthase at the ratio of 88% that had been yielded by mutating cycloartenol synthase. Due to the result of double mutant Tyr410Thr Ile481Val generated by-products 9β-lanosta-7,24-dien-3β-ol, converting a cycloartenol synthase to an accurate lanosterol synthase would require at least one additional mutant to preclude deprotonation from C-7. Therefore combining the Tyr410Thr Ile481Val double mutants with His477Asn and His477Gln mutation to construct the triple mutants. Both the result of Tyr410Thr Ile481Val His477Asn and Tyr410Thr Ile481Val His477Gln triple mutants have lanosterol, and 9β-lanosta-7,24-dien-3β-ol at the

ratio of 78: 22, and the percentages of parkeol is less than 1%. Compare to the single mutant of His477Gln, the Tyr410Thr Ile481Val mutations may block the hydride shift from C-9 to C-8 and then preclude the opportunity for His477 mutant to promote the formation of parkeol which is derived from the C-9 cation. Furthermore, Table 1.1 showed that the double mutants Ile481Val His477Asn had the maximum yield of lanosterol.[21] (Table 1.1)

AthCAS1 mutant	Cycloartenol	Lanosterol	Parkeol	9β-△7-Lanosterol	Achilleol A	Camelliol C
CAS1[I481]	99	—	1	—	—	—
CAS1[I481L]	83	1	16	—	—	—
CAS1[I481V]	54	25	21	—	—	—
CAS1[I481A]	12	54	15	—	13	6
CAS1[I481G]	17	23	4	—	44	12
CAS1[Y410T]	—	65	2	33	—	—
CAS1[H477N]	—	88	12	—	—	—
CAS1[H477Q]	—	22	73	5	—	—
CAS1[I481V/Y410T]	—	75	<1	24	—	—
CAS1[I481V/H477N/Y410T]	—	78	—	22	—	—
CAS1[I481V/H477Q/Y410T]	—	78	—	22	—	—
CAS1[I481V/H477N]	—	99	1	—	—	—

Table 1.1 Product profile of AthCAS Ile481, Tyr410 and His477 mutants.[21]

Moreover, our group also shows that the combination of random mutation coupled with in vivo selection and provided an effective method of indentifying single residue alteration could lead to the change of product specificity. We found the Y532H mutant could alter the product specificity, a new position that be indentified in the *A. thaliana* CAS active site. Furthermore, A469 and H477 residues were considered that would not in direct contact with the substrate because of their positions are not in the putative active site of CAS. H477Y mutant generated lanosterol as the dominant product and the A469 mutant yielded lanosterol and achilleol A. This is the first example of nonactive site mutations that

alter the product specificity in a triterpene synthase. Thus the A469 and H477 residues may be the nearest neighbor to the putative active site residues and provide the indirect effect on the active site structure.[34]

1.4.2 β-amyrin synthase (βAS)

The remarkable cyclization of 2,3-oxidosqualene into β-amyrin has interested organic chemists for over a half century since the biogenetic isoprene rule was proposed by Eschenmoser and co-workers. β-amyrin synthase cyclizes oxidosqualene into β-amyrin, which initially pre-folds the substrate into a chair-chair-chair conformation, and then forms the 6-6-6-5-fused tetracyclic dammarenyl C-20 cation. In contrast to the biosynthesis of lanosterol or cycloartenol, the expansion of the D-ring occurs, then the electrophilic addition of the baccharenyl cation on to the terminal double bond generates the lupenyl cation with a five-membered E-ring, which then yields the six-membered E-ring oleanyl cation by ring expansion. Finally, the oleanyl cation is produced via a hydride shift and deprotonation at C-12 to yield the 6-6-6-6-6-fused pentacyclic β-amyrin. (Fig. 1.19)
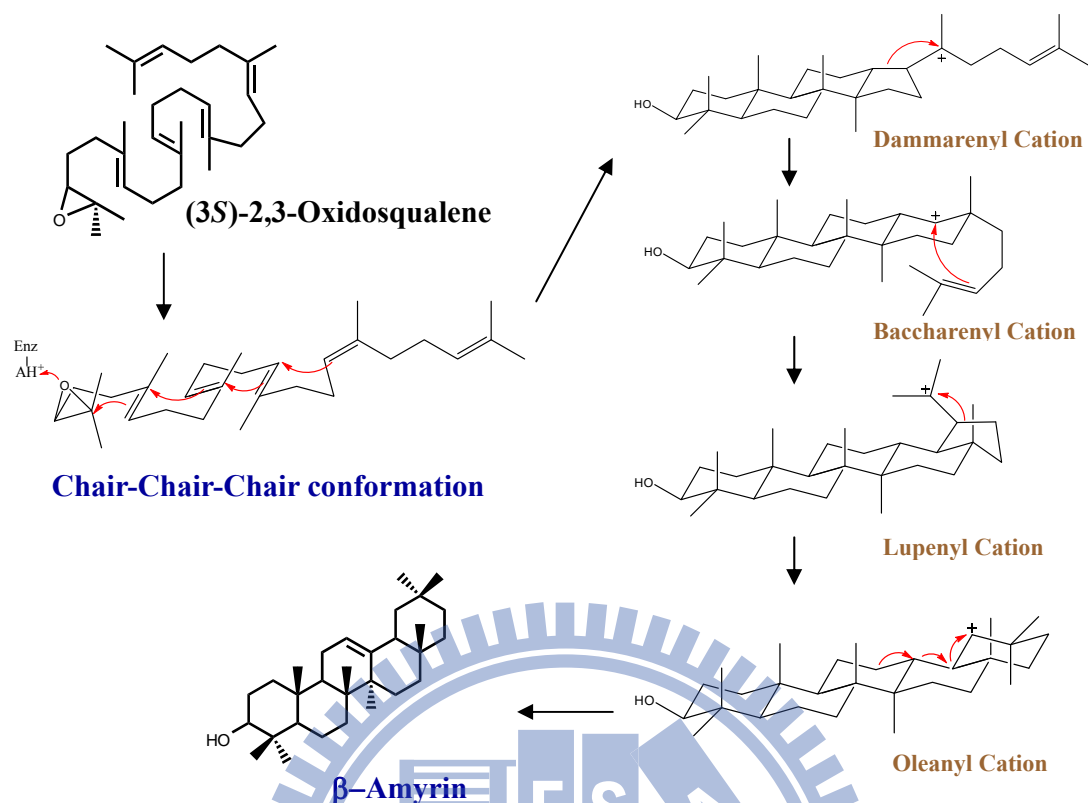
Figure 1.19 The proposed mechanism of 2,3-oxidosqualene into β-amyrin.

β-amyrin synthase (βAS) originating from several plant species, including *Pisum sativum* and *Panax ginseng*, and *Pisum sativum* β-amyrin synthase (PSY), have been cloned and functionally expressed in *Saccharomyces cerevisiae*.[35] To investigate the catalytic motifs within cyclases that form the dammarenyl cation, the Ebizuba group first generated chimeras of the *A. thaliana* lupenol synthase (LUP1) and *Panax ginseng* β-amyrin synthase (PNY). They determined the functions of the regions by using a domain swapping strategy, and only relatively small portions of the protein were found to control the production of either lupenol or β-amyrin. One chimera in which only one fourth of the protein was β-amyrin sequence but made four times as much β-amyrin as it did lupenol, and a mixed PCR method further confirmed the important region of chimeras and narrowed it down to an 80 amino acid span. Later, Ebizuba and co-workers looked within the 80 amino acid region to define specific residues responsible for the product specificity of

25

β-amyrin and lupenol synthases. Results from alignment analysis showed that Trp259 within *Panax ginseng* β-amyrin synthase (PNY), and Leu256 within *Olea europa* lupenol synthase (OEW) might control the product specificity, therefore PNY Trp259Leu and OEW Leu256Trp mutants were contructed. Lupenol occurred twice as much as β-amyrin in the product profile of the PNY Trp259Leu mutant, whereas β-amyrin was the major product in the product profile of the OEW Leu256Trp mutant. These two results demonstrate that this position plays a critical role in directing either β-amyrin or lupenol formation. Furthermore, the PNY Tyr261His mutant was also created, and the product analysis result showed that Tyr261 stabilized one of the cationic intermediates formed after the dammarenyl cation.[21,36,37] In addition, one research study for *Pisum sativum* β-amyrin synthase (PSY) showed the expansion of the D-ring could take place in the absence of the terminal double bond. Thus, the formation of the anti-Markovnikov six-membered D-ring does not depend on the terminal π-electrons. In summary, the aromatic residues within putative active site may play a crucial role for ring expansion reaction.[35]
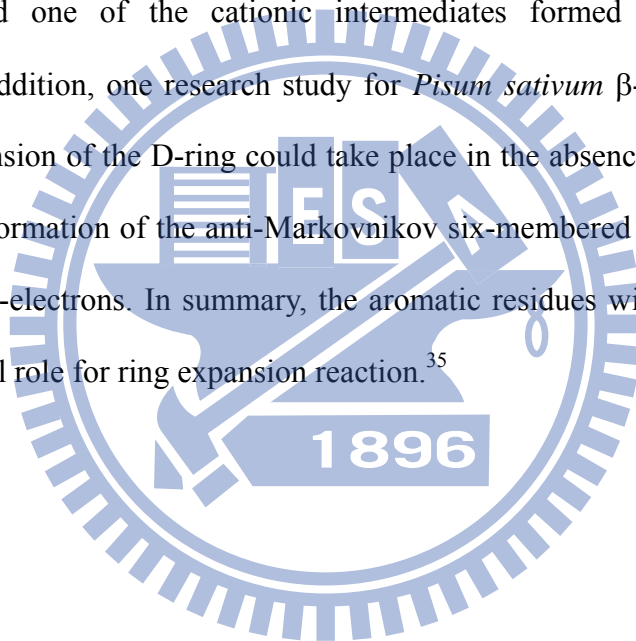
1.5 Squalene-hopene cyclase (SHC)

Squalene-hopene cyclase is a homodimeric enzyme and it is organized in two α-helical domains, which together form a dumbbell-shaped molecule. (Fig. 1.20) It contains 631 amino acids per subunit and the molecular mass are about 71.5kDa. In 1997, the X-ray crystal structure of *Alicyclobacillus acidocaldarius* SHC was first reported at 2.9 Å resolution, later refined to 2.0 Å resolution in 1999.[38,39] Moreover, Reinert *et al* reported another X-ray crystal structure which the squalene-hopene cyclase was cocrystallized with 2-azasqualene and its resolution is 2.13 Å in 2004.[40] These structures, which combine with the biological studies, provide more mechanistic insight into squalene-hopene cyclases and oxidosqualene cyclases.
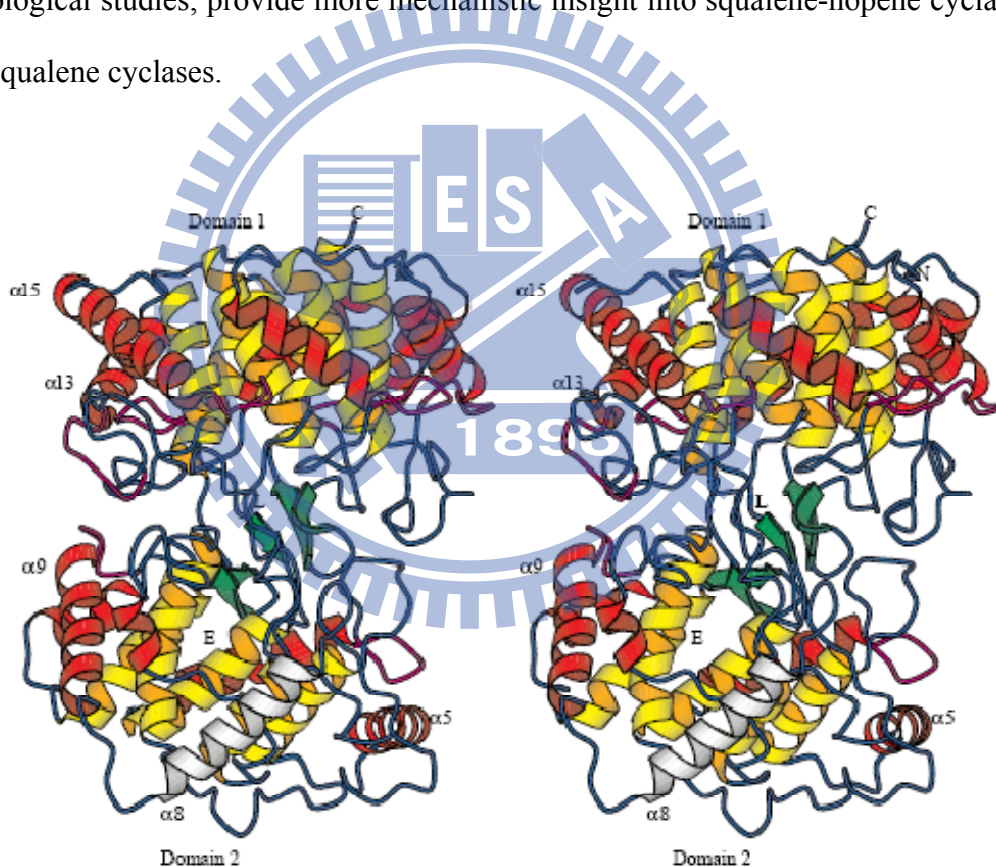
Figure 1.20 Crystal structure of *A. acidocaldarius* squalene-hopene cyclase (SHC).[38]

Different from the cyclization of oxidosqualene, squalene-hopene cyclase pre-fold squalene into an all-chair conformation, and the lower basicity of the double bond relative to the oxirane ring suggests that SHC provide a catalytic acid at least as strong as that used

to initiate the reaction of oxidosqualene. Thus, this property induce the bacterial squalene cyclase to accept oxidosqualene and its analogues as substrates for cyclization. The cyclization process of SHC is simpler than oxidosqualene cyclase, and then convert squalene into hopanyl C-22 cation with 6-6-6-6-5-fused pentacyclic ring. Without the rearrangement step, deprotonation of hopanyl C-22 cation yield hopene or addition of water to produce hopanol. (Fig. 1.21)
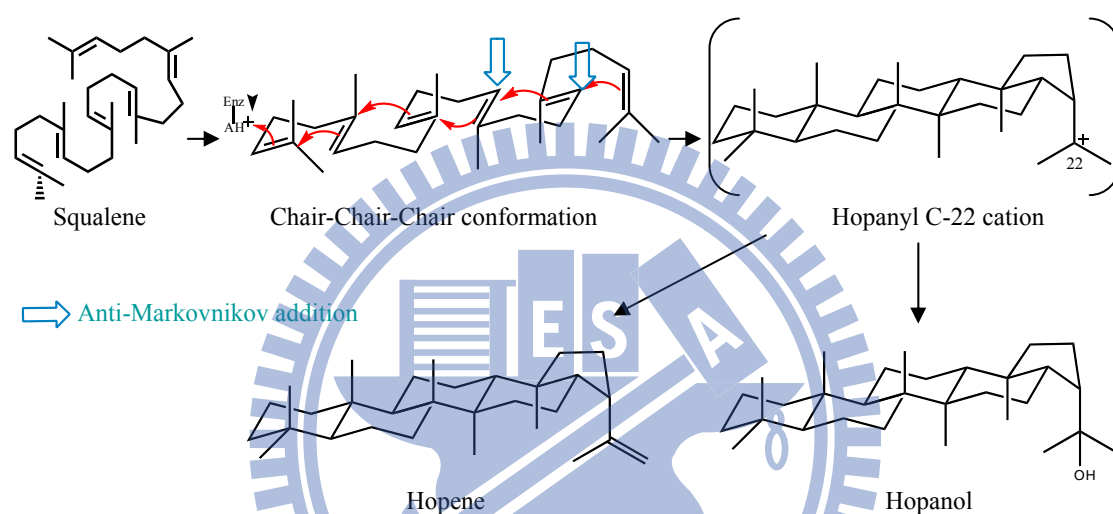
Figure 1.21 The cyclization process of squalene-hopene cyclase (SHC).

In the site-directed mutagenesis studies, the catalytic acid Asp376 residue (which corresponds to Asp456 within *Sce*ERG7) hydrogen-bonded with His451 to initiate the reaction by protonation, and then Asp376 is likely to be stabilized by His451 residue that in turn is electrostatically stabilized by the solvent accessible Glu454.[35,40] Following protonation, the process of cyclization undergo small conformational change during A to D-ring formation. Several cation-stabilizing residues have been subjected to mutational studies to evaluate the role in different stage of squalene cyclization. Substitution of Asp377 with Cys or Asn result in the formation of monocyclic product, and it suggested that Asp377 may stabilize C-10 cation, and then the other mutation of Tyr420Ala yielded

two bicyclic products and one tricyclic compound, it was the evidence of truncating the C-8 B-ring cation reaction to form the bicyclic products, indicating Tyr420 may stabilized C-8 cation intermediate within the putative active site of SHC. In addition, the Phe365Ala and Tyr609Phe mutants also induced the formation of bicyclic products, indicated that Phe365 and Tyr609 residues could stabilized C-8 cation intermediate as well as Tyr420. Furthermore, Phe601, Phe605 and Trp169 are considered to stabilize the concerted C/D ring formation and to generate the 6-6-6-5-membered tetracyclic cation intermediate by cation-π interactions. Also, the Phe605 residue is considered as a key amino acid to form the E-ring. Finally, a polarized water molecule abstracts the proton or attacks the E-ring cation, in order to terminate the reaction, and Glu45 residue has been proposed that it polarized the water molecule by its hydrogen-bond network.[4,35]
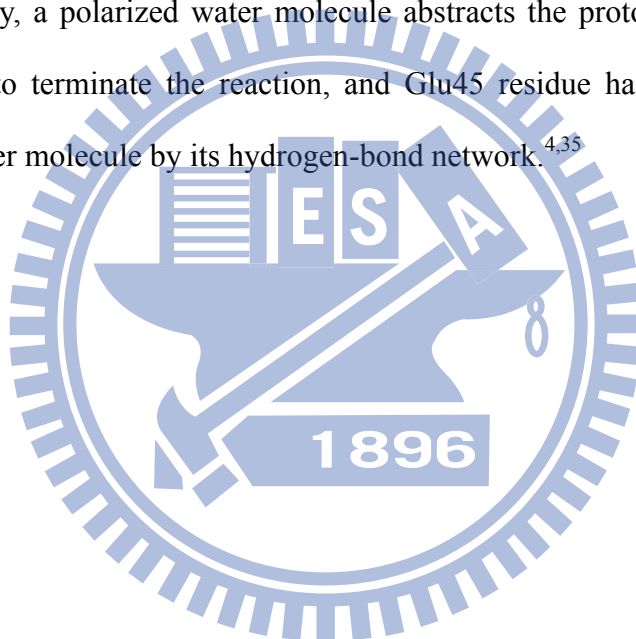
1.6 The amino acid sequence alignment of (oxido-)squalene cyclases

The amino acid sequence alignments of the enzyme cyclases could provide much information and led us understand these cyclases in depth. In the evolutionary history, some highly-conserved residues indicated that they may play some specific function and it is necessary for the organisms, and therefore it would be retained in different species, whereas some residues are not conserved or even be deleted through the variation of many species. These differences could help us figure out the cyclized mechanism of enzyme cyclases. Such as the *Sce*ERG7 showed about 40 and 37% sequence indentity to human OSC (*H. sapiens* ERG7) and *Arabidopsis thaliana* CAS (*Ath*CAS1) respectively; their similar structures, stereoselectivity and the catalytic mechanism could be investigated and deeply realize the relationship between their function and structure.[41]

In order to understand the functional residues within the putative active site of oxidosqualene cyclase, we used the program of Clustal W to produce multiple sequence alignment of the following enzymes: *H. sapiens* ERG7: P48449, *S. cerevisiae* ERG7: P38604, *A. thaliana* CAS: NP_178722, *A. acidocaldarius* SHC: BAA25185, *P. sativum* βAS: BAA97558. All of them were obtained from the Protein Data Bank (PDB) in NCBI, and the result of sequence alignment was shown below. (Fig. 1.22)

```
H.sapiens ERG7        MTEGTCLRRRGGPYKTEPATDLG--RWRLN-CERGRQTWTYLQDER--------AGREQT 49
S.cerevisiae ERG7     MTEFYSDTIG------LPKTDPR--LWRLRTDELGRESWEYLTPQQ--------AANDPP 44
A.thaliana CAS        MWKLKIAEGGS-PWLRTTNNHVGRQFWEFDPNLGTPEDLAAVEEARKSFSDNRFVQKHSA 59
P.sativum bAS         MWRLKIAEGGNDPYLFSTNNFVGRQTWEYDPEAGSEEERAQVEEARRNFYNNRFEVKPCG 60
A.acidocaldarius SHC  ----------------------------------------------------------
```

```
H.sapiens ERG7        GLEAYALGLDTKNYFKDLPKAH------TAFEGALN----GMTFYVGLQAED-GHWTGDY  98
S.cerevisiae ERG7     STFTQWLLQDPK-FPQPHPERNKHSPDFSAFDACHN----GASFFKLLQEPDSGIFPCQY  99
A.thaliana CAS        DLLMRLQFSRENLISPVLPQVKIEDTDDVTEEMVETTLKRGLDFYSTIQAHD-GHWPGDY  118
P.sativum bAS         DLLWRFQVLRENNFKQTIGGVKIEDEEEITYEKTTTTLRRGTHHLATLQTSD-GHWPAQI  119
A.acidocaldarius SHC  ---MAEQLVEAPAYARTLDRAV--------------------EYLLSCQKDE-GYWWGPL  36


H.sapiens ERG7        GGPLFLLPGLLITCHVAR---IPLPAGYREEIVRYLRSVQLP-DGGWGLHIEDKSTVFGT  154
S.cerevisiae ERG7     KGPMFMTIGYVAVNYIAG---IEIPEHERIELIRYIVNTAHPVDGGWGLHSVDKSTVFGT  156
A.thaliana CAS        GGPMFLLPGLIITLSITGALNTVLSEQHKQEMRRYLYNHQNE-DGGWGLHIEGPSTMFGS  177
P.sativum bAS         AGPLFFMPPLVFCVYITGHLDSVFPPEHRKEILRYIYCHQNE-DGGWGLHIEGHSTMFCT  178
A.acidocaldarius SHC  LSNVTMEAEYVLLCHILDR----VDRDRMEKIRRYLLHEQRE-DGTWALYPGGPPDLDTT  91


H.sapiens ERG7        ALNYVSLRILGVGPDDP---DLVRARNILHKKGGAVAIPSWGKFWLAVLNVYSWEGLNTL  211
S.cerevisiae ERG7     VLNYVILRLLGLPKDHP---VCAKARSTLLRLGGAIGSPHWGKIWLSALNLYKWEGVNPA  213
A.thaliana CAS        VLNYVTLRLLGEGPNDG-DGDMEKGRDWILNHGGATNITSWGKMWLSVLGAFEWSGNNPL  236
P.sativum bAS         ALNYICMRILGEGPDGGEDNACVRARNWIRQHGGVTHIPSWGKTWLSILGVFDWLGSNPM  238
A.acidocaldarius SHC  IEAYVALKYIGMSRDEE---PMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMV  148


H.sapiens ERG7        FPEMWLFPDWAPAHPSTLWCHCRQVYLPMSYCYAVRLSAAEDPLVQSLRQELYVEDFASI  271
S.cerevisiae ERG7     PPETWLLPYSLPMHPGRWWVHTRGVYIPVSYLSLVKFSCPMTPLLEELRNEIYTKPFDKI  273
A.thaliana CAS        PPEIWLLPYFLPIHPGRMWCHCRMVYLPMSYLYGKRFVGPITSTVLSLRKELFTVPYHEV  296
P.sativum bAS         PPEFWILPSFLPMHPAKMWCYCRLVYMPMSYLYGKRFVGPITPLILQLREELHTEPYEKI  298
A.acidocaldarius SHC  PPEIMFLGKRMPLNIYEFGSWARATVVALSIVMSRQPVFPLPERARVP--ELYETDVPPR  206


H.sapiens ERG7        DWLAQRNNVAPDELYTPHSWLLRVVYALLNLYEHHHS------AHLRQRAVQKLYEHIVA  325
S.cerevisiae ERG7     NFSKNRNTVCGVDLYYPHSTTLNIANSLVVFYEKYLRN-----RFIYSLSKKKVYDLIKT  328
A.thaliana CAS        NWNEARNLCAKEDLYYPHPLVQDILWASLHKIVEPVLMRWPG-ANLREKAIRTAIEHIHY  355
P.sativum bAS         NWTKTRHLCAKEDIYYPHPLIQDLIWDSLYIFTEPLLTRWPFNKLVRKRALEVTMKHIHY  358
A.acidocaldarius SHC  RRGAKGG----------GGWIFDALDRALHGYQKLSVHP------FRRAAEIRALDWLLE  250
```

```
H.sapiens ERG7          DDRFTKSISIGPISKTINMLVRWYVDGPASTAFQEHVSRIPDYLWMGLDGMKMQGTNGSQ 385
S.cerevisiae ERG7       ELQNTDSLCIAPVNQAFCALVTLIEEGVDSEAFQRLQYRFKDALFHGPQGMTIMGTNGVQ 388
A.thaliana CAS          EDENTRYICIGPVNKVLNMLCCWVED-PNSEAFKLHLPRIHDFLWLAEDGMKMQGYNGSQ 414
P.sativum bAS           EDENSRYLTIGCVEKVLCMLACWVED-PNGDAFKKHIARVPDYLWISEDGMTMQSF-GSQ 416
A.acidocaldarius SHC    RQAGDGSWGGIQPPWFYALIALKILDMTQHPAFIKGWEGLELYGVELDYGGWMFQASISP 310


H.sapiens ERG7          IWDTAFAIQALLEAGGHHRPEFSSCLQKAHEFLRLSQVP-DNPPDYQKYYRQMRKGGFSF 444
S.cerevisiae ERG7       TWDCAFAIQYFFVAGLAERPEFYNTIVSAYKFLCHAQF---DTECVPGSYRDKRKGAWGF 445
A.thaliana CAS          LWDTGFAIQAILATNLVE--EYGPVLEKAHSFVKNSQVLEDCPGDLNYWYRHISKGAWPF 472
P.sativum bAS           EWDAGFAVQALLATNLIE--EIKPALAKGHDFIKKSQVTENPSGDFKSMHRHISKGSWTF 474
A.acidocaldarius SHC    VWDTGLAVLALRAAGLPAD---HDRLVKAGEWLLDRQIT--VPGDWAVKRPNLKPGGFAF 365


H.sapiens ERG7          STLDCGWIVSDCTAEALKAVLLLQEK--CPHVTEHIPRERLCDAVAVLLNMRNPD----G 498
S.cerevisiae ERG7       STKTQGYTVADCTAEAIKAIIMVKNSPVFSEVHHMISSERLFEGIDVLLNLQNIGSFEYG 505
A.thaliana CAS          STADHGWPISDCTAEGLKAALLLSKVP-KAIVGEPIDAKRLYEAVNVIISLQNAD----G 527
P.sativum bAS           SDQDHGWQVSDCTAEGLKCCLLLSLLP-PEIVGEKMEPERLFDSVNLLLSLQSKK----G 529
A.acidocaldarius SHC    QFDNVYYPDVDDTAVVVWALNTLRLPD------ERRRRDAMTKGFRWIVGMQSSN----G 415


H.sapiens ERG7          GFATYETKRGGHLLELLNPSEVFGDIMIDYTYVECTSAVMQALKYFHKRFPEHRAAEIRE 558
S.cerevisiae ERG7       SFATYEKIKAPLAMETLNPAEVFGNIMVEYPYVECTDSSVLGLTYFHKYF-DYRKEEIRT 564
A.thaliana CAS          GLATYELTRSYPWLELINPAETFGDIVIDYPYVECTSAAIQALISFRKLYPGHRKKEVDE 587
P.sativum bAS           GLAAWEPAGAQEWLELLNPTEFFADIVVEHEYVECTGSAIQALVLFKKLYPGHRKKEIEN 589
A.acidocaldarius SHC    GWGAYDVDNTSDLPNHIPFCDFG--EVTDPPSEDVTAHVLECFG-------SFGYDDAWK 466


H.sapiens ERG7          TLTQGLEFCRRQQRADGSWEGSWGVCFTYGTWFGLEAFACMGQTYRDGTACAEVSRACDF 618
S.cerevisiae ERG7       RIRIAIEFIKKSQLPDGSWYGSWGICFTYAGMFALEALHTVGETYEN---SSTVRKGCDF 621
A.thaliana CAS          CIEKAVKFIESIQAADGSWYGSWAVCFTYGTWFGVKGLVAVGKTLKN---SPHVAKACEF 644
P.sativum bAS           FIFNAVRFLEDTQTEDGSWYGNWGVCFTYGSWFALGGLAAAGKTYTN---CAAIRKGVKF 646
A.acidocaldarius SHC    VIRRAVEYLKREQKPDGSWFGRWGVNYLYGTGAVVSALKAVGIDTREP----YIQKALDW 522
```

32

```
H.sapiens ERG7          LLSRQMADGGWGEDFESCEERRYLQSA--QSQIHNTCWAMMGLMAVRHPDIE--AQERGV  674
S.cerevisiae ERG7       LVSKQMKDGGWGESMKSSELHSYVDSE--KSLVVQTAWALIALLFAEYPNKE--VIDRGI  677
A.thaliana CAS          LLSKQQPSGGWGESYLSCQDKVYSNLDGNRSHVVNTAWAMLALIGAGQAEVDRKPLHRAA  704
P.sativum bAS           LLTTQREDGGWGESYLSSPKKIYVPLEGNRSNVVHTAWALMGLIHAGQSERDPTPLHRAA  706
A.acidocaldarius SHC    VEQHQNPDGGWGEDCRSYEDPAYAGKG--ASTPSQTAWALMALIAGGRAESE--AARRGV  578
```

```
H.sapiens ERG7          RCLLEKQLPNGDWPQENIAG-VFNKSCAISYTSYRNIFPIWALGRFSQLYPERALAGHP  732
S.cerevisiae ERG7       DLLKNRQEESGEWKFESVEG-VFNHSCAIEYPSYRFLFPIKALGMYSRAYETHTL----  731
A.thaliana CAS          RYLINAQMENGDFPQQEIMG-VFNRNCMITYAAYRNIFPIWALGEYRCQVLLQQGE---  759
P.sativum bAS           KLLINSQLEQGDWPQQEITG-VFMKNCMLHYPMYRDIYPLWALAEYRRRVPLP------  758
A.acidocaldarius SHC    QYLVETQRPDGGWDEPYYTGTGFPGDFYLGYTMYRHVFPTLALGRYKQAIERR------  631
```

Figure 1.22 Sequence alignment of *H. sapiens* ERG7, *S. cerevisiae* ERG7, *A. thaliana* CAS, *P. sativum* βAS and *A. acidocaldarius* SHC. The black markers mean the Ile705 and Leu734 residues within *S. cerevisiae* ERG7 and *P. sativum* βAS respectively.
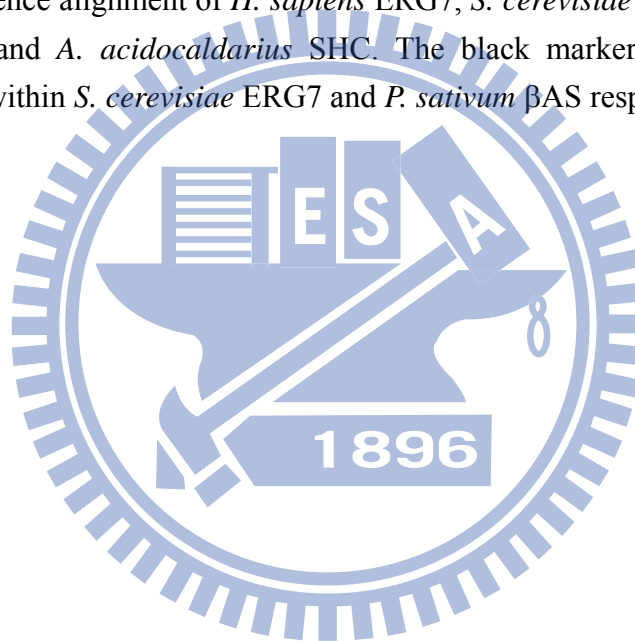
1.7 Research motive

The diversity of triterpene products, which were generated by oxidosqualene-lanosterol cyclase have fascinated scientists. Oxidosqualene-lanosterol cyclases convert (3S)-2,3-oxidosqualene into lanosterol in animals and fungi. The high stereoselectivity, species specificity, and reaction complexity led us to investigate the multiple enzyme functions or to treat as a tool for chemical synthesis. In addition, the crystal structures of *A. acidocaldarius* SHC and human OSC, and the homology model of *S. cerevisiae* ERG7 provide more insight into the structural basis for observed altered product specificity utilizing the mutants.[41]

In the sequence alignment of *Sce*ERG7, there is a highly-conserved residue, Ile705, which has hydrophobic properties. According to previous published research, the role of aromatic residues is implicated in stabilizing the high-energy cation intermediate. Different from the π-electron-rich amino acids including Tyr99, Trp232, His234, Trp390, Trp443, Phe445, Tyr510, Phe699, and Tyr707 with occur in the vicinity of the putative active site, the I705 which is hydrophobic, is a special amino acid different from the other aromatic acids. The minimum distance from Ile705 to the lanosterol molecule within the *Sce*ERG7 modeling structure is 5.61 Å. It effectively means that Ile705 is the second-tired residue of *Sce*ERG7. Although Ile705 is not in close proximity to the substrate, there are some examples where second-tired residues are also important in the active site. The His477 (a second-sphere residue) within *Ath*CAS1 is the best example, where hydrogen is bonded to Tyr410, thus it is essential for cycloartenol biosynthesis.[33] In a previous study, abnormal bicyclic products were isolated from the Leu607Lys (which corresponds to Ile705 in *S. cerevisiae*) mutation of a squalene-hopene cyclase of *Alicyclobacillus acidocaldarius*.[42] Furthermore, in 2002, Tsutomu Sato showed the steric bulk size at position 607 was a key factor during the polycyclization process. Folding squalene into a boat conformation has not been reported before.[43] For these reasons, the bulky hydrophobic residue Ile705

probably has some interaction with the substrate and may be involved in influencing the first-tired residues.

It is interesting that the B-chair/boat ring and the 17α/17β mechanistic transitions have occurred rarely in evolutionary history.[44] In 2003, Matsuda's group published a review paper, collected nearly 200 triterpene compounds and classified them into many groups. The chair-boat-chair 17α-derivatives were not discovered, while the chair-chair-chair conformation has both α and β configurations at C-17, and then some products were proposed to form via either 17α or 17β dammarenyl cation, but these were not classified specifically.[3] Subsequently, they proved that formation of the 17β-dammarenyl cation occurred in LUP1-catalyzed biosynthesis, not the 17α configuration. Although the C-17 configuration may be lost during D-ring expansion, the configurational information could be gained in C-20. They found a rule that the 17β-dammarenyl cation would always migrate from C-16 to C-20 via a five-membered D-ring expansion step, whereas the C-13 migration would only occur for the 17α-dammarenyl cation. Nearly all of the common higher plant species would form 17β tetracyclic intermediates in their catalytic processes, while 17α dammarenyl cation intermediates appeared only in bacteria or rare plant species. The multiple products of plant species include the B-chair ring or the 17α/17β skeleton, whereas the ERG7s usually yield the B-boat ring and the 17β structural compounds. For understanding the evolutionary divergence of cyclases from different species, especially the stereochemistry of enzymatic control, we choose to study Leu734 (which corresponds to Ile705 in SceERG7) in Pisum sativum β-amyrin synthase (PSY) through different conformational intermediates within oxidosqualene cyclization. By a series of site-saturated mutations on Ile705 and Leu734 residues, the structure-function relationships between these conserved amino acids in different species could be addressed in depth.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Chemicals and reagents

Acetic acid (Merck)

Anisaldehyde (Mecrk)

Acetone (Merck)

95% Alcohol (Merck)

Acetic anhydride (Sigma)

Ampicillin sulfate (Sigma)

Adenine (Sigma)

Agarose-LE (USB)

Bacto^{TM}Agar (DIFCO)

Dichloromethane (Merck)

D-Sorbitol (Sigma)

Dimethyl sulfoxide (MP Biomedicals)

DNA 10Kb Ladder (Bio Basic Inc., Tanwan)

Ethyl acetate (Merck)

Ether (Merck)

Ergosterol (Sigma)

Glycerol (Merck)

Glucose (Sigma)

G418 (Gibco)

Hemin Chloride (Merck)

Hexane (Merck)

Histidine (Sigma)
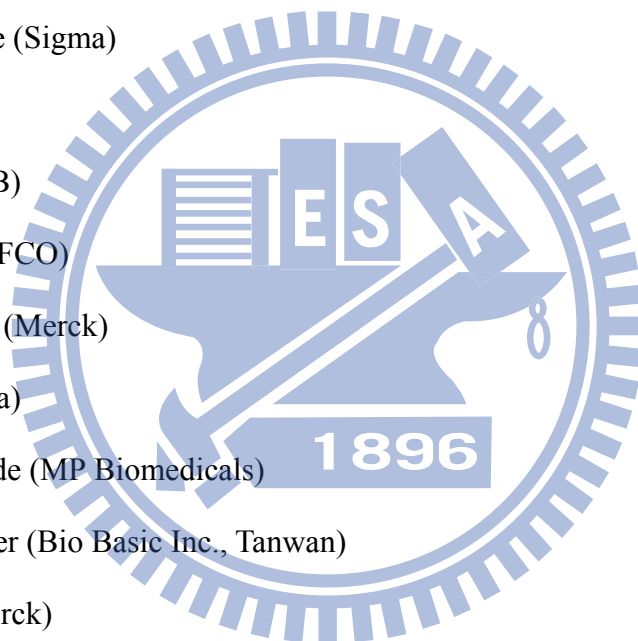
Lysine (Sigma)

LB Broth, Miller (DIFCO)

Methanol (Merck)

Methioine (Sigma)

dNTP Set, 100mM Solutions (GE Healthcare)

Potassium hydroxide (Merck)

Pyrogallol (Merck)

Pyridine (Sigma)

Primers (Bio Basic Inc.)

Restriction enzyme (New England BioLabs Inc.)

Silica gel (Merck)

Silver nitrate (Merck)

Sodium sulfate (Merck)

Sulfonic Acid (Merck)

SYBR$^®$ Green I (Roche)

TLC plate (Mecrk)

Trptophan (Sigma)

Yeast Extra (DIFCO)

Yeast Nitrogen Base w/o amino acid (DIFCO)

Uracil (Sigma)
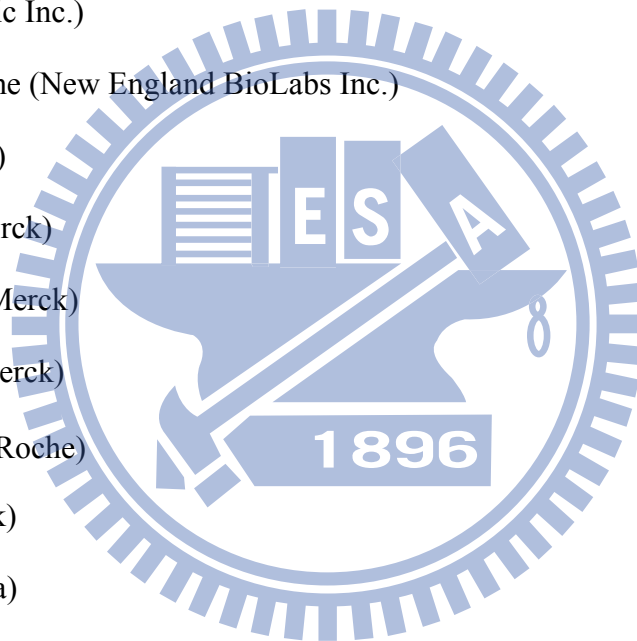
2.1.2 Kits

BigDye$^®$ Teminator v3.1 Cycle Sequencing Kit (Applied Biosystems)

GFX^{TM} PCR DNA and Gel Band Purification Kit (GE Healthcare)

Plasmid Miniprep Purification Kit (GeneMark)

QuikChange Site-Directed Mutagenesis Kit (Stratagene Inc., La Jolla, CA)

2.1.3 Bacterial, yeast strains and vectors

Escherichia coli XL-Blue (Novagen)

CBY57 (a yeast stain, MATa or MATα ERG7Δ:: LEU2 *ade*2-101 *his*3-Δ200 *leu2*-Δ1 *lys*2-801 *trp1*-Δ63 *ura3*-52 [pZS11])

TKW14C2 (a yeast stain, MATa or MATα ERG7Δ:: LEU2 *ade*2-101 *his*3-Δ200 *leu2*-Δ1 *lys*2-801 *trp1*-Δ63 *ura3*-52 *hem*1Δ::Kan^R)

Vector pRS314 (a shuttle vector with selection marker *Trp1*, New England BioLabs)

Vector pYES2 (a shuttle vector with selection marker *Ura3*, New England BioLabs)

2.1.4 Equipments

ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems)

Allegra™ 21R Centrifuge (Beckman Coulter)

Avanti® J0E Centrifuge (Beckman Coulter)

Colling Circulator Bath Model B401L (Firstek Scientific)

Centrifuges 5415R (eppendorf)

DU 7500 UV-Vis Spectrophotometer (Beckman Coulter)

Durabath™ Water Bath (Baxter)

Electrophoresis Power Supply EPS 301 (GE healthcare)

GeneAmp® PCR System 9700 Thermal Cycler (Applied Biosystems)

Hoefer® HE 33 Mini Horizontal Submarine Unit (GE Healthcare)

Kodak Electrophoresis Documentation an Analysis System 120 (Kodak)

Orbital Shaking Incubator Model-S302R (Firstek Scintific)

Pulse Controller (BioRad)

Rotary Vaccum Evaporator N-N Series (EYELA)

Steritop^{TM} 0.22μm Filter Unit (Millipore)

2.1.5 Solutions

Ampicillin stock solution (100mg/mL)

Dissolve 1g ampicillin sulfate in 10 mL ddH$_2$O. Filter through 0.22μm pore size filter and stock at -20℃.

50X TAE buffer

Dissolve Tris base 242 g, acetic acid 57.1 ml, and 0.5 M EDTA in 1 L dH2O and adjust to pH 8.5. Store it at room temperature. Dilute to 1X with dH2O and adjust pH to 7.5~7.8 before use.

50X ALTHMU solution

1g Adeine, 1.5g Lysine, 1g Tryptophan,, 1g Hisidine, 1g Methonine, 1g Uracil was dissolved in 500mL dH$_2$O and sterilized. Store at 4℃.

50X ALHMU solution

1g Adeine, 1.5g Lysine, 1g Hisidine, 1g Methonine, 1g Uracil was dissolved in 500mL dH$_2$O and sterilized. Stock at 4℃.

50X ALTH solution

1g Adeine, 1.5g Lysine, 1g Tryptophan,, 1g Hisidine, was dissolved in 500mL dH$_2$O and

sterilized. Store at 4℃.

50X ALTHU solution

1g Adeine, 1.5g Lysine, 1g Tryptophan,, 1g Hisidine, 1g Uracil was dissolved in 500mL dH$_2$O and sterilized. Store at 4℃.

50% Glucose solution

500g glucose was dissolved in 1 L dH$_2$O and sterilized.

80% Glycerol solution

80 mL glycerol was added in 20 mL dH$_2$O and sterilized. Stock at 4℃.

LB medium

25g LB Broth was dissolved in 1 L dH$_2$O and sterilized.

LB plate

25 g LB Broth and 20 g Bacto^{TM} Agar was dissolved in 1 L dH2O and sterilized. The sterile LB agar was poured and dispersed in Petri dishes before it coagulates.

G418 stock solution (1g/mL)

Dissolve 500 mg G418 in 500 μl sterile dH2O. Store it in darkness at 4 ℃.

SD medium

6.7g Yeast nitrogen base was dissolved in 1 L dH$_2$O and sterilized.

20% EA developing solution
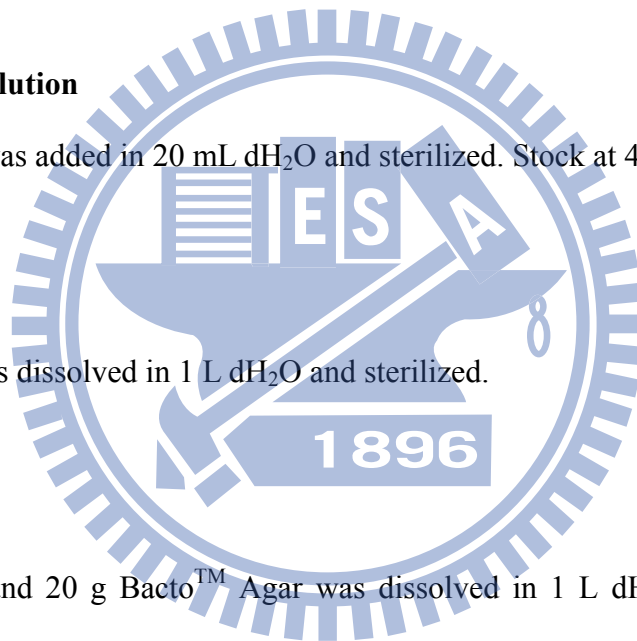
Add 20 mL ethyl acetate to 80 mL hexane and mix it.

TLC staining solution

40mL of conc. H_2SO_4 is added (slowly!) into ethanol 800mL, followed by acetic acid 12mL and anisaldehyde 16mL.

1M sorbitol solution

172.2g D-sorbitol was dissolved in 500 mL dH_2O and sterilized. Store at $4^\circ C$.

5X sequencing buffer

Dissolve 4.85g Tris base and 0.203g $MgCl_2$ in 100 mL dH_2O and adjust to pH 9. Store at $4^\circ C$.

10X SYBR Green solution

10000X SYBR Green was diluted to 10X with DMSO. Stock in darkness.

6X DNA loading dye

0.25% bromophenol blue and 30% glycerol were dissolved in ddH2O. Store at -20 $^\circ C$.

Hemin solution

0.5 g heme chloride was dissolved in 250 ml 0.2 N potassium hydroxide and thus mixes it with 250 ml 95% alcohol in aseptic condition. Store it at room temperature in darkness.

Ergosterol supplement solution

1 g Ergosterol was dissolved in 250 ml 95% alcohol and thus mixes with 250 ml Tween 80 in aseptic condition. Store it in darkness at room temperature.

ALHMU/Hemin/Ergosterol plate

0.67 g yeast nitrogen base, 2 g Bacto[TM] Agar was dissolved in 100ml dH$_2$O and sterilized. Add 2 ml 50X ALHMU solution, 4 ml 50% glucose solution, 2 ml hemin solution, 2 ml ergosterol supplement solution, and 100 μl G148 stock solution into the sterile SD medium. Then the mixture was poured and dispersed in Petri dishes before it coagulates. All of steps are in aseptic condition and stock in darkness at 4 ℃.

2.2 Methods

2.2.1 The construction of recombinant plasmids

The mutation of ERG7^{Ile705X} and PSY^{Leu734X} were performed using the QuikChange site-directed mutagenesis kit. The strategy is shown in Figure 2.1.
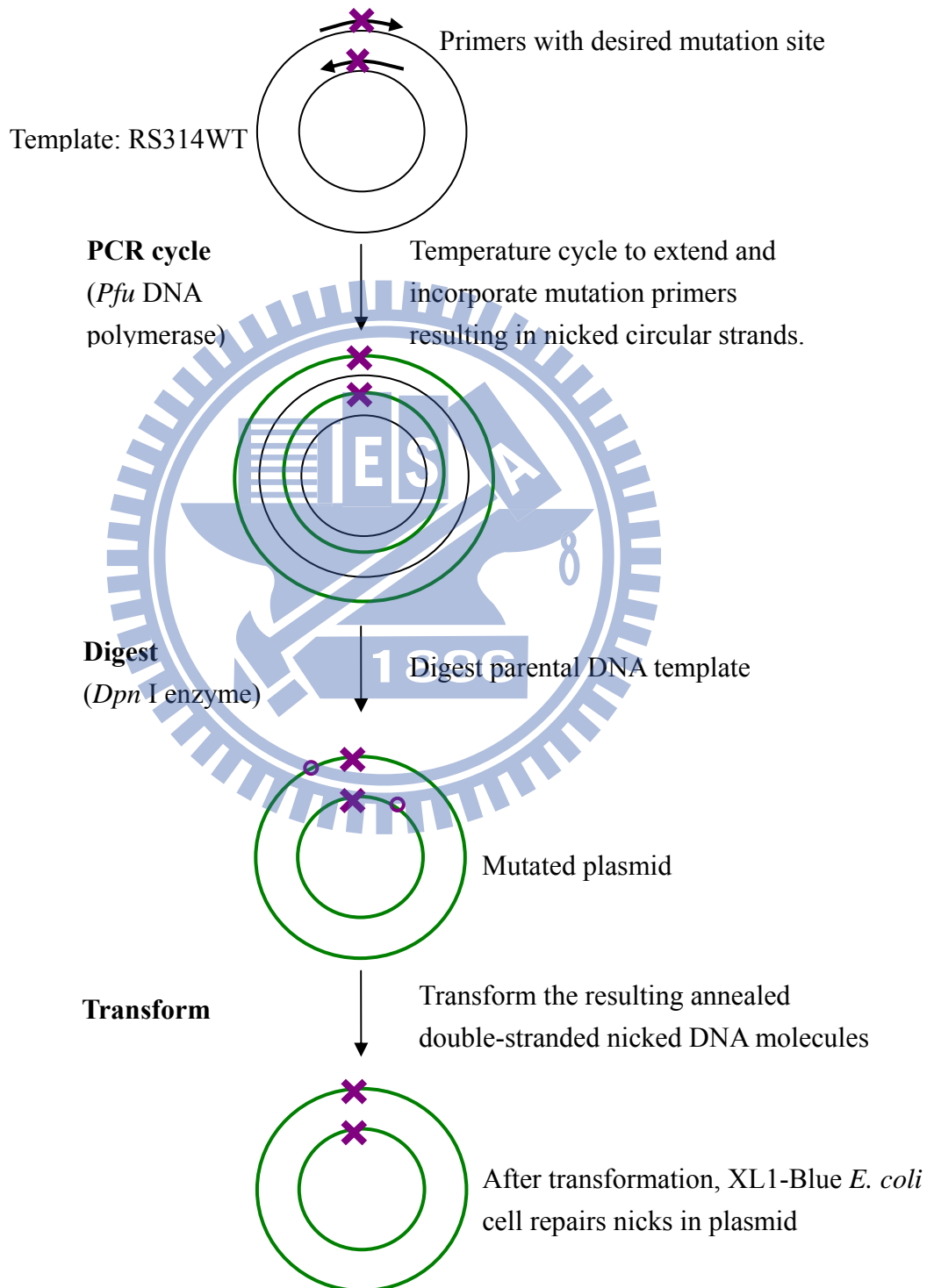
Figure 2.1 QuikChange site-directed mutagenesis strategies

(1) Primer design:

For the ERG7^{Ile705X} mutations:

YCC-YOSC-I705X-Ava I-1

5'- CAC TCT TgT gCA NNN gAA TA<u>C CCg Ag</u>T TAT CgA TTC- 3'

YCC-YOSC-I705X-Ava I-2

5'- gAA TCg ATA A<u>CT Cgg g</u>TA TTC NNN TgC ACA AgA gTg - 3'

For the PSY^{Leu734X} mutations:

YCC-PSY-L734X-Nru I-1

5'- gAA AAA TTg CAT gNN NCA CTA TCC AAT gTA <u>TCg Cg</u>A TAT ATA TC- 3'

YCC-PSY-L734X-Nru I-2

5'- gAT ATA TA<u>T CgC g</u>AT ACA TTg gAT AgT gNN NCA TgC AAT TTT TC - 3'

The gray background letters in the sequence line of primers show the target mutants, and "N" means A, T, C, G four bases, thereby 20 possible amino acids. The bold letter indicates silent mutation for *Ava* I or *Nru* I mapping analysis and construction are marked with underline. In addition, the recombination plasmids of ERG7^{F699X} have been constructed previously in our laboratory.

(2) QuikChange PCR:

Reagent	Volume (μl)
Primer1 (1000 ng/μl)	0.5
Primer2 (1000 ng/μl)	0.5
Template	0.5
dNTP (10 mM)	1.6
Pfu II buffer	2
Pfu II polymerase	0.4
DDW	14.5
Total	20

Table 2.1 QuikChange Site-Directed Mutagenesis Kit PCR composition

Segment	Cycles	Temperature (℃)	Time
1	1	95	2 min
2	18	95	30 sec
		50	1min
		68	16min
3	1	4	∞

Table 2.2 QuikChange Site-Directed Mutagenesis PCR program

(3) *Dpn* I digest parental DNA template:

The digested reaction was incubated at 37℃ for 3 hours to digest the parental supercoiled DNA.

Reagent	Volume (μl)
PCR products	8
10× NE Buffer 4	1
Dpn I	1

Table 2.3 QuikChange Site-Directed Mutagenesis PCR products diegestion

(4) Transformation into XL1-Blue and enzyme digestion:

The digested QuikChange products were added into 100 μl *E. coli* XL1-Blue competent cells of each reaction and incubated on ice for 20 min. The cells were transformed by heat shock methods for 1 min at 42 ℃, following 1 min on ice. Then the cells were transferred into 1 ml Luria-Bertani (LB) medium immediately and then shaken in 200 rpm for 1 hour at 37 ℃ incubator. Then, the cells were centrifuged at 8,000 rpm for 1 min and propagated on LB plate containing 100 μg/ml ampicillin (LBamp). Incubate these plates 16 hr at 37 ℃. Pick the colonies and culture in 3 ml LB medium containing 100 μg/ml ampicillin overnight at 37 ℃. The plasmid DNAs

were isolated by Plasmid Miniprep Purification Kit, according to the manufacturer instructions. The plasmid DNAs were then digested with *Ava* I for ERG7^{Ile705X} mutants and *Nru* I for PSY^{Leu734X} mutants to confirm the presence of the mutations.

(5) Sequencing analysis

The exact amino acid substitution at the positions of Ile705 within *Sce*ERG7 and Leu734 within *P. sativum* βAS were determined by sequencing of the DNA using ABI PRISM 3100 auto-sequencer. Nucleotide sequencing was performed using the dideoxynucleotide chain-termination method with only one forward of reverse primers. Sequencing reactions were carried out with BigDye$^{®}$ Terminator v3.1 Cycle Sequencing Kit, according to the manufacturer protocol. Briefly, each of sample was performed with 1 μl each forward or reverse primer, 2 μl plasmid DNA, 3μl 5X sequencing buffer, 0.5 μl premix and ddH2O to create a final volume of 20 μl. Each of the reaction was performed on the ABI PRISM$^{®}$ 3100 Genetic Analyzer, following the manufacturer's guidelines.

2.2.2 Preparation of competent cell (TKW14C2 and CBY57)

Pick the yeast TKW14C2 stock into the 3 ml SD medium, containing 60 μl 50X ALTHMU solution, 120μl 50% glucose solution, 60 μl hemin solution and 60 μl ergosterol solution and then incubated at 30 ℃ for five or six days. When the cell grew well, transferred it into 100 ml SD medium with the same condition and incubated at 30 ℃ for 16-18 hours. After OD$_{600}$ of it reaches 1.0~1.5, the cells were centrifuged to collect at 3,000 rpm, 10 min at 4 ℃ and the supernatant was discarded. Add about 45 mL aseptic ddH2O to resuspend the pellet and centrifuge it at 3000 rpm, 10 min at 4 ℃ and then repeated two times. Next step, adding 25 ml of 1 M D-sorbitol solution to resuspend the pellet and centrifuge it at the same condition. Finally, add n×50 μL (n is the number of samples) 1 M D-sorbitol into the pellet and resuspend it gentlemanly on ice for about 5 min.

The volume of 50μl competent cells was added into each of 1.5ml microtube with 5μl recombinant plasmids, respectively.

In addition, the preparation protocol of yeast CBY57 is the same as above mention except the yeast growth medium are ALHU solution and glucose solution.

2.2.3 Transformation of mutated plasmid into TKW14C2

The mutated plasmids of ERG7^{Ile705X} and PSY^{Leu734X} were transformed into TKW14C2, an ERG7-deficient yeast strain, by electroporation using a GenPulser with Pulse Controller (BioRad). Pipet total of electrocompetent cell with plasmid DNA onto drop and flick cuvette to settle DNA + cells mixture into bottom of cuvette. Set the conditions for transformation according to strains. For TKW14C2 cells, use 1.5k volt and the time constant should be 3-4 seconds. Dry off any moisture from cuvette outside and immediately place curette in white plastic holder. Slide holder into position and zap cells. If hearing a high constant tone, immediately add 500μL of D-sorbitol solution into cells. Aliquots of 120 μL of each culture were plated onto SD+Ade+Lys+His+Met+Ura+hemin+Erg plates, and incubated them for three to five days to select for the presence of pRS314-derived plasmids. The pRS314 and pRS314WT were transformed as negative and positive control, respectively.

2.2.4 Ergosterol supplement

Several colonies form each of plate were picked to reselect on two kinds of selective plates, SD+Ade+Lys+His+Met+Ura+hemin+Erg and SD+Ade+Lys+His+Met+Ura+hemin. The transformants were incubated three to five days at 30℃. Ergosterol supplement is a selection marker for functional complementation of cyclase in ERG7 mutants.

2.2.5 Extracting lipids and silica gel column chromatography

In the small scale incubation, the mutant transformants were grown in the 2.5 L SD liquid culture medium containing Glu+ALHMU+hemin+Erg medium at 30℃ with 220 rpm shaking for five days. The cells were harvested by centrifugation at 6000 rpm for 10 minutes. The pallets were resuspending in the solution containing 100ml ethanol, 100ml 30% KOH, and 0.2g pyrogallol. This reaction was refluxed at 110℃ for 3 hours. The hydrolysate was extracted three times with total 600ml petroleum ether, the organic phase were collected and dehydrated by sodium sulfate and concentrated using a rotary evaporator. The small amount of nonsaponofiable lipid (NSL) was dissolved CH_2Cl_2. The extract was fractionated by silica gel column chromatography using a 19:1 hexane/ethyl acetate mixture. Each of fractions was spotted on the thin-layer chromatography which developed with 4:1 hexane/ethyl acetate and the TLC plates were subjected to the stain buffer (5% Anisaldehyde, 5% H_2SO_4 in ethanol) and heated until the patterns appeared. According the TLC results, the fractions can be divided into five regions: OS (oxidosqualene), LA (lanosterol) up, LA, LA down, and ergosterol. The fractions of LA up, LA, and LA down sections were performed on GC/MS analysis for examination the triterpene products with a molecule mass of $m/z = 426$.

2.2.6 Acetylating modification and the alkaline hydrolysis reaction

In order to isolate the novel mutant products, using the acetylation modification of the triterpene alcohol fraction and this method was performed according the previous literatures.[45] (Fig.2.2) The dry triterpene alcohol fraction was first dissolved in 2ml pyridine solvent, and then 1ml acetic anhydride was added into solution. The solution was stirred overnight at room temperature. The acetylation reaction was monitored by TLC analysis. After 16 hours, 5ml of water was added to terminate the reaction and three times extraction with 10ml CH_2Cl_2 were carried out. The total organic phase was collected and

dried over with sodium sulfate, then evaporated in a rotary evaporator. The acetylation

products were separated by using argentic column chromatography and analyzed by

GC-MS.
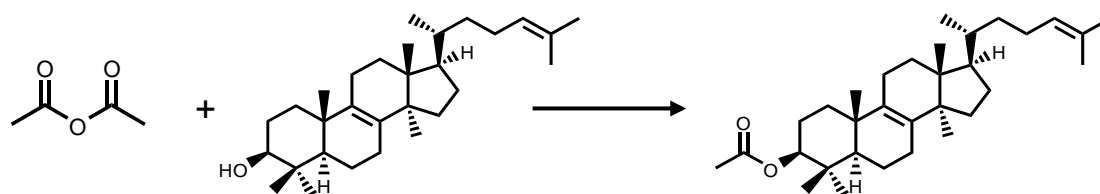
Figure 2.2 The acetylation modification

2.2.7 AgNO₃-impregnated silica gel chromatography

8.6g AgNO$_3$ and 25g silica gel dissolved in 50ml water, stirred and kept them away

from light in the oven under 110℃ for 16 hours. The gel was used to pack the column with

hexane, and then the acetylated products were fractioned by AgNO$_3$-impregnated silica gel

chromatography using 2.5~3% diethyl ether in hexane.[46] Thus, each fractions were

analyzed by GC-MS, and isolate the single product.

2.2.8 Deacetylation reaction of the modified compound

The dry acetylation triterpene fraction was dissolved in 10 ml methanol, and 0.5g

potassium hydroxide (KOH) was added into the reaction. The reaction was performed with

the closed system in the hood and stirred for 12-16 hours at room temperature. The

deacetylation reaction was monitored by TLC analysis. After 16 hours, the reactant was

dried by rotary evaporator and then 10 ml water was added to dissolve potassium

hydroxide. The deacetylated products were thrice extracted with dichloromethane. The

total organic phase was collected and dried over with anhydrous sodium sulfate (Na$_2$SO$_4$),

and then dried thoroughly in a rotary evaporator. The deacetylated products were separated

by silica gel column chromatography using 19:1 hexane/ethyl acetate mixture. The

structure of finally novel product was characterized and identified by NMR spectroscopy (^{1}H, ^{13}C, DEPT, COSY, HMQC, HMBC, and NOE).

2.2.9 GC-MS column chromatography condition

GC analyses were performed with a Hewlett-Packard model 5890 series II or Agilent 6890N chromatography equipped with a DB-5 column (30 m x 0.25 mm I.D., 0.25 μm film; oven gradient at 50 °C for 2 min, and then 20 °C per min until 300°C, held at 300°C for 20 min, 300 °C injector; 250 °C interface; 1/40 split ratio using helium carrier gas at 13 psi column head pressure). GC/MS was performed on a Hewlett-Packard model 5890 II GC (J &W DB-5MS column, 30 m x 0.25 mm I.D., 0.25 μm film; oven 280 °C, injector 270 °C, GC-MS transfer line: 280 °C) coupled to a TRIO-2000 micromass spectrometer.

2.2.10 Molecular modeling

Molecular-modeling studies were performed, using the DiscoveryStudio program with the X-ray structure of lanosterol-complexed human OSC as the template. And then using GOLD program for ligand docking, finally using SYBYL program for minimum of the energy. The homologous model structures were obtained with optimum condition. All of the molecular modeling softwares are provided by National Center for High-Performance Computing (NCHC) institution.

Chapter 3 Results and Discussion

3.1 Functional analysis of ERG7^{Ile705} within *S. cerevisiae*

3.1.1 Site-saturated mutagenesis of Ile705

The saturated mutations of Ile705 within ERG7 from *S. cerevisiae* were constructed by using QuikChange Site-Directed Mutagenesis kit, and then using the restriction enzyme (*Ava* I) mapping check. The correct mutants were digested into two parts, one was approximately 6.9 kbp, and the other was 1.1 kbp compared to the wild-type plasmid pRS314OSC which was digested into one fragment of 8 kbp. Mutated plasmids were confirmed by DNA sequencing via the ABI PRISM 3100 DNA sequencer.
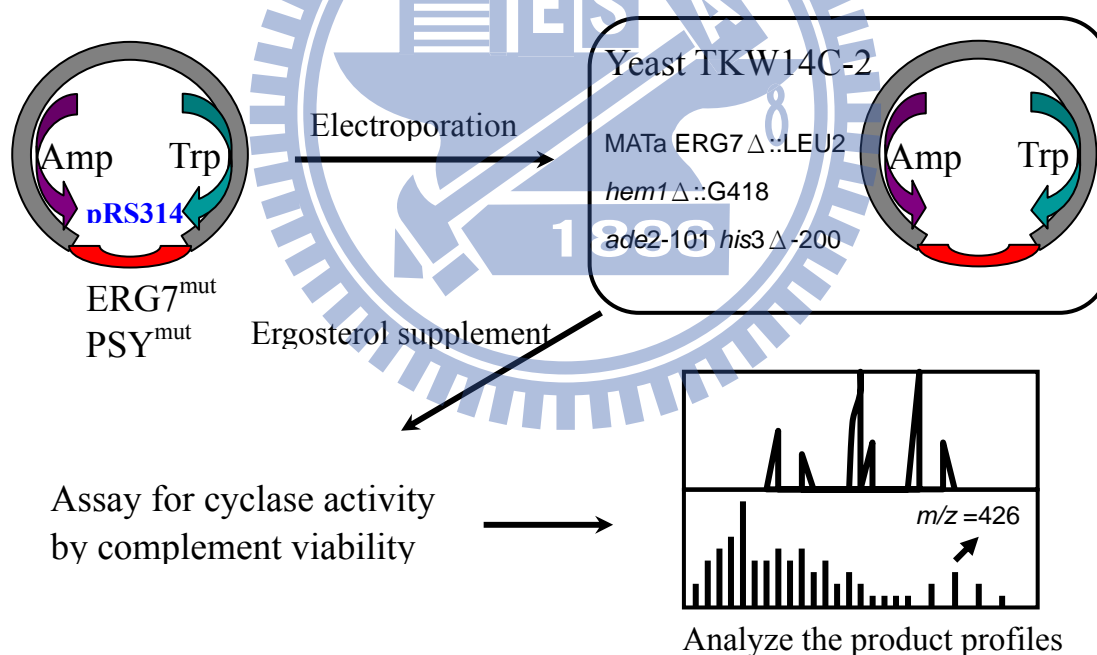
Figure 3.1 Strategies for the genetic selection method using TKW14C2 strains.

Following the mutations of the other 19 amino acids substituted at Ile705 position, these mutated plasmids were transformed into a yeast TKW14C2 strain (MATa or MATα ERG7Δ::LEU2 *hem1*Δ::G418 *ade*2-101 *his*3Δ-200 *leu*2-Δ1 *lys*2-801 *trp1*-Δ63 *ura3*-52) for the *in vivo* analysis. The yeast system, TKW14C2, is a CBY57-derived HEM1 ERG7 double-knockout mutant and is only viable when supplied with exogenous ergosterol or complemented with oxidosqualene cyclase activity derived from the ERG7[I705X] mutants. (Fig. 3.1) From the genetic selection among *S. cerevisiae* ERG7[I705X] mutants, the I705X (X=G/A/V/L/D/F/N/Q/K/T/C/M/P/S) mutants complemented the ERG7 disruption strain, TKW14C2 in the absence of exogenous ergosterol, while the others (I705W/I705H/I705E/I705R/I705Y mutants) failed. (Table 3.1) These results suggested that the different side chain substitution of I705 might influence the enzymatic activity and abolish the cyclization process of oxidosqualene, and further contributed to the understanding of the function of I705.
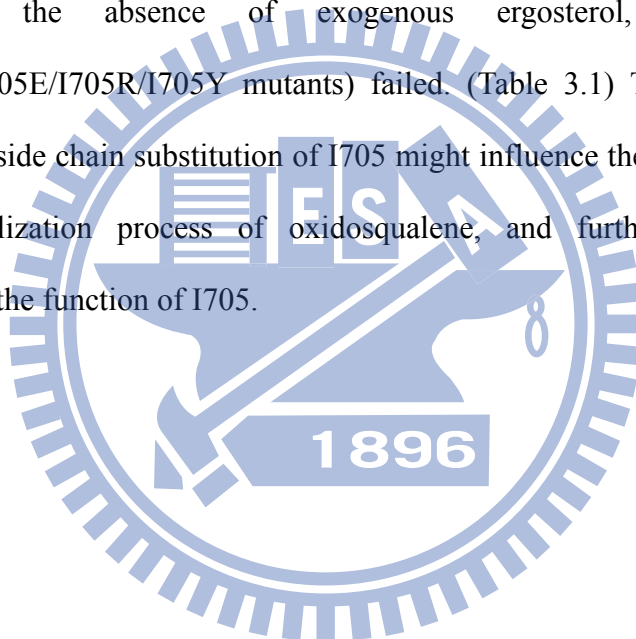
ERG7^{mut}	Restriction enzyme mapping	Sequence confirmation	Ergosterol supplement (TKW14C-2)
I705G(Gly)		V	+
I705A(Ala)		V	+
I705V(Val)		V	+
I705L(Leu)		V	+
I705F(Phe)		V	+
I705Y(Tyr)		V	—
I705W(Trp)		V	—
I705H(His)		V	—
I705D(Asp)		V	+
I705N(Asn)	*Ava* I	V	+
I705E(Glu)		V	—
I705Q(Gln)		V	+
I705K(Lys)		V	+
I705R(Arg)		V	—
I705S(Ser)		V	+
I705T(Thr)		V	+
I705M(Met)		V	+
I705C(Cys)		V	+
I705P(Pro)		V	+

Table 3.1 The site-saturated mutants of *S. cerevisiae* ERG7^{I705X} and their genetic analysis

In order to analyze the function of I705 in the oxidosqualene cyclization/rearrangement cascade, the TKW14C2[pERG7^{I705X}] mutants were incubated in 2.5 L culture medium. Products were isolated by extracting the non-saponifable lipid (NSL), applying onto silica gel column separation, and analyzing the structures of products by using gas chromatography-mass spectrometry (GC-MS). The product profiles of the *S. cerevisiae* ERG7^{I705X} mutants are listed in Table 3.2. There are eight triterpenoid products with a molecular mass of m/z = 426 including lanosterol,

(13α*H*)-isomalabarica-14*E*,17*E*,21-trien-3β-ol, (13α*H*)-isomalabarica-14*Z*,17*E*,21-trien-3β-ol, protosta-13(17),24-dien-3β-ol, (13α*H*)-isomalabarica-14(26),17,21-trien-3β-ol, 17α-protosta-20(21),24-dien-3β-ol and two novel products. The ERG7^{I705L} mutant has only one product, lanosterol, from wild-type ERG7. Twelve mutations including ERG7^{I705G}, ERG7^{I705A}, ERG7^{I705V}, ERG7^{I705D}, ERG7^{I705N}, ERG7^{I705Q}, ERG7^{I705K}, ERG7^{I705T}, ERG7^{I705C}, ERG7^{I705M}, ERG7^{I705S} and ERG7^{I705P}, produce lanosterol, (13α*H*)-isomalabarica-14*E*,17*E*,21- trien-3β-ol, (13α*H*)-isomalabarica-14*Z*,17*E*,21-trien-3β-ol, and protosta-13(17),24- dien-3β-ol. Furthermore, the ERG7^{I705G} mutant also produced a small amount of the novel product. Another mutation that produced lanosterol, (13α*H*)-isomalabarica-14(26),17,21-trien-3β-ol, 17α-protosta-20(21),24-dien-3β-ol and a large amount of novel product is ERG7^{I705F}. Finally, no products with *m*/*z* = 426 were observed in five mutations which were substituted with I705W, I705H, I705E, I705R and I705Y, respectively.
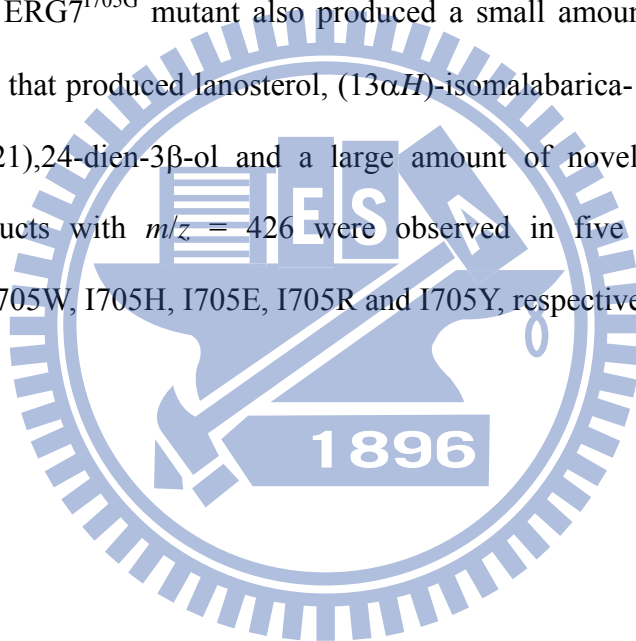
ERG7^{I705X}	Ergosterol supplement (TKW14C-2)	Product pattern (%)							
		LA	(13αH)-isomalabarica-14E,17E,21-trien-3β-ol	(13αH)-isomalabarica-14Z,17E,21-trien-3β-ol	Protosta-13(17),24-dien-3β-ol	(13αH)-isomalabarica-14(26),17,21-trien-3β-ol	17α-protosta-20(21),24-dien-3β-ol	17α-protosta-20(22),24-dien-3β-ol	Unknown B
G	+	35	23	34	6			1	1
F	+	25				21	6	42	6
A	+	78	13	8	0.3				0.7
V	+	75	10	15					
D	+	37	32	30	1				
N	+	19	41	37	3				
Q	+	25	39	31	5				
K	+	12	26.5	53.1	7.6				0.8
T	+	22	36	41	1				
C	+	72	16	11.7	0.3				
M	+	88	6	6					
S	+	12	42	44	1				1
P	+	81.9	0.1	15					3
L	+	100							
Y	—								
W	—								
H	—				No product				
E	—								
R	—								

Table 3.2 The product profiles of the *S. cerevisiae* ERG7^{I705X} mutants.

3.1.2 The identification and characterization of novel product

In order to isolate the unknown product, 70 L of mutant yeast were incubated and the non-saponifable lipids (NSL) were extracted. One of the novel products was a major compound within the culture of the mutant for $ERG7^{I705F}$. (Fig. 3.2) The products mixture which with $m/z = 426$ were acetylated for increasing the polarity, then their m/z became 468. (Fig. 3.3) According to the principle that an argentic ion would bind with a different double bond of the compound, each single compound was separated by $AgNO_3$-impregenated silica gel column chromatography with 2.5 - 3% diethyl ether in hexane as the mobile phase. By the end, through the deacetylation step, the isolated single compound (Fig. 3.2) could be identified by nuclear magnetic resonance (NMR).
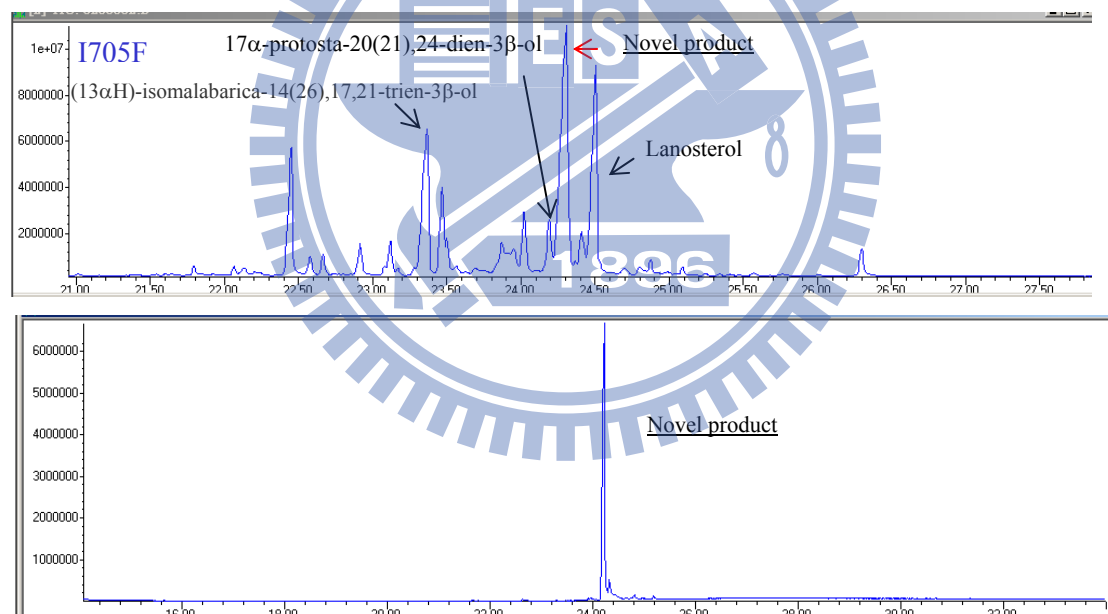
Figure 3.2 The GC data of ERG7^{I705F}. After the novel product was isolated, the complex mixture was purified to became a sole product shown in picture below.
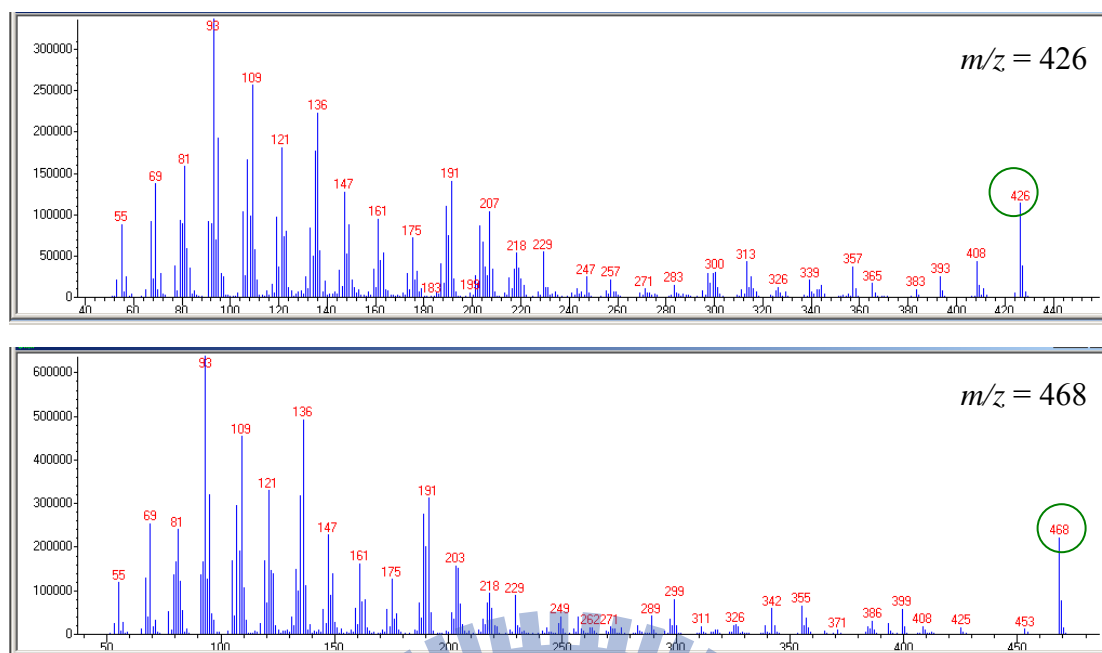
Figure 3.3 The mass spectra of novel products from the ERG7[I705F] mutant.

The unknown product (Fig. 3.4) which contains a tetracyclic scaffold with a 17α side chain and a $\triangle^{20/22}$ double bond, and its spectroscopic assignment was determined according to the following analysis of various NMR spectra (^{1}H, ^{13}C NMR, DEPT, ^{1}H-^{1}H COSY, HMQC, HMBC, and NOE), and the spectra in depth are shown in Appendix section. First, the compound showed ^{1}H NMR spectrum with three vinylic methyl singlets (δ 1.66, 1.6, and 1.52), and five methyl singlets (δ 0.95, 0.75, 0.9, 1.11, and 0.8), and two olefinic protons (δ 5.07, and 5.07). Second, the ^{13}C NMR showed the presence of two tertiary-quaternary (δ = 131.85, 124.31, and 123.91, 137.7). These results indicated the presence of a tetracyclic nucleus with a terminal hydrocarbon side chain double bond. Correlations of the HMQC, HMBC, and NOE spectra showed the following features: (1) the vinyl proton at δ 5.07 (δ_c 123.91, C-22) is coupled to carbons at 13.44 ppm (C-21), 27.77 ppm (C-23), and 51.29 ppm (C-17); (2) the vinyl proton at δ 5.07 (δ_c 131.85, C-24) is coupled to carbons at 18.18 ppm (C-27), 26.16 ppm (C-26), and 124.31 ppm (C-25); (3) the proton at δ 2.65 (δ_c 27.77, C-23) is coupled to carbons at 123.91 ppm (C-22), 124.31 ppm

(C-25), 131.85 ppm (C-24), and 137.70 ppm (C-20); (4) the proton at δ 1.59 (δ$_c$ 44.94, C-13) is coupled to carbons at 16.93 ppm (C-30), 23.41 ppm (C-11), 25.78 ppm (C-12), 50.79 ppm (C-14), 137.70 ppm (C-20)and 51.29 ppm (C-17); (5) the proton at δ 0.90 (δ$_c$ 23.2, C-19) is coupled to carbons at 33.87 ppm (C-1), 37.66 ppm (C-10), 46.72 ppm (C-9), and 48.72 ppm (C-5). These correlations established the bond connectivity between the tetracyclic nucleus skeleton and the exocyclic hydrocarbon side chain as well as the double bond positions. The distinct evidence of NOE spectra, which was observed among Me-19/Me-29, H-5/Me-18, and H-9/Me-19, confirmed the presence of chair-boat-chair conformation of the compound, and the tetracyclic structure contains the spatial NOE interaction between Me-30/H-17, and Me-30/H-9, confirmed the 17α side chain conformation.
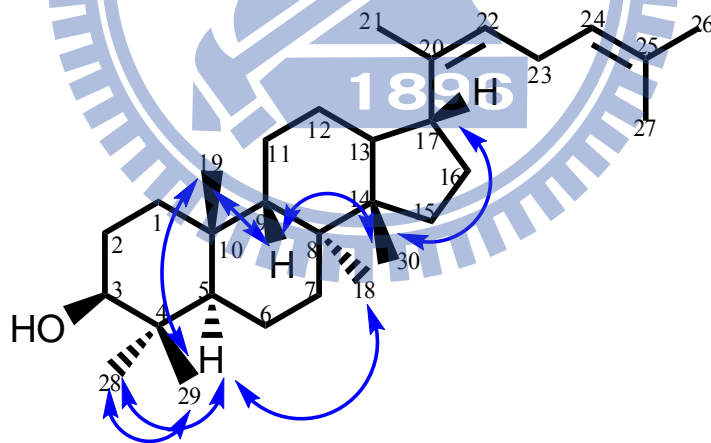
Figure 3.4 The structure and the NOE correlation of the novel compound

	$^{13}C(\delta)$	$^1H(\delta)$		$^{13}C(\delta)$	$^1H(\delta)$
C-1	33.87	1.38~1.42	C-16	27.77	1.71~1.77，1.40~1.42
C-2	30.02	1.63~1.66，1.55~1.60	C-17	51.27	2.07~2.11
C-3	79.85	3.18	C-18	22.27	1.11
C-4	39.94	—	C-19	23.2	0.9
C-5	48.72	1.41~1.46	C-20	137.7	—
C-6	19.27	1.51~1.54	C-21	13.44	1.52
C-7	35.89	1.18~1.20，1.93~1.97	C-22	123.91	5.07
C-8	40.37	—	C-23	27.77	2.65
C-9	46.72	1.45~1.50	C-24	131.85	5.07
C-10	37.66	—	C-25	124.31	—
C-11	23.41	1.43~1.46，1.20~1.22	C-26	26.16	1.66
C-12	25.78	1.44~1.47，0.96~0.98	C-27	18.18	1.6
C-13	44.94	1.59	C-28	29.61	0.95
C-14	50.79	—	C-29	16.71	0.75
C-15	32.33	1.47~1.51，1.07~1.17	C-30	16.93	0.8

Table 3.3 NMR assignments for 17α-protosta-20(22),24-dien-3β-ol for dilute CD$_2$Cl$_2$ solution

This is the second time to isolate the 17α tetracyclic triterpene alcohol derivative which has the chair-boat-chair skeleton simultaneously, while all of the other chair-boat-chair products are 17β-derivatives. The novel products, 17α-protosta-20(22),24-dien-3β-ol, and 17α-protosta-20(21),24-dien-3β-ol which was the first 17α-derivative isolated in our lab[47], are simultaneously produced by the ERG7[I705F] mutant. These data showed that I705 is an important residue that contributes to the formation of 17α-products.

3.1.3 Proposed cyclization/rearrangement pathways of TKW14C2 expressing ERG7[Ile705X]

The various products derived from ERG7[I705X] indicated that I705 is a crucial residue in the putative active site for catalytic function of oxidosqualene cyclase. Seven triterpenoid compounds include three tricyclic and four tetracyclic products. Three truncated tricyclic compounds including (13αH)-isomalabarica-14E,17E,21-trien-3β-ol, (13αH)-isomalabarica-14Z,17E,21-trien-3β-ol and (13αH)-isomalabarica-14(26),17, 21-trien-3β-ol, suggest that after oxidosqualene was pre-folded into the chair-boat-chair conformation and then following cyclization step via cation-π interactions to a 6,6,5-tricyclic C-14 cation intermediate, the three (13αH)-isomalabarica-trien-3β-ols were derived from deprotonation at either the C-15 or C-26 of the 6,6,5-tricyclic C-14 cation. Following the anti-Markovnikov expansion of the five-membered C-ring to form a 6,6,6-tricyclic C-13 cation, and then via D-ring closure to generate the protosteryl C-20 cation, the substrates through a series of methyl and hydride shifts, deprotonation at C-8 or C-9 yield lanosterol. The mutation of I705 may influence the F699 residue so that it affects the stability of the C-17 cation intermediate at or after the protosteryl cation formation, and the other parts of substrates generate the truncated triterpene, protosta-13(17),24-dien-3β-ol.[48]

Different from the previous characterization, an amazing phenomenon appeared in the I705 mutation, there are two 17α-derivatives, 17α-protosta-20(22),24-dien-3β-ol and 17α-protosta-20(21),24-dien-3β-ol, observed in the ERG7[I705F] mutant simultaneously. The speculated pathway is the unnatural D-ring closure, α-orientation of the large side-chain, and then it generates the 17α– protosteryl C-20 cation. This unnatural mechanism may arise from lack of correct amino acids in the putative active site of the cyclase, without rearrangement compared to the normal type, with deprotonation at C-21 or C-22 positions, yielding 17α-protosta-20(22),24-dien-3β-ol and 17α-protosta-20(21),24-dien-3β-ol. (Fig.

3.5)
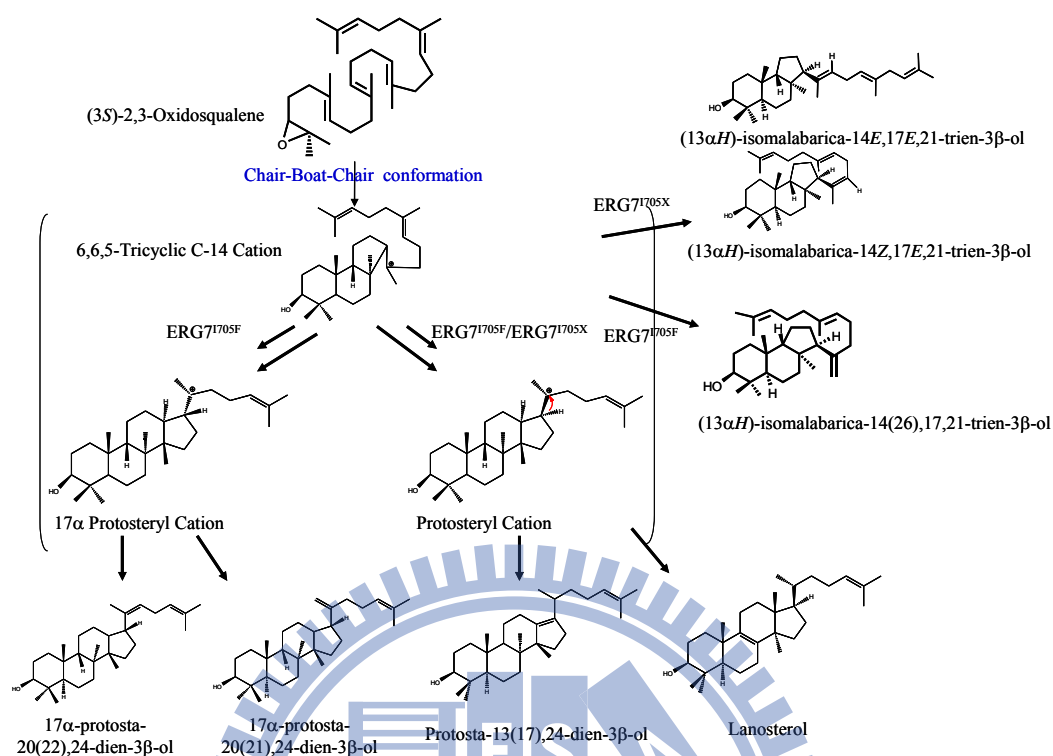
Figure 3.5 Proposed cyclization/rearrangement pathway occurred in the ERG7[I705X] site-saturated mutants

In a previous study, the near absence of 17α dammarenyl cyclases in higher plants reflects the rarity of the 17β to 17α evolutionary step.[44] Therefore, no matter what protosteryl cation or dammarenyl cation is present, the 17α configuration is unusual and rare in the cyclization process. Formation of 17α-derivatives is unnatural, and the proposed pathway via the 17α−protosteryl C-20 cation could not be defined in the mechanism of the wild-type cyclases. On the other hand, the 17α configuration intermediates did appear only in bacteria or rare plants species, indicating that the 17α configuration may exist in early evolutionary periods, which could rationally explain the rarity of the 17α−derivatives.

3.1.4 Analysis of the ERG7[Ile705X] mutants with the ERG7 homology modeling

Molecular modeling studies were performed to investigate the roles of important residues in the cyclization/ rearrangement cascade. Because of the lack of a high-resolution crystal structure of *S. cerevisiae* ERG7, the homology models of *S. cerevisiae* ERG7 and its mutated ERG7 proteins were derived from the human OSC X-ray structure, as the template, and complexed with lanosterol or 6,6,5-tricyclic C-14 cation or 17α-protosteryl cation, together with product profiles, in order to investigate the relationship between enzyme structure and product specificity. The homology model of ERG7 showed that Ile705 is a hydrophobic residue different from neighboring aromatic residues, and I705 is spatially proximal to C-14 and C-17 positions of lanosterol (Fig. 3.6a). The hydrophobic property of I705 is considered to have some interaction with the substrate, and also affect the first-tired residues. According to the product profiles of ERG7[Ile705X], the models complexed with 6,6,5-tricyclic C-14 cation and 17α-protosteryl cation, take into account the variation of the residue at 705 position. Furthermore, the product profile of ERG7[I705X] is similar to that of ERG7[F699X] mutations[47]; therefore we observe the variation of F699 for investigating the function of I705 in depth. (Table 3.4)

ERG7[I705F] and ERG7[I705G] mutants both produce 17α-protosta-20(22),24-dien-3β-ol, which is the novel product in this study. According to the homology models, the distances between ERG7[I705F] and C-14 and C-17 of lanosterol are closer than wild-type, and the distance between I705F and C-17 of 17α-protosteryl cation is also closer than ERG7[I705]. (Fig. 3.6b) The aromatic Phe705 stabilized the 17α-protosteryl cation with π-cation interactions, therefore producing both 17α-protosta-20(22),24-dien-3β-ol and 17α-protosta-20(21),24-dien-3β-ol. On the other hand, substitution of Gly with a small aliphatic side chain possibly enlarged the cavity of the active site to cause the substrate to be more flexible, and then could not form lanosterol accurately. There are many truncated products discovered in the ERG7[I705G] mutant, including a small amount of

17α-protosta-20(22),24-dien-3β-ol. Except for the 17α-product, the other truncated products are considered to form by variation of the F699 residue, where the flexible Gly possibly influences the F699 residue in this case. Formation of the 17α-protosteryl cation is rare and it was observed in F699X mutations before, so the 17α-protosteryl cation formation may be caused by the effect of F699.

Substitution of I705 into an aliphatic group such as Ala or Val produce truncated products, whereas Leu only produces lanosterol. The ratio of lanosterol decreases when the size of the side chain becomes smaller. Because Leu has a similar bulky size as Ile, the ERG7[I705L] mutant only yielded lanosterol as was observed for the wild-type. The percentage of lanosterol for ERG7[I705A] and ERG7[I705V] are about 75 - 78%, and even the ERG7[I705G] mutant only has about 35% of lanosterol shown in Table 3.2. This is the evidence that the steric bulk size at the 705 position is an important factor during catalytic cyclization. The change of the distance between I705A/V of lanosterol are both decreased, and the variation of the F699 residue within I705A/V mutants, complexed with 6,6,5-tricyclic C-14 cation, are closer than wild-type ERG7.

The obvious phenomenon within the acidic, amide and basic group such as I705D/ N/ Q/ K mutants represent smaller amounts of lanosterol compared to the aliphatic group, even without any other products yielded in the ERG7[I705E/ H/ R] mutations. The balance environments destroyed by substitution of hydrophobic residue into acidic, amide and basic group within the putative active site, leading to tendency of lanosterol formation reduce and generated some truncated products by the F699 effect. In addition, the mutated ERG7[I705S/ T/ C/ M/ P] proteins also yielded lanosterol and truncated products, similar to that observed for most ERG7[I705X] mutations. Comparing the product profiles, we discovered the side chain polarity that could affect the proportion of lanosterol. Substitutions of I705 into Cys, Met and Pro with a nonpolar side chain have higher lanosterol amounts, while Ser and Thr residues with a polar side chain have a small amount of lanosterol. The analysis of the

lanosterol formation tendency show that the acidity and polarity of the residue at the 705 position influence the environment around the substrate within the active site of the oxidosqualene-lanosterol cyclase. But the distance between I705X (X=D, N, Q, K, E, H, R, S, T, C, M, P) and lanosterol, and the variation of the F699 residue in these mutants complexed with 6,6,5-tricyclic C-14 cation are irregular. The irregularly homological results perhaps are the results of a deviation arising from the computational calculation. However, it also provided the information that the relative positions between the substrate and the cyclase were altered during the catalytic cyclization and the distance may become farther or nearer, when the cation intermediate formed.

Finally, there is no product observed in the ERG7$^{I705Y/W}$ mutants. Although Tyr and Trp are aromatic residues like Phe, the bulky size makes these two mutations difficult to complement yeast viability, causing the loss of cyclase activity. The steric bulk size may affect the F699 residue strongly, and the stabilization around F699 and the substrate destroyed, so that the catalytic processes were abolished. In conclusion, the key factors which influence the diverse product profiles and yeast viability of the I705X mutations are steric effect, acidity and side chain polarity. Variation at the I705 position certainly influences the substrate and the cyclase in significant ways.

a.   b.
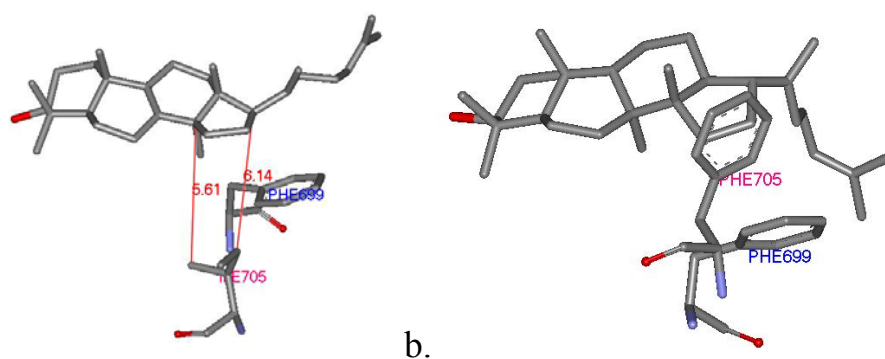
Figure 3.6 Homology models. (a) The homology model of wild-type ERG7 complexed with lanosterol, and the distance between I705 to C-14 and C-17 positions of lanosterol is 5.61 Å and 6.14 Å. (b) The homology model of ERG7^{I705F} mutant complexed with the

17α-protosteryl cation.
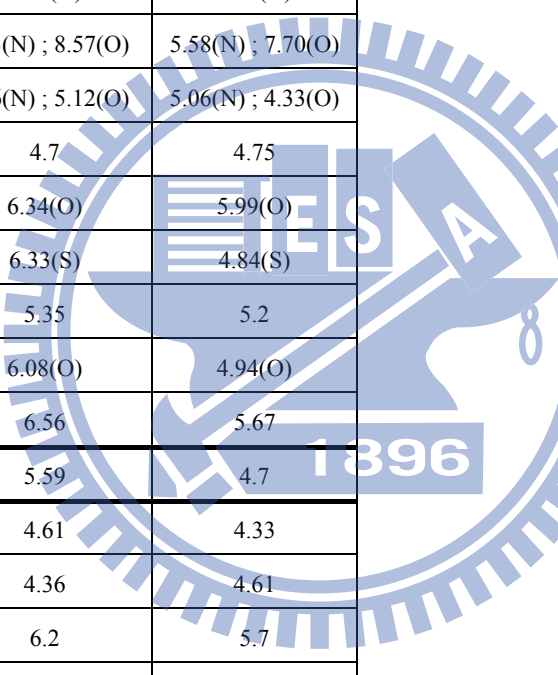
Amino acid substitution	Distance to C-17 of lanosterol (Å)	Distance to C-14 of lanosterol (Å)	Distance to C-17 of 17α-protosteryl cation (Å)	Distance to C-14 of 6,6,5-tricyclic C-14 cation (Å)	F699 distance to C-14 of 6,6,5-tricyclic C-14 cation (Å)
WT	6.14	5.61	6.42(705);4.31(699)	6.12	5.17
F	4.74	4.67	4.79(705);4.44(699)	5.12	5.47
G	8.53	7.78	7.54(705);4.46(699)	9.23	6.24
A	5.72	5.09		6.47	5.09
V	5.78	5.35		6.25	4.78
D	5.63(O)	4.72(O)		7.09	7.45
N	6.33(N) ; 8.57(O)	5.58(N) ; 7.70(O)		5.59	5.73
Q	5.06(N) ; 5.12(O)	5.06(N) ; 4.33(O)		5.45	5.93
K	4.7	4.75		4.83	6.45
T	6.34(O)	5.99(O)		6.88	5.12
C	6.33(S)	4.84(S)		5.69	5.27
M	5.35	5.2		5.77	6.9
S	6.08(O)	4.94(O)		6.51	5.23
P	6.56	5.67		6.72	5.04
L	5.59	4.7		6.12	6.57
Y	4.61	4.33		4.42	6.5
W	4.36	4.61		4.47	5.08
H	6.2	5.7		7.45	7.08
E	5.15(O)	4.54(O)		4.83	5.8
R	1.98	1.95		4.37	6.39

Table 3.4 The distance of I705 to C-14 and C-17 complexed with different ligands in the homology models, and observation of the variation of F699 within the ERG7^{I705X} mutants, using the homology model complexed with the 6,6,5-tricyclic C-14 cation.

3.1.5 Product analysis of the double mutant of ERG7$^{I705F/F699X}$

The product profile of ERG7^{I705X} is similar to that of ERG7^{F699X} mutations investigated before in our lab, and F699 had been identified to be an important plasticity

residue with its product diversity[47]. In order to prove that the hydrophobic residue at this position plays an important role that influences the first-tired residue in the putative enzymatic active site, some double mutants were constructed for study in depth. The various products of ERG7^{I705X}, 17α-protosta-20(22),24-dien-3β-ol is the only compound that ERG7^{F699X} mutations had not discovered before. The novel compound, 17α-protosta-20(22),24-dien-3β-ol, was only observed for the ERG7^{I705F} mutation, so that the combination of I705F/F699X double mutants were constructed and their products analysis will be discussed later.

In the experiment with ERG7$^{F699C/I705F}$ double mutants, as Table 3.4 describes, the single F699C mutation loses the activity of the oxidosqualene-lanosterol cyclase, causing the failed result in the absence of exogenous ergosterol, and the result of double mutants is the same as that observed for the single mutant. These data indicate that the F699 is the crucial residue while the I705 residue is critical as an "assistant" to affect the first-tired residue F699. Furthermore, in the single mutant F699M, the products yielded include lanosterol, protosta-13(17),24-dien-3β-ol, protosta-17(20),24-dien-3β-ol, (13αH)-isomalabarica-14E,17E,21-trien-3β-ol, (13αH)-isomalabarica-14Z,17E,21-trien-3β-ol, the chair-chair 6-6-5 tricyclic malabarica-14(15)E,17,21-trien-3β-ol and 17α-protosta-20(21),24-dien-3β-ol. Double mutants of I705F/F699M yield 17α-protosta-20(22),24-dien-3β-ol without 17α-protosta-20(21),24-dien-3β-ol, because F699 is crucial within the active site of cyclase, where the 17α-products might be formed by mutation of F699. With different orientations, the I705 affects the F699 residue and has a tendency to deprotonate the proton at C-22 simultaneously. Moreover, interesting data was discovered in the ERG7 $^{I705F/F699T}$ mutation shown in Table 3.5 that except for the formation of little amount of lanosterol, the single mutant of ERG7^{F699T} produced protosta-13(17),24-dien-3β-ol as the major compound. The generation of protosta-13(17),24-dien-3β-ol, 17α-protosta-20(21),24-dien-3β-ol and

17α-protosta-20(22),24-dien-3β-ol also appeared in the double mutation of I705F and F699T. This result showed that I705 is not only an important assistant residue to the first-tired residue F699, but it is also essential for contribution towards the forming of 17α-products. Furthermore, the results of homology models show the relative distance between F699M/I705F, F699T/I705F to the C-17 of lanosterol are shorter than the wild-type ERG7. However, the modeling data provide us with strong information that the I705F is influential at the C-17 position. (Fig. 3.7)

ERG7mut		Product pattern(%)									
		Lanosterol	(13αH)-isomalabarica-14(26),17,21-trien-3β-ol	17α-protosta-20(21),24-dien-3β-ol	17α-protosta-20(22),24-dien-3β-ol	Protosta-13(17),24-dien-3β-ol	Protosta-17(20),24-dien-3β-ol	Malabarica-14(15)E,17,21-trien-3β-ol	(13αH)-isomalabarica-14E,17E,21-trien-3β-ol	(13αH)-isomalabarica-14Z,17E,21-trien-3β-ol	Unknown B
I705F		25	21	6	42						6
F699M		1		1		46	10	7	17	18	
I705F	F699M	53			3	15			17	12	
F699C		No product									
I705F	F699C										
F699T		1			99						
I705F	F699T	36		7	12	45					

Table 3.5 The products analysis of double mutants between I705F and F699X

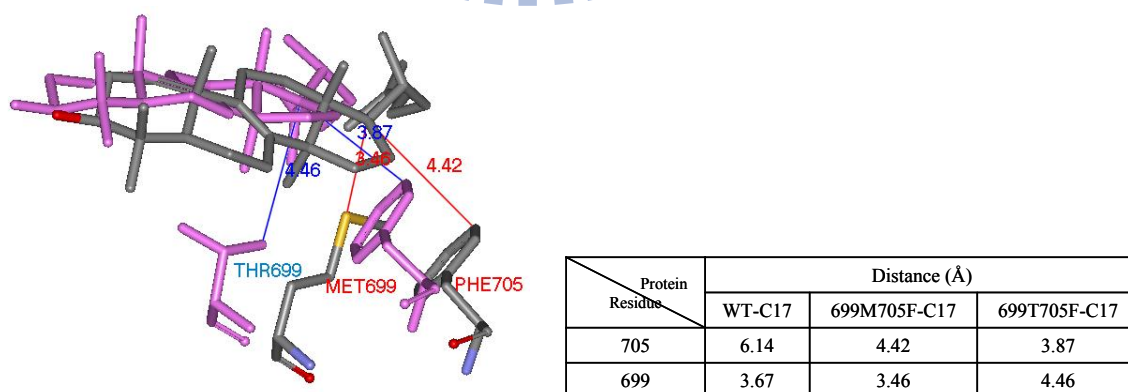| Protein Residue	Distance (Å)		
	WT-C17	699M705F-C17	699T705F-C17
705	6.14	4.42	3.87
699	3.67	3.46	4.46

Figure 3.7 The homology models of the double mutants ERG7[F699M/I705F] (gray) and ERG7[F699T/I705F] (purple) complexed with lanosterol.

3.2 Functional analysis of PSY^{Leu734} within *P. sativum*

3.2.1 Site-saturated mutagenesis of Leu734

The construction of saturated mutations of Leu734 within the β-amyrin synthase from *Pisum sativum* were done by utilizing the QuikChange Site-Directed Mutagenesis kit, then by using the restriction enzyme (*Nru* I) mapping check. The mutated plasmids were digested into 4.6 and 3.5 kbp fragments compared to the wide-type plasmid that was digested into one fragment. Similar to the construction of Ile705X, the mutated plasmid sequences were also confirmed via the ABI PRISM 3100 DNA sequencer.

The recombinant plasmids were confirmed and transformed into TKW14C2 by the same strategies as previously described in Section 3.1.1., and the genetic selection of the TKW14C2[pPSYL734X] mutants are shown in Table 3.6. Because the plant PSY gene was expressed by the yeast system, as per the previous speculation, the exogenous ergosterol was necessary for growth. After incubating them in large amounts of culture mediums, the results were represented following lipid extraction and column chromatography. Using GC-MS instrumentation for all products analysis, all of the mutations of L734X did not generate any triterpenoid structures with molecule mass of *m/z* = 426. Without any truncated compounds and β-amyrin, the result reveals the importance of the L734 residue within the putative active site of *P. sativum* βAS.

PSY^mut	Restriction enzyme mapping	Sequence confirmation	Ergosterol supplement (TKW14C-2)	Product analysis
L734G(Gly)		V	—	
L734A(Ala)		V	—	
L734V(Val)		V	—	
L734I(Ile)		V	—	
L734F(Phe)		V	—	
L734Y(Tyr)		V	—	
L734W(Trp)		V	—	
L734H(His)		V	—	
L734D(Asp)		V	—	
L734N(Asn)	*Nru* I	V	—	No product
L734Q(Gln)		V	—	
L734E(Glu)		V	—	
L734K(Lys)		V	—	
L734R(Arg)		V	—	
L734S(Ser)		V	—	
L734T(Thr)		V	—	
L734M(Met)		V	—	
L734C(Cys)		V	—	
L734P(Pro)		V	—	

Table 3.6 The site-saturated mutants of *P. sativum* PSY^L734X and their genetic selection and products analysis.

3.2.2 Experimental results of PSY^{L734X} mutants and their phenomena

β-amyrin synthase is one of the oxidosqualene cyclases that convert oxidosqualene into pentacyclic 6,6,6,6,6-fused β-amyrin, and there are many essential amino acids in the putative active site and act their specific roles. Substitution at the L734 position, a hydrophobic residue within *P. sativum* βAS, causes the production of β-amyrin to fail. As per the previous description, there is no triterpenoid compound with molecule mass of *m/z* = 426 observed for the L734 mutation. Except for the formation of β-amyrin similar to that observed for wild type, generation of truncated products provide the evidence that each residue plays a specific functional role. In this case, the phenomenon of no truncated products and β-amyrin, it could be presumed that PSY^{L734} is crucial for cyclase. Therefore, the possibility of the disrupted catalytic reaction will be discussed later.

The resolution of human OSC structure was published in *Nature*, Thoma and co-workers (2004) considered that the human OSC has a channel which is supposed to lead the substrate oxidosqualene into the hydrophobic active site. There are some residues such as Tyr237, Cys233 and Ile524 near the substrate passageway, and by changing the conformation of their side chains, the passage of the substrate subsequently could be achieved.[26] In the study of the F528 residue within *Sce*ERG7, F528 is located in the substrate entrance channel and probably influences the enzymatic activity through substrate binding. Some observations have shown that the ERG7^{F528X} mutations cannot complement the cyclase activity due to the change of pH scale or polarity in the substrate passageway, causing disruption of triterpenoids formation.[49] Nevertheless, this supposition could not be established in the case of PSY^{L734}. In the modeling picture shown in Fig. 3.8, C260 and I555 residues within *P. sativum* βAS, which correspond to Tyr237 and Ile524 in human OSC, are assumed to be the crucial amino acids that affect the substrate passageway. Different from C260 and I555, the L734 residue is located below the substrate and the distance between C260, I555 and L734 are so far, that L734 may not be one of the

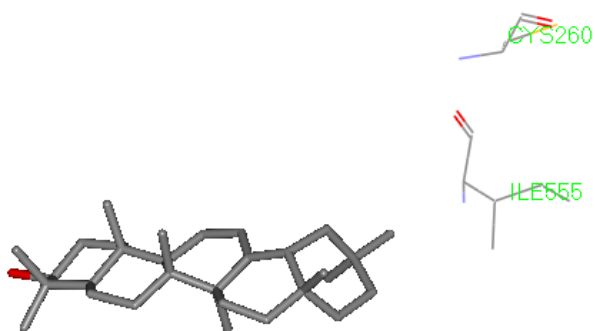important amino acids that control the substrate entrance channel.

Figure 3.8 The homology model of wild-type PSY complexed with β-amyrin. C260 and I555 residues are assumed to affect the substrate passageway within the cyclase.

On the other hand, the distance between L734 and the substrate within βAS is closer than that of I705 with lanosterol, and therefore the effect of L734 is stronger than I705 for cyclases. As shown previously in Section 3.1, the products of ERG7[I705X] mutations are diverse and special. Except for lanosterol, the formation of truncated products provided strong evidence for the importance of *Sce*ERG7[I705]. I705 is not only a crucial residue for catalytic function of the oxidosqualene-lanosterol cyclase, but also spatially close to its neighbor first-tired residues, especially F699. Because of the stronger effect in L734X mutations, the substitution of other amino acids may influence the conformation of the substrate oxidosqualene before the initial epoxide ring opening. Incorrect substrate conformation may halt catalytic cyclization, and that is the rational explanation for losing the activity of *P. sativum* βAS. In addition, the functional roles of the other amino acids

such as F728, Y736, W418, W612, F474, W257, and Y259, where their corresponding residues all have crucial roles within *Sce*ERG7, have not been confirmed by experimental results. Therefore, the direct effect for causing the loss of the activity of *P. sativum* βAS by the L734 mutations could not be explained in depth, and it is likely the case that the disruption of catalytic cyclization was influenced by the L734 mutations indirectly. Finally, we could only speculate that PSY^{L734} may stabilize the substrate conformation, but the detailed function should be investigated in the future. (Fig. 3.9)
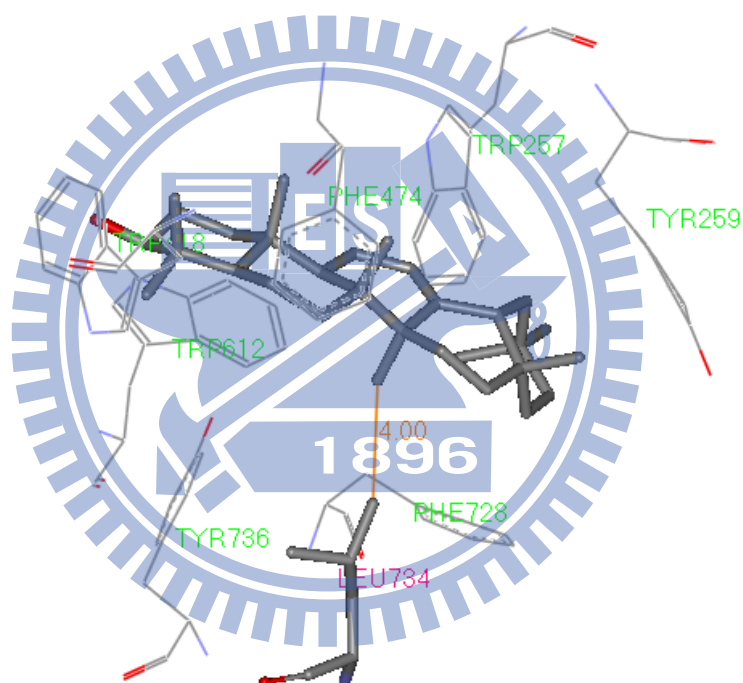

Figure 3.9 The minimum distance of Leu734 to β-amyrin is 4 Å. The possible crucial residues also shown in this modeling picture.

3.2.3 Analysis of the PSY^{L734X} mutants with the PSY homology modeling

We expected the function of L734 within the β-amyrin synthase may the same as I705 within *Sce*ERG7, because of their similar hydrophobic properties, similar sizes and relative positions between substrates and cyclases. On the contrary, no truncated products and β-amyrin were generated in L734X mutations. It is unusual that site-saturated mutagenesis of an aliphatic residue would lead to losing activities of the all mutated proteins. However, the homology models could provide us with more information about the variation of the enzyme and the substrate. Therefore, we constructed the homology models of *P. sativum* βAS and its mutated PSY proteins were derived from human OSC X-ray structure, as the template, and complexed with β-amyrin as the ligand in order to define our results reasonably.

The homology model of PSY revealed that L734 is spatially close to C-27 of β-amyrin, and the minimum distance is 4 Å. Due to the importance of the relationship between I705 and F699 within the putative active site of OSC, we observed the variation of F728, which corresponds to F699 within *Sce*ERG7, in the mutated PSY. The nearest residue of L734 is S412, therefore we also should pay close attention to the S412 residue (Fig. 3.10). In Table 3.7, the substitutions of L734 into the aromatic residues show that the distance between F728 and β-amyrin is longer, whereas S412 and β-amyrin is shorter than the wild-type protein. The results revealed the effect of the steric bulk size at the 734 position of PSY, whereby it changed the relative distance of enzyme and substrate, and caused the loss of the activity of βAS. On the other hand, the PSY^{L734X} (X= D, N, Q, E, H, K, R) mutants have the same tendency where the distance between F728 and β-amyrin became longer, and the L734X is closer to F728 residue; whereas the variation on S412 is irregular. L734D, L734E, L734H, L734K and L734R are acidic and basic groups that destroyed the acidity balance within the putative active site of cyclase. L734N and L734Q which have amide groups are polar to the substrate, and influence the neighboring residues

such as F728. Therefore the side chain polarity may also be a factor of the destruction of cyclization. The mutations of L734S, L734T, L734C, L734M and L734P also have the tendency to lengthen the relative distance between F728 and β-amyrin, and the L734X is effectively closer to the F728 residue than what is observed in wild-type. Different from *Sce*ERG7, substitution of the aliphatic residue into the other residue where the side chain has oxygen, nitrogen and sulfur atoms, would lead to the destabilization of substrate with the change of the electron density distribution. Amino acids such as Asp, Asn, Gln, Glu, His, Lys, Arg, Ser, Thr, Cys, Met and Pro conform to the former characterization, therefore the electron density distribution is also the key factor affecting the enzyme activity.

Finally, it is difficult to explain why the mutations of L734G, L734A, L734V and L734I, which have aliphatic side chains such as Leu, also have no product. Substitution with small aliphatic side chains such as Gly and Ala may make the substrate more flexible, and also influence the neighboring amino acids, and may hence destroy the stabilization of substrate. In the homology models, we could observe the relative distance of L734X (X= G, A, V, I) and β-amyrin, F728 and β-amyrin both became longer, the relative distance of S412 and β-amyrin became shorter. The effects of F728 and S412 are quite different: F728 is obvious, while S412 has just a little shift. Thus mutated L734X affects the F728 residue and causes the substrate to be unstable, disrupting the catalytic cyclization.
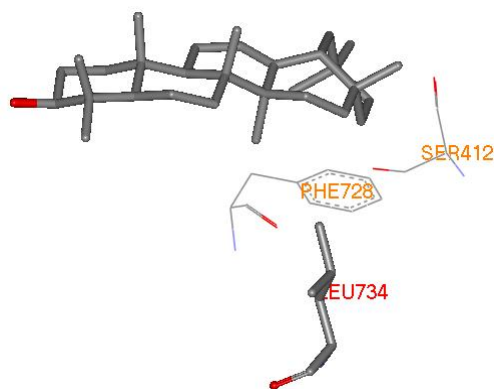
Figure 3.10 The homology model of wild-type PSY revealed that L734 is spatially close to C-27 of β-amyrin and its neighboring important residue F728. S412 is the nearest residue around L734, within the putative active site of βAS.

Amino acid substitution	The nearest distance to β-amyrin (734)	The nearest distance to β-amyrin (F728)	The nearest distance to β-amyrin (S412)	Distance between 734 and F728	Distance between 734 and S412
WT	4	2.54	3.74	4.03	2.90
G	6.30	3.23	3.60	3.75	6.73
A	5.69	3.37	3.52	3.95	3.67
V	6.50	3.54	3.11	3.98	3.83
I	4.97	3.39	3.21	3.59	4.35
F	4.48	3.21	3.41	4.11	4.06
Y	4.07	3.37	3.28	3.91	3.20
W	3.66	3.48	3.12	3.41	4.10
D	3.26	3.67	3.22	2.97	3.19
N	4.84	3.45	3.18	2.83	5.02
Q	4.96	3.78	3.38	3.72	4.28
E	4.14	3.62	3.85	3.89	2.67
H	4.85	3.45	3.23	3.92	5.13
K	4.58	3.34	3.33	3.66	3.57
R	3.28	3.48	4.22	3.88	3.50
S	5.10	3.38	3.33	3.70	4.97
T	5.70	3.78	4.10	2.91	3.14
C	6.14	3.58	3.60	3.86	3.88
M	4.14	3.37	3.62	3.59	3.98
P	6.16	3.49	3.13	4.02	4.40

Table 3.7 The variation of F728 and S412 within PSY^{L734X} mutants. The unit is Å.

Chapter 4 Conclusions

Site-directed mutagenesis is a useful molecular biological technique on the studies of enzymes. The relationships of structure-function-mechanism of specific residues within a putative active site could be understood in depth using site-directed mutagenesis for detailed analysis. In addition, the chemical compounds produced by mutated residues could provide information for its functional roles in the cyclization/ rearrangement reaction. In this study, we chose ERG7[I705] and PSY[L734] residues for detailed analyses. These two residues are located at the relative positions that I705 occupies in *Saccharomyces cerevisiae* oxidosqualene-lanosterol cyclase, corresponding to the L734 residue within *Pisum sativum* β-amyrin synthase. They both have hydrophobic properties and are highly-conserved in many species. The experimental results showed that these two residues are crucial within the cyclases. Herein, we summarize several important conclusions of our research.

4.1 Analysis of *S. cerevisiae* ERG7[I705X] mutations

(1) Construction of saturated mutations of ERG7[I705X] from *Saccharomyces cerevisiae* and the results of genetic analysis showed that most of the I705X mutations could complement ergosterol-deficient growth except for the I705Y, I705W, I705H, I705E and I705R mutants.

(2) Seven products were generated in ERG7[I705X] mutations with a molecular mass of $m/z =$ 426, including lanosterol, (13α*H*)-isomalabarica-14*E*,17*E*,21-trien-3β-ol, (13α*H*)-isomalabarica-14*Z*,17*E*,21-trien-3β-ol, protosta-13(17),24-dien-3β-ol, (13α*H*)-isomalabarica-14(26),17,21-trien-3β-ol, 17α-protosta-20(21),24-dien-3β-ol and a novel product, 17α-protosta-20(22),24-dien-3β-ol.

(3) The novel product, 17α-protosta-20(22),24-dien-3β-ol, was identified by ^{1}H and ^{13}C NMR. Large amounts of the novel product were generated in the ERG7[I705F] mutation, and then were only observed in the ERG7[I705F] and ERG7[I705G] mutants. This is the second

chair-boat-chair 17α configurational structure isolated in our lab.

(4) The similar products profile of ERG7^{F699X} such as (13αH)-isomalabarica-14E,17E,21-trien-3β-ol, (13αH)-isomalabarica-14Z,17E,21-trien-3β-ol, (13αH)-isomalabarica-14(26),17,21-trien-3β-ol, protosta-13(17),24-dien- 3β-ol and 17α-protosta-20(21),24-dien-3β-ol, indicated that I705 has a strong relationship with the F699 residue. I705 certainly affects the first-tired residue within the putative active site of oxidosqualene-lanosterol cyclase.

(5) Two unnatural 17α-derivatives, 17α-protosta-20(22),24-dien-3β-ol and 17α-protosta-20(21),24-dien-3β-ol, were generated in the ERG7^{I705F} mutant simultaneously. The double mutants of I705F and F699X provide more evidence that I705 is not only a crucial residue to influence the first-tired amino acids, but also is essential for contribution towards the formation of 17α-products.

4.2 Analysis of *P. sativum* PSY^{L734X} mutations

(1) The plant PSY gene was expressed by the yeast system, TKW14C2. Through genetic selection, the saturated mutated mutants of PSY^{L734X} from *Pisum sativum* all have the negative results, meaning the exogenous ergosterol was required for growth.

(2) By identification with GC-MS, all of the L734X mutations did not generate any triterpenoid structures with molecule mass of $m/z = 426$. Without any truncated compounds and β-amyrin, the results reveal the importance of the L734 residue within the putative active site of *P. sativum* βAS.

(3) The catalytic cyclization fails in the mutations of L734X. PSY^{L734} may stabilize the substrate conformation with its neighboring amino acids, thus mutation of L734 would cause the disruption of catalytic cyclization.

(4) The functional roles of the neighboring amino acids are not confirmed with experimental results so that the detailed function of L734 could not be completely explained.

Chapter 5 Future prospects

It is remarkable that the replacement of the conserved Ile705 of ERG7 with mutated residue generated such a diversity of products. Experimental results show that ERG7[I705] may influence the terminal cyclization and stereochemistry, and specifically switch the stereochemistry of 17β into 17α orientation, hence a new evolution of enzymatic functions. New organic methods using analogs or isotope could be designed in order to figure out the detailed mechanism via the 17α protosteryl cation intermediate. We may construct the chimeric oxidosqualene cyclase in the future, mixing the functions of multiple cyclases. Moreover, the control of chair-boat-chair and chair-chair-chair substrate pre-folding conformations by cyclases is also unsolved and is a worthy goal to pursue.

In addition, the construction of a HEM1 ERG7 ERG1 triple knockout mutant in the yeast system could be developed for the *in vivo* experiments. By the addition of the substrate, this method could prevent the interference due to the downstream enzymes and provide more detailed information for the catalytic functions of cyclases. Furthermore, we also plan a major effort on the purification of yeast OSC, whereby the resolution of yeast OSC crystal structure will provide more information about the complex cyclization/rearrangement cascade.

Chapter 6 References

(1) Connolly, J. D.; Hill, R. A. *Nat Prod Rep* **2002**, *19*, 494-513.

(2) Abe, I.; Rohmer, M.; Prestwich, G. D. *Chem. Rev.* **1993**, *93*, 2189-2206.

(3) Xu, R.; Fazio, G. C.; Matsuda, S. P. *Phytochemistry* **2004**, *65*, 261-291.

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

(5) Phillips, D. R.; Rasbery, J. M.; Bartel, B.; Matsuda, S. P. *Curr Opin Plant Biol* **2006**, *9*, 305-314.

(6) R. B. Woodward, a. K. B. *J. Am. Chem. Soc* **1953**, *75*, 2023-2024.

(7) Maudgal, R. K.; Tchen, T. T.; Bloch., K. *J. Am. Chem. Soc* **1958**, *80*, 2589.

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

(9) Tamelen, E. E. v.; Willett, J. D.; Clayton, R. B.; Lord, K. E. *J. Am. Chem. Soc* **1966**, *88*, 4752.

(10) Corey, E. J.; Russey, W. E.; Montellano, P. R. O. d. *J. Am. Chem. Soc* **1966**, *88*, 4750.

(11) Van Tamelen, E. E.; Willet, J.; Schwartz, M.; Nadeau, R. *J Am Chem Soc* **1966**, *88*, 5937-5938.

(12) Barton, D. H. R.; Jarman, T. R.; Watson, K. C.; Widdowson, D. A.; Boar, R. B.; Damps, K. *J Chem Soc, Perkin Trans 1* **1975**, 1134.

(13) Ourisson, G.; Rohmer, M.; Poralla, K. *Annu Rev Microbiol* **1987**, *41*, 301-333.

(14) Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; D. Li, X. S. *J. Am. Chem. Soc* **1997**, *119*, 1277-1288.

(15) Corey, E. J.; Cheng, C. H. B. H.; Matsuda, S. P. T.; D. Li, X. S. *J. Am. Chem. Soc* **1997**, 1289-1298.

(16) Corey, E. J.; Staas, D. D. *J. Am. Chem. Soc* **1998**, *120*, 3526-3527.

(17) Johnson, W. S.; Telfer, S. J.; Cheng, S.; Schubert, U. *J. Am. Chem. Soc* **1987**, *109*, 2517-2518.

(18) Johnson, W. S.; Lindell, S. D.; Steele, J. *J. Am. Chem. Soc* **1987**, *109*, 5852-5853.

(19) Shi, Z.; Buntel, C. J.; Griffin, J. H. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7370-7374.

(20) Abe, I.; Bai, M.; Xiao, X. Y.; Prestwich, G. D. *Biochem Biophys Res Commun* **1992**, *187*, 32-38.

(21) Segura, M. J.; Jackson, B. E.; Matsuda, S. P. *Nat Prod Rep* **2003**, *20*, 304-317.

(22) Wendt, K. U. *Angew Chem Int Ed Engl* **2005**, *44*, 3966-3971.

(23) Jenson, C.; Jorgensen, W. L. *J. Am. Chem. Soc* **1997**, *119*, 10846-10854.

(24) 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.

(25) Joubert, B. M.; Hua, L.; Matsuda, S. P. *Org Lett* **2000**, *2*, 339-341.

(26) 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.

(27) Eschenmoser, A.; Ruzika, L.; Jeger, O.; Arigoni, D. *Helv. Chim. Acta.* **1955**, *38*, 1890-1904.

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

(29) Corey, E. J.; Virgil, S. C.; Sarshar, S. *J. Am. Chem. Soc* **1991**, *113*, 8171-8172.

(30) Hess, B. A., Jr. *Org Lett* **2003**, *5*, 165-167.

(31) R, R.; J., G. *J. Am. Chem. Soc* **2003**, *125*, 12768-12781.

(32) Xiong, Q.; Wilson, W. K.; Matsuda, S. P. *Angew Chem Int Ed Engl* **2006**, *45*, 1285-1288.

(33) Lodeiro, S.; Segura, M. J.; Stahl, M.; Schulz-Gasch, T.; Matsuda, S. P. *Chembiochem* **2004**, *5*, 1581-1585.

(34) Wu, T. K.; Griffin, J. H. *Biochemistry* **2002**, *41*, 8238-8244.

(35) Abe, I. *Nat Prod Rep* **2007**, *24*, 1311-1331.

(36) Kushiro, T.; Shibuya, M.; Masuda, K.; Ebizuka, Y. *J. Am. Chem. Soc* **2000**, *122*, 6816.

(37) Kushiro, T.; Shibuya, M.; Ebizuka, Y. *J. Am. Chem. Soc* **1999**, *121*, 1208.

(38) Wendt, K. U.; Poralla, K.; Schulz, G. E. *Science* **1997**, *277*, 1811-1815.

(39) Wendt, K. U.; Lenhart, A.; Schulz, G. E. *J Mol Biol* **1999**, *286*, 175-187.

(40) Reinert, D. J.; Balliano, G.; Schulz, G. E. *Chem Biol* **2004**, *11*, 121-126.

(41) Wu, T. K.; Chang, C. H.; Liu, Y. T.; Wang, T. T. *Chem Rec* **2008**, *8*, 302-325.

(42) Schmitz, S.; ll, C. F.; Glaser, T.; Albert, K.; Poralla, K. *Tetrahedron Letters* **2001**, *42*, 883-885.

(43) Sato, T.; Sasahara, S.; Yamakami, T.; Hoshino, T. *Biosci Biotechnol Biochem* **2002**, *66*, 1660-1670.

(44) Xiong, Q.; Rocco, F.; Wilson, W. K.; Xu, R.; Ceruti, M.; Matsuda, S. P. *J Org Chem* **2005**, *70*, 5362-5375.

(45) Freimund, S.; Kopper, S. *Carbohydr. Res.* **1998**, *308*, 195.

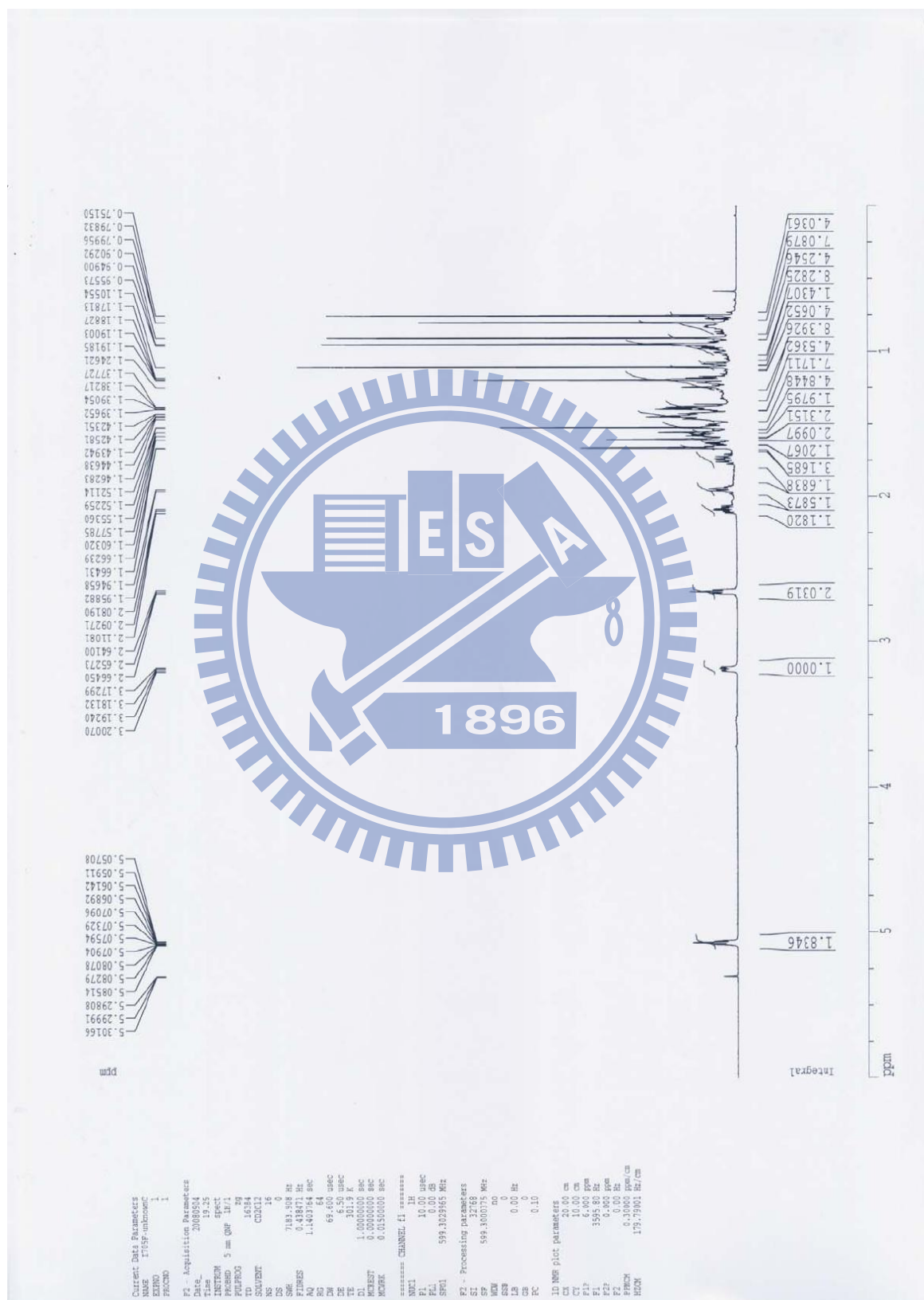(46) Freimund, S.; Kopper, S. *Carbohydr. Res.* **2004**, *339*, 217.

(47) 溫皓宇 *國立交通大學生物科技研究所* 中華民國九十六年, *碩士論文.*

(48) Wu, T. K.; Wen, H. Y.; Chang, C. H.; Liu, Y. T. *Org Lett* **2008**, *10*, 2529-2532.
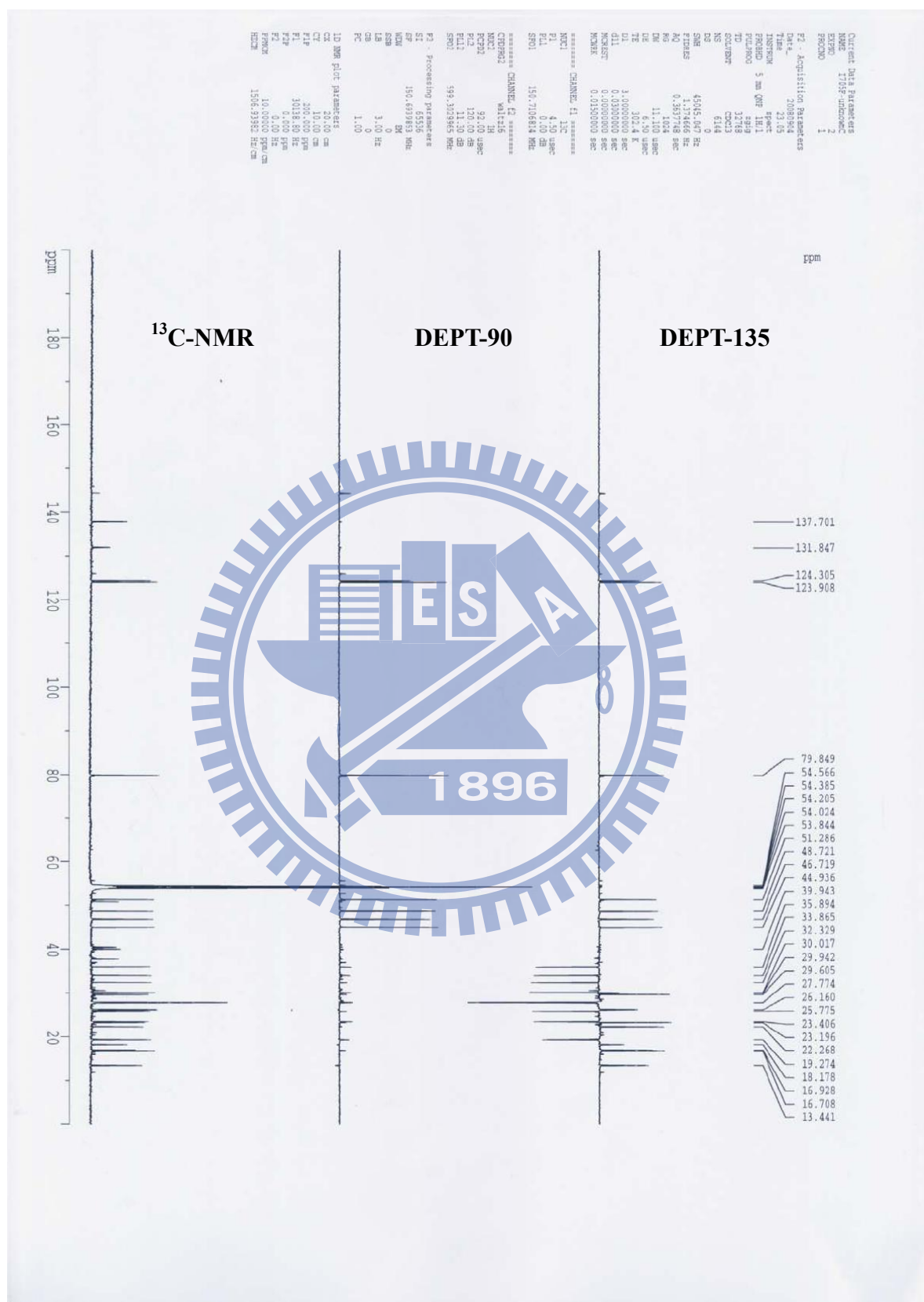
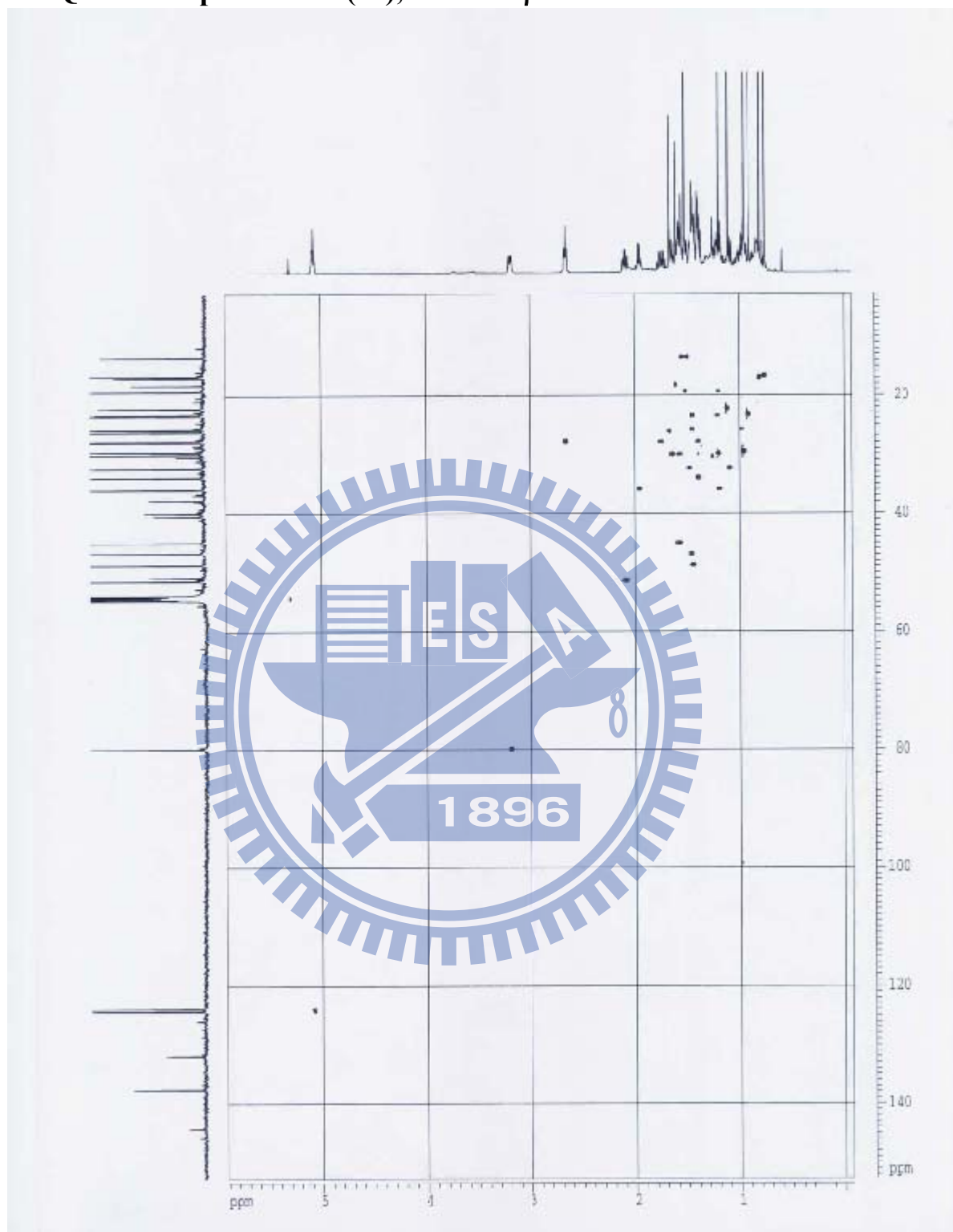(49) 謝文祥 *國立交通大學生物科技研究所* 中華民國九十七年, *碩士論文.*

Appendix

^{1}H-NMR of 17α-protosta-20(22),24-dien-3β-ol

^{13}C-NMR and DEPT of 17α-protosta-20(22),24-dien-3β-ol

^{13}C-NMR

DEPT-90

DEPT-135

137.701
131.847
124.305
123.908

79.849
54.566
54.385
54.205
54.024
53.844
51.286
48.721
46.719
44.936
39.943
35.894
33.865
32.329
30.017
29.942
29.605
27.774
26.160
25.775
23.406
23.196
22.268
19.274
18.178
16.928
16.708
13.441

HMQC of 17α-protosta-20(22),24-dien-3β-ol

HMBC of 17α-protosta-20(22),24-dien-3β-ol