

國立交通大學  
生物科技研究所  
博士論文

利用突變策略、抑制作用與 SELEX 技術  
針對氧化鯊烯環化酵素  
進行功能性之探討



**Functional Characterization of Oxidosqualene  
Cyclases Using Mutagenesis Approaches,  
Inhibition Studies, and SELEX Technology**

研究生：張程翔

**Student: Cheng-Hsiang Chang**

指導教授：吳東昆 博士

**Advisor: Prof. Tung-Kung Wu Ph.D**

中華民國九十七年七月

利用突變策略、抑制作用與 SELEX 技術針  
對氧化鯊烯環化酵素進行功能性之探討

**Functional Characterization of Oxidosqualene  
Cyclases Using Mutagenesis Approaches, Inhibition  
Studies, and SELEX Technology**

研究生：張程翔

Student: Cheng-Hsiang Chang

指導教授：吳東昆 博士

Advisor: Prof. Tung-Kung Wu Ph.D

國立交通大學  
生物科技研究所  
博士論文



A Dissertation

Submitted to Department of Biological Science and Technology

College of Biological Science and Technology

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

in

Biological Science and Technology

Hsinchu, Taiwan, Republic of China

July, 2008

中華民國九十七年七月

## 摘要

氧化鯊烯環化酵素(EC 5.4.99-) 為一群具有催化受質(3S)-2,3-氧化鯊烯((3S)-2,3-oxidosqualene)進行一系列的環化與重組反應之同源酵素。在絕大多數的物種中，氧化鯊烯環化酵素會負責專一性的生物轉化作用，進而生成多樣的固醇類或三萜類產物。受質所進行的環化機制與催化酵素本身結構間所產生的協同關係更是極度複雜與引人注目的。為了闡明受質-氧化鯊烯在機制上所進行的環化與重組反應，我們設計了許多實驗，包括分子生物學上的突變策略以及利用生物有機化學的方式來加以探討。

在一開始，我們利用“丙氨酸掃描式突變”以及“定點突變”的方式，並配合酵母菌體內的“質體交換篩選法”，來加以確定在酵母菌中氧化鯊烯環化酵素(*S.c.ERG7*) 其內部對於催化反應或結構穩定上重要的氨基酸位置。由實驗的結果發現，在我們所感興趣的酪氨酸-509 到異白氨酸-513 的這段區域上，酪氨酸-510 在突變效應的影響下產生非常迥異的催化活性。當酪氨酸-510 置換成色氨酸或離氨酸時，酵母菌的轉殖株便無法補充其體內本身的氧化鯊烯環化酵素上的缺失而導致菌株的死亡。但在換成丙氨酸時，酵母菌的突變株卻是可以存活。因此，我們進行了各轉殖株其體內非皂化酯醇的分離與鑑定，結果發現酵母菌氧化鯊烯環化酵素酪氨酸-510 的丙氨酸突變株 (*S. cerevisiae* ERG7<sup>Tyr510Ala</sup>) 可以產生單環的著醇 A (achilleol A)、四環的羊毛硬脂醇 (lanosterol) 以及 parkeol。而在致死的色氨酸或離氨酸的突變株中，酵母菌僅能產生兩個單環的產物分別為著醇 A (achilleol A) 或 camelliol C。為了更進一步的瞭解在不同的氨基酸改變下對於氧化鯊烯環化酵素酪氨酸-510 所造成突變效應，我們繼而利用“定點飽和突變”的方式來進行探討。其結果發現，不同的酪氨酸-510 突變株會產生單環、三環以及多種四環的生合成產物。此外，在原先所推測對於產物專一性非常重要的“催化氨基酸鹼性殘基對”(catalytic base-dyad)，即酪氨酸-510 與組氨酸-234 所形成的共

伴效應，可以藉由不同的酪氨酸-510 與組氨酸-234 的雙點突變來加以闡明。另一方面，我們亦利用類似的定點飽和突變結合生物有機化學的特性，發現更多對於在不同階段的環化與重組反應中相當重要的一些氨基酸位置，並分別鑑定出多個中間態產物的結構。為了仔細地瞭解這些氨基酸在酵素活性區域中的重要性，我們也建構了一系列酵母菌氧化鯊烯環化酵素的蛋白質同源模擬結構，進而合理地解釋這些關鍵性的氨基酸位置對於其單離產物間的關係。之後，我們亦利用量子力學的高斯軟體 Gaussian 03 來預測在不同的酵素突變株中，不同的產物其在理論能階上的相對應位置與其產生的趨勢。另一方面，我們將酵母菌中氧化鯊烯-羊毛硬脂醇環化酵素與阿拉伯芥中氧化鯊烯-環阿屯醇環化酵素進行一系列嵌合體酵素資料庫 (chimeric enzyme library) 的建構，期望能找出在這兩個酵素中專門與產物專一性有關的特定區域。因此，我們一共設計了十種不同的嵌合體 (chimeras)，並藉由“質體交換篩選法”來分析它們在酵母菌中的活性表現。但在所有的嵌合體轉殖菌株中，其體內的非皂化酯醇並沒有明顯的差異，這也許意味著過於粗糙的切割可能造成整體酵素在結構上的瓦解。之後，我們進一步地比較之前在三萜類合成酵素中，其嵌合體酵素資料庫的實驗數據後，我們認為在固醇類合成酵素中，產物的專一性也可能僅來自於酵素活性區域上少數幾個關鍵的功能性氨基酸。

除了分子生物學外，我們也同步進行了不同的生物化學實驗，並且將重點放在哺乳類動物中氧化鯊烯-羊毛硬脂醇環化酵素其結構與功能間的關係。我們主要是利用管柱層析、生物有機化學及不同的抑制劑修飾等實驗，對於牛肝中的氧化鯊烯環化酵素進行一系列的探討。在一開始，我們成功地由牛肝中將氧化鯊烯-羊毛硬脂醇環化酵素純化出來，並藉由串聯式質譜儀加以鑑定。之後，我們也順利地將整段序列解讀出來並將其表現在酵母菌中。而由氨基酸的比較上，我們發現牛肝中的氧化鯊烯環化酵素與其他三種哺乳類動物的氧化鯊烯環化酵素在序列上有百分之八十以上的相同度。然而，為了更進一步的瞭解其中一個非常有效的抗氧化鯊烯環化酵素之特性抑制劑-Ro48-8071 其作用機制與其明確的鍵結



位置，我們進行了光親和性的分析與設計了五個以 Ro48-8071 為骨架的螢光性衍生物。然而，由抑制酵素活性的實驗與螢光光譜特性上，我們發現化學螢光修飾會大幅減弱 Ro48-8071 的抑制活性。另外，我們利用分子入塢實驗(molecular docking) 也合理的解釋經修飾後的螢光化合物在氧化鯊烯環化酵素的活性區域中，對於酵素立體空間所產生的障礙關係以及經此修飾對於抑制能力的影響。希望未來能更進一步的研發新型式的 Ro48-8071 衍生型螢光抑制劑，除了能幫助氧化鯊烯環化酵素在結構與功能性質上進行探討，也能進一步提供在蛋白質體學上的應用，或對於在降膽固醇藥物的篩選上有所幫助。另一方面，我們也利用了一個隨機選擇性的組合式核苷酸資料庫篩選法-SELEX 對於氧化鯊烯環化酵素進行篩選，希望選擇出對其具有一定結合能力的最適體 (aptamer)。經過九次的篩選循環，我們成功的得到了多個最適體分子。針對這些最適體分子，我們除了證實了它們與牛肝中氧化鯊烯環化酵素的結合能力外並計算出其大約在 nM 範圍的親和力。然而最適體與氧化鯊烯環化酵素間的作用方式仍需更進一步的試驗來加以瞭解。期望這些與氧化鯊烯環化酵素有結合能力的分子，可以在未來不論是針對羊毛硬脂醇環化酵素或是針對膽固醇的生合成途徑皆能提供在醫療或診斷上的應用。

綜合以上所述，我們利用了不同的分子生物學實驗，包括丙氨酸掃描式突變、定點直接/飽和突變、區域置換實驗、同源模擬與量子力學等生物資訊的研究得以針對氧化鯊烯環化酵素其環化與重組反應在機制上有更深入的瞭解。另外一方面，藉由生物化學與生物有機化學的相互運用，我們更獲得許多寶貴的資訊。特別是在哺乳動物其氧化鯊烯環化酵素的取得，以及利用化學性修飾而得以針對抑制劑 Ro48-8071 其抑制作用有更進一步的瞭解。同時利用體外 SELEX 的實驗，更幫助我們對於未來針對合理地設計抗真菌類或降低膽固醇藥物的領域上開創一條展新的大道。

## Abstract

Oxidosqualene cyclases (EC 5.4.99-) constitute a family of enzymes that catalyze diverse cyclization/rearrangement reactions of (3*S*)-2,3-oxidosqualene (OS) into a distinct array of sterols and triterpenes. Notably, the product specificity among most of cyclase enzymes is species-dependent. The relationship between cyclization mechanism and enzymatic structure is extremely complex and attractive. In order to further elucidate the cyclization/rearrangement reaction of oxidosqualene cyclases, experiments including different molecular biological mutagenesis approaches as well as the bioorganic studies were accordingly carried out.

First, the alanine-scanning mutagenesis and site-directed mutagenesis coupled with *in vivo* plasmid shuffling selection were employed to identify the catalytic or structural important residues in oxidosqualene-lanosterol cyclase from *Saccharomyces cerevisiae* (*S. cerevisiae* ERG7). Among the investigated sequence segment from Thr-509 to Ile-513, Tyr-510 showed the catalytic discrepancy in the cyclase activity upon mutagenic effect. The yeast transformant failed to complement the cyclase-deficiency when this position was mutated to tryptophan or lysine residues, but still maintained the yeast viability in the *S. cerevisiae* ERG7<sup>Y510A</sup> mutant. After analysis of the nonsaponifiable lipid from *S. cerevisiae* ERG7<sup>Y510A</sup>, the monocyclic achilleol A, tetracyclic lanosterol and parkeol were identified. Moreover, two monocyclic compounds, achilleol A and camelliol C, were isolated from the lethal *S. cerevisiae* ERG7<sup>Y510K</sup> and *S. cerevisiae* ERG7<sup>Y510W</sup> mutants. In order to further investigate the mutated effects on this residue, the site-saturated mutagenesis was subsequently performed. Diverse products including monocyclic, tricyclic, and different tetracyclic products were isolated from the *S. cerevisiae* ERG7<sup>Y510X</sup> mutants. Moreover, the inherent influence on product specificity via an altered coordinative interaction between the hypothesized catalytic dyad, Tyr-510 and His-234, were further examined in more detail via construction and analysis of a different set of *S. cerevisiae* ERG7<sup>H234X/Y510X</sup> double mutations. Moreover, other catalytically important

residues and their respective premature cyclization products, involved in different cyclization/rearrangement stages, were also discovered by using similarly executed site-saturated mutagenesis, coupled with bioorganic characterization. In order to carefully explore the importance of these crucial residues within the enzymatic active site, plausible homology modeling structures were subsequently created. The diverse array of product profiles which were isolated from various mutated *S. cerevisiae* ERG7 cyclases was broadly representative. Moreover, the products' tendency in different mutated enzymes was consequently understood by using the quantum mechanics calculation of Gaussian 03. In addition, the combination of chimeric enzyme library between *Saccharomyces cerevisiae* lanosterol synthase and *Arabidopsis thaliana* cycloartenol synthase was also constructed to determine the critical functional domain responsible for the product specificity. Ten diverse domain swapping chimeras were successfully created, and their activities were subsequently confirmed via plasmid shuffling selection. No divergence of the nonsaponifiable lipid patterns was observed among these inactive chimeras, suggesting that the rough partition might disrupt the enzyme structure. After comparison with the previous experimental results from the triterpene synthases, the product specificity-determining residues among these sterol-biosynthetic cyclases might be determined by just several functional crucial residues within the enzymatic active site.

In parallel to the ongoing molecular biology approaches, we also performed a number of biochemical studies, including chromatographic purification, bioorganic characterization, and inhibition studies to examine the structure-function relationships for mammalian lanosterol cyclase. After successful purification and tandem mass characterization of bovine liver OSC, the gene encoding bovine liver OSC was subsequently determined. The deduced amino acid sequence showed >80% identity to that of the other three mammalian lanosterol synthases. The bovine liver OSC gene was also successfully cloned and functionally expressed in a yeast *erg7* disruption strain. Moreover, in order to better understand the inhibiting mechanism of one potent OSC inhibitor, Ro48-8071, as well as to solve the exact inhibitor binding site, the photoaffinity labeling and chemical fluorescent modification of Ro48-8071 was also

carried out. Several Ro48-8071-based fluorescent probes were developed and their inhibitory activity or fluorescence characteristics were analyzed. The results of chemical modification of Ro48-8071 suggested that the fluorescent Ro48-8071 derivatives dramatically reduced its inhibitory activity for purified bovine liver OSC. Moreover, the interactions between fluorescent Ro48-8071 derivative and the active site of bovine liver OSC, as well as the orientation of these probes have obviously changed, based on the molecular docking experimental data. In the future, improved site-specific fluorescent probes should be developed and applied to the chemical proteomic field for effectively screening the OSC drugs. By another approach, a randomly selected combinatorial approach, SELEX, was utilized for screening the potential OSC-binding aptamer molecules of the bovine liver OSC. After nine rounds of SELEX screening, a diverse array of aptamer candidates was isolated from the single-strand DNA library. These aptamer molecules exhibited the definitive interaction with the targeted protein and also revealed the approximate nM range affinity for bovine liver OSC. However, the binding interaction between individual aptamers and the cyclase protein should be explored in the future. These obtained OSC-binding aptamers will be applied in the pharmaceutical or diagnostic applications for lanosterol synthase as well as for studies in the cholesterol pathway in the future.

Thus, the combination of results obtained from the molecular biological approaches, including alanine-scanning, site-directed/saturated mutagenesis, domain swapping experiments, homology modeling structure, and quantum mechanics calculation, a better understanding of the cyclization/rearrangement mechanism of oxidosqualene cyclase has been achieved. In addition, the biochemical characterization coupled with the bioorganic studies toward mammalian lanosterol synthase provided valuable information, especially in obtaining the purified cyclase protein, and illustrating the inhibition mechanism of Ro48-8071. Moreover, the *in vitro* SELEX procedure opens new avenues in rational designing of antifungal agents and hypocholesteremic agents.

## 謝誌 (Acknowledgement)

每個人的一生就像一道道雷射光束般，雖然顏色、能量、強度都不盡相同，但在兩道光束交會時，卻會依倆人的緣份而綻放出繽紛的火花。宏觀地來看，雖然人與人的際遇就應像網狀結構般的雜亂無章。但，我卻在這近十年期間，與原本素昧平身的一群人，在交通大學生物科技學系吳東昆老師所領導的生物有機與分子演化實驗室內，交織出一道道最奇妙的光輝。

而這一切將會是我最難抹滅的回憶！

首先，最要感謝的是我的指導教授-吳東昆博士，願意在八年前收留一個當時正躊躇不安的大學畢業生進入他的實驗室。雖然來不及參與實驗室的草創，但我卻何其有幸地可以一路看著它成長茁壯，至今已成為人人稱羨的實驗室。常常看著滿屋子的實驗設備思考著，有著幾近外面數間實驗室總合的我們，倘若還是無法做出好的研究成果，那就真的太對不起一直為實驗室默默地付出的吳老師了。跟隨老師共同學習的這段期間，讓我感受到的是老師長期以來對於學術的熱忱與堅持。另外，他在實驗設計上的嚴謹與其對實驗結果的辯證和思維，也令我受益良多。每每困頓急墮時，老師時而的當頭棒喝，時而的經驗分享與砥礪，總能不知不覺地在我心中興起一股繼續向上的動力，不斷地激勵我繼續堅持。另外，老師在面對事情時，所表現出來的積極態度與不畏懼挫折的勇氣，更是我要好好學習的地方，我想這一路上能有老師的幫助與肯定，才是讓我能獲得學位的一大原因。

另外，從資格考開始到非論文考試，一直到最後的學位口試，一路陪著我度過的幾位老師包括李耀坤老師、張大慈老師、袁俊傑老師與鄭建中老師。感謝您們對於我在整體實驗邏輯上的慎密性，總是不吝地給予最中肯的辯證。而對於文法結構上的錯誤，也相當用心地加以簽正，並提供具體的指導及建議。老師們在口試期間所給予的討論與激盪，我想這是督促自己與檢視自己最直接的方法，也是促使自己能在研究上更摯嚴謹與完備。而老師們在學理上的雄厚底子，也是學生學習與效法的對象。

另外，生科系楊裕雄老師對於我在大學時期專題研究上的指導與支持，是讓我能順利地從化學背景轉行進入生物科技領域的關鍵。而應化系鍾文聖老師在有機化學上悉心地教授，更是讓我受益頗多。

而我的實驗之所以能順利的完成，在這一個過程中，必須要誠心地感謝許多參與其中並且不吝給予幫忙的很多人。中央研究院的陳玉如老師與廖信凱同學，對於我們在蛋白質的質譜鑑定上提供了完善且專業的合作，使得我們能順利的解決膜蛋白定序上的難題。應化系刁維光老師、駱立揚學長與張智煒同學，在超快雷射技術上以及理論化學上的輔助，更提供我們對於實驗結果上有了另一個角度之思維方式。生科系林玉淳學妹與陳彥甫學弟，在生物資訊學上的幫助與不厭其煩的指導，得以讓我拓展自己的研究領域。而應化系陳奎百同學、吳曜杉學弟與何孟寰學弟，對於我在有機合成上也給予了許多寶貴的建議與幫忙。而交大貴儀中心的李蘊明小姐在氣相層析儀上的輔助與專業的測量，也是讓我們收穫良多。另外特別要感謝清大貴儀中心的彭菊蘭小姐，您在 NMR 光譜上提供了我最實質的幫助。除了悉心地指導我使用 NMR 外，也非常體恤我們，總是額外地犧牲自己的假期來幫助我們進行光譜的測量，我們的期刊得以順利發表，您功不可沒。



而在我初入實驗室時，那種惶恐忐忑的心情至今尚令我記憶猶存。感謝當初幸如學姐、安堡學長與逢弦學長總是能以輕鬆風趣的方式來帶領我，使我能很快的進入狀況，對於實驗的本身也無私地傾囊相授，使我收穫良多。而乾吉學長與素華學姐，您的經驗傳承也在潛移默化中讓我對於我的未來有更明確的領悟。此外，我要特別感激喬盈對於我在實驗上與生活上的幫忙，我們曾共同地為了純化蛋白質而長期的焚膏繼晷，那種刻苦銘心的一再堅持與嘗試，真的讓我印象深刻。也謝謝妳在論文其中文部份的潤飾。而在生活上，妳也是一直給予我最溫暖的支持，讓我得以堅強地唸完博士，真的謝謝妳在這一段期間的鼓勵。而青穎與任民是跟我一起進入實驗室的同學，感謝你們跟我共同地分擔當時的大小庶務。你們現在都有好的工作與學業，身為同學的我也與有榮焉。

再來則是要感謝震宇學弟，也是因為你的推薦才讓我選擇進入吳老師實驗室。在實驗上，你總是跟我一起分擔實驗室的工作，也是當時我看棒球的最佳伙伴。貼心的你總是會適時地給予我久違的關心與問候。另外，媛婷學妹則是跟隨我最久的弟子，反應敏捷的妳總是給我最多的激盪，與妳共同建立的實驗流程，以及絞盡腦汁後所鑑定出來的化合物，我想是我們留在實驗室的寶藏，期待妳可以進一步地發揚光大，記得要適時的肯定自己-妳是很優秀的。裕仁學弟則是我見過最為聰明的孩子，跟我辛苦地完成專題研究也不時的給予建議，並幫助我建立了一系列的基因資料庫，期望你可以早日突破困境完成博士學位。

吳老師實驗室與我同屬 Cyclase Project 的每一個學弟妹，也都是我在博士班生涯的好幫手。分生部份的媛婷、美婷、文暄、皓宇、怡伶、文祥與采婷，你們辛苦地培養與分離實驗來源，並與我共同地進行光譜分析，才使我們得以成功的鑑定出幾個關鍵結構，也才能發表不錯的結果，其箇中的辛勞是不可言喻的。喬盈、震宇、佳宜、大景、皓宇、裕仁與柏輝更是不辭辛勞並無願無悔的陪同我徹夜地固守純化的崗位。我們一同守護著這些實驗室得來不易的珍寶，有時令人洩氣沮喪的結果，有時歡欣鼓舞的收割，這一切一切的滋味，你們都陪我走過。雖然遺憾的是一直沒有好的結果可以報導，但我深深的覺得純化組在我心中一直是最勞苦功高的。

而其他分屬不同計畫的學弟妹，也分別給予許多有形或無形的幫助。其中裕國一直是老師不可獲缺的另一個支柱，你在細胞與抗體技巧上的專業也是我協尋幫忙與請益的對象，特別感謝你與文鴻幫忙建構出我們的單株抗體。晉豪則是幫我在儀器設備上做了無法取代的分憂，尤其是在電腦資訊上的長才，也常幫助我化解因為我本身沒有 update 自己資訊技術，而無法解決老師問題的窘境。文鴻，謝謝你替我分擔了許多庶務管理上的責任，你是一個不可多得，並執著努力且具有責任感的人。雖然偶遇困境，但希望你能自我突破，以後不管做什麼都能大放異彩。登源則是一個幽默風趣，但又不擾其內心沉著與冷靜，並充滿智慧的孩子。一直以來，你常與我在實驗室外對於實驗有令人印象深刻的討論。生活上，你是我的室友也給予了很多的幫助，謝謝你。聖慈，妳是一個對自我有高度期許的孩子，對於實驗室的營運也富含想法，謝謝妳提出的許多寶貴建議，也盼妳能帶領這一個實驗室走出另一個高鋒，對於研究也希望妳能百尺竿頭。Mili is a kindly friend in our lab. Thank your for your joining, concern, and friendship. 另外，謝謝衣鵬平日與我的討論，使我獲益良多，期許妳能早日發表論文，實驗也越來越順遂。

怡親小妹，妳是個認真有想法的孩子，雖然有時後會想太多。但妳積極進取與上進獨立的心態卻是無法抹煞的。謝謝妳跟大景、文暄、文祥在這一年多來所帶來的歡樂，學長也預祝妳在未來博士生涯可以有一個好的開始、好的過程與順利的結果，並常保愉快。另外還有一些已經畢



業的碩士班學弟妹，其中，宏城在電腦模擬上的指導，庭翊對於實驗上的執著認真，宏明所帶來的歡樂以及其他許多的學弟妹，一直都是我向上提升的一股正面力量。另外，目前還在實驗室努力的良璋、天昶、育勳、禕庭與亦諄，你們在許多雜事上的分擔也是辛勞，期待你們能再多放點心思在實驗室上，讓它越來越好，謝謝你們。

而在生科所這一路走來，我遇到的許多貴人，你們總是在最適當的時機給我幫忙，使我不致於停滯不前。其中，盈蓉學姐是我最後一個階段的戰友，不時給我打氣與鼓勵，也很阿薩力的幫忙修改我的英文論文，謝謝妳！也希望妳能一直都很開心並舞藝精進。靖婷學姐則是給予了許多實驗技巧上的指導與幫忙，尤其是放射性實驗的部份真得讓我受益匪淺。美惠學姐與建龍學長則是在我初入研究所時，總是不吝的與我分享經驗與促膝長談，使我能早日的下定決心。另外，也因為因緣際會的關係，我可與你們同屆畢業，我也殊感開心。恩仕學長與天木學長則是在我專題研究時的指導學長，謝謝你們在當時的悉心教授，讓我有一個好的開始也明瞭如何伴演學長的角色。繼鋒學長與文亮則是我談天的對象，他們努力的方式也令我深感慚愧。另外，生科系辦的郭珍佑小姐、呂聖鈴小姐、郭淑卿小姐、吳佳文小姐以及葉連發先生總是在許多事務上對我給予幫助，謝謝你們。而我們的師母-賴美伶小姐一直以來，總是在實驗上與經驗上給予我許多的寶貴建議，您親切與開朗的態度也常伴演我們與老師間的潤滑劑，謝謝您。

大學同學伯凱是我在生物科技方面的嚮導員，謝謝你總是很熱情地丟給我很多東西，希望你身體越來越好並早日榮獲學位。春發、秉恒、志豪、哲一則是我一直以來的好友，另外明妮、姿惠、怡珊、智華對於我的博士班生活也常保關心。小黃家族的每一個成員都是非常在乎我的人，也是我這段期間能夠放鬆休憩的地方，你們的關懷我都充滿感激，也都是我心中的力量。

最後，更要感謝的是我的家人-我的父母親、兩個弟弟還有我的奶奶以及已經過世的爺爺。父親的身教與言教是督促我向上努力的最大後盾。您當初為了我的學業，拖著疲憊的身體參與考試並通勤上班，至今的記憶我依然清晰難忘。母親的關心則是我心中最大的暖流，是我永遠的避風港，我終於可以告訴您：妳的傻兒子終於拿到學位了。兩個聰明的弟弟則是我的驕傲，你們一直都比我優秀也很有成就，是身為大哥的我能安心唸書的原因，謝謝你們。爺爺與奶奶永遠都是最疼我們的人，也是我心中最為牽掛的人。

在交大十二年的光陰要感謝的人實在是太多了，深怕不慎有所遺漏。若有為之，真的甚感抱歉！感謝上蒼賜與的這一切。感謝所有關心我、陪伴著我的人，僅能以小小的版面真的不足以聊表我最深忱的謝意。未來的我會更加的努力，也在此給與您們最大的祝福。

願大家都能平安、順心、健康、愉悅。

程翔

2008/7/25

# Table of Contents

<b>Chapter 1. General introduction.....</b>	<b>1</b>
1.1 Overview of oxidosqualene cyclase .....	1
1.2 Historical hypothesis of cyclization mechanism.....	6
1.2.1 The mechanistic and stereochemical insights of oxidosqualene cyclization.....	6
1.2.2 The theoretical enzyme-templated cyclization reactions .....	15
1.3 The biological studies of 2,3-(oxido)squalene cyclase enzymes .....	21
1.3.1 Enzymology of 2,3-(oxido)squalene cyclase enzymes .....	21
1.3.2 Molecular cloning of 2,3-oxidosqualene cyclase.....	23
1.3.3 Crystallization and structural characterization of cyclase.....	27
<b>Chapter 2. Theis organization .....</b>	<b>33</b>
<b>Chapter 3. Mutagenesis approach to investigate the putative active-site residues from     oxidosqualene cyclase.....</b>	<b>35</b>
3.1 Introduction .....	35
3.2 Results and Discussion. ....	42
3.2.1 Generation of the alanine-scanning mutations or the site-directed mutations of <i>S. cerevisiae</i> ERG7 gene.....	42
3.2.2 Principle of plasmid shuffle methodology .....	43
3.2.3 Screening of inactive ERG7 mutants by using the plasmid shuffle method .....	45
3.2.4. Lipid extraction, column chromatography, and product characterization .....	46
3.2.5 The construction of novel gene disruption strain TKW14 for analysis of inactive mutated cyclase.....	49
3.2.6 Characterizing the mutant products in the novel gene disruption strain .....	51
3.2.7 Site-saturated mutagenesis approach to investigate the functional importance of the critical Tyr-510 residues.....	55
3.2.8 Homology modeling illustration of the functional critical residues .....	64
<b>Chapter 4. Construction of chimeric oxidosqualene cyclase gene to study the product specificity     .....</b>	<b>72</b>
4.1 Introduction. ....	72
4.2. Results and Discussion. ....	76
4.2.1 Construction of a chimeric library between oxidosqualene-lanosterol cyclase and oxidosqualene-cycloartenol synthase.....	76
4.2.2 Functional analysis of chimeric enzymes library via plasmid shuffle and the product characterization .....	80
<b>Chapter 5. Mechanistic insights into oxidosqualene cyclization through homology modeling and     quantum mechanism calculation.....</b>	<b>86</b>
5.1 Introduction. ....	86
5.2 Results and Discussion. ....	87
5.2.1 Sequence alignment, secondary structure prediction, and homology modeling .....	87
5.2.2 Geometrical verification of the generating homology models .....	91
5.2.3 Comparison between homology models and mechanistic hypothesis.....	95
5.2.4 Correlation of the OSC model to the experimental mutagenesis studies .....	106
5.2.5 Quantum mechanism calculation for illustration of the lanosterol synthesis.....	121
5.3 CONCLUSION AND FUTURE PERSPECTIVE. ....	125

<b>Chapter 6. Purification, tandem mass characterization, and molecular cloning of oxidosqualene lanosterol cyclase enzyme from bovine liver .....</b>	<b>130</b>
6.1 Summary.....	130
6.2 Introduction.....	131
6.3 Results and Discussion.....	135
6.3.1 Purification of bovine liver OSC.....	135
6.3.2 Identification of OSC by tandem mass spectrometry.....	139
6.3.3 Molecular cloning of OSC from bovine liver.....	142
6.3.4 Genetic complementation of recombinant bovine liver OSC in yeast erg7 deficient strain: .....	147
6.3.5 Mass spectrometry determination of the photoaffinity labeling bovine liver OSC with Ro48-8071.....	147
6.3.6 Monoclonal antibody preparation and Western blotting analysis.....	150
<b>Chapter 7. Development and applization of fluorescent Ro48-8071 derivative probes to study oxidosqualene-lanosterol cyclase.....</b>	<b>154</b>
7.1 Summary.....	154
7.2 Introduction.....	155
7.3 Results and Discussion.....	157
7.3.1 Synthesis of newly Ro48-8071-based fluorescent probes.....	157
7.3.2 The inhibitory activity of the Ro48-8071-based fluorescent probes.....	162
7.3.3 Spectroscopy of these newly Ro48-8071-based fluorescent probes.....	165
7.3.4 Homology modeling and ligand docking studies of newly Ro48-8071-based fluorescent probes.....	170
7.3.5 Time-resolved fluorescence studies of Ro48-8071-based fluorescent probes.....	173
<b>Chapter 8. Identification of DNA aptamers for bovine liver oxidosqualene-lanosterol cyclase via SELEX.....</b>	<b>176</b>
8.1 Introduction.....	176
8.2 Results and Discussion.....	184
8.2.1 In vitro selection of DNA aptamer for bovine liver OSC.....	184
8.2.2 Gel mobility shift assay of the binding between DNA aptamer and OSC.....	187
8.2.3 Gel mobility shift assay provide alternative selection strategy for SELEX.....	190
8.2.4 Molecular cloning and DNA sequencing analysis of individual aptamer.....	191
8.2.5 The binding characterization of the individual aptamer molecules.....	199
8.3. Conclusion.....	200
<b>Chapter 9. Experimental section.....</b>	<b>201</b>
9.1 Materials:.....	201
9.1.1 General bacterial strains and molecular cloning/expression vectors:.....	201
9.1.2 Enzymes, chemicals, equipments, and reagents:.....	201
9.2 General experimental procedures:.....	206
Chapter 3.....	206
Chapter 4.....	210
Chapter 5.....	211
Chapter 6.....	213
Chapter 7.....	219
Chapter 8.....	221

## Table of Figures

Figure 1.1 The product diversity of (oxido)squalene cyclase in different species.....	2
Figure 1.2 The cyclization cascades of oxidosqualene to lanosterol and the end product, cholesterol in the human cholesterol biosynthetic pathway.....	3
Figure 1.3 The cyclized products are classified according to the stereochemistry configuration of carbonocationic intermediates.....	4
Figure 1.4 The postulated cyclization mechanism of oxidosqualene .....	5
Figure 1.5 Squalene cyclization model for lanosterol production.....	7
Figure 1.6 The acyclic azasqualenes mimic the C-8 and C-20 carbonium ion intermediates.....	9
Figure 1.7 Structure of potent inhibitors of oxidosqualene cyclase enzymes.....	9
Figure 1.8 Some of the substrate derivatives inhibitors.....	10
Figure 1.9 The geometrical isomers of substrate analogs.....	11
Figure 1.10 Minor products with a 6,6,5,4 tetracyclic scaffold and 6-6-5 scaffold were isolated from the enzymatic transformation with various analogs.....	13
Figure 1.11 Cyclization of oxidosqualene analogs produce 17 $\beta$ -protosteryl derivatives .....	13
Figure 1.12 The hypothesized cyclization cascade of oxidosqualene he acyclic .....	14
Figure 1.13 Johnson proposed three cation-stabilizing auxiliaries in the cyclase. ....	15
Figure 1.14 Hypothetical model based on the cyclization mechanism. ....	17
Figure 1.15 Potential modes of mechanism-based inactivation.....	19
Figure 1.16 Affinity labeled cyclase fragment with mechanistic suicide inhibitors.....	19
Figure 1.17 Structures of various squalene and dodecyl derivatives of thiol-modifying agents.....	20
Figure 1.18 Overall structure of <i>A. acidocaldarius</i> squalene-hopene cyclase (SHC).....	30
Figure 1.19 Crystal structure of human OSC .....	31
Figure 1.20 Local view of crystallized structure of human OSC.....	32
Figure 3.1 Thin-layer chromatographic analysis of the nonsaponifiable lipid extract.....	48
Figure 3.2 Acetylation with acetic anhydride/ pyridine was used to modify the polarity of alcohol moiety.....	48
Figure 3.3 The homology recombination strategy for exchanging the HEM1 gene with a drug geneticin resistance gene .....	50
Figure 3.4 GC analyses of the nonsaponifiable lipid (NSL) products .....	53
Figure 3.5 The X-ray crystallized structure of (a) SHC or (c)OSC and the respective homology modeling structures of (b) wild-type ERG7 and (d) human OSC .....	67
Figure 3.6 Local view of the homology modeled <i>S.cerevisiae</i> ERG7 structure .....	68
Figure 3.7 Local view of superimposition of homology modeled <i>S.c.</i> ERG7 structure and different single-point mutations or double-positions mutations. ....	70
Figure 3.8 Local views of Asp-456, Trp-443, Lys448, and Phe445 positions in <i>S. cerevisiae</i> ERG7 structure. ....	71
Figure 4.1 The analysis of the preliminary NSL extract from respective ten chimeras on the TLC plate. ....	82
Figure 4.2 The GC-FID spectrums of the partially purified NSL from the TKW14 strain transformed with the respective chimera enzymes.....	82
Figure 4.3 Sequence alignment and secondary structure elements comparison between $\beta$ -amyrin synthase and lupeol synthase or lanosterol synthase and cycloartenol synthase.....	85
Figure 5.1 Multiple sequence alignment and secondary structure elements comparison .....	90

Figure 5.2 The 3D-1D profiles of the homology modeling structure of yeast ERG7 (Upper plot) and human OSC crystal structure (Downer plot).....	93
Figure 5.3 The Ramachandran plot of the homology model of yeast S.c.ERG7 .....	94
Figure 5.4 Superimposition of the S.c.ERG7 homology modeling structure with crystal structures .....	96
Figure 5.5 Local views of the superimposition of homology modeled S.c. ERG7 structure with human OSC crystal structure .....	97
Figure 5.6 Local views of the superimposition of homology modeled S.c. ERG7 structure with SHC crystal structure.....	100
Figure 5.7 The hypothetical enzyme-triggered steric control for the formation of the boat B-ring and Local view of S.c.ERG7 homology modeling structure. ....	103
Figure 5.8 The different orientation of Trp232 from S.c.ERG7 and Phe601 from A.a.SHC involved in the formation of C ring. ....	104
Figure 5.9 Local view of the homology modeling S.c.ERG7 structure.....	106
Figure 5.10 The chemical structure of the possible producing cationic reactive intermediates or various annulated products. ....	108
Figure 5.11 The isolated respective products from various S.c ERG7X mutates in this thesis .....	109
Figure 5.12 Stereo representations of <i>S. cerevisiae</i> wild-type ERG7 homology modeling structure ...	112
Figure 5.13 The steric representation of various S.c ERG7 mutated homology modeling structures. ....	114
Figure 5.14 Putative functional residues in the $\pi$ -electron pocket of S.c ERG7, in which Phe-445 is near the B/C ring junction.....	117
Figure 5.15 Local view of the homology model of S.c. ERG7.....	118
Figure 5.16 The partially relative energy profiles among various intermediate stages, derived from quantum mechanics calculations.....	124
Figure 6.1 Silver-stained SDS–PAGE gel and Coomassie-stained non-denaturing PAGE gel of bovine liver OSC .....	138
Figure 6.2 Tandem mass spectrum of the doubly charged tryptic peptide from purified bovine liver OSC protein. ....	140
Figure 6.3 Sequence alignment of trypsin-digested peptides from bovine liver OSC observed by tandem mass spectrometry. ....	141
Figure 6.4 The nucleotide and deduced amino acid sequences of cDNA from bovine liver OSC .....	144
Figure 6.5 Multiple sequence alignments of oxidosqualene-lanosterol cyclase.....	146
Figure 6.6 TLC analysis of cyclase activity before and after incubated with Ro 48-8071 .....	149
Figure 6.7 Peptide coverage of bovine liver OSC from LC-MS/MS analysis.....	150
Figure 6.8 Crystal structure of human OSC in complex with lanosterol And Homology modeling structure of bovine liver OSC .....	152
Figure 6.9 (a) SDS-PAGE and (b) Western blotting analysis of purified bovine liver OSC.....	153
Figure 7.1 TLC analysis of cyclase activity from which OSC was pre-incubated with different Ro48-8071-based derivatives.....	164
Figure 7.2 The wavelength scanning of fluorescent probes in aqueous solution.....	166
Figure 7.3 The fluorescence spectrum of Ro4-DPA fluorescent probe in aqueous solution, ethanol, and DMSO.....	168
Figure 7.4 Fluorescence characteristic of Ro48-8071-based fluorescent probes on either (a) TLC plate, or (b)in the aqueous and in the DMSO solution.....	169
Figure 7.5 Fluorescence spectra of Ro4-DPA with addition of bovine liver OSC.....	169
Figure 7.6 The molecular ligand docking results of the (a) Ro4-NA1, (b) Ro4-NA2, (c) Ro4-BP3, and (d) Ro4-BP4 for the bovine liver OSC active site cavity. ....	172
Figure 7.7 The molecular ligand docking results of the Ro4-DPA for the bovine liver OSC active site cavity.....	173
Figure 7.8 (a) The laser induced fluorescence decay or (b) the anisotropy analysis of the Ro4-NA1	



fluorescent probes.....	175
Figure 8.1. The schematic diagram for the selection of DNA aptamer to target molecule, i.e. the general SELEX procedure .....	179
Figure 8.2. PCR amplicon of the resultant DNA which was eluted from the surface of the microtiter plates during SELEX-3rd-cycle .....	186
Figure 8.3 Gel mobility shift assay of different amount of DNA probe interacted with the fixed bovine liver OSC .....	188
Figure 8.4 Gel mobility shift assay of fixed amount of DNA probe with the various amount of bovine liver OSC .....	188
Figure 8.5 Gel mobility shift assay of competition experiment by using the unlabeled DNA probe. ..	189
Figure 8.6 (a), (b) Gel mobility shift assay of DNA molecules isolated from microtiter plate for bovine liver OSC .....	190
Figure 8.7 Sequence alignments of the respective aptamer clones.....	193
Figure 8.8 The representative predicted secondary structures of aptamer DNAs from the SELEX-9th-cycle via the mfold server prediction .....	196
Figure 8.9 Gel mobility shift assay of individual DNA aptamer molecules for bovine liver OSC.....	197
Figure 8.10 The concentration dependent OSC-binding response was revealed on the SPR analysis of aptamer-125-9-13.....	199

## Table of Schemes

Scheme 3.1 The respective <i>A. thaliana</i> CAS1 mutants changed their product specificity from cycloartenol to lanosterol.....	37
Scheme 3.2 Screening of inactive ERG7 mutants using a plasmid shuffle method.....	44
Scheme 3.3 Oxidosqualene-lanosterol cyclase mutants converted oxidosqualene to a variety of monocyclic and deprotonated compounds .....	54
Scheme 3.4. Proposed cyclization/rearrangement pathway of oxidosqualene in TKW14C2 expressing ERG7 <sup>Y510X</sup> site-saturated mutagenesis.....	60
Scheme 4.1 The schematic representation of the partial chimera library between $\beta$ -amyrin synthas and lupeol synthase enzymes.....	74
Scheme 4.2 The schematic representation of the entire chimerical library between oxidosqualene-lanosterol cyclase gene (OSC) and oxidosqualene-cycloartenol cyclase gene (CAS) .....	78
Scheme 4.3 The diagrammatic explanation for the PCR-based construction strategy .....	79
Scheme 5.1 The hypothesized cyclization mechanism of oxidosqualene or squalene .....	98
Scheme 5.2 The proposed model of Tyr-99 for the stereochemical control.....	119
Scheme 6.1 Three major photoreactive groups .....	134
Scheme 6.2 Formation of covalent adducts from photoexcited benzophenone group.....	134
Scheme 7.1 Synthesis routine of the aromatic boronic pinacol ester.....	158
Scheme 7.2 The scheme shows the first published Suzuki coupling, which is the palladium-catalyzed cross coupling between organoboronic acid and halides .....	159
Scheme 7.3 Derivation reaction of R048-8071 with DPA or other four boronic pinacol ester compounds .....	160
Scheme 7.4 The chemical structures of five newly R048-8071-based fluorescent probes.....	161



## Table of Tables

Table 1.1 OSCs that have been cloned and functionally characterized in yeast .....	26
Tabel 3.1 Complementation result of cyclase mutants in a ERG7 knockout strain, CBY57, by using plasmid shuffling counter selection. ....	46
Table 3.2 Product ratios of the <i>Saccharomyces cerevisiae</i> OSC single mutants.....	54
Table 3.3 The product profiles of <i>S. cerevisiae</i> TKW14 expressing the ERG7 <sup>Y510X</sup> site- saturated mutants and two ERG7 <sup>H234WY510X</sup> mutants.....	56
Table 5.1 Statistical occurrence of $\phi$ , $\psi$ angles in the PDB bank compared with the $\phi$ , $\psi$ angles distribution in the homology modeling structure of yeast S.c.ERG7.....	94
Table 6.1 Purification and properties of 2,3-oxidosqualene cyclase from vertebrates, higher plants, and yeast .....	137
Table 6.2 The purification of oxidosqualene-lanosterol cyclase from bovine liver .....	137
Table 7.1 The inhibitory activity assay of the newly R $\alpha$ 48-8071-based fluorescent probes for bovine liver OSC .....	163
Table 8.1 The comparison of functional characteristics between aptamers and antibodies.....	180
Table 8.2. The concentration of OSC or DNA library load into the each SELEX round. ....	186
Table 9.2.4.1 Primers sequence in the chimerical library construction.....	211

## List of Abbreviations

<b>A, Ade</b>	<b>Adenine</b>
<b>APS</b>	<b>Ammonium Persulfate</b>
<b>bp</b>	<b>Base Pair</b>
<b>BSA</b>	<b>Bovine Serum Albumin</b>
<b>cDNA</b>	<b>Complementary DNA</b>
<b>dH<sub>2</sub>O</b>	<b>Distilled Water</b>
<b>ddH<sub>2</sub>O</b>	<b>Double Distilled Water</b>
<b>DEPC</b>	<b>Diethylpyrocarbonate</b>
<b>DMF</b>	<b>Dimethylformamide</b>
<b>DMSO</b>	<b>Dimethylsulfoxide</b>
<b>DNA</b>	<b>Deoxyribonucleic acid</b>
<b>dsDNA</b>	<b>Double-stranded DNA</b>
<b>dNTP</b>	<b>Deoxynucleoside triphosphate</b>
<b>DPA</b>	<b>4-(4,5-diphenyl-1H-imidazol-2-yl) phenylboronic acid</b>
<b>DTT</b>	<b>Dithiothreitol</b>
<b>EDTA</b>	<b>Ethylenediaminetetraacetic acid</b>
<b>EMSA</b>	<b>Electrophoresis Gel mobility shift assay</b>
<b>ELISA</b>	<b>Enzyme-Linked Immunosorbent Assay</b>
<b>GC</b>	<b>Gas Chromatography</b>
<b>H, His</b>	<b>Histidine</b>
<b>HAP</b>	<b>Hydroxyapatite</b>
<b>HCl</b>	<b>Hydrogen Chloride</b>
<b>HPLC</b>	<b>High Performance Liquid Chromatography</b>
<b>IPTG</b>	<b>Isopropyl-1-thio-<math>\beta</math>-D-galactopyranoside</b>
<b>kb(s)</b>	<b>kilobase(s)</b>
<b>kDa(s)</b>	<b>kiloDalton(s)</b>
<b>KPi</b>	<b>Potassium phosphate</b>



<b>L, Lys</b>	<b>Lysine</b>
<b>LA</b>	<b>Lanosterol</b>
<b>LB</b>	<b>Luria-Bertani</b>
<b>LiOAc</b>	<b>Lithium Acetate</b>
<b>mAb</b>	<b>monoclonal antibody</b>
<b>M, Met</b>	<b>Methionine</b>
<b>MS</b>	<b>Mass Spectromerty</b>
<b>NaCl</b>	<b>Sodium Chloride</b>
<b>NaN<sub>3</sub></b>	<b>Sodium Azide</b>
<b>NMR</b>	<b>Nuclear Magnetic Resonance</b>
<b>NaOH</b>	<b>Sodium Hydroxide</b>
<b>NC</b>	<b>Nitrocellulose</b>
<b>OS</b>	<b>Oxidosqualene</b>
<b>pAb</b>	<b>polyclonal antibody</b>
<b>PBS</b>	<b>Phosphate Buffered Saline</b>
<b>PCR</b>	<b>Polymerase Chain Reaction</b>
<b>PEG</b>	<b>Polyethylene Glycol</b>
<b>PMSF</b>	<b>Phenylmethysufonyl fluoride</b>
<b>PPh<sub>3</sub></b>	<b>Triphenylphosphine</b>
<b>PVDF</b>	<b>Polyvinylidene difluoride</b>
<b>RT</b>	<b>Room Temperature</b>
<b>rpm</b>	<b>revolutions per minute</b>
<b>SDS-PAGE</b>	<b>Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis</b>
<b>SELEX</b>	<b>Systematic Evolution of Ligands by Exponential Enrichment</b>
<b>SPR</b>	<b>Surface Plasmon Resonance</b>
<b>ssDNA</b>	<b>Single-stranded DNA</b>
<b>T, Trp</b>	<b>Tryptophan</b>
<b>TCA</b>	<b>Trichloroacetic Acid</b>
<b>TEMED</b>	<b>N,N,N',N'-tetramethylethylenediamine</b>
<b>TFA</b>	<b>Trifluoroacetic Acid</b>
<b>THF</b>	<b>Tetrahydrofuran</b>
<b>TLC</b>	<b>Thin Layer Chromatography</b>
<b>Tris base</b>	<b>Tris(hydroxymethyl)-aminomethane</b>
<b>TX-100</b>	<b>Triton X-100</b>
<b>U, Ura</b>	<b>Uracil</b>
<b>UV/Vis</b>	<b>Ultraviolet/Visible</b>
<b>wt</b>	<b>Wild type</b>
<b>X-gal</b>	<b>5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside</b>
<b>YNB</b>	<b>Yeast Nitrogen Base</b>



## List of Genes or Proteins

<i>ADE2</i>	Phosphoribosylaminoimidazole Carboxylase gene
<i>bAS</i>	$\beta$ -Amyrin Synthase
<i>b. t. LSS</i>	<i>Bos Taurus</i> 2,3-oxidosqualene-lanosterol cyclase gene; lanosterol synthase gene
<i>B. l. OSC</i>	<i>Bos Taurus</i> 2,3-oxidosqualene-lanosterol cyclase; lanosterol synthase
<i>CAS1</i>	Cycloartenol Synthase gene
<i>CAS</i>	Cycloartenol Synthase
<i>A. t. CAS1</i>	<i>Arabidopsis thaliana</i> Cycloartenol Synthase
<i>ERG1</i>	Squalene Epoxidase gene
<i>ERG7</i>	Oxidosqualene Cyclase gene; Lanosterol Synthase gene
<i>S. c. ERG7</i>	<i>Saccharomyces cerevisiae</i> lanosterol synthase
<i>ERG11</i>	Lanosterol 14- $\alpha$ -demethylase gene
<i>HIS3</i>	Imidazoleglycerol-Phosphate Dehydratase gene
<i>Kan<sup>R</sup>/Neo<sup>R</sup></i>	Neomycin Phosphotransferase; Kanamycin/neomycin resistance gene; G418 Resistance gene
<i>LUP1</i>	Lupeol synthase 1 gene
<i>LUS</i>	Lupeol synthase
<i>LEU2</i>	$\beta$ -isopropylmalate Dehydrogenase (IMDH) gene
<i>PNY2</i>	<i>Panax ginseng</i> $\beta$ -Amyrin Synthase gene
<i>PSY</i>	<i>Pisum sativum</i> $\beta$ -Amyrin Synthase gene
<i>OSC</i>	Oxidosqualene-Lanosterol Cyclase
<i>PNK</i>	T4 polynucleotide kinase
<i>SE</i>	Squalene Epoxidase
<i>TRP1</i>	Phosphoribosylanthranilate Isomerase gene
<i>URA3</i>	Orotidine-5'-Phosphate (OMP) Decarboxylase gene



# *Chapter 1*

## *General Introduction*

### 1.1 Overview of oxidosqualene cyclase

Oxidosqualene cyclases (OSCs) (EC 5.4.99.-) constitute a family of enzymes which catalyze the common substrate, oxidosqualene, into vast skeletal diversity of cyclization products, acting as the biosynthetic precursors for sterols, related membranous components, steroid hormones, and other secondary metabolites. Up to date, over one hundred different cyclic triterpene alcohols are respectively formed via their individual oxidosqualene cyclases.<sup>1</sup> The species-dependent product specificity of cyclase is well-known and fascinating. For example, oxidosqualene is solely transformed into tetracyclic lanosterol in animals and fungi, whereas a variety of polycyclic triterpenoid products, including cycloartenol, lupenol or  $\beta$ -amyrin are produced in higher plants or algae. In another example, squalene cyclase converts the precursor of oxidosqualene, squalene, into pentacyclic hopene or hopanol, which act as their corresponding membranous triterpenoid lipid in lower plants or bacteria (Figure 1.1).

This kind of single protein capable of multiple functional enzymatic transformations is one of the most impressive naturally existing reaction-phenomena. The oxidosqualene cyclization reaction is remarkable for its complexity. One acyclic substrate with only one stereogenic center is converted into polycyclic products containing multiple rings and several chiral centers. The conversion from the linear oxidosqualene into polycyclic products contains at least ten steps of covalent bond cleavage and formation. The cyclization cascade comprises one oxirane ring opening,

four or five stereochemically controlled ring formations, four nucleophilic hydride or methyl group migrations, and the final deprotonation reaction. Although the end product, cholesterol, produced in the human cholesterol biosynthetic pathway is required by another eighteen enzymatic reactions, the oxidosqualene cyclases are undoubtedly responsible for the most fascinating step (Figure 1.2).

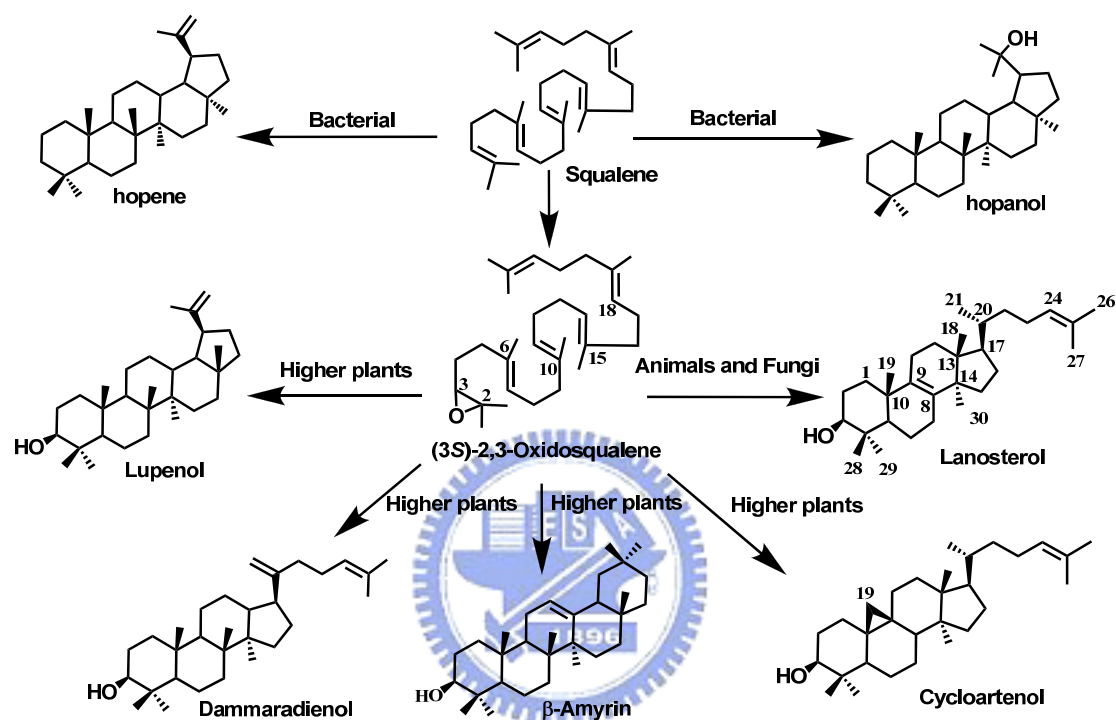


Figure 1.1 The product diversity of (oxido)squalene cyclase in different species.

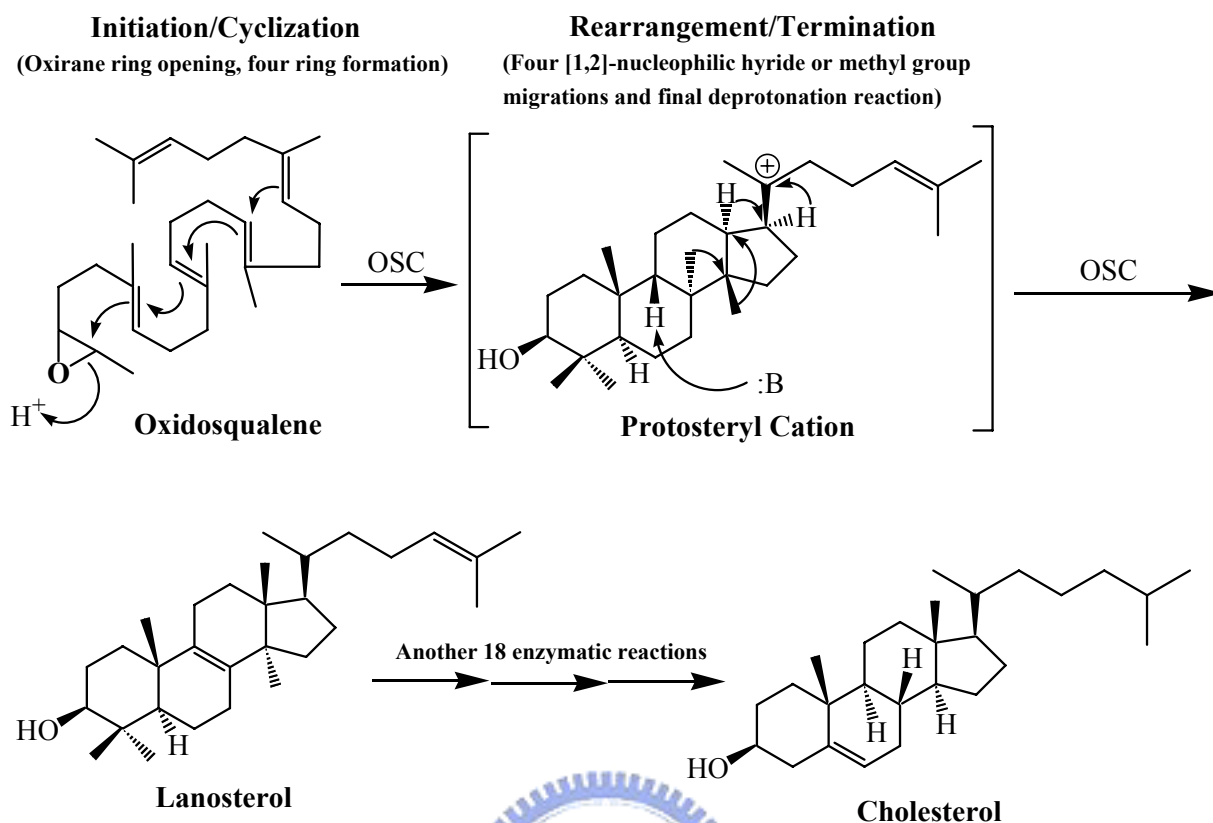


Figure 1.2 The cyclization cascade of oxidosqualene to lanosterol and the end product, cholesterol, produced in the human cholesterol biosynthetic pathway.

These species-dependent cyclization processes could be further classified according to the respective cationic intermediates, *i.e.* the protosteryl cation or the dammarenyl cation. The main difference between these two major intermediate cations is the stereochemical configuration of the second cyclohexyl B-ring. In the protosteryl pathway, the “boat” six-membered B-ring is produced, whereas the “chair” cyclohexyl B-ring is found in the dammarenyl cationic pathway (Figure 1.3). Thus, after the formation of protosteryl cation via the “chair-boat-chair” pre-folded conformation, the substrate is triggered to yield either lanosterol or cycloartenol followed by a series of [1,2]-methyl/hydride groups shifts, and the final alternative elimination. In parallel, the “all-chair” dammarenyl cation processes the similar rearrangement, ring-expansion, and final elimination to give dammaradienol, lepenol,  $\alpha$ -amyrin or  $\beta$ -amyrin in higher plants or algae (Figure 1.4).<sup>2</sup>



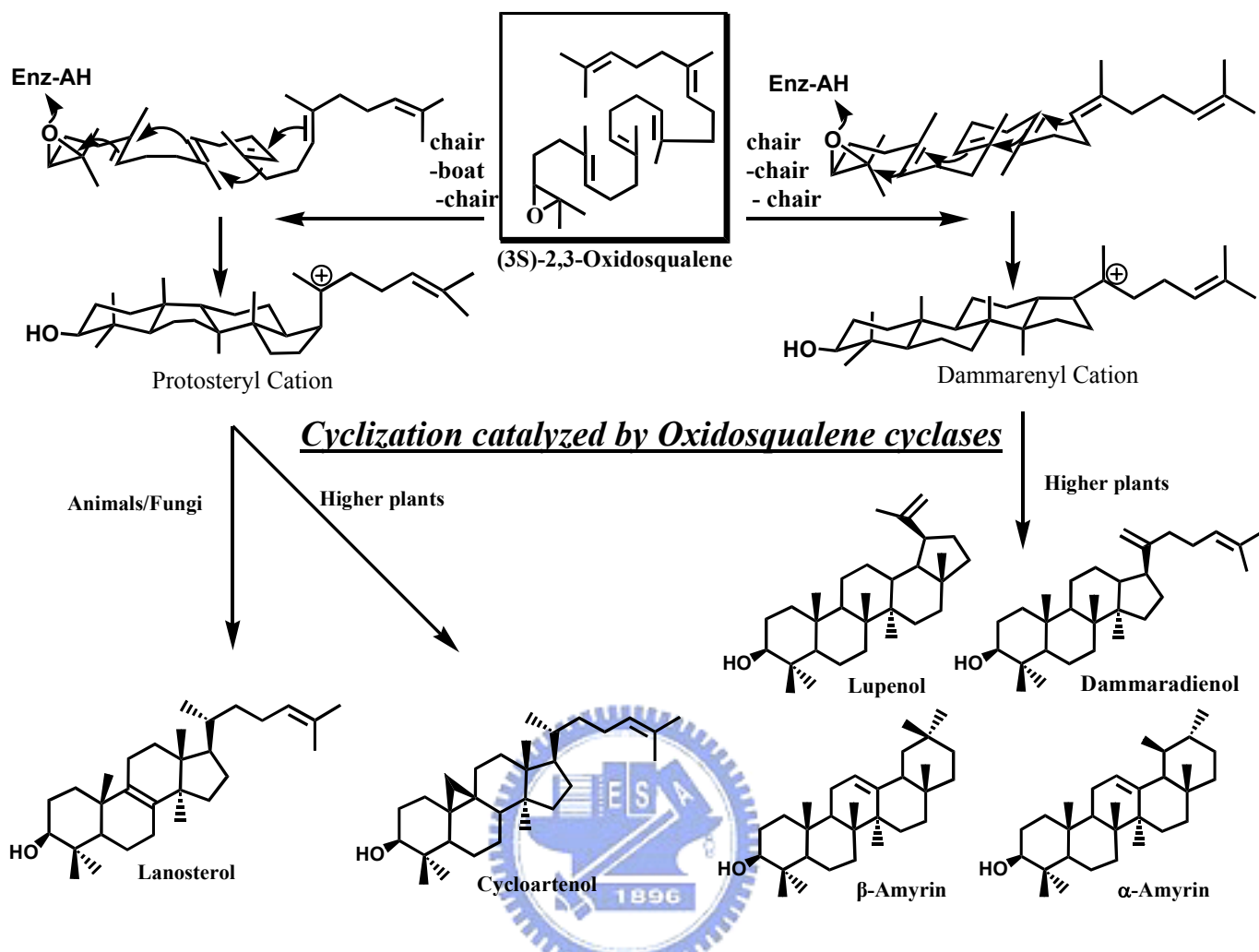


Figure 1.3 The cyclic products are classified according to the stereochemical configuration of carbocationic intermediates.

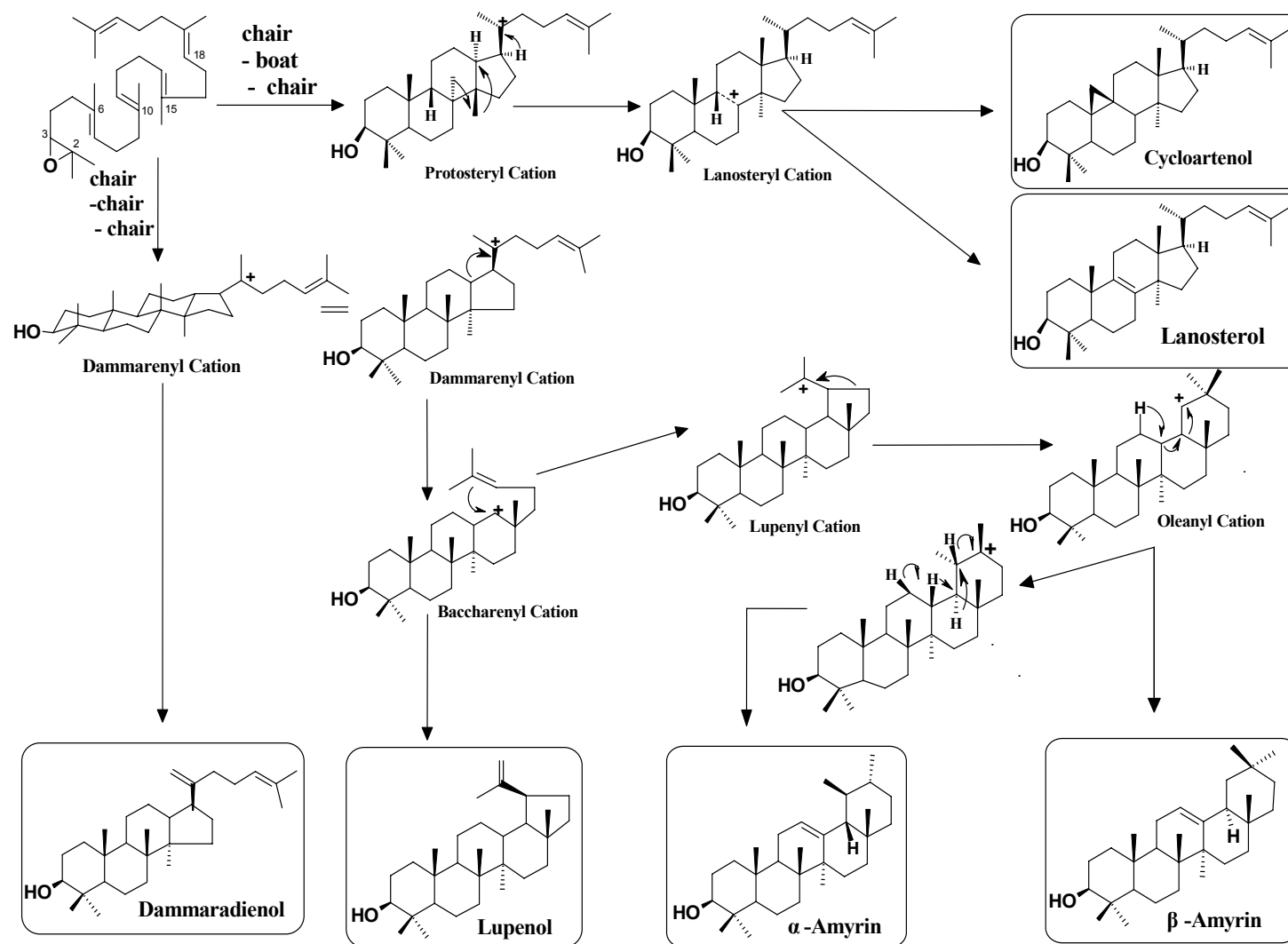


Figure 1.4 The postulated cyclization mechanism of oxidosqualene into diverse products via different cationic intermediates. The respective products are shown in frame

## 1.2 Historical hypothesis of cyclization mechanism

### 1.2.1 The mechanistic and stereochemical insights of oxidosqualene cyclization

From the view of chemical reaction, diverse products should be obtained after the intramolecular cyclization reaction of the flexible linear polyolefinic substrate, oxidosqualene. The highly exothermic process during the formation of novel carbon-carbon bonds from the nucleophilic attack of olefinic  $\pi$ -electrons on carbocationic atoms also cause the difficulty in controlling the reaction tendency in the non-enzymatic cyclization of oxidosqualene. However, the biological system utilized the species-dependent oxidosqualene cyclases to rigorously trigger the cyclization reaction for generations of specific biosynthetic products. The substrate folding manners, the substrate-enzyme interactions, even the processes of substrate immigration or products released are thought to be subject to regulation. Many organic and bioorganic experiments have been designed to understand the relationships between the cyclization mechanism and the cyclase enzyme itself. Because of the intrinsic difficulty, the cyclization mechanism of squalene or oxidosqualene has intrigued scientists for over half of century. After the first pioneering model was established in 1934, different radioisotope feeding experiments were further used to examine the biosynthetic source of cholesterol.<sup>3, 4</sup> With the successful structural determination of lanosterol, Woodward and Bloch proposed a putative model for lanosterol production from squalene, and also confirmed its biological importance (Figure 1.5).<sup>5-8</sup> The following experiments established that the oxidosqualene rather than the squalene is the direct precursor of lanosterol, while the 3*S* isomer was further proved to be the exclusive precursor.<sup>9-11</sup>

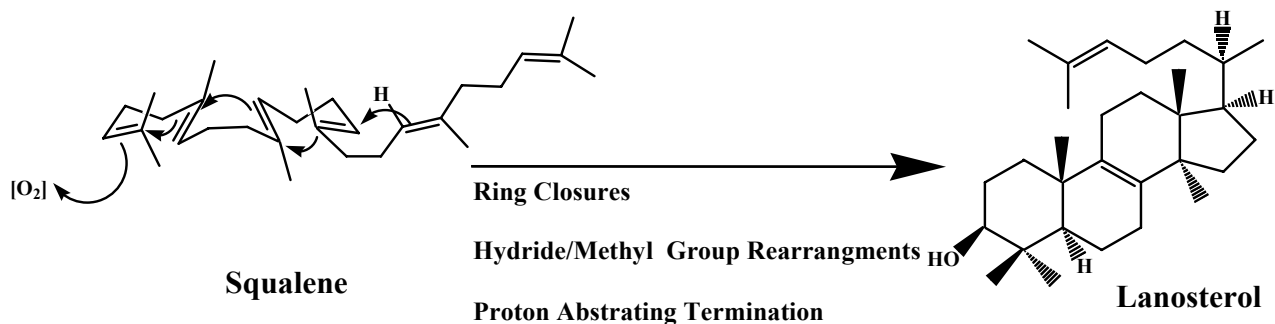


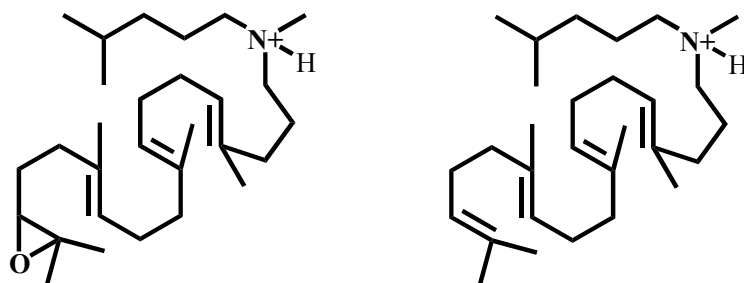
Figure 1.5 Squalene cyclization model for lanosterol production, proposed from Woodward and Bloch.<sup>6</sup>

Subsequently, various enzymatic incubations of chemically synthesized substrate analogs or inhibitors provided much valuable information regarding the mechanism of oxidosqualene cyclization. For example, a series of structural analogs were utilized to elucidate the conformational requirement or the molecular geometry of the substrate within the enzymatic active sites. These chemical derivatives not only indicated the importance of internal methyl groups/unsaturated double bonds for the oxidosqualene cyclization, but also illustrated the enzymatic toleration for the modification on the oxidosqualene molecule. For example, substrate derivatives, which are modified in the distal terminus, including the double epoxide rings derivative (2,3-22,23-dioxidosqualene), the saturation of terminal double bond derivative (22,23-dihydro-2,3-oxidosqualene), and distal functional modification derivatives (22,23-dihydro-22-methylene-2,3-oxidosqualene or 30-acetyl-2,3-oxidosqualene), were all successfully cyclized via an enzyme triggered reaction.<sup>12-15</sup> Similarly, neither the hydroxyl, trimethylsilyl, nor methylidene modification at C-29 methyl group appear to influence the cyclase activity (Figure 1.8).<sup>16</sup> Nevertheless, the substrate derivatives with saturation modification of the internal double bond, either 10,11-dihydro-2,3-oxidosqualene or 14,15-dihydro-2,3-oxidosqualene, are not suitable substrates for the cyclization reaction. Cyclase also only partially catalyzed

the double demethyl analogs, 10,15-demethyl-2,3-oxidosqualene, into 6,6,5,4 tetracyclic product, suggesting the functional importance of internal methyl group (Figure 1.8).<sup>17,18</sup> The modification near the epoxide terminus of substrate, such as 2,3-iminosqualene, 2,3-cis-1' nor-2,3-oxidosqualene, 2,3-trans-1' nor-2,3-oxidosqualene, or 1,1'-bisnor-2,3-oxidosqualene were used to illustrate the enzymatic tolerance (Figure 1.8).<sup>19-21</sup> The necessity of epoxide ring for the OSC-catalyzed reaction was further supported via two acyclic azasqualenes analogs in Figure 1.6.<sup>22</sup> The 19-aza-18,19,22,23-tetrahydrosqualene-2,3-epoxide and its non-epoxide ring analogs, 19-aza-18,19,22,23-tetrahydrosqualene, were applied for the cyclase mediated reaction. Only 19-aza-18,19,22,23-tetrahydrosqualene-2,3-epoxide exhibited the inhibitory activity for the oxidosqualene cyclase, indicating the importance of the epoxide ring. In other experiments, a series of substrate analogs or ammonium ions that mimic the transition state in the different stages of cyclization were also synthesized and used to illustrate their inhibition effect (Figure 1.7).<sup>19, 23-26</sup>

The conformational requirements of oxidosqualene for the cyclase-triggered reaction were further observed by using the different acyclic azasqualene derivatives (Figure 1.9). For example, two geometrical isomers, (6*E* or 6*Z*)-10-aza-10,11-dihydrosqualene 2,3-epoxide, were used to interact with cyclase.<sup>27,28</sup> Only the 6*E*-isomer which corresponds to the natural all *E*-form conformation in oxidosqualene was functionally active. Moreover, one pair of isomers, (18*Z* or 18*E*)-29-methylidene-2,3-oxidosqualene were also utilized as the mechanistic inhibitors for oxidosqualene cyclase.<sup>29</sup> The (*Z*)-form methylidene derivative showed the potent and irreversible inhibitory activity, whereas the (*E*)-form revealed no inhibition. The large difference in the IC<sub>50</sub> value between these isomers indicated the steric requirement for the protersteryl cation in the enzymatic cyclization process. These geometrical isomers, *E*-form or *Z*-form derivatives, confirmed that the natural substrate should be

folded with all *E*-form conformation except for the carbon-18, and be cyclized via *all-trans* geometry for the cyclization reaction in the enzyme active site.<sup>27-29</sup>



**19-aza-18,19,22,23-tetrahydrosqualene-2,3-epoxide and its analogues without epoxide**

Figure 1.6 Acyclic azasqualenes mimic the C-20 carbonium ion intermediates.

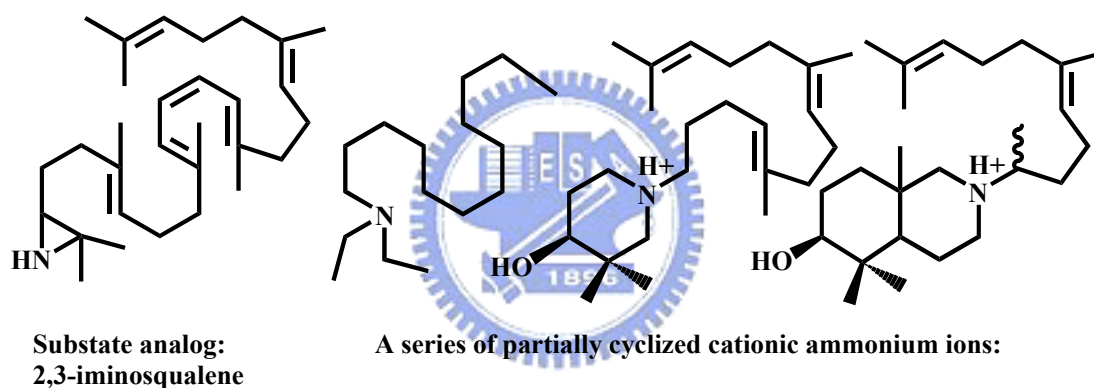


Figure 1.7 Structure of potent transition state inhibitors of oxidosqualene cyclase.



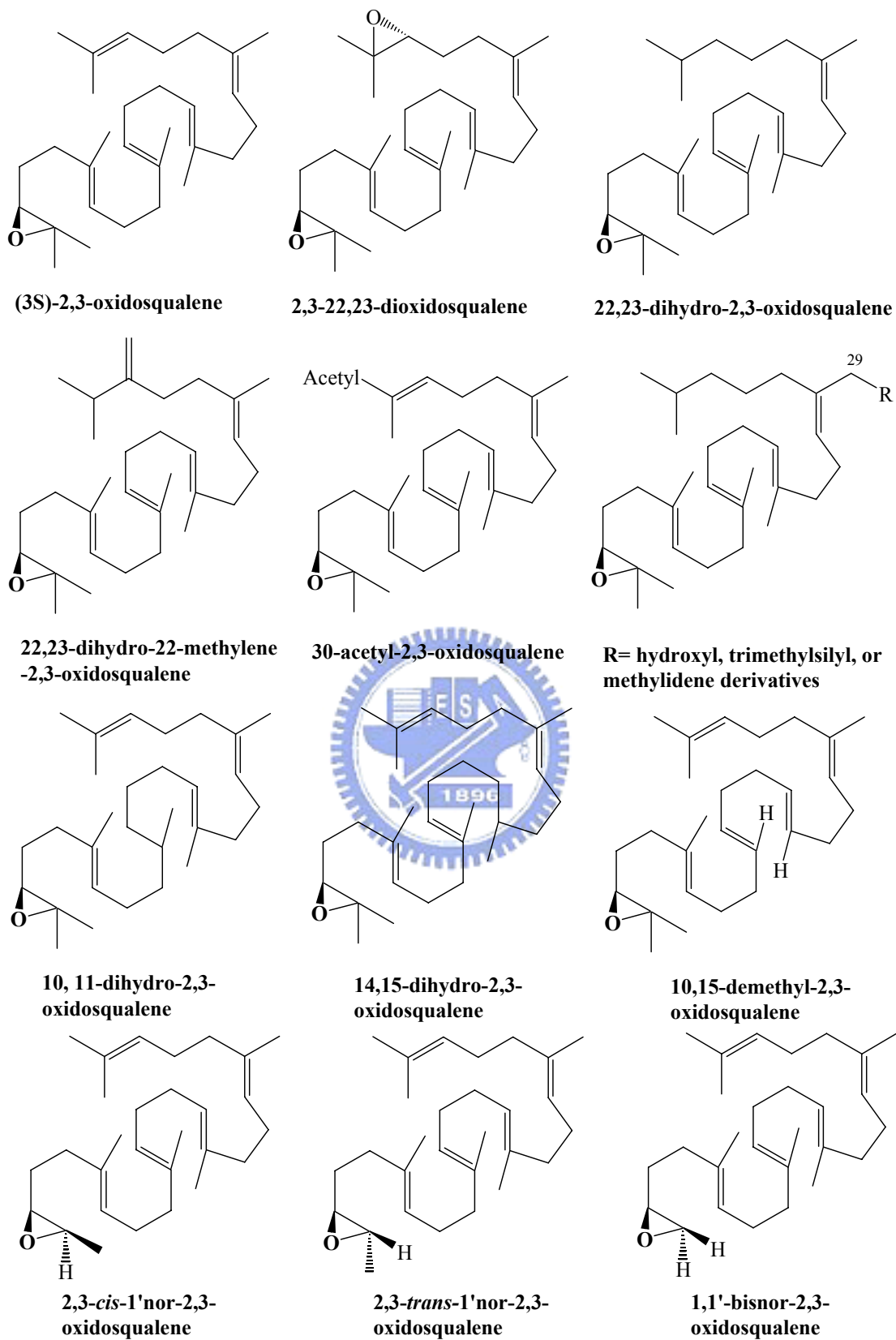
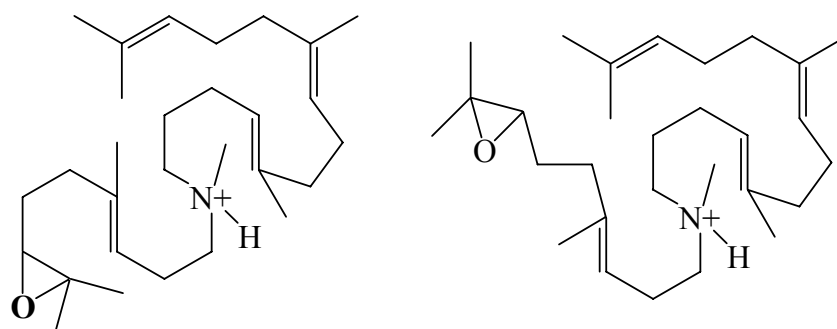
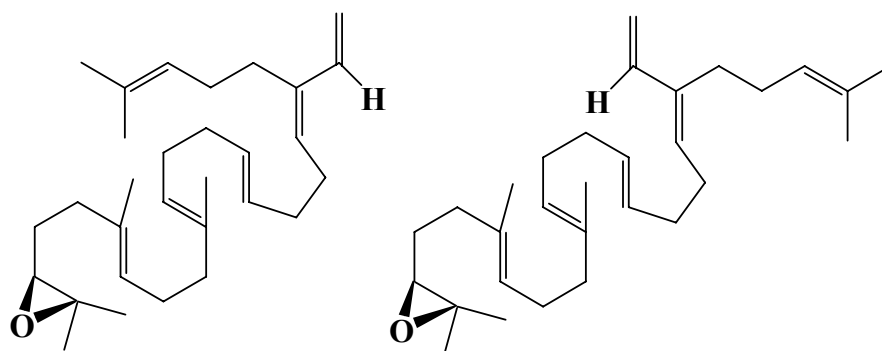


Figure 1.8 Some substrate derivative inhibitors.



**(6E or 6Z)-10-aza-10,11-dihydro-squalene 2,3-epoxide**



**(18Z or 18E)-29-methylidene-2,3-oxidosqualene**

Figure 1.9 The geometrical isomers of substrate analogs.

The argument about a stepwise or a concerted mechanism for the oxidosqualene cyclization has also been considered as the result of direct experimental evidence. After the initial epoxide ring opening, the anchimeric assistance from the neighboring double bond (C-6/C-7) has been observed for the formation of the A-ring.<sup>30,31</sup> Subsequently, the non-concerted pathway including a series of conformational rigid and partial cyclic cationic intermediates have been proposed for the remaining process of cyclization cascade.<sup>32,33</sup> The bicyclic to pentacyclic triterpenoid products corresponding to the truncated cyclization intermediates were isolated from different natural sources such as *P. resins* or *T. pyriformis*, indicating the possibility of a stepwise mechanism for the oxidosqualene cyclization.<sup>34,35</sup> The nonstop, concerted cyclization hypothesis, *i.e.*, conversion from oxidosqualene into lanosterol process without mono-, bi-, or tricyclic carbocationic intermediates, was further against by

Corey and coworkers in 1995.<sup>36</sup> A minor product with a 6-6-5-4 tetracyclic scaffold was isolated from the enzymatic 20-oxa-2,3-oxidosqualene transformation (Figure 1.10). Thus, Corey proposed that the cyclization of 2,3-oxidosqualene for the tetracyclic protosteryl cation must proceed through a discrete tricyclic 6-6-5 Markovnikov carbocationic intermediate and a sequential ring expansion process with the formation of an anti-Markovnikov cyclohexyl C-ring intermediate.<sup>36</sup> Several truncated cyclization products with a 6-6-5 scaffold obtained from the enzymatic incubation with substrate analogs further supported the evidence for the C-ring expansion (Figure 1.10).<sup>37-39</sup> Moreover, Corey and coworkers also illustrated the stereochemistry of C-20 during the oxidosqualene cyclization process.<sup>40, 41</sup> The 20-oxa-2,3-oxidosqualene and the (20*E*)-20,21-dehydro-2,3-oxidosqualene were respectively synthesized and used in the enzymatic transformation to produce protosteryl derivatives with 17 $\beta$  side chain conformation (Figure 1.11). These experimental results excluded the previously hypothesized 17 $\alpha$  side chain conformation, and also facilitated the understanding of a natural *R*-form configuration at C-20. After the formation of the 17 $\beta$  side chain, the smaller rotation angle (60°) of the side chain is required for producing the natural *R*-form configuration at C-20. The hindered rotation angle (120°) was required while the side chain conformation appeared with the 17 $\alpha$  conformation. Therefore, after decades of efforts by chemists with the representative chemical studies on the respective stages of cyclization processes, the broadly accepted mechanistic hypothesis for the formation of lanosterol is summarized in Figure 1.12.



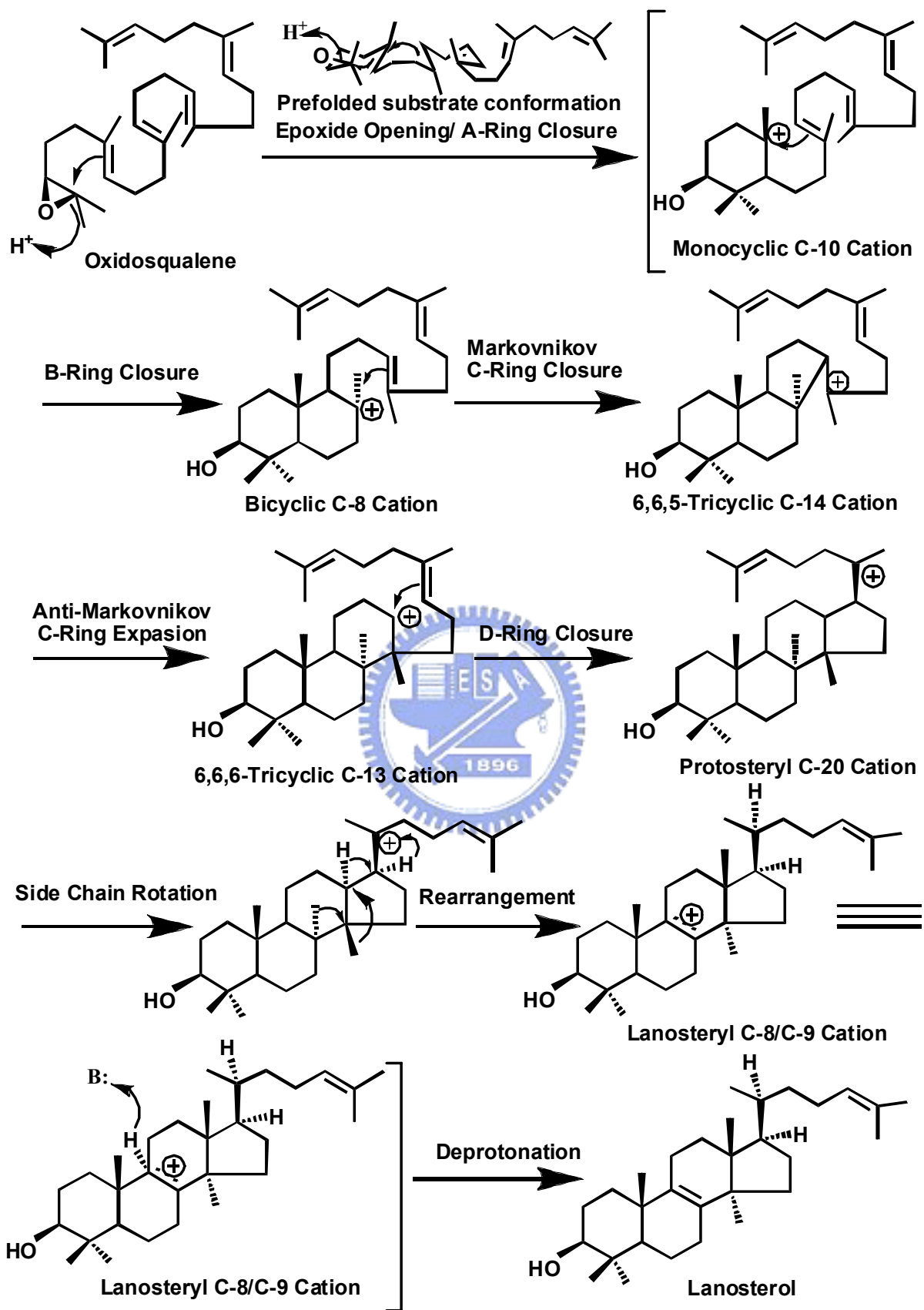


Figure 1.12 The hypothesized cyclization cascade of oxidosqualene.

## 1.2.2 The theoretical enzyme-templated cyclization reactions

In addition to chemical studies on the cyclization mechanism, the critical issues concerning the enzyme itself historically have been poorly understood. After observation of the common structural feature of triterpene molecules, Ourisson proposed a phylogenetic hypothesis for the cyclases.<sup>42</sup> It was proposed that a primitive ancestral cyclase containing several essential catalytic residues exists, yet some of the critical active site residues have undergone changes, indicating evolutionary divergence for different species.<sup>43,44</sup> Moreover, according to the studies of substrate specificity and the stereochemistry of the cyclization reaction, Johnson has proposed another theoretical model for illustrating the functional role of cyclase enzyme. One Lewis acid residue, which acts as the proton donor for epoxide ring opening, and three axial negative charge residues facing the highly energetic intermediate cations were proposed for the enzymatic cation-olefin cyclization.<sup>45,46</sup> These negative charge residues might lower the activation energy for the formation of the boat B-ring and the anti-Markovnikov closure of C-ring (Figure 1.13).

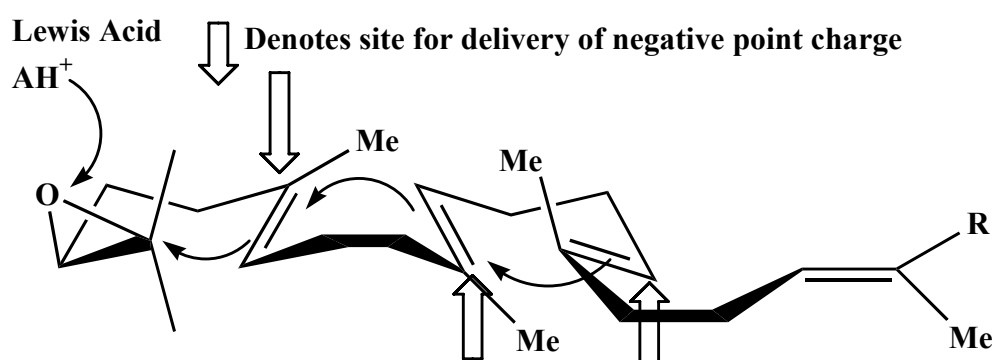


Figure 1.13 Johnson proposed three cation-stabilizing auxiliaries in the cyclase.<sup>45, 46</sup>



In addition, the “Q-W motif theory” was the hypothetical result, according to the careful comparison of the amino acid sequence’s constitution in the bacterial cyclases as well as in the eukaryotic cyclases.<sup>47,48</sup> Sixteen amino acid repeats, Arg/Lys Gly/Ala X<sub>2-3</sub> Tyr/Phe/Trp Leu X<sub>3</sub> Gly X<sub>2-5</sub> Gly X Trp, were highly conserved within the carboxyl region of cyclase gene. One of the Q-W motifs was further labeled with a mechanistic inhibitor, 29-methylidene-2,3-oxidosqualene.<sup>49,50</sup> These Q-W motifs were considered for stabilizing the electron-deficient cationic intermediates and for the structural importance in the cyclase enzyme.<sup>51</sup> In addition to the “Q-W motif theory”, Griffin proposed the “*Aromatic hypothesis*” for the cyclase enzymatic active site.<sup>52</sup> They suggested that the electron-rich indole rings from tryptophan residues or the phenyl groups from tyrosine/phenylalanine residues, which occurred frequently in the Q-W motifs, are more suitable for the role of stabilizing delivery force than those negative point charges postulated by Johnson. In this “*Aromatic hypothesis*”, two electron-rich residues are assumed to occupy the position above or below the high energy cationic intermediate in each stage of cyclization processes (C-2, C-8, C-13, and C-20).<sup>52</sup> The cation- $\pi$  interaction in the enzymatic active site is theoretically responsible for stabilizing these electron-deficient cationic intermediates. Therefore, according to these theoretical models, cyclases from different origins might be very similar in their major structural conformation. However, for the purpose of the phylogenetic derivation, small but critical mutational changes exist in these cyclase enzymes, especially in the region spatially near the carbocationic intermediates (Figure 1.14).

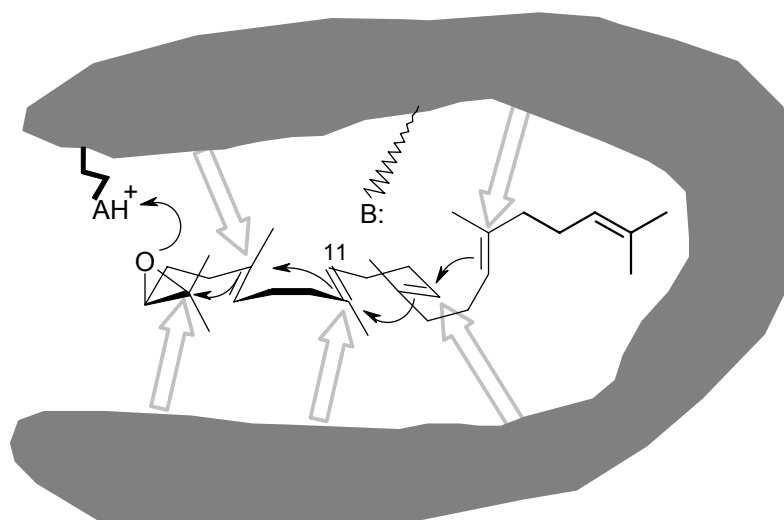


Figure 1.14 Hypothetical model based on the cyclization mechanism. The arrows in gray indicate the proposed stabilizing delivery force for the carbonium ions.

On the other hand, some catalytically or structurally important amino acids were further identified via an affinity labeling strategy of several substrate analogs in Figure 1.15 and Figure 1.16. For example, the DCTAE motif, a highly conserved region in cyclases, was specifically labeled with the first mechanism-based irreversible inhibitor, 29-methylidene-2,3-oxidosqualene, (29-MOS), implicating the functional importance of this region in the oxidosqualene cyclization process (Figure 1.15).<sup>49,53</sup> The tritium labeled 29-MOS was further used to identify the alkylated site (aspartate-456) in the rat OSC.<sup>50</sup> Similar experimental results were also observed in the squalene-hopene cyclase from *A. acidoterrestris*, suggesting the crucial influence of this region on cyclase activity.<sup>54,55</sup> Another three mechanistic inhibitors, including 6-desmethyl-2,3-oxidosqualene, 10-desmethyl-2,3-oxidosqualene, and 10,15-desdimethyl-2,3-oxidosqualene which were expected to mimic the monocyclic or bicyclic reactive intermediates, were specifically labeled on the fragment from Leu-489 to Lys-512 in *S. cerevisiae* ERG7.<sup>50,56</sup> Moreover, 20-oxa-2,3-oxidosqualene and other three substrate analogs, which mimic the cationic intermediate during the formation of D-ring, revealed covalent labeling at His-234 and Trp-232 residues in the

cyclase enzyme (Figure 1.16).<sup>50, 56</sup> Another mechanism-based irreversible inhibitor, 18-thia-2,3-oxidosqualene, was also used to covalently modify the active site residues of OSC (Figure 1.15).<sup>57</sup> These suicide inhibitors provide the clues in concerning with the functional role of these amino acids participating in the substrate binding or the stabilization of unstable cationic intermediates.

In addition, *p*-chloromercuribenzenesulfonic acid (PCMBs) and N-ethylmaleimide (NEM) also exhibited strong inhibitory activity for cyclases, indicating that the cysteine residues are critical for the catalytic function in OSC or SHC enzymatic active site.<sup>58-62</sup> Moreover, some of squalene skeleton-based thiol-modifying reagents were further developed for site specific inhibition (Figure 1.17). Among them, the maleimide-type squalene derivative was the most efficient one for mammalian OSC inhibition.<sup>63</sup> The Ellman-type squalene derivatives also exhibited obvious inhibitory activity. In contrast, the dodecyl derivatives of thiol-modifying agents showed the complete loss of inhibition. These experiments suggested that the squalene skeleton of thiol-modifying reagents act as the carrier moiety for effectively leading these inhibitors into the enzymatic active site.<sup>64, 65</sup> At the saturated condition, squalene could only protect the cyclase against the inhibition from the squalene skeleton based inhibitor but not the Ellman reagent, 5,5'-dithio-bis-(2-nitrobenzoic) (DTNB), or other general thiol-modifying reagents, indicating that a different set of thiol groups exist in the enzyme active site.<sup>63, 65</sup> Cysteine-345 of squalene-hopene cyclase from *A. acidocaldarius* acting as one of the constriction residues in the substrate access was thus further examined.<sup>66</sup> The quadruple mutant, bearing Cys-435 as its sole cysteine residue, was successfully modified by the CNDT-squalene and caused the blockage of the substrate entrance. Trypsin digestion and MS analysis additionally demonstrated this covalent modification and showed that this cysteine is critical in the enzymatic substrate

entrance channel.

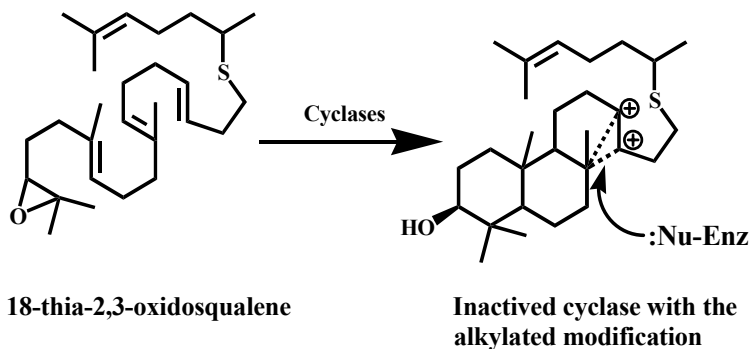
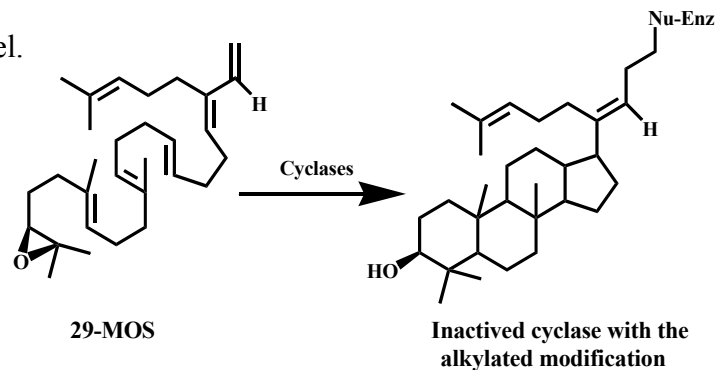


Figure 1.15 Potential modes of mechanism-based inactivation

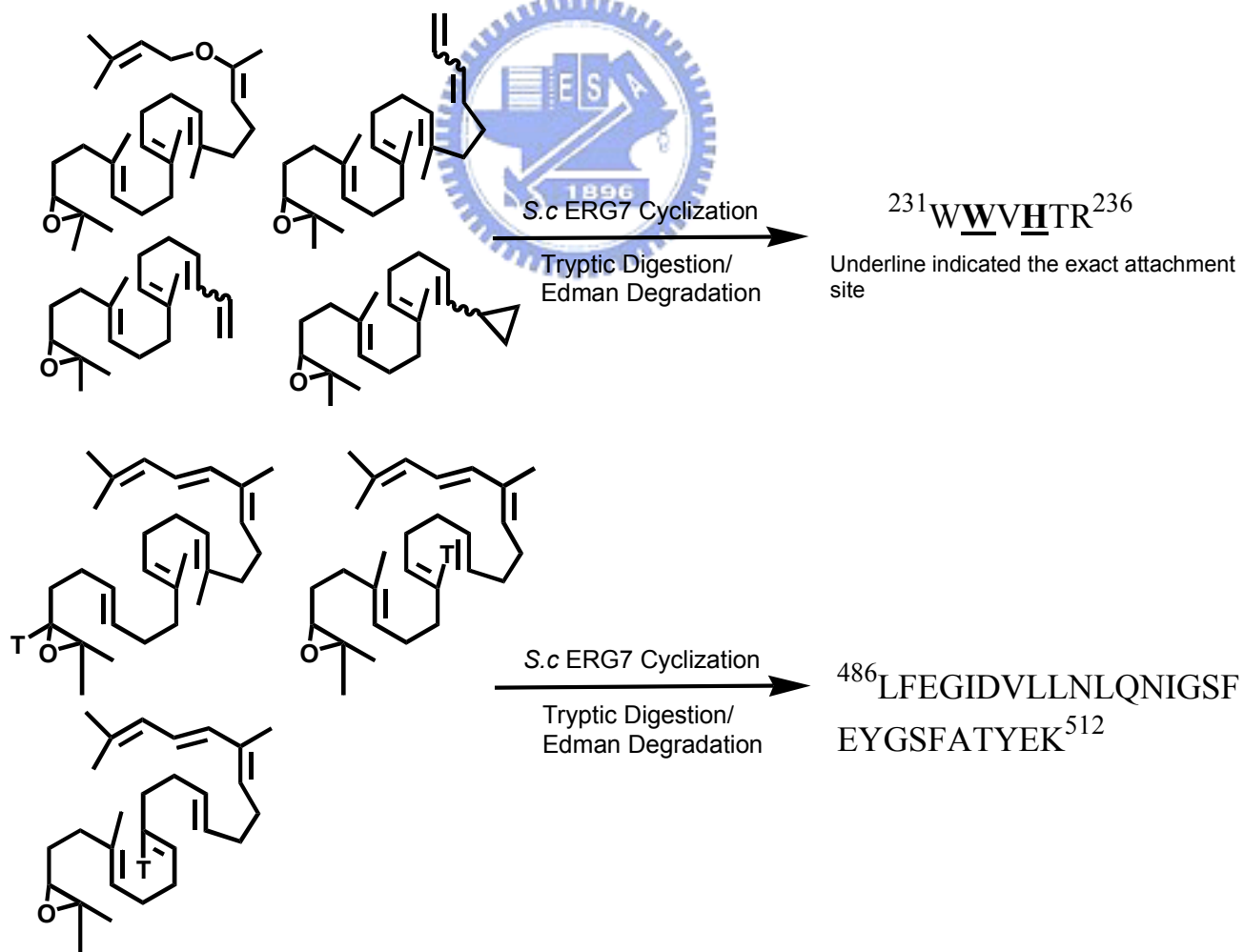


Figure 1.16 Affinity labeled cyclase fragment with mechanistic suicide inhibitors.

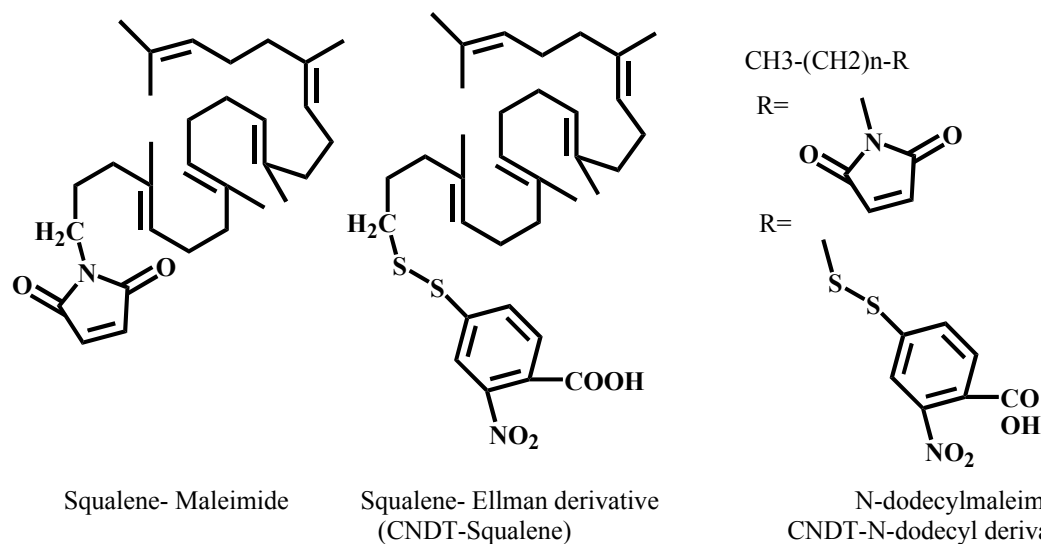


Figure 1.17 Structures of various squalene and dodecyl derivatives of thiol-modifying agents.

In conclusion, the inhibition analysis of these substrate analogs that mimic the cyclic intermediate cations at the C-2, C-8, C-10, C-13, and C-20 position or other mechanism-based irreversible inhibitors, provided the direct examination for the previously postulated hypotheses. Accordingly, the cyclization reaction of oxidosqualene must proceed through a series of conformational rigid carbonium ionic intermediates.<sup>64</sup> Moreover, these experimental results also supported the hypotheses of the “Johnson negative charge model” and the “Griffin aromatic hypothesis” that assumed that the intermediates are stabilized via several suitable external forces in the enzymatic active site. All of these substrate derivatives used in the inhibition studies indicated that the precise control between cyclase itself and substrate conformation are indispensably essential, especially in the terminal epoxide ring or the internal methyl groups/double bonds.

In taking a different approach, after the affinity labeling strategy and the preliminary mutagenesis analysis of 76 highly conserved positions in the *Saccharomyces cerevisiae* lanosterol synthase (*S. cerevisiae* ERG7), the catalytically important residues, including Asp-456, His-146, His-234, and Met-532, were first

identified as the essential component for the cyclase catalysis.<sup>56</sup> The putative model of these essential residues assisting in the ERG7-catalyzed cyclization was thus proposed. The Asp-456 residue was implied as the candidate for the role of Lewis acids to initiate the epoxide ring.<sup>30</sup> However, the precise role of these functional residues in controlling the cyclization process, or whether there are more catalytically important amino acids involved in different stages of the complex cyclization/rearrangement, could not be defined at that time. Most importantly, no residue has been found to facilitate the substrate pre-folding, cyclization cascade, or the rearrangement processes from those studies solely based on the mechanism-based analogs or early mutagenesis studies.

### 1.3 Biological studies of 2,3-(oxido)squalene cyclases

#### 1.3.1 Enzymology of 2,3-(oxido)squalene cyclase enzymes

Despite the fact that the chemical mechanistic studies on the cyclization reaction has provided detailed understanding about the oxidosqualene cyclase-mediated reaction, the enzymology for the cyclase enzymes remained incompletely characterized for decades. The cyclases were thought to be difficult for purification because of their membrane-associated properties. Moreover, the cyclase enzymes from different species usually exhibited the distinct tolerance range for detergent or salt concentrations. Attempts to purify the active cyclases always suffered from the difficulty in obtaining the soluble enzyme. Even under detergent solubilization, cyclase enzymes were usually unstable and rapidly lost activity. These factors lead to difficulty in finding the optimal conditions for cyclase enzyme purification. Fortunately, several oxidosqualene cyclase enzymes have been successfully purified to homogeneity ranging from vertebrate, yeast, and plant sources, by using different



strategies.<sup>2</sup>

The breakthrough in the purification of the oxidosqualene cyclase was first carried out in higher plants.<sup>67-70</sup> An isoelectric focusing electrophoresis coupled with hydroxylapatite, DEAE-cellulose, and gel filtration chromatographies were applied to separate the cycloartenol synthase and  $\beta$ -amyrin synthase from *R. japonica-hara*, according to their distinctive pI value. The active enzyme was effectively solubilized and purified to homogeneity in the presence of the nonionic detergent Triton X-100. By using a similar purification procedure, the rat and pig liver cyclase enzymes have been purified in different purification fold.<sup>60, 71-73</sup> Among these purification strategies, lauryl maltoside was also alternatively used for solubilizing the cyclase enzyme.<sup>73</sup> The simplified procedures, via combination of hydroxylapatite and DEAE-cellulose chromatographies, have been successfully used to obtain the mammalian oxidosqualene cyclase with molecular weight of 75 kDa in pig or 78 kDa in rat.<sup>71</sup> The first potent mechanism-based irreversible suicide inhibitor of cyclase, 29-MOS, was then used to specifically label the different vertebrate OSCs and gave the expected molecular weight ranging from 70 to 80 kDa (78kDa for rat; 73 kDa for dog; 75 kDa for pig; and 73 kDa for human). These results demonstrated that highly similar cyclases exist among different mammals.

The lanosterol synthase from yeast *S. cerevisiae* was also purified by using DEAE chromatography, coupled with an affigel-15 column. The purified cyclase exhibited a rather small size (26 kDa) than other mammalian oxidosqualene cyclase enzymes.<sup>74</sup> The incorrect purification strategy or the contaminated protein collection might be misapplied with regard to this yeast cyclase purification. The later molecular cloning experiment of *S. cerevisiae* oxidosqualene-lanosterol cyclase has determined that the exact molecular weight of yeast cyclase is about 80 kDa.<sup>75</sup> In other studies, the cyclases from pea *P. sativum* have been purified with determined molecular

weight of 55 kDa for cycloartenol synthase, and 35 kDa for  $\beta$ -amyrin synthase. These plant cyclases also required the non-ionic detergent for their optimal activity.<sup>62, 67, 69</sup>

### 1.3.2 Molecular cloning of 2,3-oxidosqualene cyclase

In general, purification of OSCs in sufficient quantity for amino acid sequence determination or antibody production has proven to be difficult, mainly caused by the trace amount of purified protein or the problems in obtaining the soluble enzyme. Thus, attempts to isolate cDNA clones or genes encoding cyclase enzymes by this way have met with only limited success. However, the utilization of internal amino acid sequence information from purified cyclase enzymes has been successfully used to isolate cDNA for some animal oxidosqualene-lanosterol cyclases. For example, the cDNA encoded rat OSC (*R. norvegicus*) was determined via this methodology.<sup>76, 77</sup> An open reading frame of 2,202 nucleotides including the stop codon was identified, which encoded an 83.3 kDa protein with 733 amino acids. Moreover, Corey also cloned and characterized the human lanosterol synthase by using hybridization techniques in a *H. sapiens* liver cDNA library.<sup>78</sup> The deduced amino acids showed 83% identity to that of rat cyclase.

The gene encoding the *C. albicans* OSC was first obtained by Kelly *et al.* on the basis of its ability to complement the oxidosqualene-cyclase deficient (*erg7*) strains in *S. cerevisiae*.<sup>79</sup> The gene consists of 2,187 nucleotides and encodes a predicted protein of 728 amino acids with molecular weight of 83.7 kDa. On the other hand, the *ERG7* gene of *S. cerevisiae* oxidosqualene-cyclase was also reported by using the genetic complementation as well as the oligonucleotides hybridization.<sup>80</sup> The complementary DNA and the derived amino acid sequences of this gene have also been determined. The open reading frame of 2,196 nucleotides was identified, which encoded 731 amino acids with molecular weight of 83.4 kDa.<sup>52</sup>

Because cycloartenol or other triterpene alcohols have never been found in yeast, molecular cloning of a plant triterpene synthases gene via genetic complementation may appear to not be feasible in *S. cerevisiae*. However, accumulating the high levels of 2,3-oxidosqualene and the absence of lanosterol in the oxidosqualene-cyclase deficient strains (*erg7*) could facilitate the synthesis and identification of novel cyclization products, which were generated via heterologous expression of triterpene synthases. In 1993, a yeast mutant strain lacking lanosterol synthase was transformed with an *A. thaliana* cDNA library and assayed for cyclase activity via TLC analysis. Corey thus isolated a cycloartenol synthase gene possessing the ability to cyclize oxidosqualene for the cycloartenol production. This gene encoded an 86 kDa protein and a 2,271-bp open reading frame with significant homology to OSC from *C. albicans* or SHC from *A. acidocalcarius*.<sup>81</sup>

From the availability of increasing DNA sequence information from individual OSCs, the highly conserved regions could be identified. The genes encoding oxidosqualene-cyclases proteins from yeast, human, and rat have been cloned via PCR based amplification methodology, by using the degenerated oligonucleotide primers which were designed according to these highly conserved regions.<sup>75, 82</sup>

While several OSC cDNAs involved in sterol biosynthesis have been cloned from various organisms, the triterpene synthase genes were rarely reported. However, even though the reaction products are different between squalene-hopene cyclase and the mammalian OSC, several amino acid sequences fragments, such as the QW motif or the DCTAE motif are highly conserved. Additionally, from the view of the product diversity, sterol biosynthetic OSCs and triterpene synthases might come from a common ancestral gene. Thus, triterpene synthase genes were expected to possess these highly conserved amino acid sequence regions. Accordingly, degenerated oligonucleotide primers based on these regions were designed to isolate the triterpene

synthase genes from *P. ginseng*.<sup>83</sup> Two cDNAs, one being the  $\beta$ -amyrin synthase gene, whereas the other encoded cycloartenol synthase, were thus revealed from heterologous expression in the yeast strain. These two cyclases exhibited 60% amino acid identity to each other. In the same way, the cDNA encoding the lupeol synthase from *O. europaea* and *T. officinale* were identified in 1999.<sup>84</sup> They shared significant sequence identity to each other (78%), whereas only about 58% identity was observed after comparing with that of the known lupeol synthase in *A. thaliana*. The multifunctional enzyme which produces diverse triterpene products has also been cloned from *A. thaliana*.<sup>85</sup>

In the last ten years, at least fifteen triterpene cyclase genes have been cloned and sequenced. These oxidosqualene-cyclase genes which have been functionally characterized in yeast are listed in Table 1.1.<sup>86</sup>



Table 1.1 OSCs that have been cloned and functionally characterized by expression in yeast

<b>Organism</b>	<b>Enzyme</b>	<b>Abbreviation</b>	<b>EMBL/Gene Bank AC</b>
<i>Rattus norvegicus</i>	lanosterol synthase	<i>OSC</i>	U31352
<i>Saccharomyces cerevisiae</i>	lanosterol synthase	<i>ERG7</i>	U04841
<i>Candida albicans</i>	lanosterol synthase	<i>ERG7</i>	L04305
<i>Arabidopsis thaliana</i>	cycloartenol synthase	<i>CASI</i>	U02555
<i>Pisum sativum</i>	cycloartenol synthase	<i>PSX</i>	D89619
<i>Panax ginseng</i>	cycloartenol synthase	<i>PNX</i>	AB009029
<i>Panax ginseng</i>	$\beta$ -amyrin synthase	<i>PNY</i>	AB009030
<i>Panax ginseng</i>	$\beta$ -amyrin synthase	<i>PNY2</i>	AB014057
<i>Pisum sativum</i>	$\beta$ -amyrin synthase	<i>PSY</i>	AB034802
<i>Pisum sativum</i>	Mix amyrin synthase <sup>a</sup>	<i>PSM</i>	AB034803
<i>Olea europaea</i>	lupeol synthase	<i>OEW</i>	AB025343
<i>Taraxacum officinale</i>	lupeol synthase	<i>TRW</i>	AB025345
<i>Arabidopsis thaliana</i>	lupeol synthase <sup>b</sup>	<i>ATLUP1</i>	U49919
<i>Arabidopsis thaliana</i>	Multifunctional <sup>c</sup>	<i>ATLUP2</i>	AF003472

<sup>a</sup> Primary products:  $\alpha$ - and  $\beta$ -amyrin (6:4) and other minor triterpenes

<sup>b</sup> Primary product: lupeol and other at least five minor triterpenes

<sup>c</sup> Primary products:  $\beta$ -amyrin,  $\alpha$ -amyrin, and lupeol (55:30:15)

### 1.3.3 Crystallization and structural characterization of cyclase

Recently, the X-ray structure of squalene-hopene cyclase from *A. acidocaldarius* has been determined.<sup>51, 87</sup> This three-dimensional structure of triterpene cyclase facilitated the understanding of relationships between the enzymatic active sites and the hypothesized mechanistic models for the first time. The polypeptide chain of squalene-hopene cyclase is organized in a dumbbell shape of homodimer constituted of two sets of  $\alpha$ - $\alpha$  helices bundles (Figure 1.17).<sup>88</sup> The enzymatic active site of squalene-hopene cyclase is located deeply in the large hydrophobic cavity. The catalytic center of this membrane-bound enzyme is accessible through a hydrophobic channel connecting with the non-polar part of the cytoplasmic membrane. The active sites are mainly lined by several highly conserved aromatic residues, which were expected to stabilize the individual intermediate carbocations and to determine simultaneously the folding manner of the substrate during the cyclization reaction. Asp-376 together with another two Asp residues (Asp-374 and Asp-377) and one His residue (His-451) are believed to initiate the cyclization cascade via protonating the terminal double bond. The termination reaction site was presumably represented by a polar amino acid (Glu-45) associated with a water molecule for effectively abstracting the proton or hydroxylating the cationic position (Figure 1.18).<sup>51, 87-89</sup>

After determination of the crystal structure of squalene-hopene cyclase from *A. acidocaldarius*, the structure-based mutagenesis could be carried out. The isolated diverse early-terminated products from different squalene-hopene cyclase mutants represented the stabilizing interaction between enzymatic active site residues and the respective unstable carbocationic intermediates. Several putative cation-stabilizing residues, including Glu-45, Trp-169, Ile-261, Gln-262, Pro-263, Trp-312, Phe-365, Asp-374, Asp-376, Asp-377, Tyr-420, His-451, Trp489, Tyr495, Gly-600, Phe-601, Phe-605, Leu-607, Tyr-609, and Tyr-612 have been examined for their importance in



different stage of squalene cyclization.<sup>51, 54, 88, 90-101</sup> The complete mutagenesis analysis of these functional residues participating in the SHC-catalyzed cyclization process from squalene to hopene/hopanol was summarized recently.<sup>102, 103</sup> Obviously, these mutagenesis studies not only provided the direct evidence about the previously hypothetic intermediates, but also elucidated the detailed interactions between enzyme and reactants from the view of biological triterpene synthetic system. Interestingly, in contrast to squalene-hopene cyclase, there were only few functional residues have been illustrated for the importance in facilitating the formation of lanosterol or cycloartenol within the corresponding oxidosqualene-cyclase enzymatic active site, solely based on the structure of the bacterial squalene cyclase.<sup>30, 56, 104</sup>

In 2004, Thoma *et al.* successfully solved the crystal structure of human OSC complexed with the lanosterol molecule and created an important milestone for understanding the triterpene synthase family.<sup>105</sup> A monotopic protein containing a membrane-inserted plateau with 25 Å diameter and a substrate entrance channel was observed in this crystallized structure (Figure 1.19). On the basis of the X-ray crystallographic analysis of human OSC structure, Thoma proposed an elegant exposition for illustrating the lanosterol synthase-triggered cyclization. At the top of the cavity in human OSC crystal structure, Asp-455 was hydrogen bonded with the lanosterol hydroxyl group, consistent with the proposed role of Asp-455 as the general acid that donates a proton to the epoxide oxygen of the substrate to initiate the cyclization cascade. The side chain of Asp-455 was also observed to form a hydrogen bond with Cys-456 and Cys-533 for enhancing its acidity. The reprotonation of Asp-455 could be achieved by the bulk solvent in a small polar channel through the hydrogen-bonded solvent network with Glu-459, spatially near the Asp-455 (Figure 1.19).

Moreover, the hydrophobic enzymatic active site was observed to be lined by

abundant highly electron-rich aromatic residues as in the case of previous SHC structure. Among them, Trp-387, Phe-444, Tyr-503, and Trp-581, which are oriented around the A/B ring of lanosterol molecule, might stabilize the tertiary intermediate cation at C-6 or C-10 through the cation- $\pi$  interaction. The energetically unfavorable boat conformation of the B-ring was thought to be stabilized via the spatially mediated pressure from Tyr-98. The unstable secondary anti-Markovnikov cyclohexyl carbocation at C-14 which is produced during the C-ring cyclization should be stabilized via the side chain of His-232 and Phe-696 at the appropriate orientation. Moreover, Thoma also suggested that His-232 and Tyr-503 are possible candidates for the role of a general base that accepts the C-9 proton to terminate the cyclization/rearrangement reactions cascade (Figure 1.20).

The availability of the long-awaited crystal structure provides an opportunity for altering the catalytically important residues and consequently modifying the product patterns. These experimental results could further verify the cyclization mechanism. The more detailed structure-based illustration for the site-directed mutagenesis studies or for the homology modeling structural constructs will be described in the following chapters.

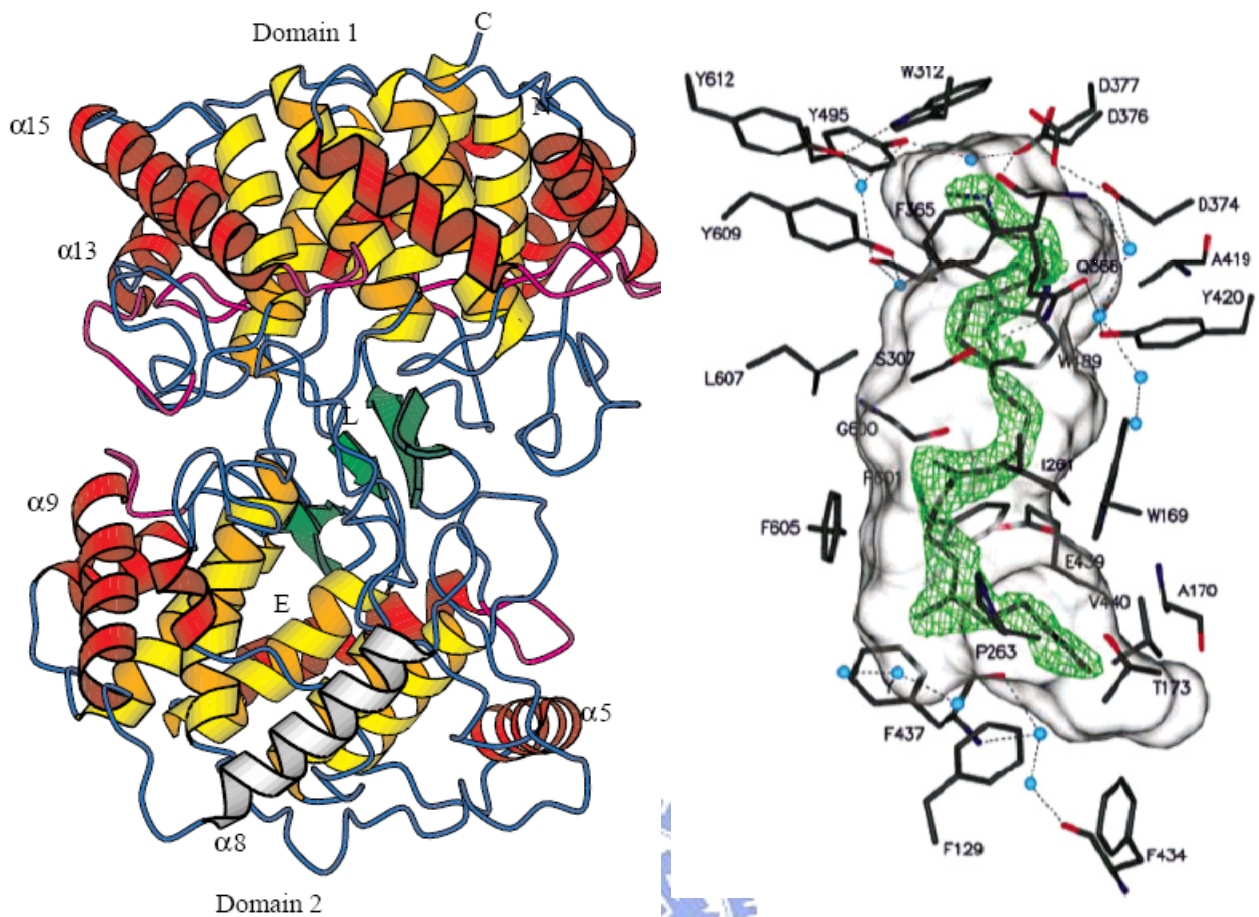


Figure 1.18 (Right) The overall structure of *A. acidocaldarius* squalene-hopene cyclase (SHC).<sup>89</sup> N and C: NH<sub>2</sub>- and COOH-termini; L: the position of the competitive inhibitor LDAO; E: the entrance of the active site channel; Color code: internal (yellow) and external (red) barrel helices;  $\beta$  structure (green); QW-motifs (purple); and helix  $\alpha$ -8 is the membrane-binding region (white). (Left) The putative active site cavity in the squalene-hopene cyclase.

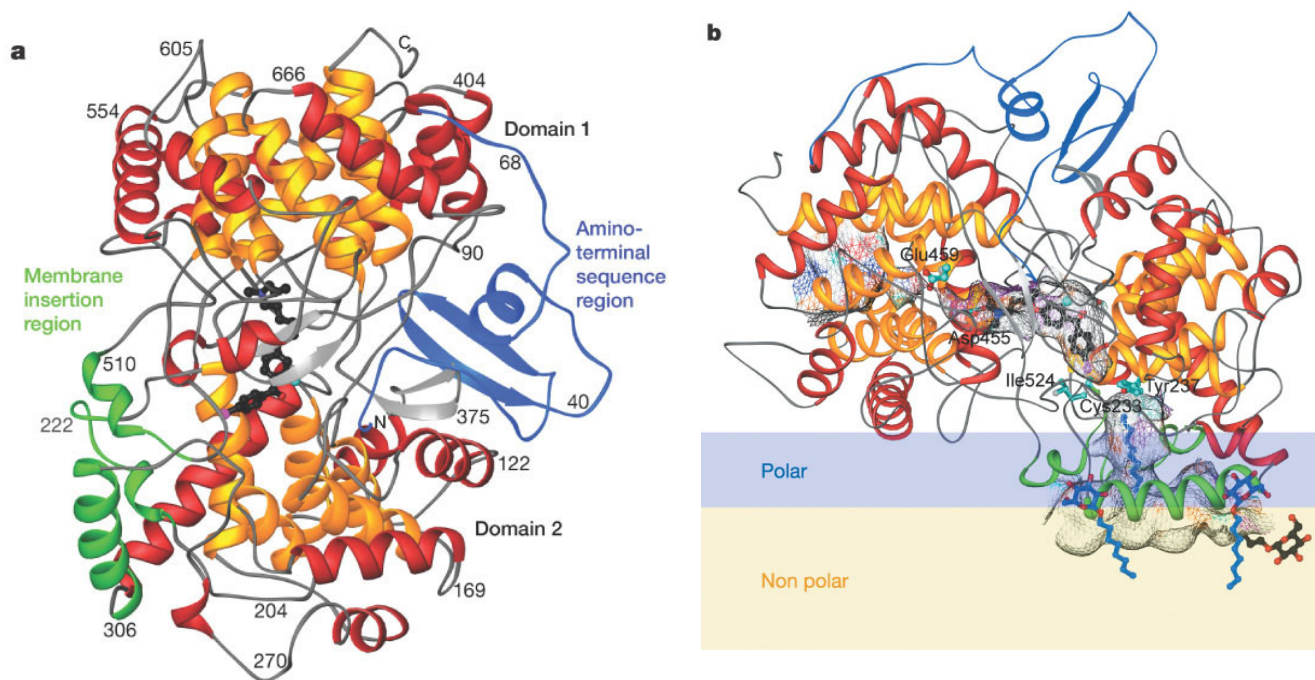


Figure 1.19 Crystal structure of human OSC<sup>105</sup> (a) Ribbon diagram of human OSC. The C and N termini and several sequence positions are labeled. The inner barrel helices are colored in yellow. The bound inhibitor, Ro48-8071 (black), indicates the location of the active site. (b) The orientation of OSC is relative to one leaflet of the membrane, Ro48-8071 binds in the central active-site cavity.<sup>105</sup>

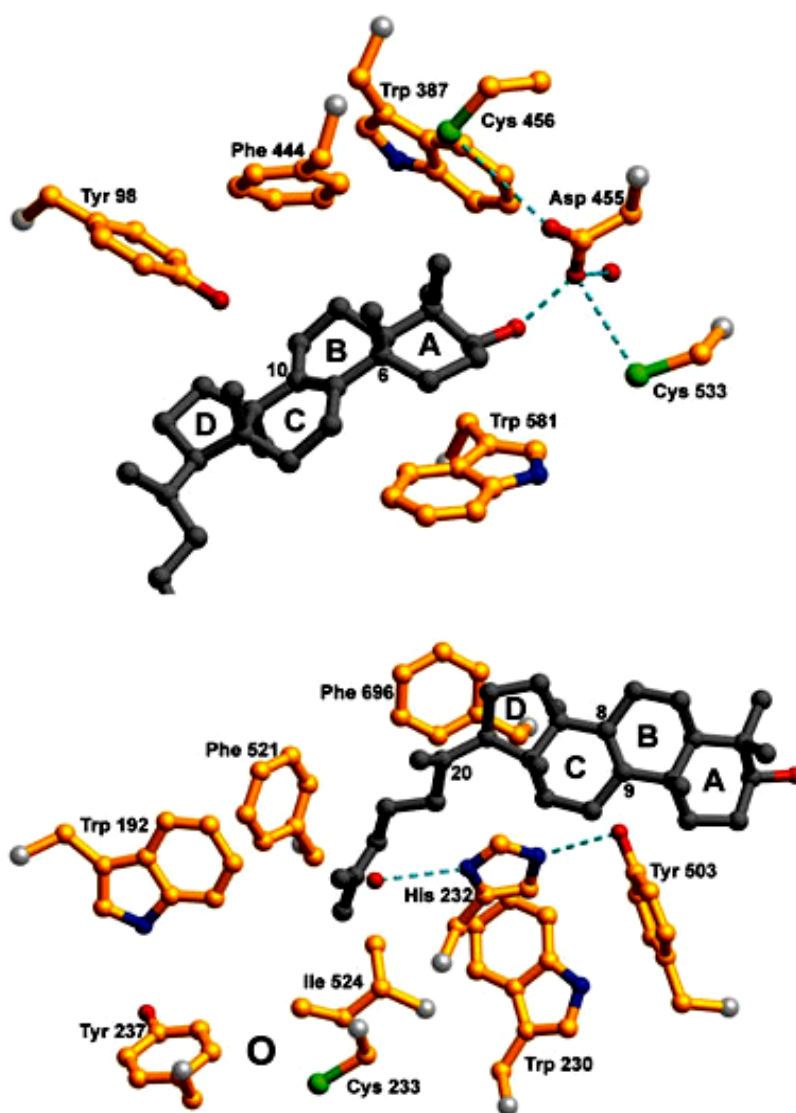


Figure 1.20 Local view of crystal structure of human OSC.<sup>105</sup>

## Chapter 2

### Thesis Organization

As previously described, the relationship between the cyclization mechanism and cyclase enzyme itself is extremely interesting. The hypothesis concerning the possible genetic evolution from the “primitive” ancestor to modern respective oxidosqualene cyclases has been proposed. Putative models further support the concept that specific motifs in the cyclase enzymes may account for the functional differences. Recently, the X-ray structures of squalene-hopene cyclase from *A. acidocaldarius* and oxidosqualene-lanosterol cyclase from *H. sapiens* have been elucidated. These three-dimensional structures provided the basis for understanding the variability and biochemical characteristics of the triterpene cyclases.

Due to the unique role in the biosyntheses of ergosterol, cholesterol, and phytosterol, OSCs have become potent targets for the development of antifungal, anticholesterolemic, and phytotoxic agents. Searching for enzyme inhibitors is a rational and efficient approach for the drug discovery. Another purpose in designing enzyme inhibitors is to obtain the information on the mechanism and the active site residues involved in catalysis.

There are several research issues, either from mutagenesis approaches or from bioorganic characterization, described in this thesis.

In the mutagenesis approach, we investigated the structure-function relationships of oxidosqualene cyclase. First, we used the alanine-scanning and site-directed/saturated mutagenesis to elucidate the specific amino acid residues responsible for the cationic intermediates stabilization, hydride/methyl groups' rearrangement, and the deprotonation reaction. Second, the chimeric enzyme library



between different cyclases was constructed using domain swapping technology to determine the functional domains responsible for the product specificity. In addition, we also tried to utilize the homology modeling studies and quantum mechanic calculation for reasonably illustrating the mutated effect.

Regarding to bioorganic characterization, we first tried to purify the oxidosqualene-lanosterol cyclase from bovine liver. The information obtained from the peptide sequencing was used to clone the OSC gene from a bovine liver cDNA library. After functional expression of the OSC enzyme, the potent inhibitor, Ro48-8071, was used as a photoaffinity labeling probe to elucidate the possible binding mode. Second, several Ro48-8071-based fluorescent probes were developed. The inhibition studies and the fluorescence spectrometric examination by using these newly developed probes were carried out. Third, an *in vitro* selection technology, systematic evolution of ligands by experimental enrichment (SELEX), was utilized to search the potential OSC-binding aptamers. The biochemical and biophysical properties of the identified aptamers were further analyzed. These OSC-binding aptamers may provide the opportunity for diagnostic applications or therapeutic purposes.

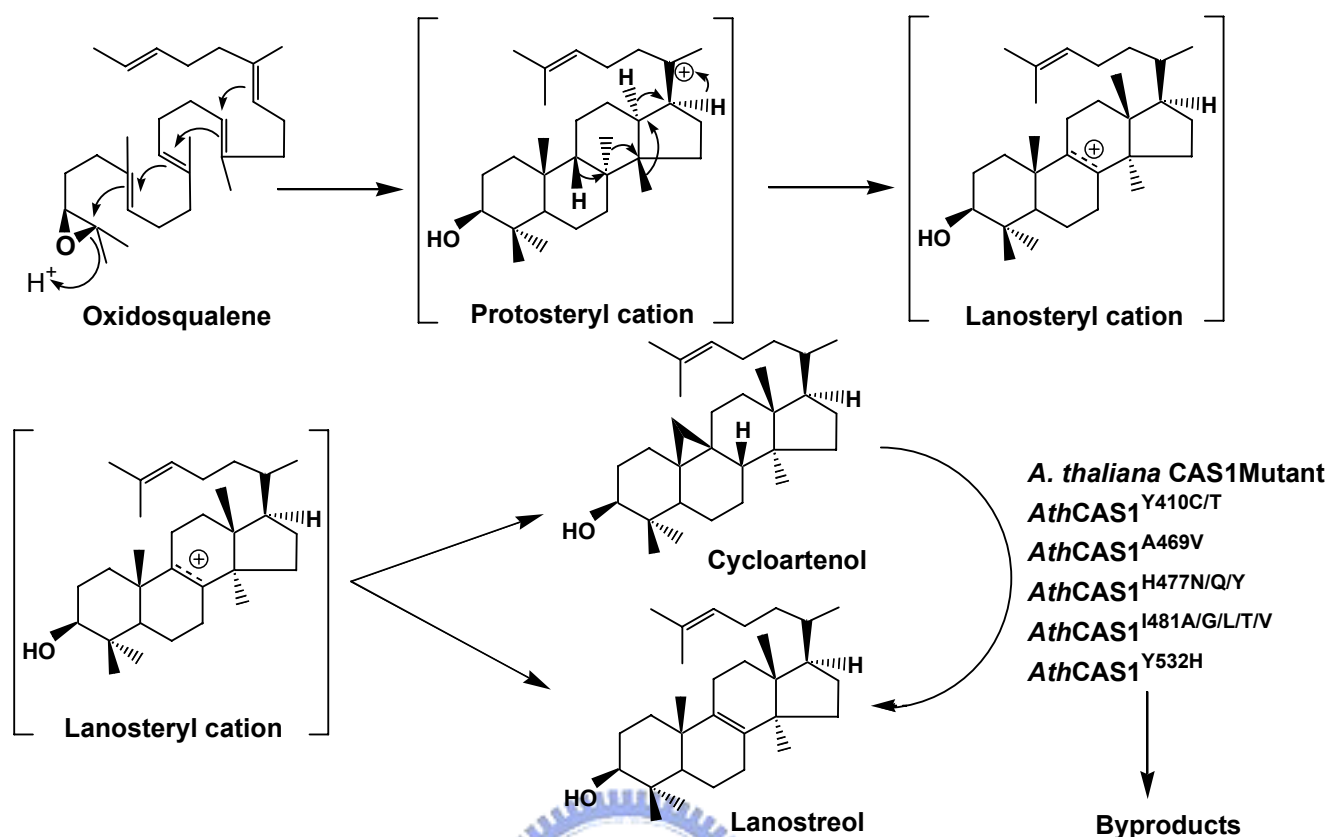
## Chapter 3

### *Mutagenesis Approach to Investigate the Putative Active-Site Residues from Oxidosqualene Cyclase*

#### 3.1 Introduction

The product specificity and the stereoselectivity of the oxidosqualene-catalyzed cyclization reactions are believed to be rigorously controlled via the ingenious mediation between substrate conformations and the enzyme active site residues.<sup>2</sup> For example, cycloartenol and lanosterol are synthesized by cycloartenol synthase and lanostreol synthase, respectively. The cyclization processes between these two cyclase-catalyzed reactions are almost identical except for the final deprotonation step. The direct proton abstraction at C-8/C-9 position or the alternative cyclopropyl ring closure at C-9/C-19 position resulted in the formation of either lanosterol or cycloartenol, respectively. In spite of the highly parallel reaction, cycloartenol synthase and lanostreol synthase share only 36% sequence identity. Accordingly, the amino acid difference between these two cyclases might probably imply the role in determining the catalytic distinction. In order to elucidate these catalytically important residues that direct the specific deprotonation reaction between cycloartenol synthesis and lanosterol synthesis, the genetic complementary characteristic of yeast provides an excellent system for analyzing the heterologous expression of mutated cycloartenol synthase. Because that lanosterol synthase (ERG7) is the sole oxidosqualene cyclase in the yeast cell, the *erg7* deficient strain could not independently grow without exogenous addition of ergosterol. Accordingly, if some of the heterologous cycloartenol synthase mutants acquired the lanosterol synthase activity to effectively

convert oxidosqualene into lanosterol, yeast transformants could grow without exogenous ergosterol. Identifying these product-specificity altering residues would illustrate the molecular distinction of sterol biosynthetic cyclases between yeast and plant cells, which specifically guides the unique cyclization products, respectively. A random mutated cycloartenol synthases library was generated by either chemical mutation with hydroxylamine mutagen, or the spontaneous genetic mutation in a DNA repair-system deficient strain.<sup>106</sup> In parallel, the similar strategy based on the spontaneously generated cycloartenol synthases mutants in either *A. thaliana CASI* gene or *D. discoideum CASI* gene were also independently carried out.<sup>107, 108</sup> Five distinct single-point mutations of oxidosqualene-cycloartenol synthase from *A. thaliana* have been identified to alter the product specificity from cycloartenol to lanosterol.<sup>106</sup> These five single-point mutations are Tyr410Cys, Ala469Val, His477Tyr, Ile481Thr, and Tyr532His. In addition to lanosterol, Ala469Val, Ile481Thr, and Tyr532His mutants also produced a monocyclic product, achilleol A. A structural homology model showed that Tyr-410, Ile-481, and Tyr-532 were located within the active site of cycloartenol synthase, implying that these three amino acids might contact with the substrate.<sup>106</sup> The functional importance of Tyr-410 and Ile-481 in oxidosqualene-cycloartenol synthase were also reported by Mutsuda and his colleagues.<sup>107, 109, 110</sup> Moreover, the second-sphere residue His-477 was further described for its effect on the cycloartenol synthase activity via electrostatic interaction with the side chain of Tyr-410.<sup>111</sup> The schematic representation of functional residues which changed the product specificity in *A. thaliana* oxidosqualene-cycloartenol synthase is listed in the Scheme 3.1.



Scheme 3.1 The *A. thaliana* CAS1 mutants changed their product specificity from cycloartenol to lanosterol.

Subsequently, the corresponding amino acids relative to the above demonstrated product specificity-determining residues were chosen to examine their catalytic role in the lanosterol synthase. Two functional residues Val-454 and Thr-384, the corresponding residues of Ile-481 and Tyr-410 in *A. thaliana* CAS1, exhibited the catalytic signification for the lanosterol production, respectively. The steric effect of Val-454 participating in the substrate folding and facilitating the cyclization process was illustrated from either Val454Ala or Val454Gly mutated substitution.<sup>112</sup> The spatially smaller substitution might enlarge the active site cavity, perturb the substrate conformation, and result in the premature termination at monocyclization step to isolate the monocyclic achilleol A.<sup>112</sup> Interestingly, the recently discovered putative camelliol C synthase from *A. thaliana*, which catalyzes the oxidosqualene into the predominant monocyclic camelliol C, was found to be strictly conserved by the

smaller alanine residue in the corresponding position of Val-454, further implying the steric importance of this residue for the polycyclic triterpene biosynthetic cyclase.<sup>113</sup> Moreover, the Thr384Tyr mutation of *S. cerevisiae* ERG7, which exchanged the highly conserved threonine residue in lanosterol synthase for the tyrosine residue in cycloartenol synthase, slightly altered the control strength for the deprotonation reaction and resulted in generation of alternatively deprotonated by-products, including parkeol and lanost-24-ene-3 $\beta$ ,9 $\alpha$ -diol.<sup>114</sup>

Among these product specificity-determining residues, the functional role and catalytic importance of Tyr-532 within oxidosqualene-cycloartenol synthase or its corresponding residue Tyr-510 within oxidosqualene-lanosterol cyclase still was ambiguous. First, the tritiated suicide inhibitors were specifically labeled with the Tyr-510 residue of *S. cerevisiae* ERG7, and produced several truncated cyclization products, indicating the functional significance of Tyr-510 in the oxidosqualene cyclization process.<sup>56</sup> Moreover, the corresponding residue Tyr-532 of the *A. thaliana* CAS cyclase was responsible for the product specificity from cycloartenol into lanosterol.<sup>106, 108</sup> In parallel, various abortive cyclization products isolated from the derivative mutations at the corresponding Tyr-420 in *A. acidocaldarius* SHC further validated the importance of this critical tyrosine residue.<sup>95, 101</sup>

Structural illustration either from homology models or crystal structures also provides some insights into this puzzling residue. The Tyr-532 of the *A. thaliana* CAS enzyme was observed as one of the hydrogen-bonding network residues that guides the proton abstraction direction from C-19 of cycloartenol to the active-site acidic residue for the reprotonation reaction.<sup>111</sup> Moreover, from the recent structural determination of human oxidosqualene-lanosterol cyclase, Thoma proposed another speculation on the role of Tyr-503 (relative to Tyr-510 in the *S. cerevisiae*) that directly assists the deprotonation reaction at the lanosteryl C-8/C-9 cation. The

phenolic oxygen of Tyr-503 was found to form a hydrogen-bond with the proposed catalytic base residue, His-232 (relative to His-234 in the *S. cerevisiae*), for abstracting the proximal proton.<sup>105</sup> This hydrogen-bonded His:Tyr catalytic dyad has also been illustrated from the *A. thaliana* CAS enzyme.<sup>111, 115</sup> In order to further substantiate the functional role of Tyr-510 position as well as the neighboring residues in the yeast oxidosqualene-lanosterol cyclase, the amino acid region from Thr-509 to Ile-513 (<sup>509</sup>TYEKI<sup>513</sup>) of *S. cerevisiae* ERG7 were originally subjected to the alanine-scanning mutagenesis, *in vivo* functional studies, and *in vitro* product analysis. Moreover, the results of multiple sequence alignment in all of the known oxidosqualene cyclases showed that tyrosine residue was conserved at the corresponding position in both lanosterol synthase and cycloartenol synthase, which triggers the substrate via the “pre-chair-boat-chair” conformation. Conversely, lupeol synthase and  $\beta$ -amyrin synthase that generate the tetracyclic dammarenyl cation via the “pre-chair-chair-chair” conformation strictly maintain the tryptophan residue at the relative position. This correlation implies the importantly catalytic discrepancy of this residue in all of the cyclase enzymes. Thus, due to the observation of conservative difference between ERG7/CAS1 and  $\beta$ -amyrin/lupeol synthase, an additional Tyr510Trp substitution was also carried out to investigate the possible effect on product diversity. Moreover, in order to examine the influence of different electronic environment on the cyclase activity, the Tyr510Lys mutant was also prepared.

However, in order to completely comprehend the relationship of His:Tyr hydrogen-bonding catalytic dyad between the Tyr-510 and His-234, as well as to understand that how the subtle changes in the ERG7 active site dramatically impact on the intrinsic product specificity, we generated the site-saturated mutations on the Tyr-510 from *S. cerevisiae* ERG7 and expressed the respective ERG7<sup>Y510X</sup> mutations as its sole oxidosqualene cyclase to characterize each of mutated enzymes.<sup>116</sup> The

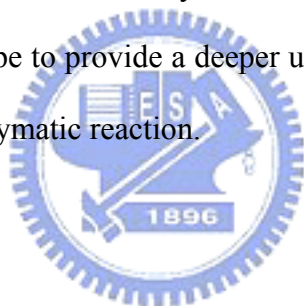
power of site-saturated mutagenesis analysis that exchanged all other nineteen proteinogenic amino acids for the specific residue could provide the unbiased opportunity for fully elucidating the influence upon the active site residue substitution from steric effect, orientation, size, polarity, or electrostatic propriety. Accordingly, we previously have performed the site-saturated mutagenesis experiment on the His-234 residue of *S. cerevisiae* ERG7. The isolated diverse product profile indicated that His-234 of ERG7 might play a key role in stabilizing various carbocationic intermediates and also guiding the deprotonation reaction.<sup>117</sup> We also established that the Trp-232 and Phe-445 might act as the catalytically important residues for the *S. cerevisiae* ERG7-catalyzed reaction, by using the similar site-saturated mutagenesis.<sup>118, 119</sup>

Three previously mentioned product specificity determining residues (Ala-469, His-477, and Ile-481) of cycloartenol synthase were observed to locate within a region of fifteen amino acid residues upstream to the putative active site Asp-483 of *A. thaliana* CAS1 enzyme. Based on the multiple sequence alignment, we speculated that the corresponding region of Ala469-Ile481 from other cyclase enzymes might also be involved in determining the oxidosqualene cyclization cascade. Thus, a series of amino acids residues <sup>441</sup>GAWGFSTKQGYT<sup>453</sup> from *S. cerevisiae* ERG7 were subjected to the alanine-scanning mutagenesis to examine their catalytic role. Among these single-point mutations, the Trp443Ala, Phe445Ala, and Lys448Ala substitutions failed to complement the *erg7* disruption, indicating the functional importance of these three residues in the oxidosqualene cyclization reaction. In order to explore the role of these residues, the mutation-induced products profiles of Trp443Ala, and Lys448Ala mutations were further analyzed via expression of the individual ERG7 mutants in a novel *hem1/erg7* double knockout yeast strain. Moreover, in order to further understand the functional role of Trp-443, site-saturated mutagenesis on



Trp-443 for another nineteen amino acids were recently carried out.<sup>120</sup> The possible catalytic role of Trp-443, Phe-445, and Lys-448 within the oxidosqualene-lanosterol cyclase was further characterized using the homology modeling strategy.

In this chapter, we will examine the detailed mutational effect of several functionally important residues, especially on the Tyr-510, Trp-443, and Lys-448 residues of *S. cerevisiae* ERG7 in the oxidosqualene cyclization cascades. Other functionally critical residues will be described in the next chapter from the homology modeling structure insight or the quantum mechanistic illustration. We will also illustrate the design and construction of the novel *hem1/erg7* double knockout yeast strain, TKW14, for effectively facilitating the analysis of the inactive oxidosqualene-lanosterol cyclase mutants. By utilization of these molecular biological mutational approaches, we hope to provide a deeper understanding about the complex cyclization/rearrangement enzymatic reaction.



## 3.2 Results and Discussion.

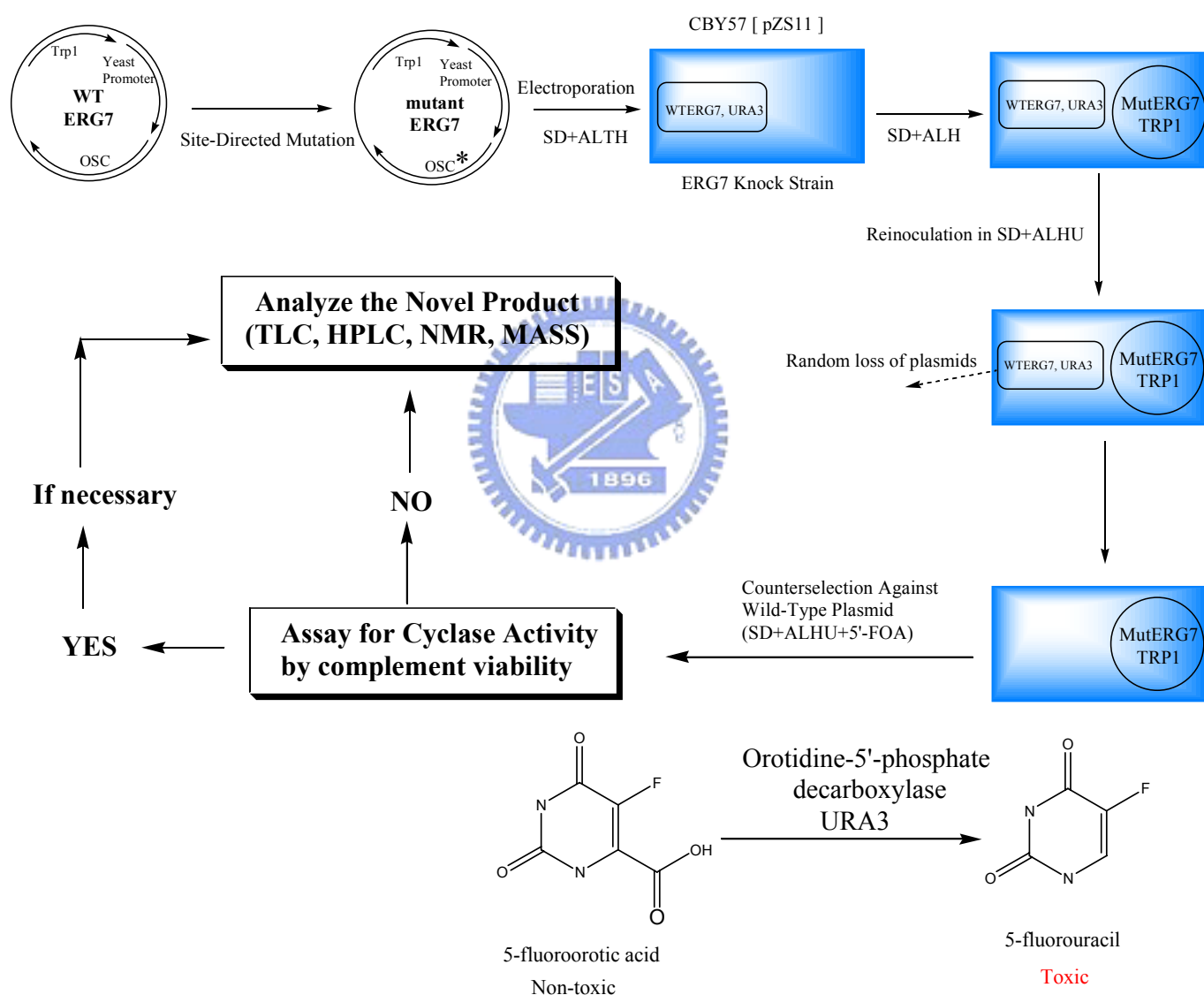
### **3.2.1 Generation of the alanine-scanning mutations or the site-directed mutations of *S. cerevisiae* *ERG7* gene**

The alanine-scanning mutagenesis approach was first used to exchange the corresponding amino acid residues from the region of Thr-509 to Ile-513 in the *S. cerevisiae* *ERG7* for the alanine residue. One of the parallel residues in *A. thaliana* CAS1 (Tyr-532) revealed the crucial role in the cyclization/rearrangement mechanism.<sup>106</sup> Theoretically, the alanine residue substitution that only alters the functional group beyond the  $\beta$ -carbon should not dramatically change the structure of protein or destroy the skeletal backbone. An additional tryptophan substitution of Tyr-510 was also included to investigate the possible effect on product diversity based on the conservative difference between “chair-boat-chair” and “chair-chair-chair” conformational pathways. Moreover, the Tyr510Lys substitution was further used to examine the possible product distinction caused by the electronic environment. All of the selected residues of *S. cerevisiae* *ERG7* were converted to alanine or other substitutions by using the PCR-based mutagenesis with the corresponding mutagenic primer pairs. A silent mutation was also concomitantly introduced to easily screen the desired mutants according to a restriction endonuclease analysis. The presence of the mutations was further verified by using sequencing determination. The Trp443Ala and Lys448Ala mutations were also constructed by the similar strategy.

### 3.2.2 Principle of the plasmid shuffle methodology

Plasmid shuffle involves the exchange of a plasmid bearing the wild type gene with a plasmid bearing the mutation of interest. Plasmid shuffle is a powerful tool for rapid functional analysis of the mutated gene of interest. Initially, a haploid yeast cell containing the lethal genomic target gene disruption was prepared and transformed with a plasmid bearing the wild-type gene for viability. Thus, the survival or the death of the yeast strain will rigorously depend on the presence of the wild-type gene plasmid. Any random loss of the wild-type gene plasmid would result in the death of the yeast cell. Moreover, the wild-type gene plasmid also possesses a selectable/counters selectable genetic marker for removing the wild-type gene plasmid exclusively. The *URA3* gene encoded orotidine-5'-phosphate decarboxylase which converts the 5'-fluoroorotic acid (5'-FOA) into the cell-toxic 5'-fluorouracil, was commonly used as the selectable/counters selectable marker. Thus, the yeast cells without *URA3* activity are resistant to the 5'-FOA. By using the counterselection against the wild-type gene plasmid possessing the *URA3* activity, the functional importance of the mutated gene of interest could be rapidly analyzed. After transformation of the plasmid bearing the mutated gene, the yeast transformants would contain two kinds of plasmids and maintain its viability due to the presence of first wild-type gene plasmid. Subsequently, the extra additional uracil, which is the end product of the *URA3*-involved biosynthetic pathway, could facilitate the partial loss of the first wild-type gene plasmid with a frequency up to 1% per generation. Counterselection against the wild-type gene plasmid with 5'-FOA was further used to confirm that the yeast cells must only retain the second transformed mutated gene plasmid. Therefore, the survival or the death of yeast transformants now depends on the functional activity of the mutated enzyme. If the mutated enzyme abolished the enzymatic activity and lost the complementary activity for the first wild-type plasmid,

the yeast transformants are expected to die. Conversely, if the functional activity was not dramatically influenced by the mutation, the yeast transformants would still survive on the counterselection plate with 5'-FOA. This selectable/counterselectable plasmid shuffle method provides a powerful tool for analyzing the mutated *S. cerevisiae* ERG7. The strategy of plasmid shuffle methodology is shown in the Scheme 3.2.



Scheme 3.2 Screening of inactive ERG7 mutants using a plasmid shuffle method. Genome type of CBY57[PZS11] is (*MAT a/α ERG7Δ::LEU2 ade2-101 his3-Δ200 leu2-Δ1 lys2-801 trp1-Δ63 ura3-52*) [PZS11]. The abbreviation, including A, L, T, H, and U in different cultural media represent the nutrition additive: Adenine, Lysine, Trptophan, Histidine, and Uracil for the yeast cell viability, respectively.

### 3.2.3 Screening of inactive ERG7 mutants by using the plasmid shuffle method

The genome type of haploid yeast CBY57 is (MAT a/ $\alpha$  *ERG7* $\Delta$ ::*LEU2 ade2-101 his3- $\Delta$ 200 leu2- $\Delta$ 1 lys2-801 trp1- $\Delta$ 63 ura3-52*). The yeast CBY57[pZS11], bearing a genomic disrupted *erg7* gene and a pZS11 plasmid with the wild-type *S. cerevisiae* *ERG7* gene, was transformed with these pRS314-derived *ERG7* mutation plasmids by electroporation. The genetic selectable markers involved in the pRS314-derived plasmids and pZS11 plasmid are *TRP1* gene and *URA3* gene, respectively. The plasmid shuffle strategy illustrated above was used to analyze the effect of mutation on the complementary ability to the *erg7* deficiency. The positive and negative control with the pRS314*ERG7*WT and pRS314 were also carried out to confirm the accuracy of the plasmid shuffle. Yeast transformants were first selected on the SD+Ade+Lys+His plates, and then re-selected on the SD+Ade+Lys+His+Ura +5'-fluoroorotic acid plates to elucidate the complementation effects. In addition, the plasmids whose encoded mutated cyclase failed to complement the cell viability were isolated and sequenced to further confirm the identity of the mutations. Among the results of plasmid shuffle method, all of the alanine-scanning mutations from Thr509-Ile513 produced colonies on the counter-selection plates with 5'-FOA, indicating that Thr509Ala, Tyr510Ala, Glu511Ala, Lys512Ala, and Ile513Ala mutations were active and did not dramatically influence the *ERG7* activity. In contrast, the Tyr510Lys and Tyr510Trp mutations failed to complement the *erg7* disruption, suggesting that different amino acid substitutions on Tyr-510 might impair the *ERG7* activity below the threshold level needed to support yeast cell growth and result in halting the cyclization/rearrangement reaction or altering the product specificity. Similarly, cyclase mutants bearing the Trp443Ala and Lys448Ala substitutions also failed to complement the *erg7* deficiency (Table 3.1).

Tabel 3.1 Complementation result of cyclase mutants in an *erg7* knockout strain, CBY57, by using plasmid shuffling counter selection. The symbol “+” indicated the mutated cyclase can complement the knockout strain. The symbol “x” indicated the inactive mutants.

### Complementation of :

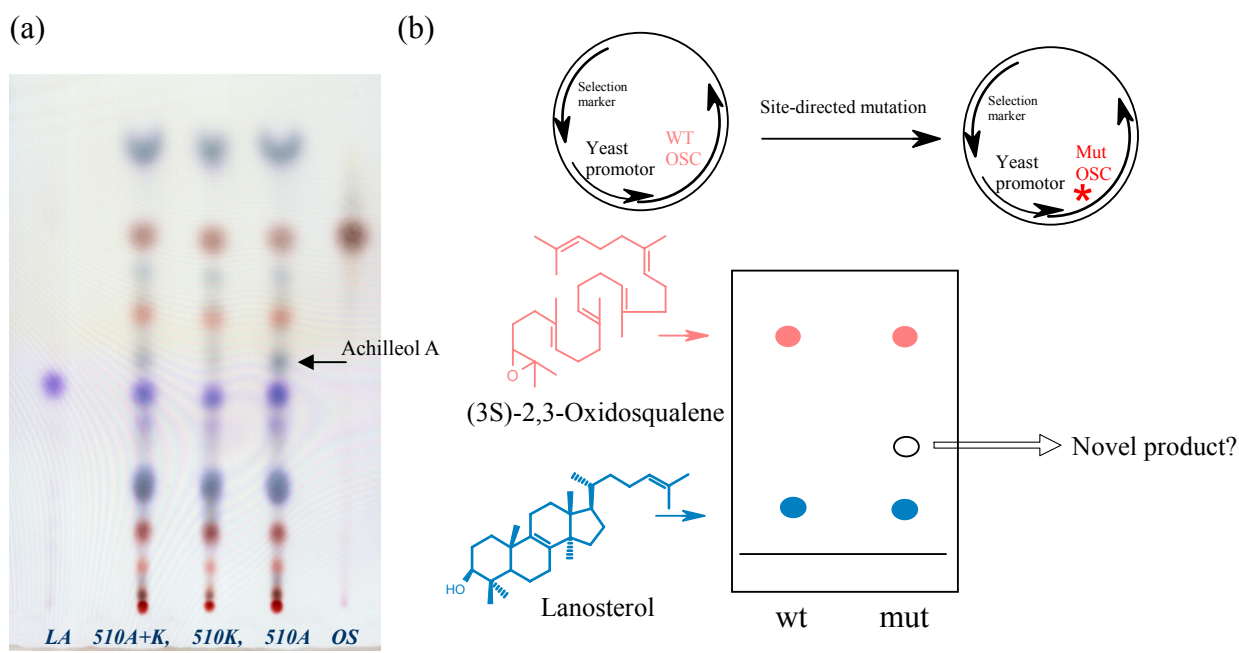
Mutation	CBY57
Wildtype(pRS314ERG7WT)	+
None(pRS314)	x
Thr-509-Ala	+
Tyr-510-Ala	+
Glu-511-Ala	+
Lys-512-Ala	+
Ile-513-Ala	+
Tyr-510-Lys	x
Tyr-510-Trp	x
Trp-443-Ala	x
Lys-448-Ala	x

#### 3.2.4. Lipid extraction, column chromatography, and product characterization

The crude nonsaponifiable lipids from yeast transformants with Tyr510Ala mutation were first prepared, spotted on the thin-layer chromatography (TLC plate), developed with 20% EA/Hexane, and then observed via anisaldehyde staining. Accordingly, consistent with functional complementation results, the lanosterol-positioned spot was expectably present on the TLC plate, while another obvious spot migrating between oxidosqualene and lanosterol was also observed (Figure 3.1). In order to examine the chemical structure of these triterpene products, the nonsaponifiable lipid extract from the large scale of cultural yeast was further prepared for the silica gel chromatographic purification. Monocyclic triterpene achilleol A which migrates between oxidosqualene and lanosterol was characterized via the GC-MS and NMR spectroscopic analyses. The detailed spectroscopic results of monocyclic achilleol A are shown in the Appendix 3.1. On the other hand, for the

lanosterol-positioned product, GC-MS spectroscopic analysis revealed that there are simultaneously two compounds with identical molecular mass ( $m/z = 426$ ) and same  $R_f$  value on the TLC plates. Repeated chromatographic purification exhibited that these two products were indistinguishable from one another in TLC analysis, silica gel chromatography, or even the high performance liquid chromatography (HPLC). Thus, the acetylation reaction with acetic anhydride/pyridine was subsequently carried out to modify the polarity of alcohol moiety (Figure 3.2 and Appendix 3.2). However, this acetylated modification does not allow for successful separation of these two indistinguishable NSL. Fortunately, the chemical assignments of these two compounds with either intact or acetylated form were finally confirmed from the  $^1\text{H-NMR}$ ,  $^{13}\text{C-NMR}$ , and different DEPT-NMR spectroscopic analyses, respectively (Appendix 3.3). For example, the distinct chemical shifts for the methyl proton groups and vinyl protons of lanosterol ( $\delta$  0.687 and 0.810 ppm) and parkeol ( $\delta$  0.637, 0.729 and 5.225 ppm) were obviously observed and compared with the available literature data.<sup>110, 121</sup> The  $^{13}\text{C-NMR}$  spectra of the acetylated products also displayed all 32 lanosteryl acetate and parkeyl acetate signals, and thus confirmed the identities of lanosterol and parkeol. GC-MS analysis of these compounds showed that the NSL extract of *S. cerevisiae* ERG7<sup>Y510A</sup> mutant accumulated achilleol A, lanosterol, and parkeol in a 27:39:34 ratios. The identities of resulting products were confirmed via co-injection with the authentic samples in the GC-MS analysis. In contrast, neither  $9\beta$ -lanosta-7,24-dien- $3\beta$ -ol nor any other oxidosqualene cyclization products with consistent molecular weight could be detected from the nonsaponifiable lipid extract of Tyr510Ala mutation.





Analysis of nonsaponifiable lipid by TLC

Figure 3.1 (a) Thin-layer chromatographic analysis of the nonsaponifiable lipid (NSL) extract from Tyr510Ala and Tyr510Lys mutants. Lanosterol was present as expected, while a significant amount of the second NSL (achilleol A) migrating between oxidosqualne and lanosterol was observed in the Tyr510Ala mutant. The authentic lanosterol and oxidosqualene were also included. (b) The scheme for illustrating the analysis of novel product generated from the mutated ERG7 on the TLC plate.

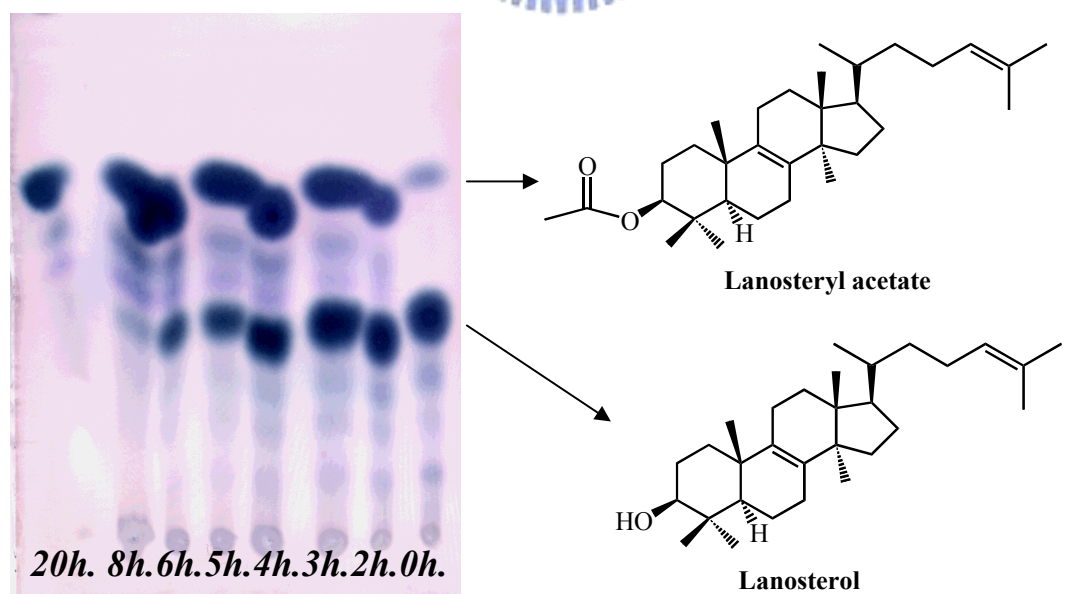


Figure 3.2 Acetylation with acetic anhydride/pyridine was used to modify the polarity of alcohol moiety. After stirring for 20 h, most of product with higher R<sub>f</sub> value was observed on the TLC plate.

### 3.2.5 The construction of novel gene disruption strain TKW14 for analysis of inactive mutated cyclase

In order to unambiguously determine the functional role of inactive mutations, a novel cyclase-deficient strain, TKW14, derived from CBY57[pZS11] was generated. Accordingly, the pZS11 plasmid bearing the wild-type *ERG7* was removed and the viability of yeast cells was mediated via the exogenous addition of the ergosterol. As it has been suggested that mutations in the heme biosynthetic pathway could allow the uptake of sterol from the media under aerobic conditions, efforts were thus made to incorporate a *hem1* gene disruption into the CBY57[pZS11] strain.<sup>122, 123</sup> Using homologous recombination technology, we have successfully exchanged the *HEM1* gene with a drug Geneticin (G418) resistance gene in the CBY57[pZS11] strain. In order to confirm the accurate genetic recombination in the *HEM1* gene, the genomic DNA of yeast transformants was extracted and verified via PCR based amplification with the *HEM1* gene terminus sequence primers. Moreover, the restriction enzyme mapping with *NcoI*, a unique restriction site for the Geneticin resistance gene (G418<sup>R</sup>) but not for *HEM1* gene, was further used for genetic verification (Figure 3.3). After the successful genetic knockout of the *HEM1* gene in the CBY57[pZS11] strain, the pZS11 plasmid was removed from the yeast cell. Based on the similar principle of plasmid shuffle, the presence of 5'-FOA in the selection plates would completely eliminate yeast containing the pZS11 plasmid. The viable yeast colonies which bear the genomic *hem1/erg7* double knockout, eliminate the pZS11 plasmid, and maintain cell viability via the uptake of exogenous ergosterol from the medium were renamed as the TKW14 strain. The genotype of novel haploid yeast TWK14 is (MATa or MATα *ERG7*Δ::*LEU2 ade2-101 lys2-801 his3-Δ200 leu2-Δ1 trp1-Δ63 ura3-52 HEM1*Δ::*kanR*).

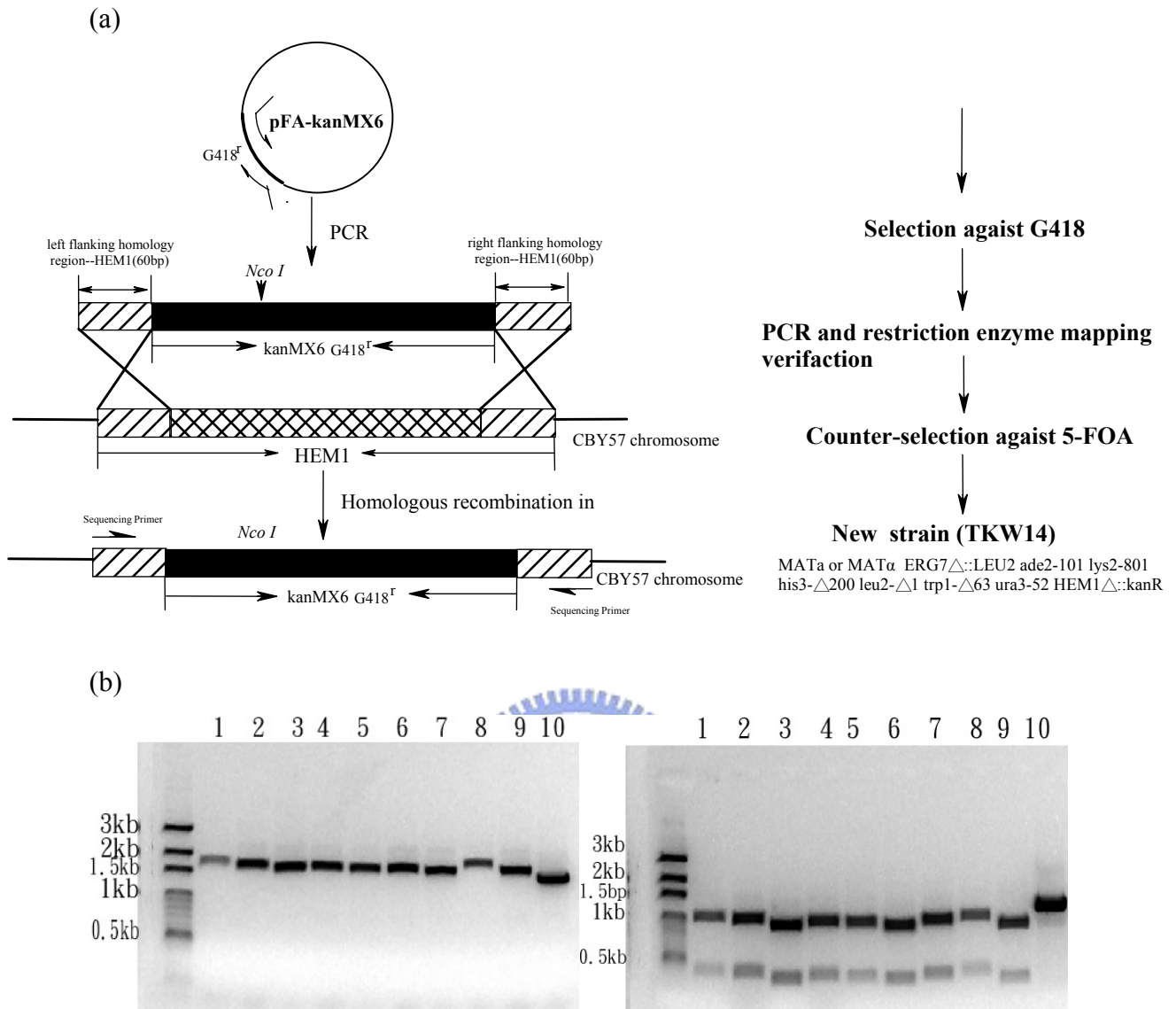
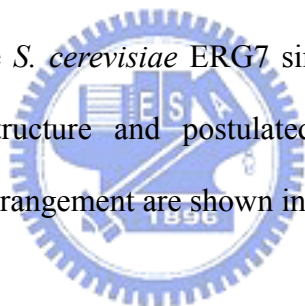


Figure 3.3 (a) The homologous recombination strategy for exchanging the *HEM1* gene with a drug geneticin resistance gene ( $G418^R$ ) in the *CBY57*[pZS11] strain. (b) The PCR amplicons (1.5 kb) of the genomic DNA with specific primers (left), and the result of restriction enzyme mapping with *NcoI*, a unique mapping site for Geneticin resistance gene ( $G418^R$ ), with the expected sizes of 1 kb and 500 bp (right). Lane 1~Lane 9 indicated the individual colonies from selection plates, Lane 10 showed the host *CBY57* strain, as a negative control.

### 3.2.6 Characterizing the mutant products in the novel gene disruption strain

The Tyr510Lys and Try510Trp mutated plasmids were each transformed into the novel TKW14 strain, respectively. As expected, the Tyr510Lys and Try510Trp substitutions failed to maintain the cell viability in the absence of the exogenous ergosterol. This result further supported the inactive characteristic from the previously mentioned plasmids shuffle experiments. Both mutants were then grown in the liquid media and harvested. Preliminary NSL extract showed that both mutations were less efficient than the native ERG7 and gave fewer amounts of NSL per liter. The chromatographic results with no lanosterol-positioned products isolation also demonstrated the inactive activity of these two mutations. In addition, two non-lanosterol products, which migrate between oxidosqualene and lanosterol, were detected from TLC and GC-MS analysis. After co-injection with the authentic samples, one of the triterpene alcohols was determined to be achilleol A (Figure 3.4). However, attempt to separate the other NSL always presented difficulties due to its indistinguishable characteristic with the monocyclic achilleol A. Thus, the detailed analyses of the partially purified mixture with the ratio of 1:0.88 were carried out via different GC-MS,  $^1\text{H-NMR}$ , and  $^{13}\text{C-NMR}$  spectra (Appendix 3.4). A multiplet signal at 5.216 ppm on  $^1\text{H-NMR}$  suggested that the second NSL product might contain a trisubstituted olefin distinct from that in the side chain. In addition, the chemical shifts for the region of methyl groups on  $^1\text{H-NMR}$  showed that  $\delta = 0.691$  and  $1.001$  ppm are the signals of achilleol A and the  $\delta = 0.806$  and  $0.946$  ppm belonged to the other NSL. Similarly, the  $\delta = 108.379$  and  $147.212$  ppm from  $^{13}\text{C-NMR}$  spectrum are the unique chemical shifts of achilleol A, whereas the  $\delta = 118.257$  and  $137.083$  ppm are the characteristic peaks of the second NSL. Fortunately, a natural monocyclic triterpene, camelliol C, which was extracted from sasanqua oil (*Camellia sasanqua*) exhibited a identical and consistent spectroscopic result with ours in GC-MS,

$^1\text{H-NMR}$ , and  $^{13}\text{C-NMR}$  analyses.<sup>124</sup> This result confirmed that the identity of the second compound is camelliol C, which was identified from the ERG7 mutations for the first time.  $^1\text{H-NMR}$  and GC-FID analysis of the non-lanosterol-positioned products showed that Tyr510Lys and Tyr510Trp mutants produced achilleol A and camelliol C in the ratio of 86:14 and 94:4, respectively. Interestingly, the achilleol A and camelliol C were also produced from the other inactive mutants of Trp443Ala and Lys448Ala. The analysis in the new TKW14 strain showed that these two mutants produced different ratio of monocyclic triterpene products from that of Tyr510 mutants. These results indicated that both Trp443Ala and Lys448Ala mutations might compromise the lanosterol production, halt the cyclization reaction at the monocyclic stage, and result in the accumulation of achilleol A and camelliol C as its end product. The ratio of product from the *S. cerevisiae* ERG7 single-point mutants are listed in Table 3.2. The chemical structure and postulated mechanism of the relevant oxidosqualene cyclization/rearrangement are shown in Scheme 3.3.



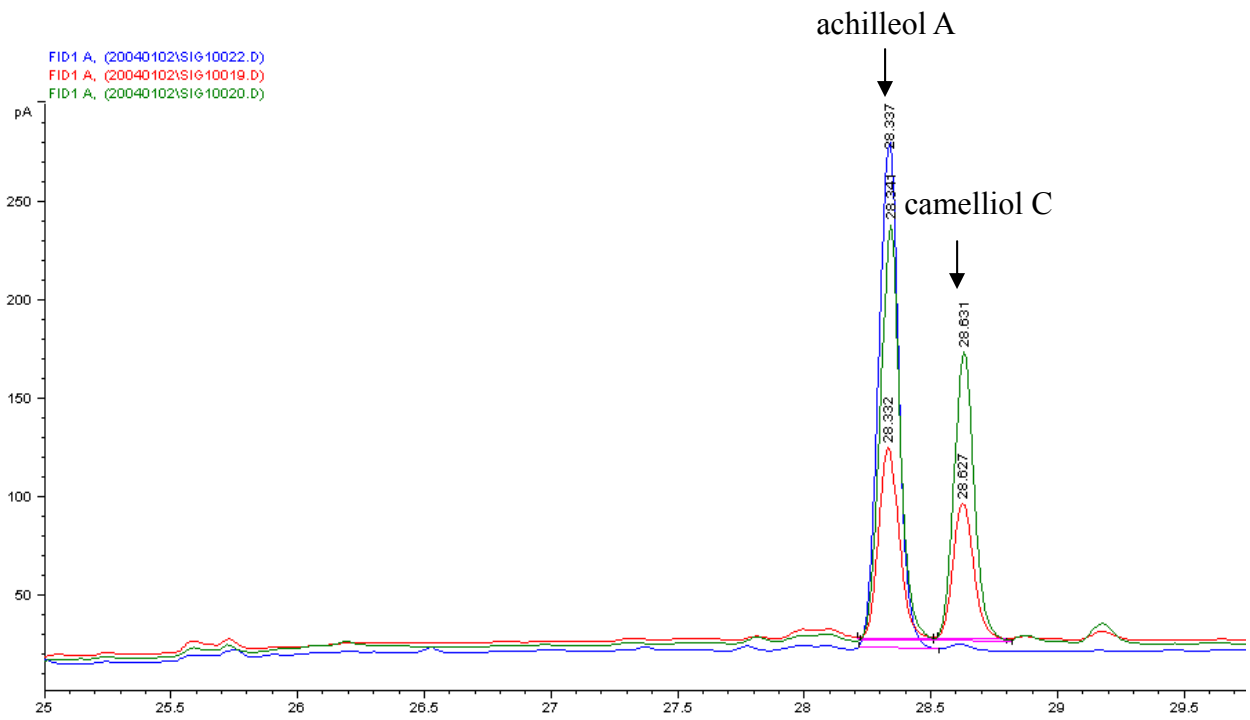
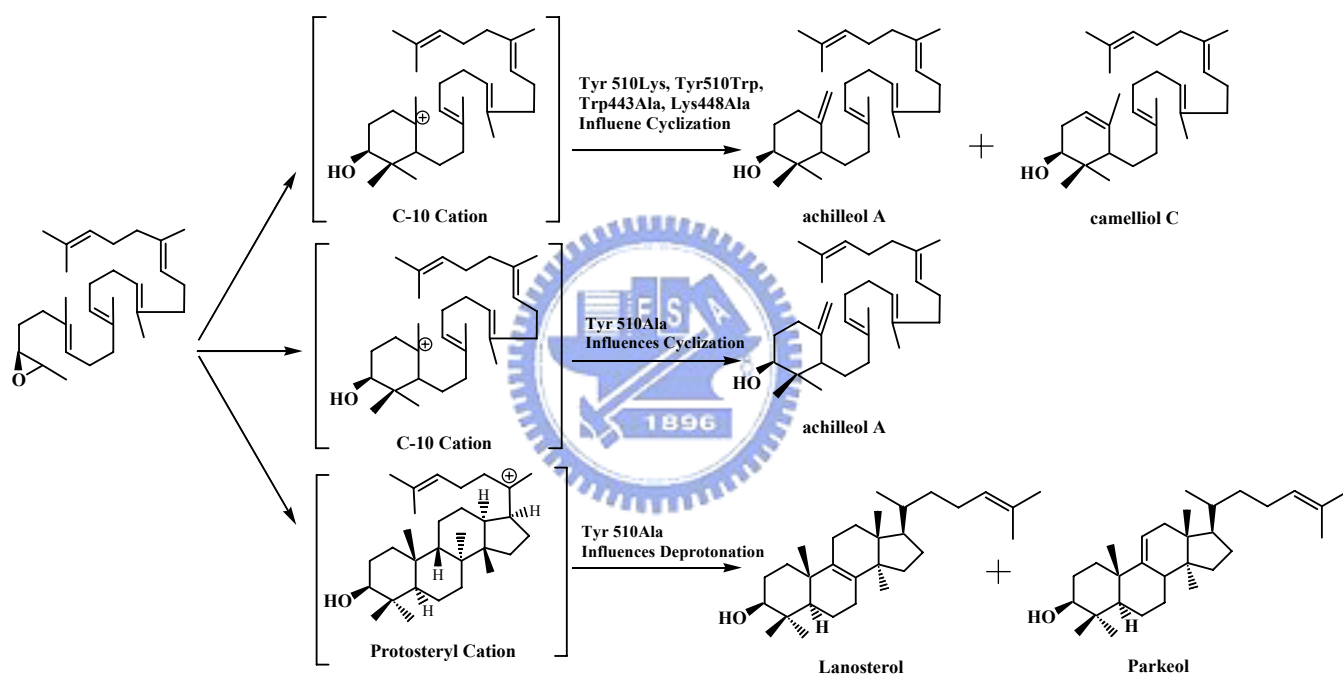


Figure 3.4 GC analyses of the nonsaponifiable lipid, derived from the yeast  $ERG7^{Y510K}$  mutant. Two non-lanosterol-positioned products, which migrate between oxidosqualene and lanosterol, were detected. The blue color line indicates the authentic achilleol A from  $ERG7^{Y510A}$  mutant. The red color line indicates the partially purified NSL from  $ERG7^{Y510K}$  mutant. The co-injection of the authentic samples is shown in the green color line.

Table 3.2 Product ratio from the various *S. cerevisiae* ERG7 single-point mutants

	achilleol A	camelliol C	Lanosterol	Parkeol
Tyr510Ala	27	-	39	34
Tyr510Lys	86	14	-	-
Tyr510Trp	94	4	-	-
Trp443Ala	68	32	-	-
Lys448Ala	70	30	-	-



Scheme 3.3 Oxidosqualene-lanosterol cyclase mutants from *S. cerevisiae* converted oxidosqualene to a variety of monocyclic and deprotonated compounds, including achilleol A, camelliol C, lanosterol, and parkeol.



### 3.2.7 Site-saturated mutagenesis approach to investigate the functional importance of the critical Tyr-510 residues

These saturated mutations of Tyr-510 in oxidosqualene-lanosterol cyclase were respectively created by using QuikChange Site-Directed Mutagenesis kit and confirmed by DNA sequencing using the ABI PRISM 3100 DNA sequencer.<sup>117-119</sup> These recombinant plasmids were then electroporated into the yeast TKW14 strain for analyzing the functional activity, as previously described.<sup>117-119</sup> The plasmids shuffle were also carried out with CBY57[pZS11] strain. From the functional complementary assay among these various *S. cerevisiae* ERG7<sup>Y510X</sup> mutants, the Tyr510Lys, Tyr510Arg, Tyr510Thr, Tyr510Pro and Tyr510Trp substitutions failed to complement the *erg7* deficiency in the absence of exogenous ergosterol. This result suggested that the different substitution of Tyr-510 might influence the enzymatic activity and abolish the cyclization process of oxidosqualene.

The nonsaponifiable lipid extracts from each mutant were subsequently characterized. The product profiles of *S. cerevisiae* ERG7<sup>Y510X</sup> mutants are listed in Table 3.3. Accordingly, the ERG7<sup>Y510X</sup> mutants produced diverse product profile ranging from monotonous to polycyclic compounds with molecular weight of  $m/z = 426$ . Based on the GC-MS spectrum as well as the NMR data, all of the compounds including achilleol A, camelliolC, (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien-3 $\beta$ -ol, lanosterol, and parkeol were respectively identified.<sup>115-117, 125</sup>

Table 3.3 The product profiles of *S. cerevisiae* TKW14 expressing the ERG7<sup>Y510X</sup> site-saturated mutants and two ERG7<sup>H234WY510X</sup> mutants.

amino acid substitution	(13 $\alpha$ H)-isomalabarica-				
	achilleol A	camelliol C	14(26),17,21-trien-3 $\beta$ -ol	lanosterol	parkeol
Gly	15	—	17	27	41
Ala	27	—	—	39	34
Val	4	—	17	38	41
Leu	—	—	26	74	—
Ile	4	—	11	78	7
Asp	9	—	11	40	40
Asn	5	—	9	9	77
Glu	6	—	27	18	49
Gln	5	—	26	14	55
His	45	—	24	5	26
Lys	87	13	—	—	—
Arg	—	—	—	—	—
Ser	—	—	4	40	56
Thr	—	—	—	—	—
Cys	—	—	67	33	—
Met	1	—	15	43	41
Phe	4	—	10	50	36
Trp	94	6	—	—	—
Pro	—	—	—	—	—
Tyr	—	—	—	100	—
product profiles of ERG7 <sup>H234XY510X</sup> double mutants					
H234W/Y510V	2	—	8	90	—
H234W/Y510W	99	1	—	—	—
H234Y/Y510A <sup>117</sup>	—	—	—	100	—
Y510W	94	6	—	—	—
H234W <sup>117</sup>	—	—	—	—	100

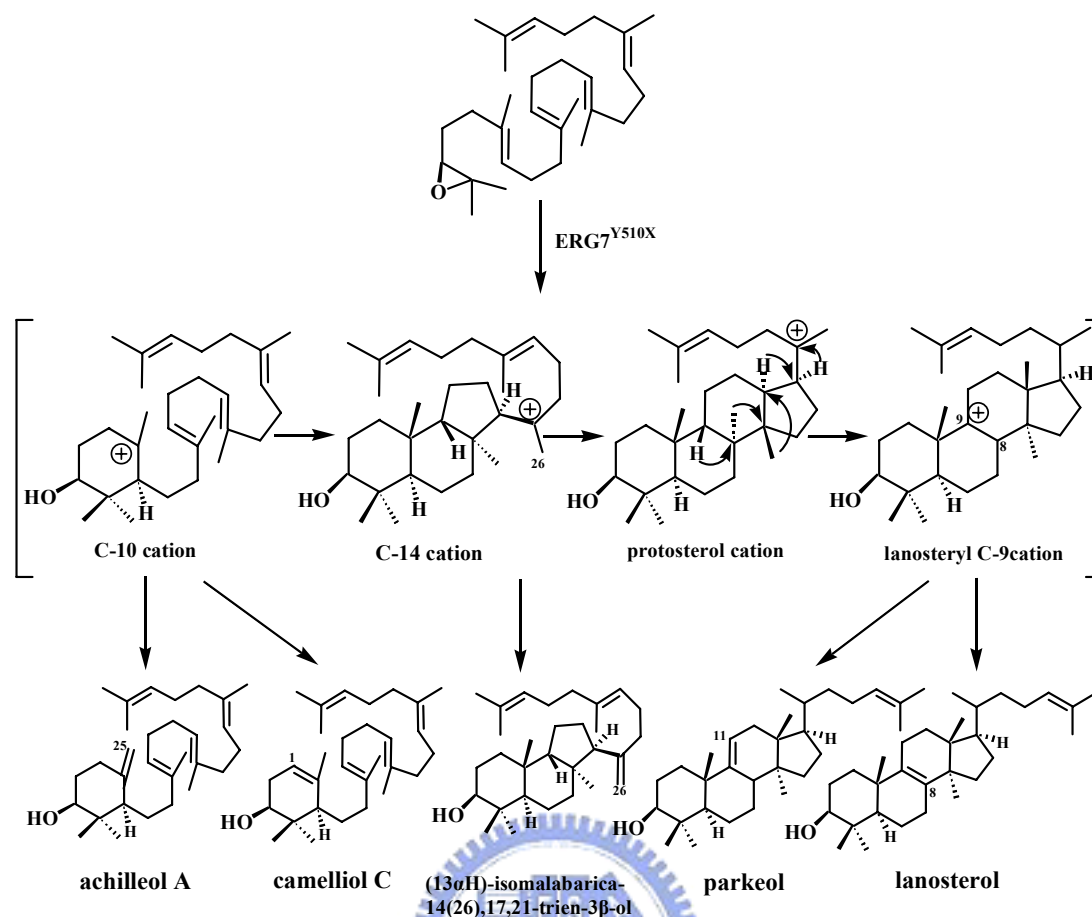
For the products analysis of Tyr-510 inactive mutants, replacement of tyrosine with arginine, proline and threonine failed to generate any product with molecular mass of  $m/z = 426$ , consistent with the genetic selection results. No product production might be caused by the incorrect protein folding of *S. cerevisiae* ERG7 from the individual amino acid substitution on the Tyr-510 position. The strongly basic arginine and the conformational restricted proline might directly disrupt the protein structure. Interestingly, in contrast to the multiple products profile from other polar residue substitutions, the exact reason for no product extraction from Tyr510Thr mutant remained unclear. Moreover, another two nonviable mutants, ERG7<sup>Y510K</sup> and ERG7<sup>Y510W</sup> generated trace amount of monocyclic product, either achilleol A or camelliol C, as the previous section mentioned.<sup>125</sup>

For the viable mutants, similar product profiles with different ratio were observed. In most of the viable mutants, lanosterol or its alternatively deprotonated product, parkeol, distributed predominantly over the entire of product profiles (near or over 70%), except for the ERG7<sup>Y510H</sup> or ERG7<sup>Y510C</sup> mutants. This finding suggested that the cyclization/rearrangement cascade of oxidosqualene in the TKW14[pERG7<sup>Y510X</sup>] strains is almost unaffectedly complete until the final deprotonation step. Therefore, the position of Tyr-510 or its corresponding residue might be critical for the final deprotonation reaction as well as for the product specificity.<sup>106, 108</sup> The nonpolar, aliphatic amino acid substitution, especially in the ERG7<sup>Y510L</sup> or ERG7<sup>Y510I</sup> mutants, might provide the selective pressure for the precisely deprotonated control to produce the abundant lanosterol. In contrast, the dominant production of parkeol in ERG7<sup>Y510S</sup>, ERG7<sup>Y510N</sup>, ERG7<sup>Y510E</sup>, or ERG7<sup>Y510Q</sup> mutants might be caused by their polar or negative charge characteristic for the alternative deprotonation reaction. Interestingly, the longer side chain substitution in either ERG7<sup>Y510E</sup> or ERG7<sup>Y510Q</sup> mutants further influenced the cyclization progression

and resulted in the generation of considerable quantities of premature (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien-3 $\beta$ -ol, which was previously observed from the *S. cerevisiae* ERG7<sup>Y510F</sup> or ERG7<sup>Y510H</sup> mutants by Mutsuda and his coworkers and also independently isolated from our previous *S. cerevisiae* ERG7<sup>H234X</sup> mutants.<sup>117,</sup>  
<sup>126</sup> The tricyclic byproduct provided direct mechanistic evidence for the formation of crucial chair-boat 6-6-5 tricyclic Markovnikov cationic intermediate in the ERG7-catalyzed oxidosqualene cyclization reaction. Interestingly, the production of (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien-3 $\beta$ -ol was observed from most of the *S. cerevisiae* ERG7<sup>Y510X</sup> mutants. The different product ratio compared with that of previously established ERG7<sup>Y510F</sup> or ERG7<sup>Y510H</sup> mutants, might be attributed to the different yeast analysis system used, or the physiological metabolism of lanosterol. The accurate quantitation of product ratio from *S. cerevisiae* ERG7<sup>Y510X</sup> mutations without any background interference should be carefully carried out by using the *in vitro* assay system or by using the over-expression system with the purified cyclase.<sup>105</sup>  
However, the result described herein could provide qualitative analysis for the functional role of Tyr-510 position in the *S. cerevisiae* ERG7. Moreover, the metabolized lanosterol in TKW14[pERG7<sup>Y510X</sup>] strain could facilitate the chemical characterization of the trace of novel triterpene alcohol that might be obscure due to the abundant lanosterol in the yeast host strain. Distinctively, the relatively dominant proportion of (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien-3 $\beta$ -ol in the product profiles of ERG7<sup>Y510C</sup>, might come from the powerful nucleophilic property of cysteine that disrupted the transient dipole interaction between carbocationic intermediate and the enzymatic active site, which retarded the tetracyclic ring formation and resulted in the early termination. Without other tricyclic alternatively deprotonated products might be due to steric position of the  $\beta$  facing C-14 methyl group that is close to H-9 $\beta$  proton and directly be abstracted from the closest catalytic base residue.

Additionally, the monocyclic achilleol A and camelliol C which is different in the alternative proton abstraction were also indisputably observed from most of the ERG7<sup>Y510X</sup> mutants. The higher accumulation of achilleol A over that of camelliol C or the production of achilleol A whenever camelliol C is produced might be due to the kinetically favored double-quaternary double bond deprotonation or the stereochemical control from the enzymatic active site. The bulky indole ring of tryptophan, the imidazole group of histidine, and the positive charged lysyl group of lysine might have dramatically influenced the progression of cyclization and resulted in the production of large amount of monocyclic products. The cyclization process might be retarded either from the enzymatic perturbation or the substrate misfolding. Interestingly, the ERG7<sup>Y510H</sup> mutation which produces abundant monocyclic products but also with lanosterol for supporting the yeast survivability might be caused by the relatively small electrostatic change in the active site cavity or the slighter influence on the cyclization process, compared to that of the lysine or the tryptophan substitutions.

In consideration of the generated product profiles for various amino acid substitutions of the Tyr-510 position, the functionally important role of this residue on the oxidosqualene cyclization/rearrangement cascade could be elucidated. Scheme 3.4 shows the proposed oxidosqualene cyclization/rearrangement cascade occurred in the *S. cerevisiae* ERG7<sup>Y510X</sup> site-saturated mutants. The epoxide ring was first opened via a proper catalytic acid activation and the first monocyclic C-10 cation was immediately generated. The progression of cyclization was partially blocked due to the steric influence from the substitution of an original tyrosine residue for the bulky group and resulted in the production of achilleol A and camelliol C.



Scheme 3.4 Proposed cyclization/rearrangement pathway of oxidosqualene in TKW14 expressing ERG7<sup>Y510X</sup> site-saturated mutants.

During the further B- and C-ring cyclization, oxidosqualene adopted a “chair-boat” conformation and a Markovnikov-favored 6-6-5 ring closure to produce a tricyclic C-14 cation. The C-26 methyl proton from this cationic intermediates was directly abstracted to yield (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien-3 $\beta$ -ol in most of *S. cerevisiae* ERG7<sup>Y510X</sup> site-saturated mutants. After successful C-ring expansion coupled with D-ring annulation, a tetracyclic protosterol C-20 cation with the natural C-20 *R*-form and C-17  $\beta$ -side chain configuration was generated. Interestingly, none of the ERG7<sup>Y510X</sup> site-saturated mutants was found to affect the following backbone rearrangement stage of two hydrogen groups and two additional methyl group shifts until the lanosteryl C-9 cation generation. The skeletal rearrangement byproducts have been isolated from other functionally important residues, but not in our ERG7<sup>Y510X</sup>

site-saturated mutations.<sup>117, 118</sup> Without generation of any 9 $\beta$ -lanosta-7,24-dien-3 $\beta$ -ol further demonstrated that the cationic rearrangement really arrives at the lanosteryl C-9 position.<sup>109, 119</sup> Finally, the deprotonation reactions alternatively occurred at either the C-8 or C-11 position to form the lanosterol or parkeol in the *S. cerevisiae* ERG7<sup>Y510X</sup> site-saturated mutants, respectively.

How the various single amino acid substitutions of ERG7<sup>Y510X</sup> generated diverse products with different ratio are complicated and poorly understood. According to the crystal structure of human OSC, Tyr-503 (the corresponding residue for yeast Tyr-510) is hydrogen bonded to His-232 which was considered as the only closest basic residue for the final deprotonation step.<sup>56, 105, 127</sup> In addition, our previous study on the site-saturated mutagenesis of His-234 (the corresponding residue for human His-232) showed that the substitutions of the His-234 position might influence the stabilization of the C-13 and C-20 positions as well as the C-14 Markovnikov tertiary cation during the rings formation.<sup>115, 117</sup> Different substitutions of ERG7<sup>H234X</sup> affected the steric and/or electrostatic interactions between the cationic intermediate and the side chain of active site residues, and resulted in the distinct products ratio. For example, the substitution of His-234 with small nonpolar hydrophobic residues facilitated the production of polycyclic products such as parkeol, protosta-12,24-dien-3 $\beta$ -ol, and protosta-20,24-dien-3 $\beta$ -ol, but interfered with the monocyclic achilleol A formation.<sup>115</sup> Moreover, substitution for the tyrosine or phenylalanine residue on His-234 induced the steric hindrance or electrostatic repulsion to the Tyr-510 residue, relocated the possible proton acceptor, and resulted in the production of achilleol A and other altered products. In contrast, lanosterol was the sole product when TKW14 expressed the ERG7<sup>H234Y/Y510A</sup> mutant, which might be due to the released electrostatic repulsion and the returned active site environment in this double substitution mutation.<sup>115</sup>



In parallel, the fascinating and compatible result was also observed in the products profiles of  $ERG7^{Y510X}$  described herein. Accordingly, the substitutions of larger bulky or basic amino acids on Tyr-510 tended to produce monocyclic products, whereas the small or acidic amino acids substitutions tended to produce polycyclic products. Thus, the impact on the hypothetical hydrogen-bonding basic dyad between His-234 and Tyr-510 could be imagined from the observation of individual site-directed mutation. For example, the substitution of Tyr-510 with small residues might slightly influence the position of His-234 and generate the alternative polycyclic deprotonation products. This transient disturbance increased when Tyr-510 was changed into the polar or acidic amino acids including Asp, Asn, Glu, and Gln. The electronic density or the polar group influenced the orientation of His-234 for stabilizing the C-14 cation and resulted in the isolation of tricyclic byproduct. Additionally, the substitution of Tyr-510 with basic group or the large amino acid such as His, Lys, Arg and Trp might cause the steric or electronic repulsion to His-234 residue. These mutants thus increased the production of monocyclic compounds. To our surprise, the  $ERG7^{Y510W}$  mutant produced almost monocyclic but no tricyclic or tetracyclic products. In contrast, the previous  $ERG7^{H234W}$  mutant generated almost one hundred percentage of parkeol without any other cyclization products.<sup>117</sup> These findings suggested that the spatial influence between these two important residues might disrupt the proper orientation of substrate or the coordinated interaction among other functional residues in the enzyme active site cavity. In order to investigate the coordinative action in these two hypothetical hydrogen-bonding dyads, the  $ERG7^{H234W/Y510V}$  and  $ERG7^{H234W/Y510W}$  double mutants were further created and analyzed (Table 3.3). The nearly opposite result of the product profiles ratio revealed that the side chain substitution of either Tyr-510 or His-234 position, especially in the huge and bulky tryptophan residue, would dramatically influence the cyclization

reaction. The abundant lanosterol was observed in the  $ERG7^{H234W/Y510V}$  mutant. The similar tetracyclic scaffold but with the alternatively deprotonated site, comparing with the previous parkeol production in the  $ERG7^{H234W}$  mutation, supported that the smaller valine group substitution might release the steric pressure and adjust the orientation of the substrate for the proper deprotonation reaction. On the other hand, the tryptophan residue introduced in the histidine position of  $ERG7^{H234W/Y510W}$  mutant produced almost one hundred percent of monocyclic product as  $ERG7^{Y510W}$  mutant. The functional importance of Tyr-510 for the stabilization of C-10 cation, His-234 for the stabilization of C-13 or C-20 cations, and their coordinative action for the deprotonation reaction could be elucidated carefully from the analysis of the product profiles described above. However, the altered hydrogen-binding networks, the newly generated catalytic base for the deprotonation reaction, even a new cyclization pathway, or novel enzymatic activity caused by mutagenesis cannot be excluded.

In conclusion, a site-saturated mutation on the Tyr-510 position as well as on the simultaneous His-234/Tyr-510 double positions provides critical evidence for the functional role of Tyr-510 involved in the complicated  $ERG7$  catalyzed oxidosqualene cyclization. The diverse products, including truncated monocyclic, tricyclic and the altered deprotonation products, suggested the catalytic role of this residue in affecting the cationic intermediate stabilization for the cyclization stage and for final deprotonation step. Without any truncated rearrangement products also supported that the function of Tyr-510 is not crucial for the hydride/methyl groups' migration. In addition, the isolation of (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien- 3 $\beta$ -ol implied the destabilization of C-14 cationic intermediate caused by the spatially influenced His-234 residue. The different combination of  $ERG7^{H234X/Y510X}$  double mutations further supported the coordinative action of this intrinsic His-234:Tyr-510 hydrogen-bonding network.

### 3.2.8 Homology modeling illustration of the functional critical residue

Due to the lack of high resolution crystal structure of wild-type *S. cerevisiae* ERG7 protein or the respective mutated cyclases in complex with those isolated products, it is very difficult to clarify the important role of these functional residues. The subtle change within the enzymatic active site could be imagined through the substitution of the critical functional residues. Moreover, the topography of other coordinated residues, the substrate folding manner, or the lifetime of different transition-state intermediates might further be influenced. In order to verify the importance of the critical amino acids involved in the oxidosqualene cyclization reaction as well as to understand the different mutational effects of specific active site residues substitution, the homology modeling structures of wild-type *S. cerevisiae* ERG7 and the respective mutated cyclase were created based on the X-ray structure of *A. acidocaldarius* SHC or human *Homo sapiens* OSC (Figure 3.5).<sup>105</sup> Good agreements in the distribution of secondary structure, stereochemical quality verification, and their 3D profile supported that *A. acidocaldarius* SHC or human *Homo sapiens* OSC are both appropriate templates for generating the homology models of oxidosqualene-cyclases. The further detailed examination of the homology modeling structures construction, the geometrical correctness verifications, and the structure-based illustration of other recently identified functionally important residues will be discussed in the next chapter. In this section, we only focus the discussion specifically on the Tyr-510, Trp-443, and Lys-448 residues. According to the homology modeling structure of *S. cerevisiae* ERG7, the hydroxyl group of Tyr-510 is located at a distance of about 5.4 Å to the C-10 cation of lanosterol and also near to the A/B ring fusion of lanosterol. Tyr-510 also shows a distance of about 4.6 Å to the C-6 cation of the monocyclic intermediate. This distance agreed with the dipoles for observed aromatic amino acid residues at the distance of 3.5 to 5.5 Å in the

lanosterol-bound OSC complex (Figure 3.6).<sup>115</sup> Thus, the mutagenic change on the position of Tyr-510 might directly affect the cyclization cascade to produce achilleol A or the deprotonation reaction to produce lanosterol and parkeol. These diverse products were possibly generated from the partial disruption of transient interactions between carbocationic intermediates and the hydroxyl group or the  $\pi$  electrons of the aromatic ring of Tyr-510. The enlarged active site cavity also allowed the perturbation of the flexible substrate. The premature cyclization for the monocyclic achilleol A or camelliol C isolation might be caused by the transient migration of the monocyclic C-6 carbocationic intermediate to the active site catalytic base and thus disrupt the B-ring formation (Figure 3.6). Moreover, in consideration of the hypothetical role of Tyr-510 in facilitating the formation of lanosterol despite the loss of the deprotonating hydroxyl group, the impaired hydrogen-bonding network might be observed among the different substitutions of the Tyr-510. Moreover, the isolation of tricyclic (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien-3 $\beta$ -ol also implied the affected His-234 residue in these *S. cerevisiae* ERG7<sup>Tyr510X</sup> mutations. Based on the homology modeling structure of the oxidosqualene cyclase, the residues of Tyr-510 occupies the appropriate position to stabilize C-6 and C-10 cation intermediates, whereas His-234 is positioned to stabilize the Markovikov tertiary cation at C-14. Accordingly, the  $\pi$ -electron density distributions, active site cavity size, as well as the orientation of His-234 were slightly altered due to the substitution of the Tyr-510 residues with other amino acids. This modified hydrogen-bonding network thus caused the steric influence to arrest the D-ring formation and also affected the stability of the C-14 cation intermediate. The spatially impaired His-234:Tyr-510 catalytic hydrogen-bonding network was further examined through the observation of the homology modeling structures in the various *S. cerevisiae* ERG7<sup>H234X/Y510X</sup> double mutations (Figure 3.7). Interestingly, this kind of hydrogen-bonding pair in which

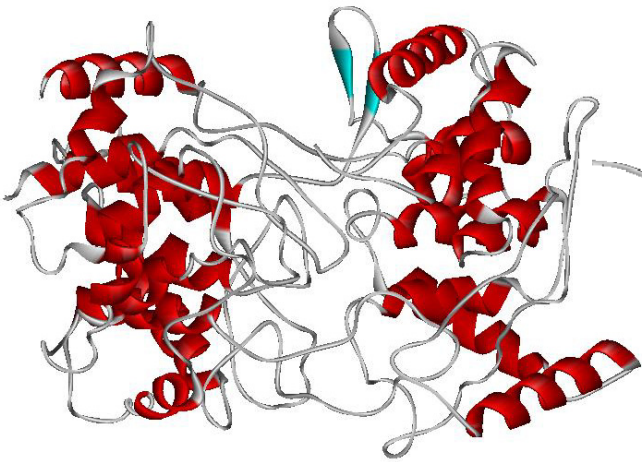
either residue acts as the general base and to influence the other one for the deprotonation reaction has been observed in the *A. thaliana* CAS1 enzyme. The Tyr-410 and His-257 residues were examined as the coordination partners for the C-19 angular methyl proton abstraction.<sup>111</sup> The further illustration of the coordinative interaction between the His-234:Tyr-510 dyad from the homology modeling study will be examined in the next chapter.

In addition, the Trp443Ala and Lys448Ala mutations which are located at the upstream of the putative active site also produced monocyclic achilleol A and camelliol C. The multiple sequence alignment analysis showed that the Trp-443 are highly conserved in most of cyclases, whereas Lys-448 is changed to acidic aspartic acid residue in SHC or changed to alanine residue in CAS enzyme. In the homology modeling analysis, the Trp-443 is positioned spatially opposite to the Asp-456, above the molecular plane of lanosterol and shows the far distance to the substrate (10.3 Å between oxygen of tryptophan and C-10 of lanosterol; 10.1 Å between oxygen of tryptophan and C-2 of lanosterol) (Figure 3.8). However, the Trp-443 is located in a flexible region and might act as the neighbor to the putative active site residues. Trp-443 might stabilize the highly energetic cationic intermediate through the indirect steric influence. Substitution of tryptophan with alanine disrupted the steric or cation- $\pi$  electrostatic interaction between the substrate and enzyme. Thus, the elongation of the cyclization cascade was inhibited and halted at monocyclic stage. On the other hand, Lys-448 is located at a flexible loop region between two helices motifs (Figure 3.8). Lys-448 might also exhibit the interaction with two amino acids, Phe-426 and Asn-332 and thus hold the correct conformation in the helices association. Therefore, replacing of Lys-448 with alanine might disrupt the electrostatic interaction among the structure-based functional residues, further influence the enzymatic active site environment, and result in the formation of the

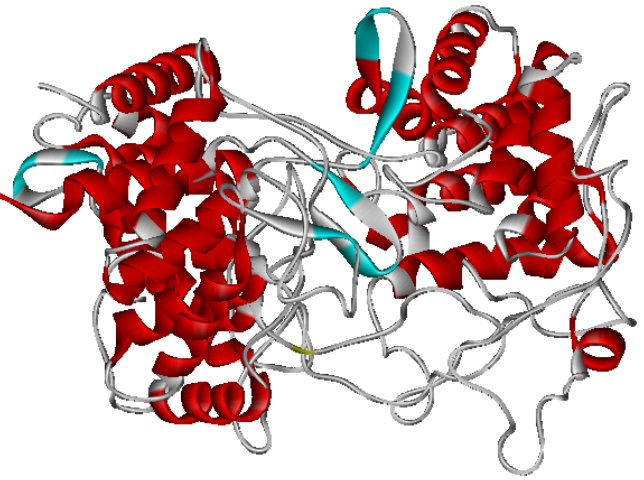


monocyclic products.

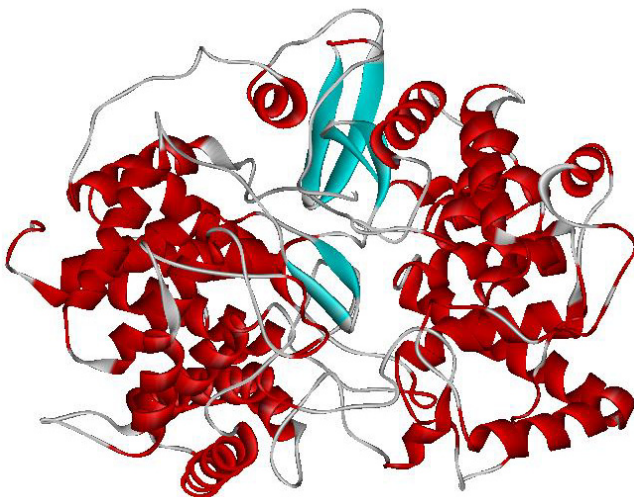
(a)



(b)



(c)



(d)

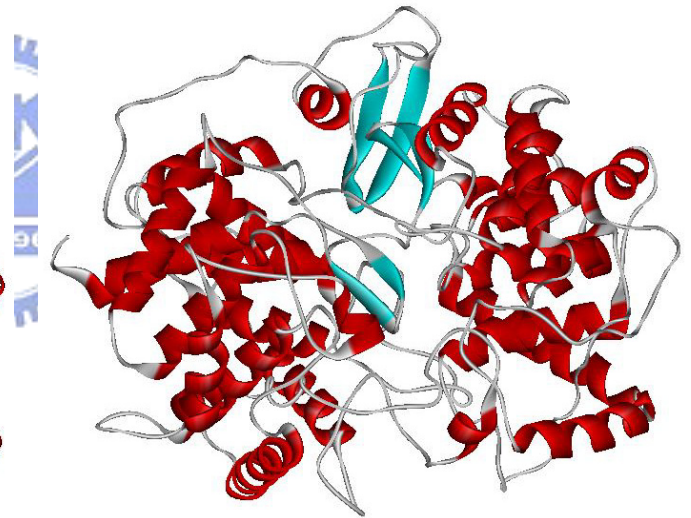


Figure 3.5 The X-ray crystal structures of (a) *A. acidocaldarius* SHC or (c) human *Homo sapiens* OSC, and the homology modeling structures of wild-type *S. cerevisiae* ERG7 based on the structures of (b) SHC or (d) human OSC, respectively.

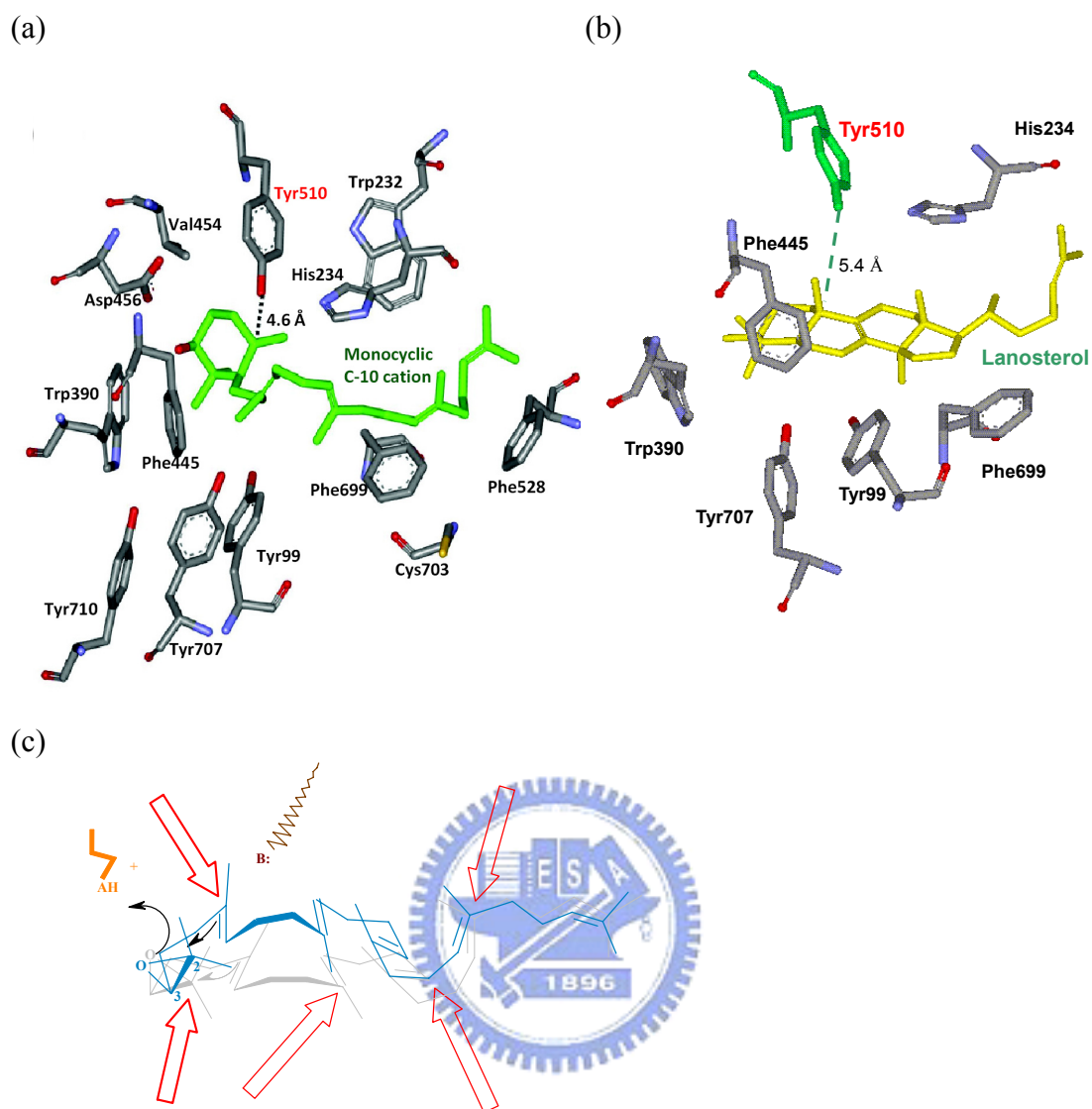
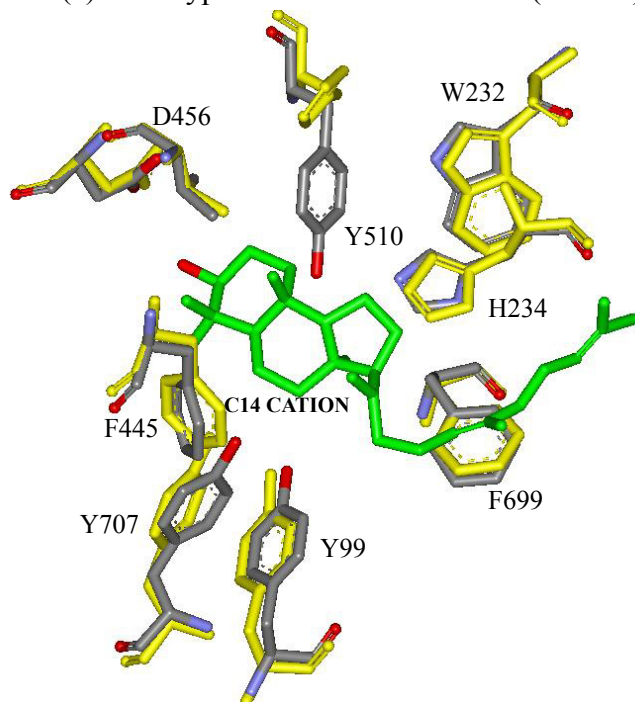


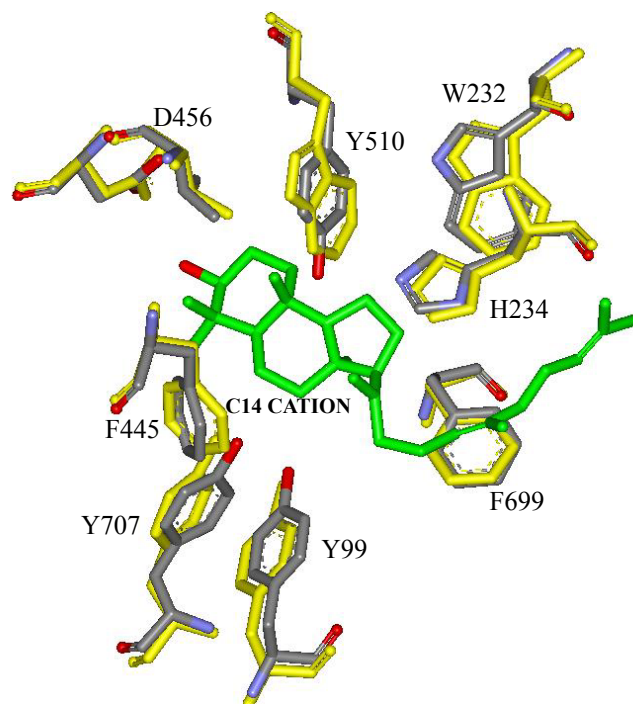
Figure 3.6 Local view of the homology modeled *S. cerevisiae* ERG7 structure. Putative active site residues (stick representation) are included. (a) The substrate analog, 2-azasqualene, is shown in orange. The distance between C-6 of 2-azasqualene and phenolic oxygen of Tyr-510 residue is exhibited with a dotted green line. (b) Lanosterol is included in the homology model, the distance between C-10 of lanosterol and phenolic oxygen of Tyr-510 residue is also shown. (c) The possibly specific function of the Tyr-510 for the stabilization of the monocyclic cationic intermediate.



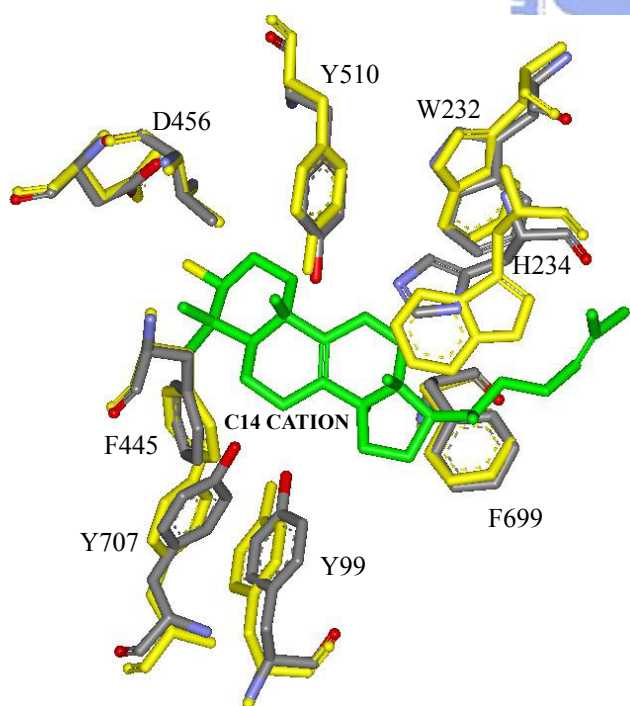
(a) Wild-type ERG7 and ERG7<sup>Y510H</sup> (Yellow)



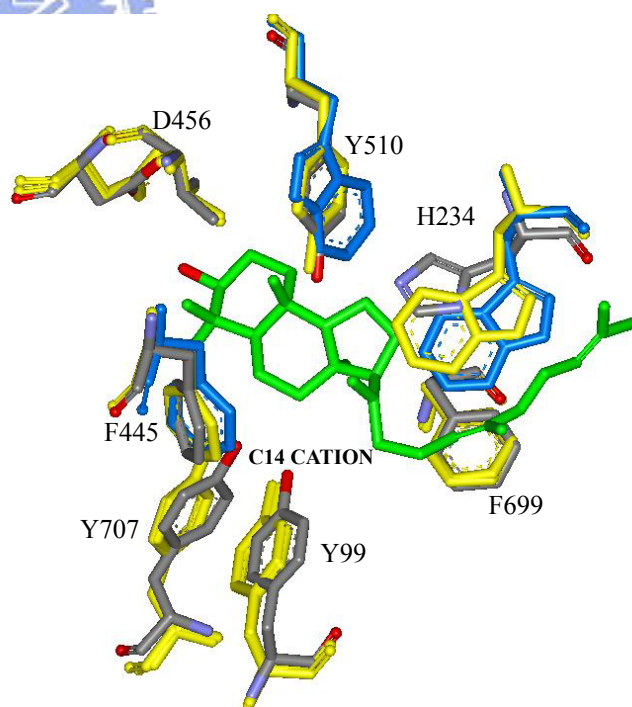
(b) Wild-type ERG7 and ERG7<sup>Y510W</sup> (Yellow)



(c) Wild-type ERG7 and ERG7<sup>H234W</sup> (Yellow)



(d) Wild-type ERG7, ERG7<sup>H234W</sup> (Yellow), and ERG7<sup>H234W Y510W</sup> (Blue)



(e) Wild-type ERG7, ERG7<sup>H234WY510V</sup> (Yellow)  
and ERG7<sup>H234W Y510A</sup> (Blue)

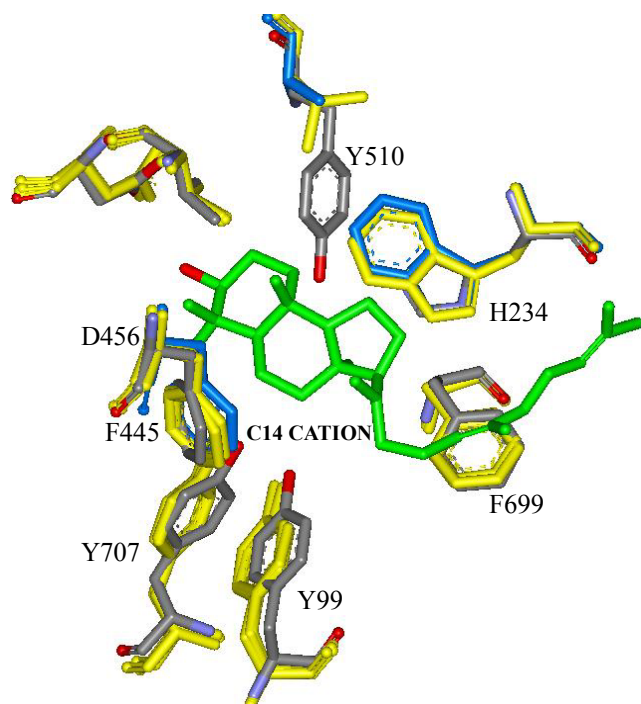


Figure 3.7 Local view of superimposition of homology modeled *S. cerevisiae* ERG7 structure and different single-point mutation's structure or double-positions mutation's structure. The intermediate C-14 cation is also included. Putative active site residues in the wild-type ERG7 are shown in stick representation, while different point mutations are displayed in yellow or blue color. In order to simply the figure, the appellations of amino acids residues are only listed with those in the wild-type ERG7. Accordingly, the orientation of His-234 and the modified hydrogen-bonding network were slightly altered due to the substitution of the Tyr-510 residues and/or His-234 residues with other amino acids.

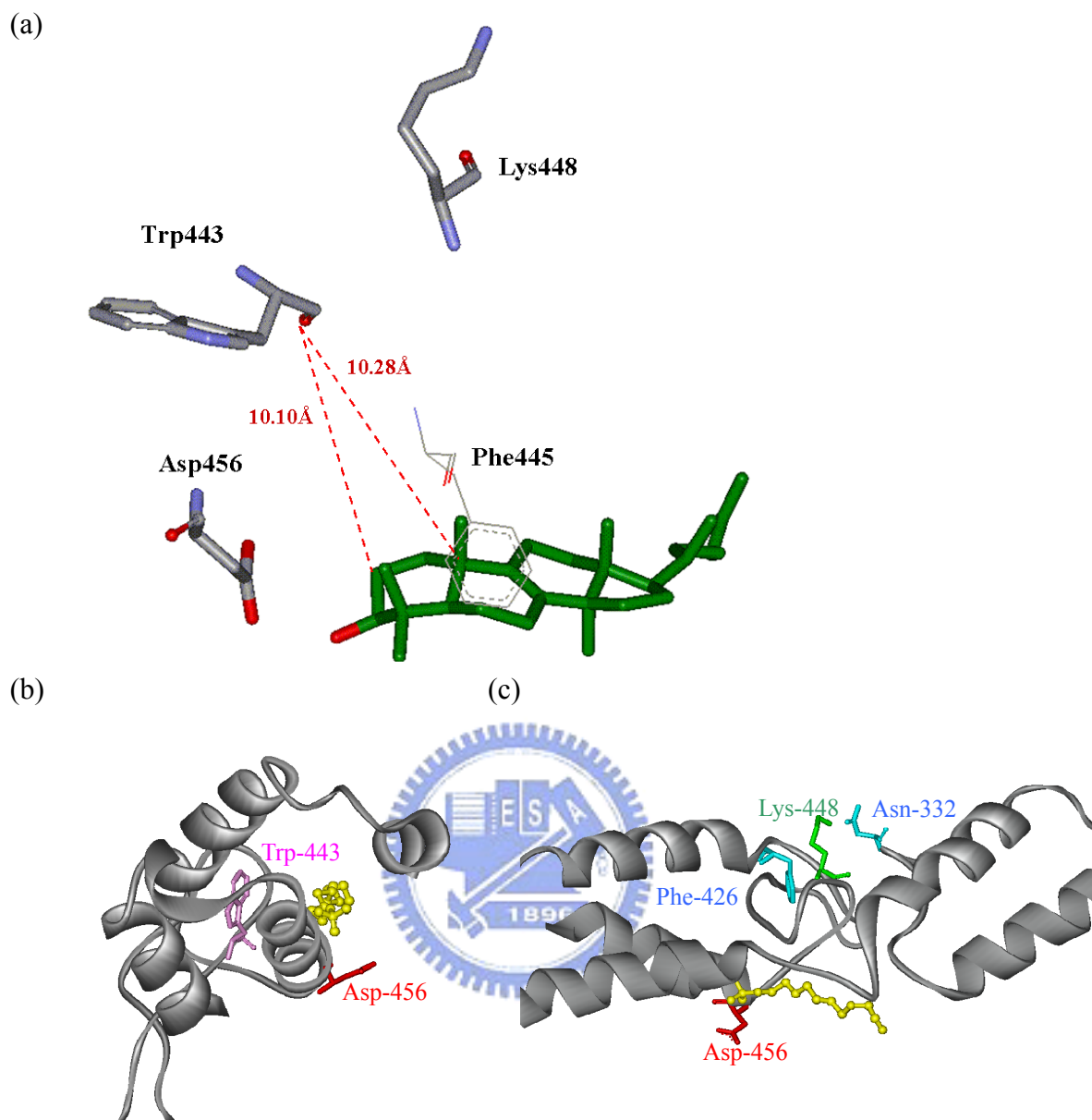


Figure 3.8 Local views of Asp-456, Trp-443, Lys-448, and Phe-445 positions in the *S. cerevisiae* ERG7 structure.

## Chapter 4

# ***Construction of Chimeric Oxidosqualene Cyclase Gene to Study the Product Specificity***

### 4.1 Introduction

To understand how each triterpene synthase controls the complex cyclization reaction, and leads to the defined polycyclic product is a major issue for illustrating the diversity of the natural triterpenes and for future protein engineering. The highly parallel mechanism among the reactions of oxidosqualene cyclases suggested that the skeletal diversity of sterols and triterpenes could be attained by only small modifications on the structure of the enzymes. Moreover, the results of molecular site-directed/saturated mutagenesis studies in previous chapters indicated that cyclase enzymes may maintain the general protein structure, but develop diverse product specificities via subtle change of the shape of the active site and/or the position of the crucial functional amino acids.

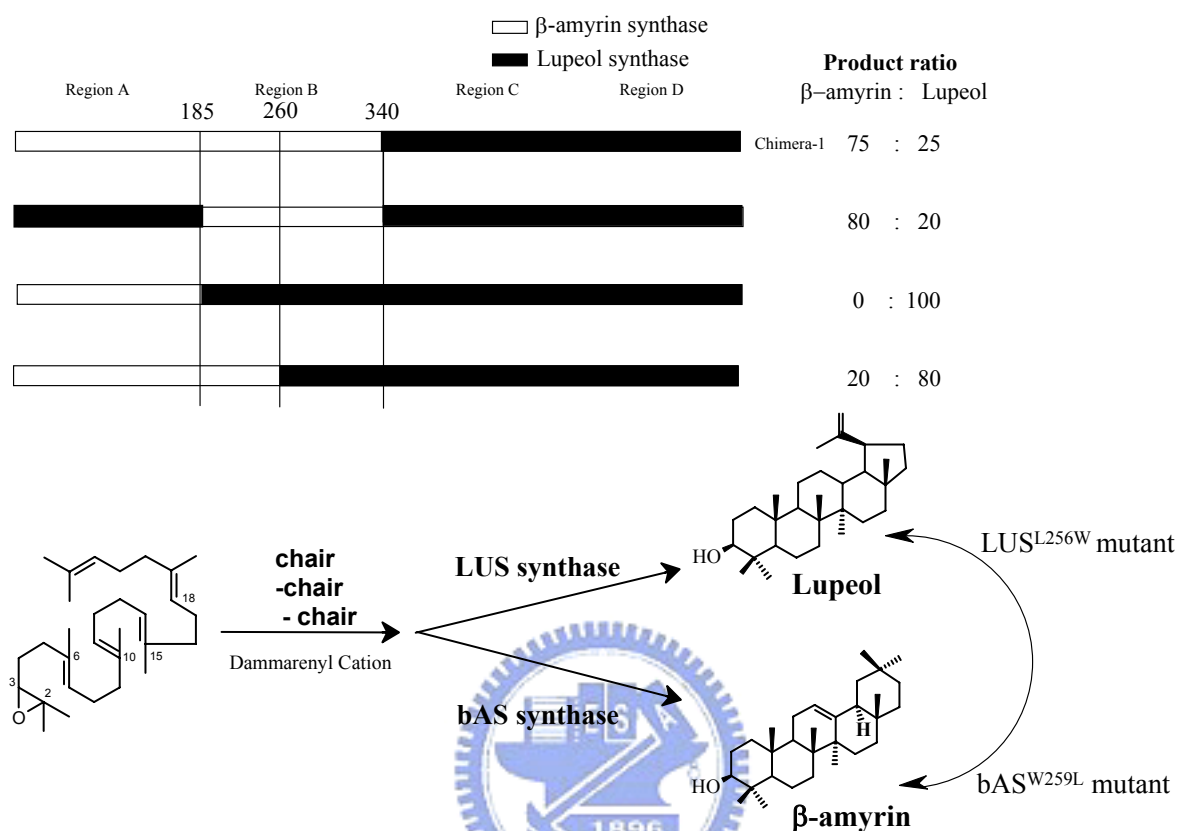
In order to investigate in depth the enzymatic control on the product specificity, we set up a domain swapping study between oxidosqualene-lanosterol cyclase (ERG7) and oxidosqualene-cycloartenol synthase (CAS1) to examine which functional region is responsible for the product specificity. The feasibility of this approach has been demonstrated by Chappell *et al.* for two sesquiterpene cyclases and Kushiro *et al.* for two triterpene-producing cyclases.<sup>128, 129</sup>

From the results of a domain swapping experiment published in 1999, a series of chimeric enzymes were created between  $\beta$ -amyrin synthase from *Panax ginseng* and lupeol synthase from *Arabidopsis thaliana*. Interestingly, the product ratio derived

from these newly constructed chimeric enzymes narrowed down the responsible region for the products' specificity. For example, the Chimera-1, from which the *N*-terminal half is  $\beta$ -amyrin synthase and the *C*-terminal half is lupeol synthase, produced  $\beta$ -amyrin and lupeol in a 3:1 ratio. In the following constructs, they established several possible combinations of these two cyclases. Among these chimeric enzymes, one chimera whose region-B (the second quarter from the *N*-terminus) is  $\beta$ -amyrin synthase whereas the other three quarters (the *N*-terminal quarter and half of *C*-terminus) is lupeol synthase produced  $\beta$ -amyrin and lupeol in the ratio of 4:1, indicating the importance of region-B in the  $\beta$ -amyrin production. Moreover, the critical role of this region-B in product specificity was further proven in another chimera which produced one hundred percent of lupeol when its region-B was exchanged into the lupeol synthase gene from original Chimera-1 construct. (*i.e.* only the *N*-terminal quarter is  $\beta$ -amyrin synthase while other parts is lupeol synthase) This product determining region was finally reduced into eighty amino acids region within the region-B. The construction of the chimeric enzyme library between  $\beta$ -amyrin synthase and lupeol synthase, as well as the resulting products ratio is shown in the Scheme 4.1.

After the successful domain swapping experiment between two triterpene synthases, Kushiro consequently found a functionally important residue within the region-B in controlling the products specificity.<sup>130</sup> Site-directed mutagenesis was carried out on the Trp-259 position of  $\beta$ -amyrin synthase in order to exchange it for the corresponding leucine residue of lupeol synthase. This single-point mutation dramatically altered the enzymatic specificity and produced lupeol as its major product. In parallel, the Leu-256 of lupeol synthase was exchanged for the tryptophan residue and thus exclusively engineered lupeol synthase into the functional  $\beta$ -amyrin synthase. Therefore, the Trp-259 of  $\beta$ -amyrin synthase or Leu-256 of lupeol synthase

was identified to be the product-specificity-determining residue in controlling  $\beta$ -amyrin or lupeol formation, respectively (Scheme 4.1).



Scheme 4.1 (Above) The schematic representation of partial chimeric library between  $\beta$ -amyrin synthase and lupeol synthase. The products ratio generated from individual chimeras are included. (Bottom) Single amino acid substitution found in the region-B exclusively changes the product specificity from  $\beta$ -amyrin into lupeol or from lupeol into  $\beta$ -amyrin.

According to the results of domain swapping chimeric enzymes construction and the site-directed mutagenesis studies, the functional role of important regions or amino acid residues in regulating product specificity among these triterpene-producing cyclases has been observed. In order to study the relationship between structural modifications and product alternations among the sterol-producing cyclases, we also built a series of domain swapping chimera between oxidosqualene-lanosterol cyclase (ERG7) and oxidosqualene-cycloartenol synthase



(CAS1). We used the restriction enzyme digestion to generate four different clones (Chimera-1~4). Chimera-1 was constructed to include two putative DCTAE active site motifs, where ERG7's DCTAE region is in the *N*-terminal region, whereas CAS1's DCTAE region is in the *C*-terminal region. The construction of Chimera-3 is similar to Chimera-1 except the eliminated DCTAE motifs and the inverted connecting sequence; *i.e.* the *N*-terminal region is CAS1 and the *C*-terminal region is ERG7. Moreover, Chimera-2 and Chimera-4 are the truncated ERG7 and CAS1 protein, respectively. Subsequently, we divided the entire sequence into four regions of approximately equal length. Region A contains two well-conserved QW motifs, while region B contains a functionally important region WWVHTR (231-236) of *Saccharomyces cerevisiae* oxidosqualene-lanosterol cyclase. Region C contains the DCTAE motif, while region D contains several conserved QW motifs. By using a PCR based strategy, another six chimeric enzymes between oxidosqualene-lanosterol cyclase (ERG7) and oxidosqualene-cycloartenol synthase (CAS1) were obtained (Chimera-5~ Chimera-10). The complementation effect of these chimeric enzymes in the yeast *erg7* deficiency strain were examined, and their non-saponifiable lipid were further extracted and analyzed by using TLC plate analysis, GC-MS, and NMR spectra as previously described in the Chapter 3. The results of domain swapping experiments should provide some clues for better understanding of the complex oxidosqualene cyclization reaction. The detailed construction of entire chimerical enzymes between oxidosqualene-lanosterol cyclase (ERG7) and oxidosqualene-cycloartenol synthase (CAS1) and the resulting data will be discussed in the following section.

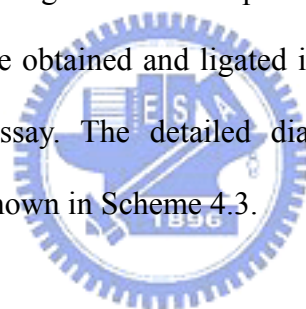


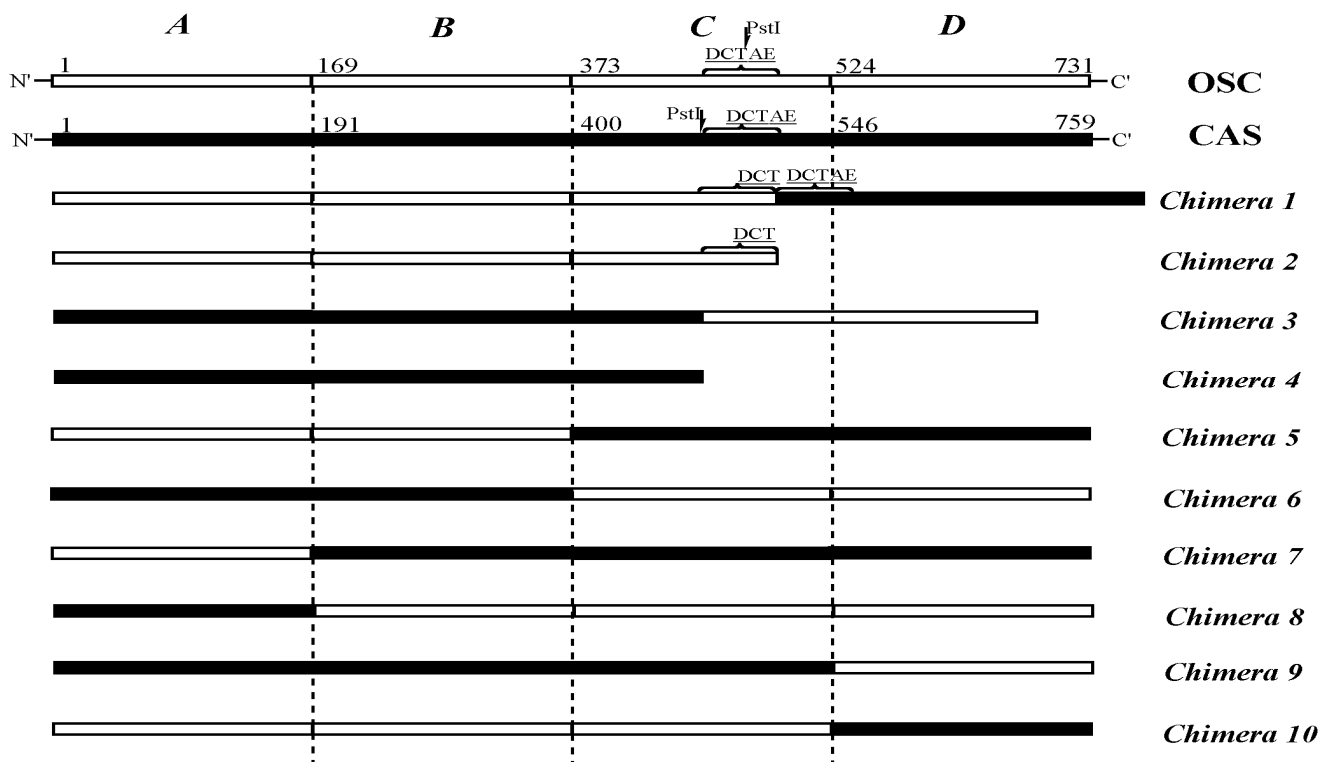
## 4.2. Results and Discussion

### 4.2.1 Construction of a chimeric library between oxidosqualene-lanosterol cyclase and oxidosqualene-cycloartenol synthase

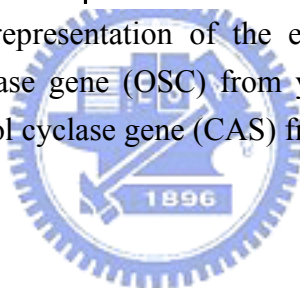
The oxidosqualene-lanosterol cyclase gene (*ERG7*) from yeast *Saccharomyces cerevisiae* and oxidosqualene-cycloartenol cyclase gene (*CAS1*) from *Arabidopsis thaliana* have been successfully cloned into the pRS314 vector to create the pTKERG7RS314WT and pTKP-5 plasmids, respectively. This yeast-*E.coli* shuttle vector, pRS314, is a *TRP1* centromeric plasmid which has been inserted with the promoter region of the *ERG7* gene for facilitating genetic selection, as well as the heterogeneous protein expression. The construction of the pTKP-5 plasmid and functional expression of the active *CAS1* protein in yeast has been described previously.<sup>106</sup> In order to create the diverse hybrid *ERG7/CAS1* chimeric library, the full length of *ERG7* gene and *CAS1* gene were divided into several genetic blocks. Scheme 4.2 illustrates the construction map of the chimeric library. The *Pst*I cutting site is located within the region of the putative DCTAE active site of oxidosqualene-lanosterol cyclase *ERG7* gene, whereas *Pst*I cuts the oxidosqualene-cycloartenol cyclase *CAS1* gene just in front of the DCTAE region. Thus, the construction of Chimera-1~ -4 could be carried out via a coupled enzymatic digestion and religation process, respectively (Scheme 4.3). Consequently, Chimera-1 has two putative active site motifs from which the *N*-terminal half is *ERG7* gene and the *C*-terminal half is *CAS1* gene, whereas Chimera-3 lose the DCTAE motif and also encodes the *vice versa* hybrid *CAS1/ERG7* sequence. Moreover, Chimera-2 or Chimera-4 exhibits the truncated *ERG7* cyclase protein or the *CAS1* cyclase protein, respectively.

In addition, the Chimera-5 ~ -10 was constructed by using a two-stage PCR strategy (Scheme 4.3). First, two set of primer pairs, two terminal primers coupled with another two hybrid primers, were respectively used to amplify two truncated PCR fragments from either the *ERG7* gene or the *CAS1* gene. Notably, the sequence of these two designed hybrid primers are based on the switch point and the corresponding sequence of the opposite cyclase gene, *i.e.* the 5'-half sequence of the hybrid primer are based on the corresponding sequence of the opposite cyclase gene, whereas the 3'-half sequence of the primer are designed according to the downstream gene from the switching point. Thus, these two truncated PCR fragments could be assembled in a complementary fashion due to their fused opposite gene. After the nested PCR reaction by the original terminal primer pairs, the full length of the chimeric hybrid gene could be obtained and ligated into the pRS314 plasmid for the following cyclase activity assay. The detailed diagrammatic explanation of the construction strategy is also shown in Scheme 4.3.

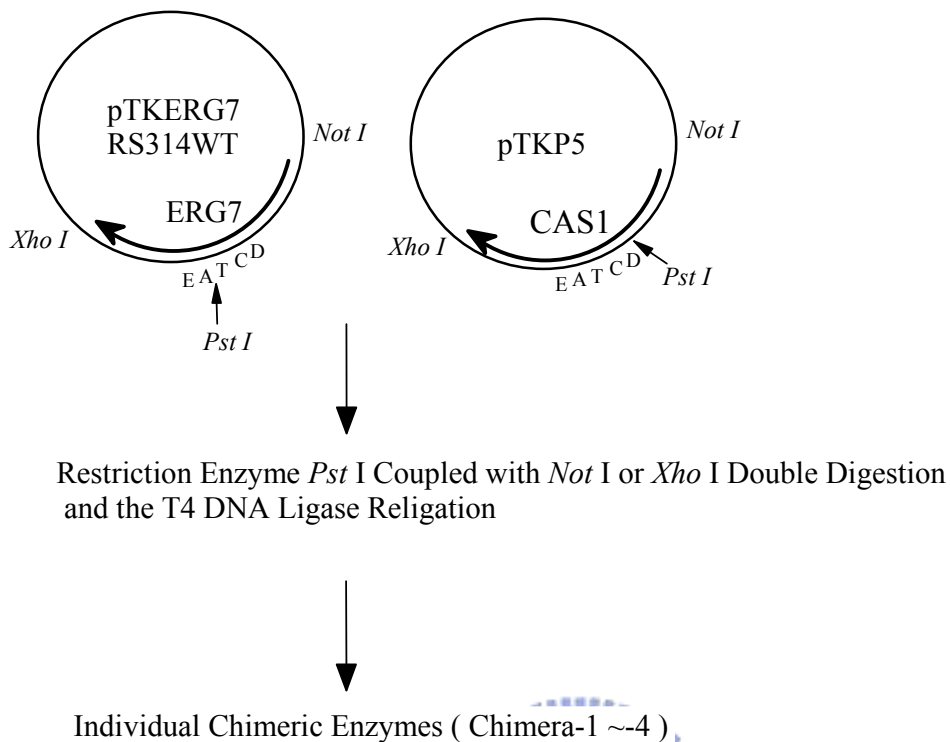




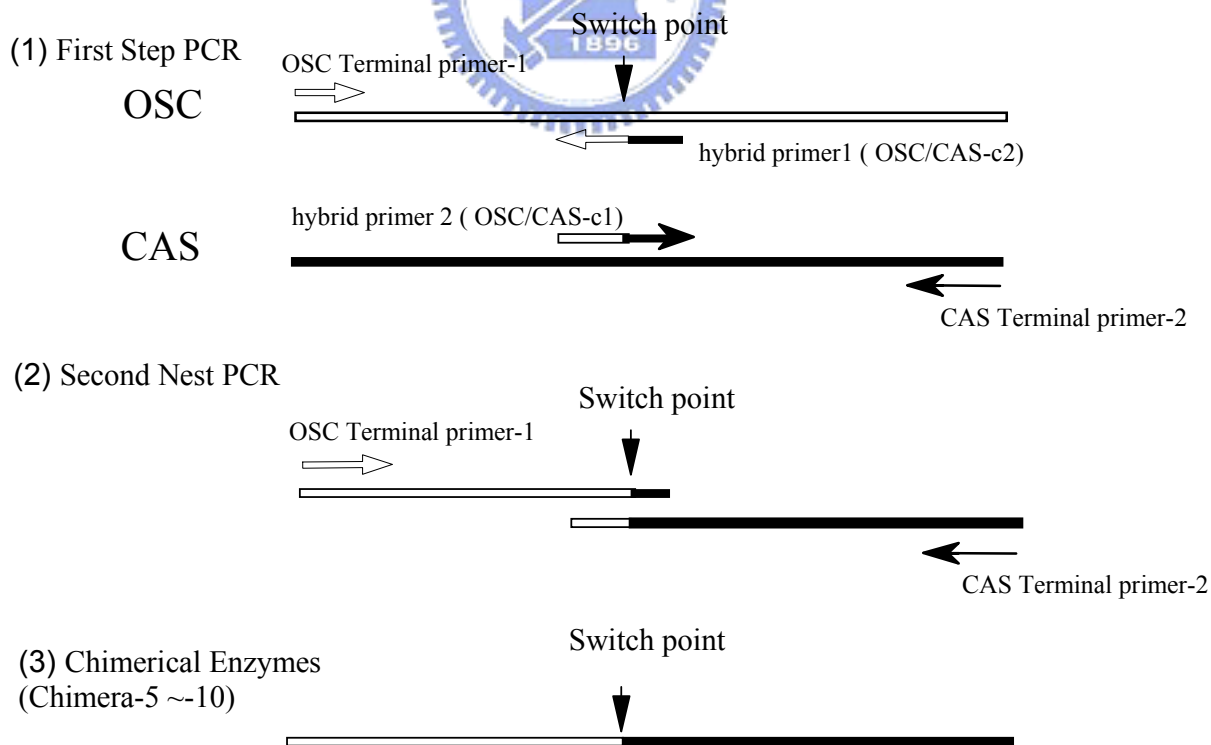
Scheme 4.2 The schematic representation of the entire chimeric library between oxidosqualene-lanosterol cyclase gene (OSC) from yeast *Saccharomyces cerevisiae* and oxidosqualene-cycloartenol cyclase gene (CAS) from *Arabidopsis thaliana*



### Chimera-1 ~ Chimera-4 construction



### Chimera-5 ~ Chimera-10 construction



Scheme 4.3 The diagrammatic explanation for the PCR-based construction strategy

#### 4.2.2 Functional analysis of chimeric enzymes library via plasmid shuffle and the product characterization

The previously mentioned plasmid shuffle methodology was used to analyze the complementation effect of the chimeric enzyme library. Surprisingly, the functional analysis assay revealed that all of the chimeric enzymes failed to complement the *erg7* disruption in the CBY57[pZS11] strain. This finding suggested that these swapped domains might be involved in an activity-related role. The suppression of ERG7 activity or altering product specificity thus could be imagined in the yeast transformants with the chimeric enzyme library.

Therefore, the resulting nonsaponifiable lipids (NSL) of individual chimera were analyzed. As shown in Figure 4.1, the nondistinctive product pattern was observed among these chimerical strains. In order to examine the resulting products in more detail, the abundant ergosterol was further eliminated from the preliminary NSL extract. Because ergosterol shares similar retention time with that of monocyclic triterpene on the GC-FID spectrum and thus makes the trouble to analyze the products. The purification of preliminary NSL was carried out on the silica gel chromatography to collect the respective NSL products whose R.f. value is between oxidosqualene and ergosterol. The collected lipids were then dissolved in CH<sub>2</sub>Cl<sub>2</sub> and subjected to the GC-FID analysis (Figure 4.2).

The detailed comparison of products pattern of individual chimerical enzymes from the GC-FID spectrums further indicated that these chimeras presented the identical products. The dramatic exchange of functional domains might disrupt the overall three-dimensional structure. Moreover, the rough partition of these two cyclase genes might directly influence the active site structure or indirectly disrupt the structurally relative helical domains. Hence, these chimeric enzymes might lose the catalytic ability for the oxidosqualene cyclization.

Although, the actual factors that caused the nondistinctive product pattern remain unknown, the comparison of the respective enzymatic characteristics with previous experimental results might provide some clues. The domains swapping approach has been successfully applied to understand the responsible region for the product specificity between  $\beta$ -amyrin synthase and lupeol synthase.<sup>128, 129</sup> That might suggest that these two groups of enzymes (oxidosqualene-lanosterol cyclase/cycloartenol synthase vs. lupeol synthase/ $\beta$ -amyrin synthase) should be studied from different mutagenesis strategies. Obviously, the sterol precursor either lanosterol or cycloartenol arose early during the eukaryotic evolution. Thus, these sterol-producing cyclases are distributed broadly among all kinds of eukaryotes. Interestingly, several cyclase enzymes responsible for the same sterol precursor production possess the high sequence diversity, indicating that the product specificity might be regulated by few conserved motifs. Thus, the large-scale domains swapping not only change the product specificity-determining motifs but also disrupt the structure-dependent domains. In contrast, cyclases that mediate the triterpene alcohol production might evolve more recently. These triterpene alcohol-producing cyclases are much related, usually with >80% identity, even for generation of different biosynthetic products. Hence, the large-scale domains exchanging among these enzymes might only produce minor change in the primary amino acid sequence. In addition to the sequence similarity, triterpene synthases with their weaker strength in product-specificity control than that of the sterol-producing cyclases let these triterpene synthases be more flexible and suitable for domains swapping studies. Many triterpene synthases have multiple functions for synthesis of different triterpene products in the sole enzymatic cyclization. The detailed multiple sequence alignment and the predicted secondary structural elements among these cyclases are shown in Figure 4.3.

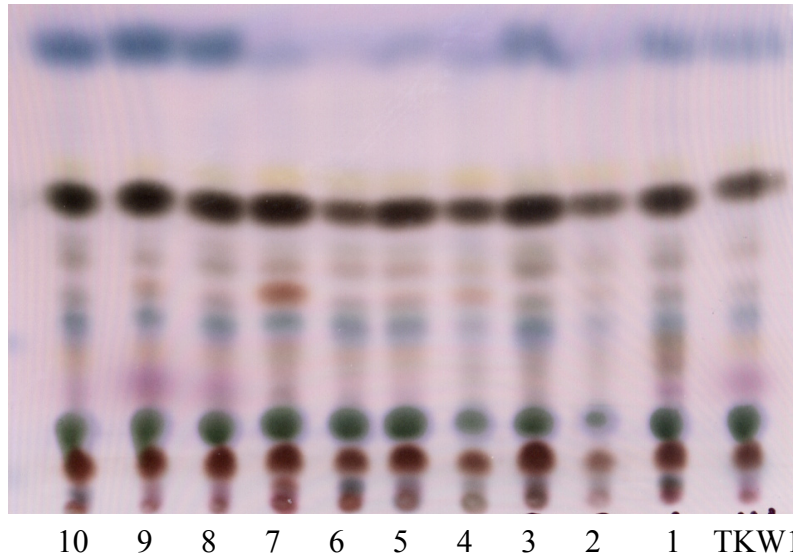


Figure 4.1 TLC analysis of preliminary NSL extract from various chimeras. Lane-1~10 indicated the NSL extract from individual chimera mutants; Lane-TKW14 indicated the negative control.

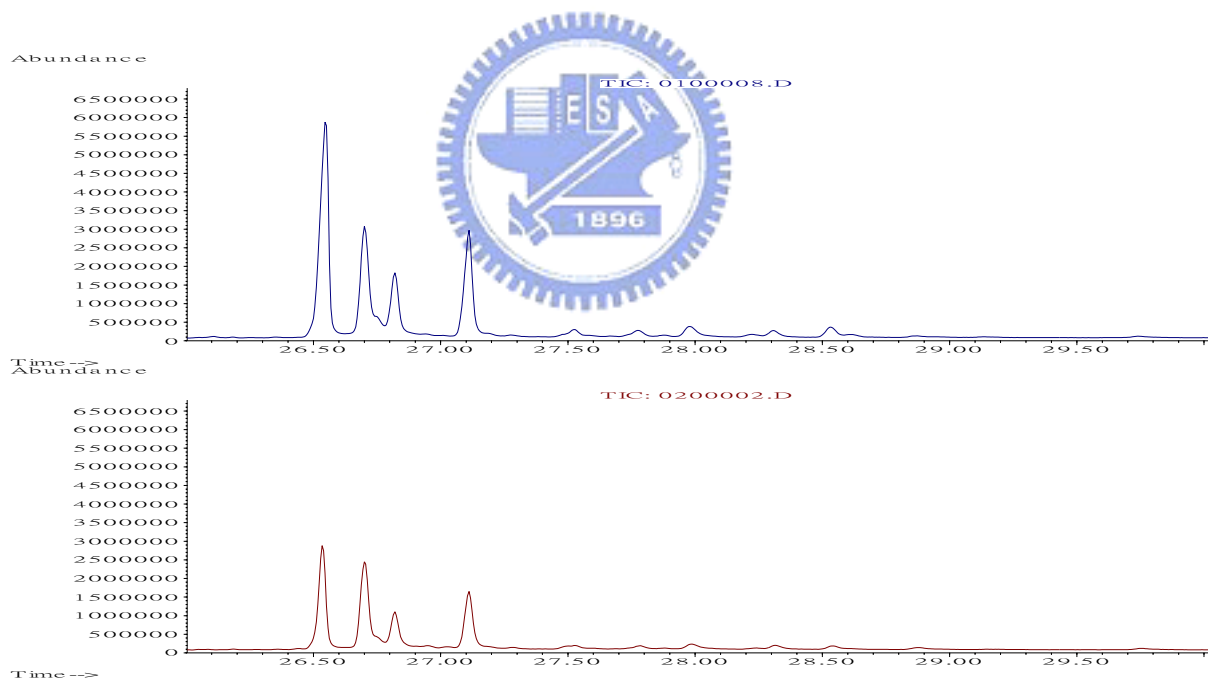


Figure 4.2 GC-FID spectra of the partially purified NSL from the TKW14 strain transformed with the representative chimera enzymes. The blue line shows the product pattern from sole host TKW14 strain, and the red line indicates the products from one of chimeras. All of the chimeras share the identical product pattern.



[A]

A.t.LUS MWKLIKIGKNGEDPHLFSNNFVGRQTKFD-HKAGSPEERAAVEEARRGFLDNRFVVKG 59  
.....  
P.g.bAS MWKLIKIAEGNKNDYLYSTNNFVGRQTFDPDYVASPGELEEVEQVRRQFWDNRYQVKP 60  
.....

A.t.LUS CSDLLWRMQFLREKKFEQGIPOLKATNIEEITTYETTINALRRGVRYFTALQASDGHWPGE 119  
.....  
P.g.bAS SGDLLWRMQFLREKNFROTIPQVKGVDDEAVTYEAATTTLRRAVHFFSALQASDGHWPGE 120  
.....

A.t.LUS ITGPLFFLPPIIFCLYITGHLEEVFDAEHRKEMLRHIYCHQNEGGWGLHIESKSVMFCT 179  
.....  
P.g.bAS NSGPLFFLPPIVMCVYITGHLDVFPAEHRKEILRYIYCHQNEGGWGLHIEGHSVMFCT 180  
.....

▼L185B186  
A.t.LUS VINYICLRMLGENPE---QDACKRARQWILDRCGVIFIPSWGKEWLSILGVYDWSGTNPT 236  
.....  
P.g.bAS TISYICMRILGEGPDGGVNNACARGRKWILDHCSVTAIPSWGKTWLSILGVYEWTCGNPM 240  
.....

▼L257B260  
A.t.LUS PPELLMLPSFLEIHPGKILCYSRMVSIPMSYLYKRFVGPITPLILLREELYLEPYEEI 296  
.....  
P.g.bAS PPEFWILPSFLPMHPAKMWCYCRMVYMPMSYLYKRFVGPITPLILQLREELYGOPYNEI 300  
.....

▼L337B340  
A.t.LUS NWKKSRRLYAKEDMYAHPLVQDLLSDTLQNFVEPLLTRWPLNKLVREKALQLTMKHIHY 356  
.....  
P.g.bAS NWRKTRRVCAKEDIYYPHPLIQDLLWDSLYVLTPEPLLTRWPFNKLR-EKALQTMKHIHY 359  
.....

A.t.LUS EDENSHYITIGCVEKVLCLACWVENPNGDYFKKHLARIPDYMWVAEDGMMQSFQCQLW 416  
.....  
P.g.bAS EDENSRYYITIGCVEKVLCLVCWVEDPNGDYFRKHLARIPDYIWWVAEDGMMQSFQSQEW 419  
.....

A.t.LUS DTGFAIQALLASNLDPDETDDALKRGHNYIKASOVRENPSGDFRSMYRHISKGAWTFSDRD 476  
.....  
P.g.bAS DTGFSIQALLDSDLTHEIGPTLMKGHDFIKKSOVKDNPSGDFRSMYRHISKGSWTFSDQD 479  
.....

A.t.LUS HGQVSDCTABALKCCLLLSMMSADIGGQKIDDEQLYDSVNL LLSLQSGNGGVNAWEPSR 536  
.....  
P.g.bAS HGQVSDCTABGLKCCLIFSTMPEEIVGKKIKPERLYDSVNV LLSLQQRKNGGLSAWEPAG 539  
.....

A.t.LUS AYKWLLELLNPTEEFMANTMVEREVECTSSVIOALDLFRKLYPDHRKKEINRSLEKAVQFI 596  
.....  
P.g.bAS AQEWLLELLNPTEEFADIVIEHEHYVECTSSAIOALVLFKLYPCHRKKEIDNFTTNAVRYL 599  
.....

A.t.LUS QDNQTPDGSWYGNWGVCFIYATWFALGGLAAAGETYNDCLAMRNGVHFLLTQRDDGGWG 656  
.....  
P.g.bAS EDIQMPDGSWYGNWGVCFIYGSWFALGGLAAACKTYYNCAAVRKAVEFLLKSNDDGGWG 659  
.....

A.t.LUS ESYLSCSEQRYIPSEGERSNLVQTSWAMMALIHGQAERDLTPLHRAAKLIINSQLENGD 716  
.....  
P.g.bAS ESYLSCPCKVYVLEGNRSNLVHTGWALMGLIHSEQAERDPTPLHRAAKLIINSQMEDGD 719  
.....

A.t.LUS FPQOEIVGAFMNTCMLHYATYRNTFPLWALAEYRKVV FIVN--- 757  
.....  
P.g.bAS FPQOEISGVFMKN CMLHYAAYRNIIYPLWALAEYRRVPLPSLGT 763  
.....

[B]

S.c.ERG7 -----MTEFYSDTIG--LPKTDPRWR---LRTDELGRESW---EYLTPQQA 40  
A.t.CAS1 MWKLIKIAEGGSPWLRRTTNNHVGROQWFEFDPNLGTPEDLAAVEEARKSFSDNRFVQKHSAD 60  
.....  
S.c.ERG7 NDPPSTFTQWLLQDPKFPQPHPERNKHSPDFSAFDACHNGASEFKLLQEPDSGIFECQYK 100  
A.t.CAS1 LLMRLQFSRENILISPVLPQVKIEDTDDVTEEMVETTLLKRGLEDYSTIQAHG-GHWPGDYG 119  
.....  
S.c.ERG7 GPMFMTIGYVAVNYIAG---IEIPEHERIELIRYIVNHTAHPVDGGWGLHSVDKSTVFGTV 157  
A.t.CAS1 GPMFLLPG IITLSITGALNTVLSHQHKQEMRRYLYNHQN-EDGGWGLHIEGPTSMFGSV 178  
.....  
S.c.ERG7 **▼E169C191**  
LNYVILRLLG-LPKDHP-VCAKARSTLLRLLGGAIGSPHWGKIWLSALNLYKWEQVNPAPP 215  
A.t.CAS1 LNYVTLRLLGEGPNDGDGDMKGRDWILNHGGA TNITSWGKMWLSVLGAFEWSCNNPLPP 238  
.....  
S.c.ERG7 ETWLLPYSLEMPHGRWVWVHTRGVYIPVSYLSLVKESCPMTPLLEELRNEIYTRFDKINF 275  
A.t.CAS1 ETWLLPYFLPIHPGRMWCHCRMVYLPMSYLYGKRFVGPITSTVLSLRKELFTVYHEVNW 298  
.....  
S.c.ERG7 SKNRNTVCGVDLYYPHSTTLNLAN-SLVVFEKYLNRNRFIYSLSKKKVY---DLIKTELO 331  
A.t.CAS1 NEARNLCAKE DLYYPHPLVQDILWASLHKIVEPVLMRWPGANREKKAIRTAIEHTHYEDE 358  
.....  
S.c.ERG7 **▼E373C400**  
NTDSLCLAPVNOAFCALVTLIEGVDSEAFORLOQREKDALFHGPGGMTIMGTVGQVOTWD 391  
A.t.CAS1 NTRYICIGPVNKVLNMLCCWVED-PNSEAFKLHLPRIHDFWLAEDGMKMQSYNGSOLWD 417  
.....  
S.c.ERG7 CAFAIQYFFVAGLAERPEFYNTIVSAYKFLCHAQFDTECVPGS--YRDKRKGAWGFSTK 448  
A.t.CAS1 TGFAIQAILATNLVE--EYGPVLEKAHSFVKNSQVLEDCPGDLNYWYRHISKGAWGFSTA 475  
.....  
S.c.ERG7 **▼Pst 1**  
TQGYTVADCTAEAIKAIIMVKNSEVVFSEVHHMTSSERLFEIGIDVLLNLQNI GSFYGSFA 508  
A.t.CAS1 **▼Pst 1**  
DHGWPISDCTAEGKKAALLSKVPEK-EIVGEPTDAKRLYEAVNVIISLQAD----GGLA 530  
.....  
S.c.ERG7 **▼E524C546**  
TYEKIKAPLAME TLNPAE VFGDIMVEYPYVECTDSSVLGLTYEHKYFD-YRKEEIRTRIR 567  
A.t.CAS1 TYELTRSYPWLELINPAE VFGDIVIDYPYVECTSAAIQALISFRKLYPGHRKKEVDECIE 590  
.....  
S.c.ERG7 IAIEFIKKSQLEP DGSWYGSWGCFTYAGMFALEALHTVGETYENSSTVRKGCDFIVSKQM 627  
A.t.CAS1 KAVKFIESIQAADGSWYGSWAVCFTYGTWFGVKGLVAVGKTLKNSPHVAKACEFLLSKQQ 650  
.....  
S.c.ERG7 KDCGGGESMKSSE--LHSYVDSEKSLVVQTAWALIALFAEYPNKE--VIDRGIDLLKNR 683  
A.t.CAS1 PSGGWGESYLSQCDKVYSNLDGNRSHVVNTAWAMLALIGAQAEVDRKPLHRAARYLINA 710  
.....  
S.c.ERG7 QEESGEWKVESVEGVFNHSCALEYPSYRFLFPIKALGMYSRAYETHTL- 731  
A.t.CAS1 QMENGDFPQQEIMGVFNRCMITYAAYRNI FPIWALGEYRCQVLLQOGE 759  
.....

Figure 4.3 Sequence alignment and secondary structure comparison between [A] *Panax ginseng*  $\beta$ -amyrin synthase<sup>83</sup> (P.g.bAS) and *Arabidopsis thaliana* lupeol synthase<sup>85</sup> (A.t.LUS), and [B] *Saccharomyces cerevisiae* lanosterol synthase<sup>52</sup> (S.c.ERG7) and *Arabidopsis thaliana* cycloartenol synthase<sup>81</sup> (A.t.CAS1). Orange boxes indicate the  $\alpha$ -helices, and black dotted lines denote the  $\beta$ -strands. The prediction of secondary structure elements were carried out via bidirectional recurrent neural networks (SSpro, and PHDsec program).<sup>131, 132</sup> The consensus sequences are boxed in black. The frames (with an inverted triangle and the amino acid number) indicate the position of the switching gene, from which the Symbol-E, C, L, and B represent the lanosterol synthase, cycloartenol synthase, lupeol synthase, and  $\beta$ -amyrin synthase, respectively. The PstI denotes the restriction cutting site of the restriction enzyme *PstI*.



## Chapter 5

# *Mechanistic Insights into Oxidosqualene Cyclization through Homology Modeling and Quantum Mechanism Calculation*

### 5.1 Introduction

The functional role of cyclase enzymes in the oxidosqualene cyclization process is believed extensively from four aspects. First, OSCs trigger the substrate, oxidosqualene, to adopt the prefolded conformation in the enzyme active site cavity before the initiation reaction. Second, the rigidly held, partially cyclized carbocationic intermediates are sequentially produced and stabilized by the cation- $\pi$  interaction in the proper enzymatic active site. Third, the different rearrangement or ring expansion products are formed via the species-dependent cyclases catalyzed reactions, respectively. Most importantly, the cyclases could prevent the early deprotonated truncation or the nonspecifically water nucleophilic additional termination. Thus, the product specificity and the stereoselectivity are rigorously mediated via these cyclases triggered reactions. In contrast, the tricyclic structure with a 3-hydroxylated **6-6-5** ring configuration was isolated from the nonenzymatic, lewis acid catalyzed cyclization reaction of oxidosqualene in presence of stannic chloride in benzene.<sup>133</sup> The spontaneous methyl group migrations were also observed in this nonenzymatic cyclization reaction. Thus, the importance of enzymatic control in the biological system is proven more directly for prevention of chemical tendency toward five-membered C-ring and also for requirement of cyclohexyl C-ring formation. Moreover, the highly exothermic cyclization reaction of the flexible oxidosqualene is precisely regulated via the enzyme mediated reaction to avoid producing diverse

products. The conserved Q-W motif domains, constitute of several aromatic acid repeats in the enzyme outer barrel helices, were believed to stabilize this highly exergonic cyclization reaction.<sup>48, 51</sup>

The mechanistic difference and sequences diversity among different cyclases imply that the structural insight is very critical for understanding the enzymatic mediated cyclization. In order to further clarify the importance of OSC enzyme in the oxidosqualene cyclization reaction, we performed the homology modeling studies toward the yeast *Saccharomyces cerevisiae* ERG7 enzyme. The generated yeast homology models coupled with bioinformatic ligand docking experiments provide the molecular basis for examining the isolated truncated/early-terminated products from different yeast ERG7 mutants. Moreover, in order to examine the relationship between the mutated enzyme and the thermodynamic tendency of diverse products, the quantum mechanical calculations were also carried out by using the Gaussian G03 program. The relative energies calculation among these truncated/early terminated products coupled with the homology modeling structural elucidation would provide complementary but comparative insights for fully elucidating the mechanisms of oxidosqualene cyclization.

## 5.2 Results and Discussion.

### 5.2.1 Sequence alignment, secondary structure prediction, and homology modeling

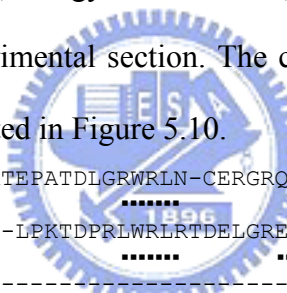
In order to verify the template selection and compare the diversity among the individual cyclases, multiple sequence alignment of *Homo sapiens* OSC (*H. sapiens* OSC), *Saccharomyces cerevisiae* ERG7 (*S. cerevisiae* ERG7), and *Alicyclobacillus acidocaldarius* SHC (*A. acidocaldarius* SHC) was carried out.<sup>134</sup> Pairwise sequence

alignment of yeast *S. cerevisiae* ERG7 and human OSC showed 40% identity, whereas 18% identity was calculated between yeast *S. cerevisiae* ERG7 and SHC. Approximately 26% sequence identity was also observed between human OSC and *A. acidocaldarius* SHC, as previous literature has reported.<sup>127</sup> The multiple sequence alignment revealed that the highly conserved active site residues exist among these three cyclases (Figure 5.1). Due to the same reaction mechanism, the homology modeling construction of *Saccharomyces cerevisiae* ERG7, based on the crystal structure of *Homo sapiens* OSC, could be expected to be reliable. Moreover, in order to elucidate the mechanistic difference between OSC and SHC at the molecular level, the homology modeling studies of OSC based on the SHC crystal structure is meaningful.

The secondary structural elements from crystal structures of *Alicyclobacillus acidocaldarius* SHC and *Homo sapiens* OSC were also included in the multiple sequence alignment and compared with the predicted secondary structural elements of *Saccharomyces cerevisiae* ERG7, respectively (Figure 5.1).<sup>131, 132, 135</sup> As could be expected, the predicted secondary structure of *Saccharomyces cerevisiae* ERG7 showed similarity to the other two cyclases. The main secondary structure elements,  $\alpha$ -helices, matched perfectly in the analogous regions among these cyclase genes. Moreover, short  $\beta$ -strands assigned in the yeast ERG7 also displayed highly parallel alignment to the other two genes, except in the *N*-terminal region. The similar protein folding among these cyclases is theoretically predicted according to the distribution of secondary structure elements (Figure 5.1). Thus, *Alicyclobacillus acidocaldarius* SHC or *Homo sapiens* OSC would be the appropriate template for the generation of a homology modeling structure of yeast ERG7.

The homology models of the *Saccharomyces cerevisiae* ERG7 were created by using the Insight II Homology program utilizing X-ray crystal structures either from

*Alicyclobacillus acidocaldarius* SHC (PDB entry 1UMP), or from *Homo sapiens* OSC (PDB entry 1W6K) as the primary template. The generated models were applied into the energy minimization calculation module using the Sybyl 7.0 software until energy convergence. The different molecules considered including protosteryl cation, lanosteryl cation, or other truncated cyclization/rearrangement products were docked into the homology models by using GOLD 3.0 software. Finally, the ligand-protein complexes were also applied into the energy minimization calculation by using the Sybyl 7.0. The reliability of the generated homology models or the specific interaction between ligands and the specific residues within the enzyme active site will be discussed in the following sections. The detailed procedure for generation of homology modeling structures, energy minimization, and ligand docking experiments will be presented in the experimental section. The chemical structures of all ligands for GOLD 3.0 software are listed in Figure 5.10.



```

HsOSC  MTEGTCLRRRGGPYKTEPATDLGRWRLN-CERGRQTWTYTLQDERAGREQTGLEAYALGLD  59
          .....
ScOSC  MTEFYSDTIG-----LPKTDPRLLWRLRTDELGRESWEYLTPQQAANDPPSTFTQWLLQD  54
          .....
AaSHC  -----
          .....

HsOSC  TKNYFKDLPKAH-----TAFEGALNGMTFYVGLQAED-GHWTGDYGGPFLFLPGLLITC  112
          .....
ScOSC  PK-FPQPHPERNKHSPDFSAFDACHNGASFFKLLQEPDSGIFPCQYKGPMTIGYVAVN  113
          .....
AaSHC  ---MAEQLVEAP-----AYARTLDRAVEYLLSCQKDEGYWWGPLLSNVTMEAEYVLLC  50
          [α1] ..... [β1]... [α2] .....

HsOSC  HVARIPLPAGYREEIVRYLRSVQLP-DGGWGLHIEDKSTVFGTALNYVSLRILGVGPD  171
          .....
ScOSC  YIAGIEIPEHERIELIRYIVNTAHPVDGGWGLHSVDKSTVFGTVLNIVILRLLGLPKDHP  173
          .....
AaSHC  HILDR-VDRDRMEKIRRYLLHEQRE-DGTWALYPGGPPDLDTTIEAYVALKYIGMSRDEE  108
          [α3] ..... [α4] ..... [α5] .....

HsOSC  DLVRARNILHKKGGAVAIPSWGKFWLAVLNVSWEGLNTLFPPEMWFPPDWAPAH PSTLWC  231
          .....
ScOSC  VCAKARSTLLRLGGAIGSPHWGKIWLSALNLYKWEVNPAPPETWLLPYSLPMHPGRWWV  233
          .....
AaSHC  PMQKALRFIQSQGGIESSRVFTRMWLALVGEYPWEKVPMVPPEIMFLGKRMP LNIYEFGS  168
          [α6] ..... [α7] .....

HsOSC  HCRQVYLPMSYCYAVRLSAAEDPLVQSLRQELYVEDFASIDWLAQRNNVAPDELYTPHSW  291
          .....
ScOSC  HTRGVYIPVSYLSLVKFSCPMTPLEELRNEIYTKPFDKINF SKNRNTVCGVDLYPHST  293
          .....
AaSHC  WARATVVALS---IVMSRQPVFPLPERAR----VPELYETDVP PRRRGAK----GGGGW  216
          ..... [α8] .....

```





## 5.2.2 Geometrical verification of the generating homology models

In order to verify the quality of the yeast ERG7 homology modeling structure, some geometrical check programs were used and their protein 3D profiles were also generated via the NIH MBI Laboratory Servers in UCLA (<http://nihserver.mbi.ucla.edu/>).

### 5.2.2.1 Protein volume evaluation.

**Volume Z-scores**, the atomic volume deviation values computed from 64 highly resolved and well-refined protein crystal structures, were used to assess the quality of the protein crystal structures as well as the homology models. Volume Z-score root mean square deviation (**Z-score rms**) also represents the resolution of a structure or models by inversely *i.e.* Z-score rms decreasing means the resolution improving. These two “Volume-based structure validation” provide an independent methodology to identify the problems in specific regions within a protein structure.<sup>136</sup> The result from the program “**PROVE**” (**PROtein Volume Evaluation**) calculation showed that the percentage of the buried outliers in the *S. cerevisiae* ERG7 homology model is smaller than 1.0%. This small distribution revealed a good satisfaction in this homology modeling structure.

### 5.2.2.2 VERIFY\_3D program verification.

“**VERIFY\_3D program**” that plots an averaged 3D-1D score within a sequence position window was developed and provided the evaluation for the correctness of the homology modeling structure.<sup>137</sup> The high score value in the 3D-1D profile was originally observed from the perfect match between the atomic coordination in the well known protein structures (3D) and their own sequences (1D). The 3D-1D profiles of the *S. cerevisiae* ERG7 model and human OSC crystal structure respectively showed that nearly ninety percent residues have an average 3D-1D score >0.2, indicating that these amino acids are in the preferable environment (88.25% in *S.*

*cerevisiae* ERG7 model, whereas 97.11% in human OSC crystal structure as illustrated in Figure 5.2). The higher score value for the human OSC structure calculation revealed that the better overall packing was existed for the crystal structure as previously reported other homology modeling experiments.<sup>127</sup>

#### 5.2.2.3. PROCHECK, the stereochemical quality program of protein structures.

The Ramachandran plots originally showed the phi-psi torsion angles for all residues in the structure, except those at the chain terminus. According to the statistical occurrence of the phi-psi torsion angles in well-known structure, Thornton *et al.* developed a methodology to examine the quality of modeling structure.<sup>138</sup> Because of the atypical dihedral angle distributions, proline and glycine residues are always excluded. The result from the program “PROCHECK” for the *S. cerevisiae* ERG7 homology model showed that almost all of the residues were assigned to either the CORE, ALLOWED, or GENEROUS region, whereas only four residues (0.6% of overall population) were located in the disallowed region (Figure 5.3). This result is better than that of previous human OSC modeling structure based on the SHC crystal structure.<sup>127</sup> Accordingly, from the observation of 121,879 residues in the 463 known structure in the PDB bank, 1.3% of overall residue population belonged to the disallowed region (Table 5.1).<sup>139</sup> Especially, these four residues (Phe-71, Glu-328, Phe-426, or Glu-526) are neither located in the putative active site nor be proposed to participate in the enzyme function.

Additionally, the “good satisfactory” were also assigned for most of the resulting sections according to the “WHAT\_CHECK” calculation, a suite of programs to assess the "stereochemical quality" of a given protein structure. Therefore, the homology modeling structure of the yeast ERG7 would be a reliable model to provide useful mechanistic information from the structural aspects.

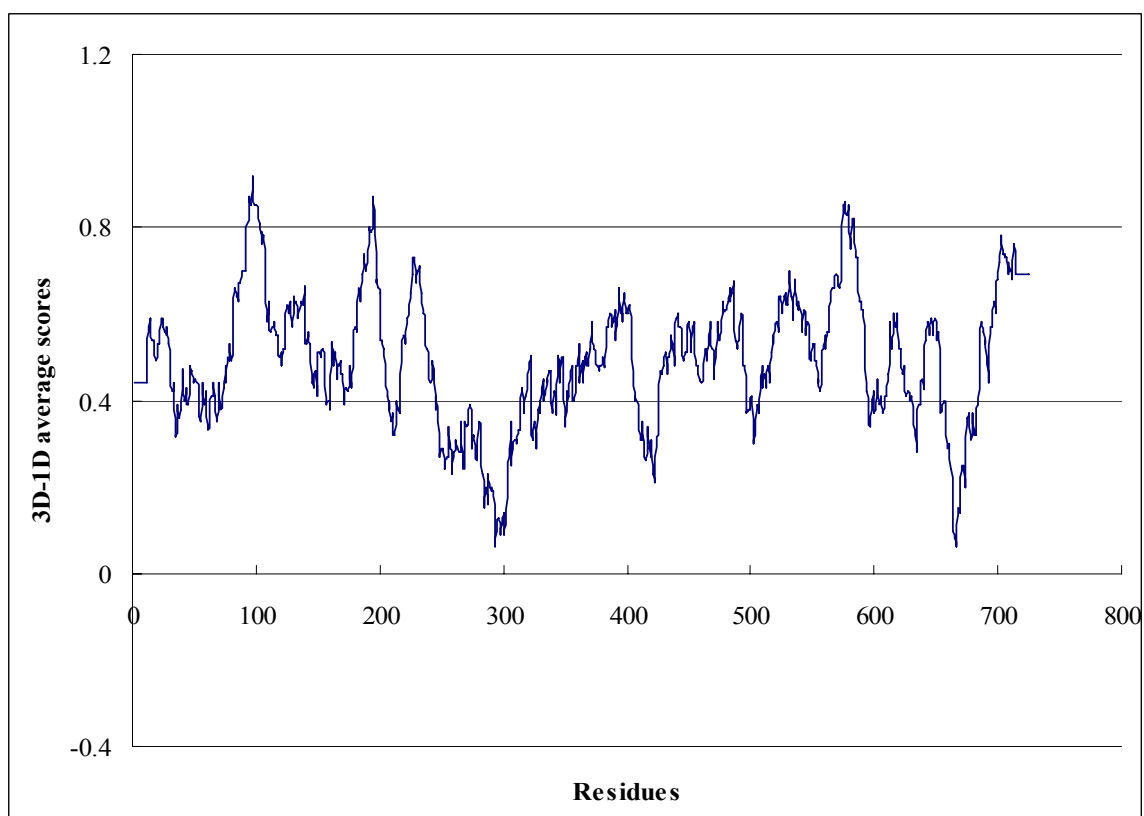
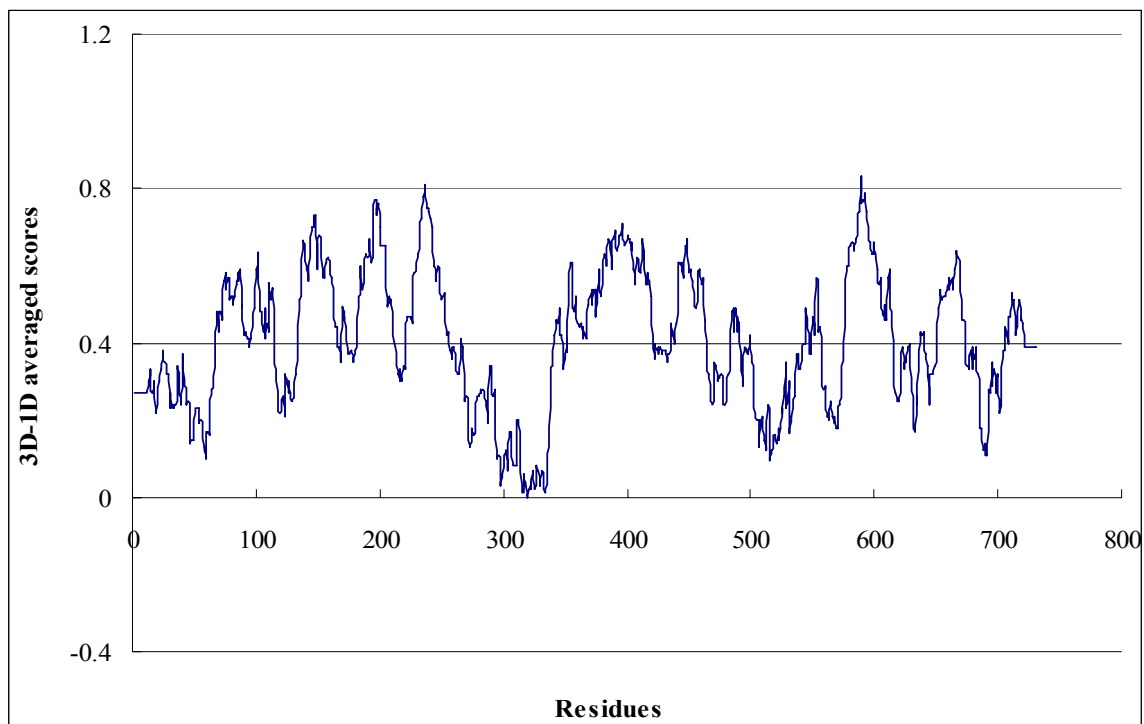


Figure 5.2 The 3D-1D profiles of the homology modeling structure of yeast ERG7 (Upper plot) and human OSC crystal structure (Downer plot).

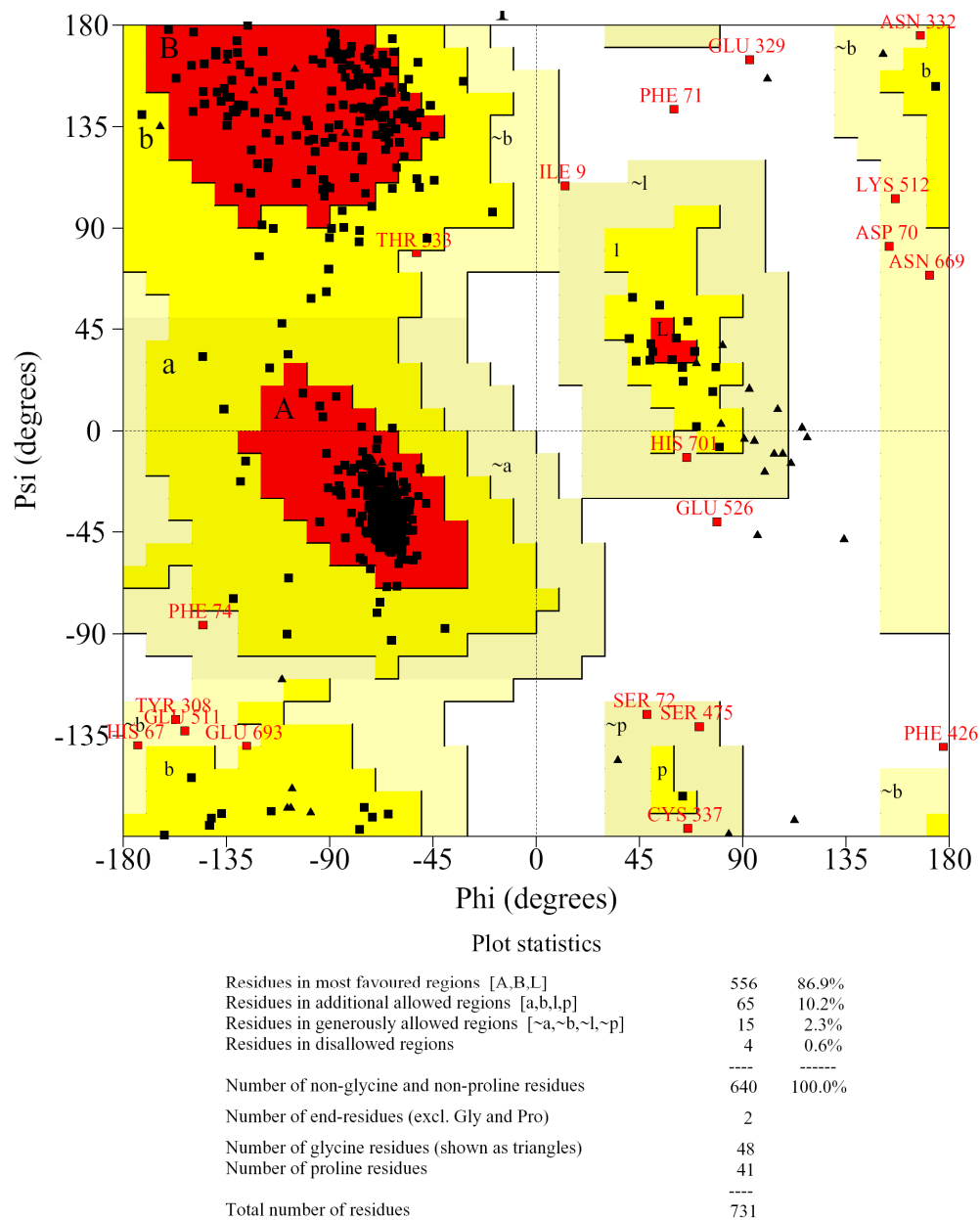


Figure 5.3 The Ramachandran plot of the homology model of yeast ERG7.

Table 5.1 Statistical occurrence of  $\phi$ ,  $\psi$  angles in the PDB bank compared with the  $\phi$ ,  $\psi$  angles distribution in the homology modeling structure of yeast *S. cerevisiae* ERG7.<sup>127, 139</sup>

Region	Occurrence in PDB bank <sup>139</sup>	in <i>S. cerevisiae</i> ERG7
	(%)	(%)
<b>Core</b>	81.9	86.9
<b>Allowed</b>	14.8	10.2
<b>Generous</b>	2.0	2.3
<b>Outside</b>	1.3	0.6

### 5.2.3 Comparison between homology models and mechanistic hypothesis

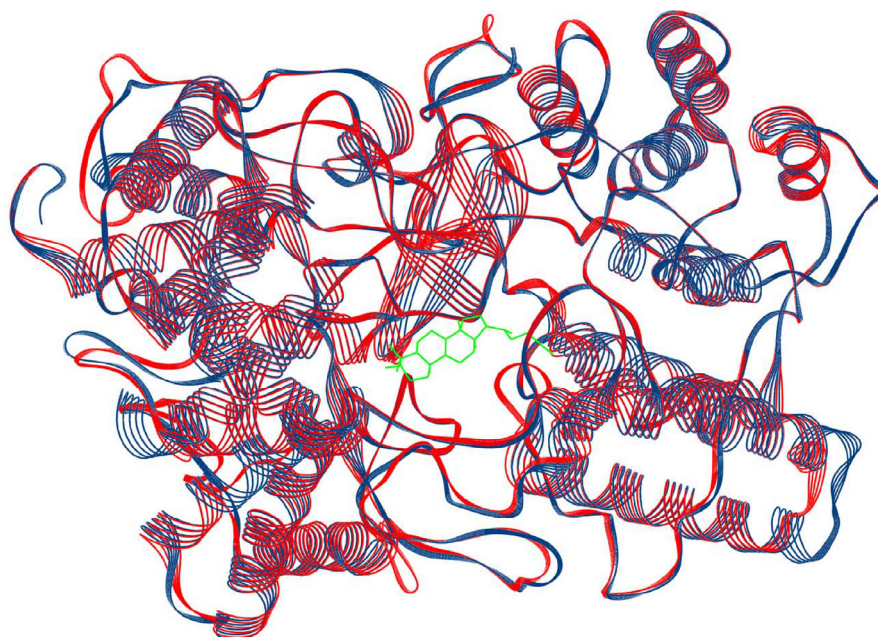
The preliminary superimposition of the *S. cerevisiae* ERG7 homology modeling structure with either crystal structure of squalene-hopene cyclase or the human oxidosqualene-lanosterol cyclase was performed by using WebViewer software, respectively. As shown in Figure 5.4, the homology model of *S. cerevisiae* ERG7 exhibited the high similarity with those two crystal structures, indicating that the overall skeleton coincidence was conserved among these mechanistic similar cyclases. Due to the redundant sequence alignment, the *N*-terminal region of *S. cerevisiae* ERG7 was specifically omitted in the superimposition of the *S. cerevisiae* ERG7 homology model with the squalene-hopene cyclase crystal structure. This *N*-terminal region has been proven recently to fill the space between two  $\alpha$ - $\alpha$  barrel domains and act as the stabilizing role in the human OSC structure.<sup>105</sup> Moreover, Ruf's group has identified several residues at a distances within 5 Å to lanosterol molecule in the human OSC crystal structure.<sup>105</sup> These putative active site residues were also highly conserved in the yeast ERG7 from the observation of the multiple sequences alignment (Figure 5.1). The superimposition of the *S. cerevisiae* ERG7 homology modeling structure with the crystal structure of human oxidosqualene-lanosterol cyclase also exhibited the similar orientation among these putative active site residues (Figure 5.5).

On the other hand, for reasons of the inherent differences in cyclization reaction between OSC and SHC: such as using squalene instead of oxidosqualene as substrate, the vicinal importance of histidine residue for assisting the catalytic acid residues, the opposite substrate folding manner during the B-ring formation, the cyclohexyl D/E-ring formation, and the loss of the rearrangement step, led the elucidation of these diversity via structural insight would be plausible and meaningful. The hypothesized mechanism for oxidosqualene or squalene cyclization is listed in the



Scheme 5.1 and will be discussed in the following text.

(A)



(B)

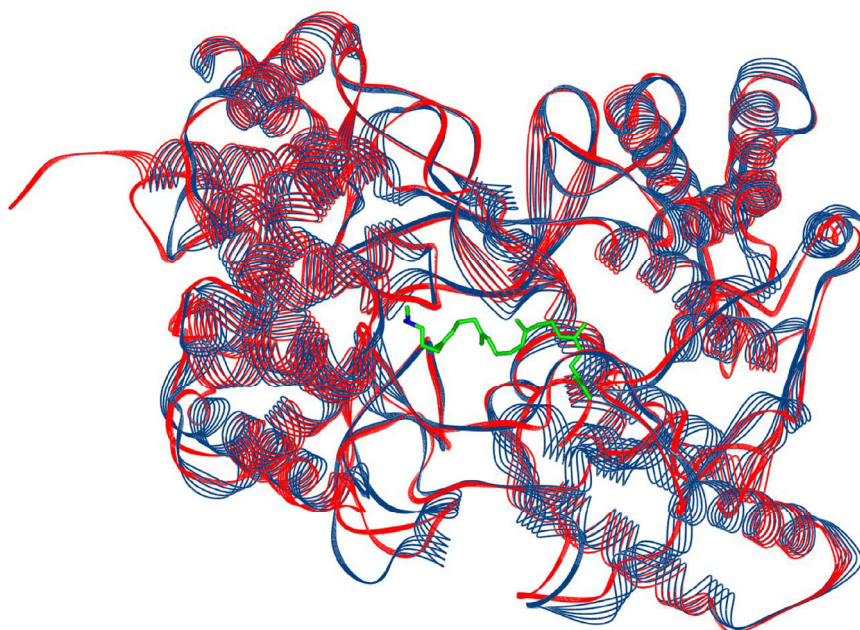


Figure 5.4 (A) Superimposition of the *S. cerevisiae* ERG7 homology modeling structure (red) with crystal structure of human OSC (blue), and the docked product, lanosterol (green), is included. (B) Superimposition of the *S. cerevisiae* ERG7 homology modeling structure (red) with crystal structure of *A. acidocaldarius* SHC (blue), and the inhibitor, 2-azasqualene, is shown in green.



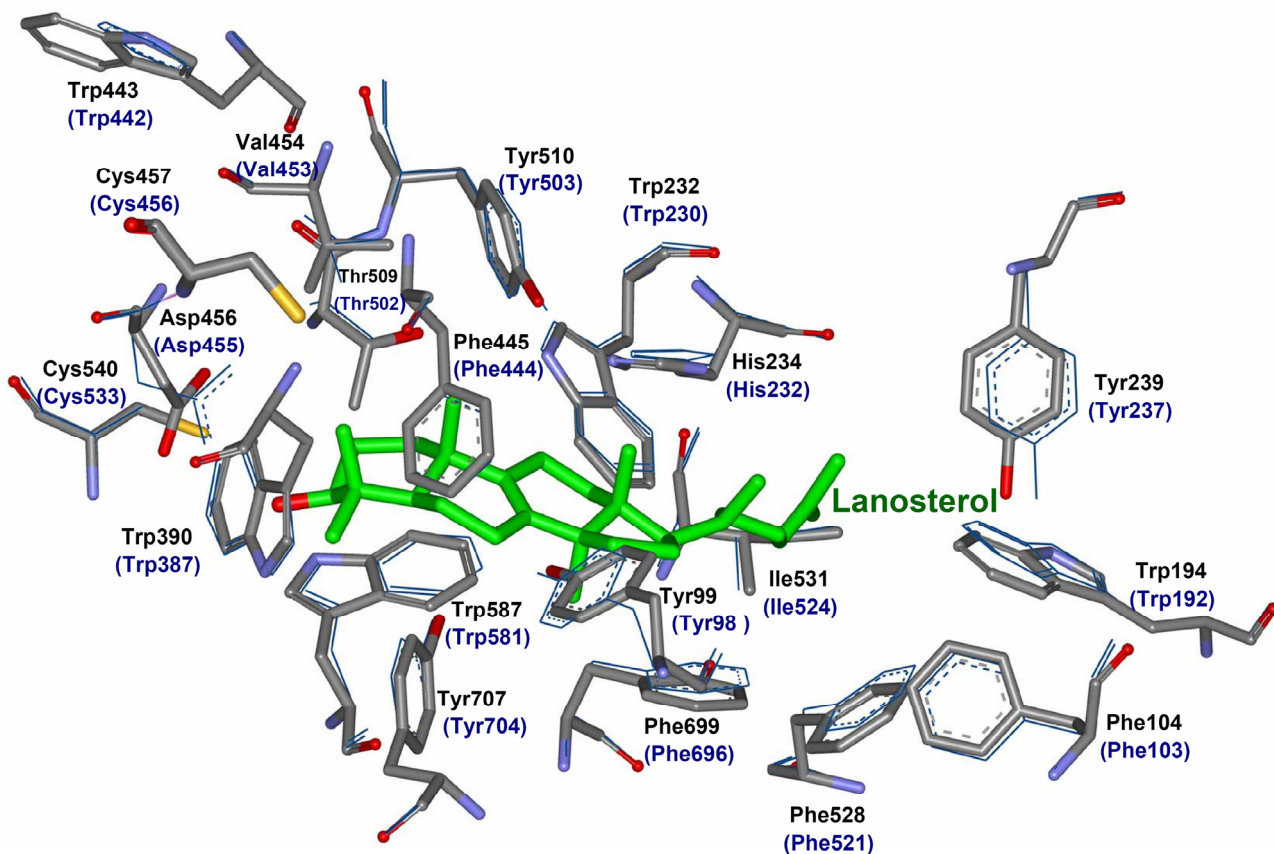
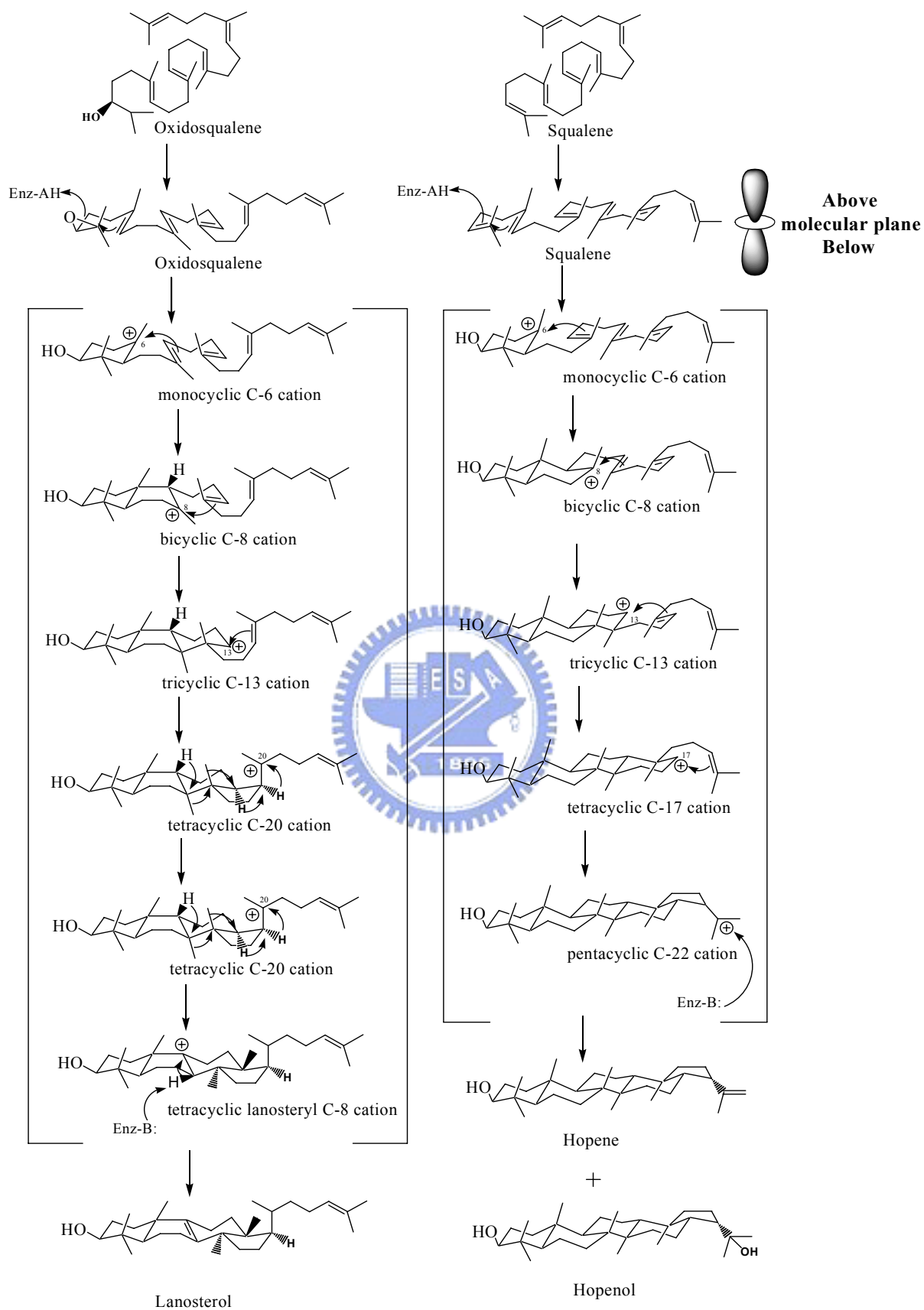


Figure 5.5 Local views of the superimposition of homology modeled *S. cerevisiae* ERG7 structure (shown in stick representation and the corresponding residues are marked with black color) with human OSC crystal structure (shown in line representation with blue color). Lanosterol (green) is also included.



Scheme 5.1 The hypothesized cyclization mechanism of oxidosqualene (left) and squalene (right).

First, because of the highly consistency in the A-ring formation, most of the putative amino acids participating in this process were highly conserved from the observation of superimposition of the *S. cerevisiae* ERG7 homology model with the squalene-hopene cyclase crystal structure except two residues: Asp-374 is changed to Val-454 and Asp-377 is changed to Cys-457 (*S.c.*ERG7 numbering). The cysteine substitution has been proposed for the functional role in activating the catalytic Asp-456 (*S. cerevisiae* ERG7 numbering) during the initiation reaction, whereas the valine substitution provided the steric importance for the proper substrate folding.<sup>105,</sup>  
<sup>112</sup> Obviously, most of the aromatic amino acids near the first tertiary C-6 cation, such as Phe-445, Tyr-510, Trp-587, Tyr-707 (*S.c.*ERG7 numbering) were observed in parallel to their corresponding amino acids in SHC structure, respectively (Figure 5.6).

The first mechanistic differences between OSC and SHC mediated cyclization reaction is the substrate folding manner during the B-ring formation. The pre-chair-boat-chair conformation of OS is adopted in the OSC active site, whereas the squalene is prefolded with all-chair conformation during its cyclization process. Therefore, the orientations of the resulting tertiary cation (C-8 cation) and the connecting methyl group are respectively either located above the molecular plane of squalene, or below the molecular plane of oxidosqualene during the B-ring formation. The energy unfavorable B-ring conformation in the oxidosqualene cyclization reaction was proposed to be mediated by the enzymatic active site residues.<sup>105</sup> One amino acid deletion below the molecular plane and one amino acid insertion above the molecular plane of oxidosqualene were mentioned in the OSC active site (Figure 5.7).<sup>127</sup> The steric pressure from Tyr-98 or the electrostatic interaction from Lys-331 in human OSC were also proposed to assist this energetically unfavorable conformation formation during the oxidosqualene cyclization.<sup>105, 127</sup>

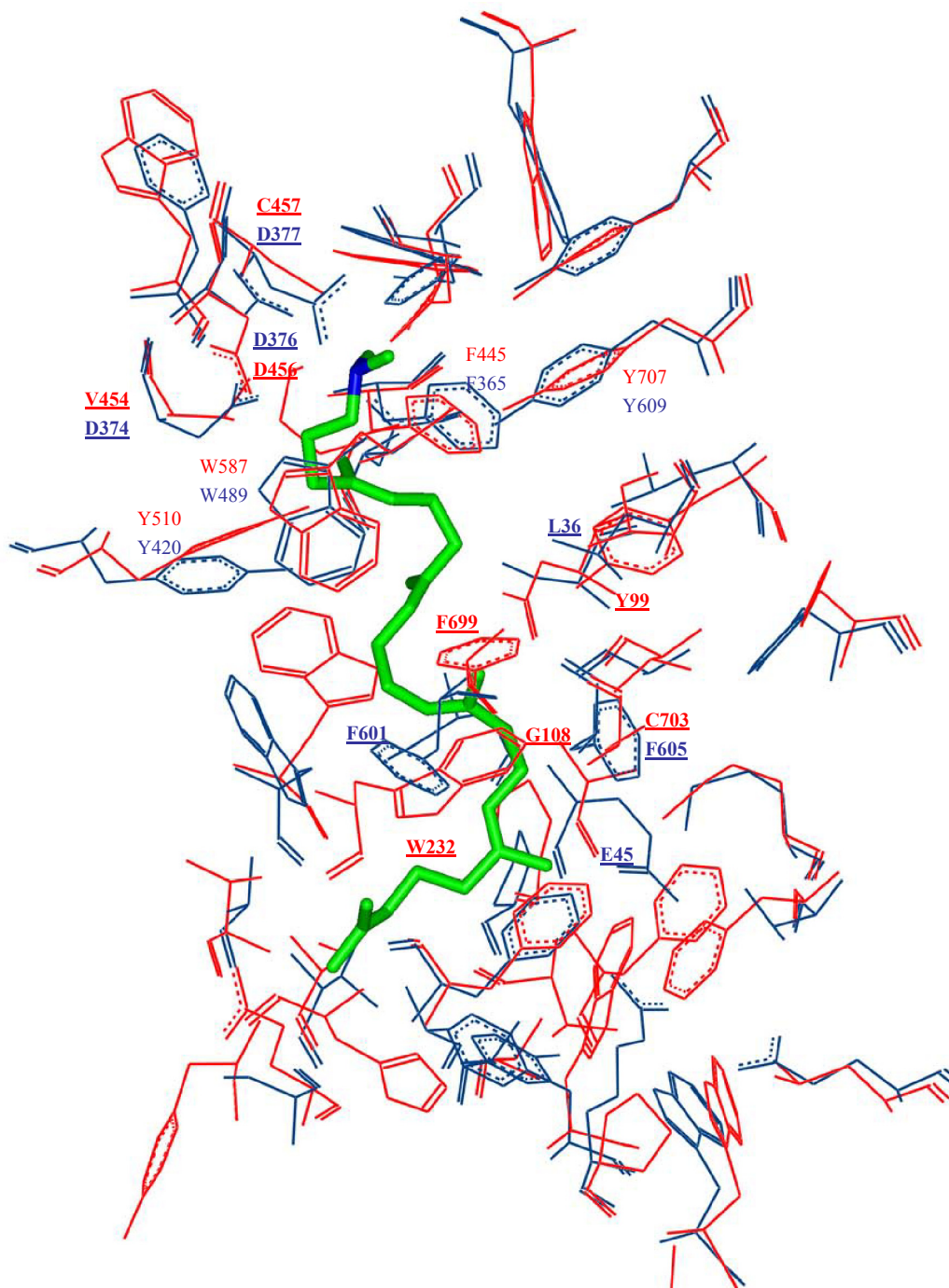


Figure 5.6 Local views of the superimposition of homology modeled *S. cerevisiae* ERG7 structure (red) with SHC crystal structure (blue). 2-azasqualene is shown in green. The active site residues that mentioned in the text are marked in bold and underlined.

From the homology modeling illustration of the *S. cerevisiae* ERG7, the Tyr-99 (*S. cerevisiae* ERG7 numbering) which is positioned above the C-8 atom might push the methyl group below into the molecular plane (Figure 5.7). Moreover, from the multiple sequence alignment, the same position is replaced with the leucine residue (Leu-36) in the SHC active site (Figure 5.1). The smaller steric hindrance of leucine residue might facilitate the formation of the chair form conformation during squalene cyclization. On the other hand, one amino acid deletion (relative to the Thr-601 in SHC) in the ERG7 active site might promote the position of Phe-699 (*S. cerevisiae* ERG7 numbering) closer to the methyl group during the B/C ring formation and also create the less hinder space between Phe-699 and Tyr-707 (*S. cerevisiae* ERG7 numbering) for the orientation of the resulting methyl group. In contrast, the insertion of threonine residue in the SHC might promote the Phe-601 to interact with the deeper cation that was produced in the following squalene cyclization cascade (Figure 5.7).

However, the electrostatic interaction from Lys-331 was never observed in the yeast oxidosqualene cyclase homology modeling structure. The negative charge residue, Asp-334, is located at the corresponding position according to the multiple sequences alignment result. The electrostatic interaction proposed from the human OSC crystal structure might not be suitable for illustrating the formation of the boat B-ring in the yeast ERG7.<sup>127</sup> Additionally, from the observation of the *S. cerevisiae* ERG7 homology model, one amino acid insertion above the molecular plane does not influence the steric space enormously. This insertion might just promote Tyr-99 in the proper orientation for the formation of boat B-ring during the oxidosqualene cyclization.





Figure 5.7 (A) The hypothetical enzyme-triggered steric control for the formation of the boat cyclohexyl B-ring. One-residue insertion above and a one-residue deletion below the molecular plane of the substrate within the enzymatic active site were proposed. (B) Local view of *S. cerevisiae* ERG7 homology modeling structure. The functional residues, Asp-456, Tyr-99, Phe-699, and Tyr-707 are listed in the stick representation. The truncated cyclization intermediate C-14 cation is shown in gray, and the lanosterol molecule is shown in green. (C) The local relative spatial position of amino acid residues near to the B/C ring junction and the partial sequence alignment of cyclases for illustrating the position of one amino deletion in the *S. cerevisiae* ERG7.

For the C-ring formation, both squalene and oxidosqualene adopt the pre-chair conformation in the enzyme active site. The differences between squalene and oxidosqualene cyclization in the C-ring formation are the reverse orientation of methyl group and the resultant cation. This conformational difference is primarily caused by the previously reversed B-ring conformation. Thus, the methyl group connected to the C-14 is either above the molecular plane of oxidosqualene or below the molecular plane of squalene as previously mentioned in the B-ring closure. The homology model and the crystal structure showed that the Trp-232 from *S. cerevisiae* ERG7 and Phe-601 from SHC might provide the stabilizing  $\pi$ -cation interaction toward the resultant cation in different orientation. Moreover, the Phe-699 from *S. cerevisiae* ERG7 might also participate in the interaction with the resultant cyclohexyl cation due to its proper position just below the C/D ring plane (Figure 5.6, and Figure 5.8).

The final ring closure in the oxidosqualene cyclization is the formation of cyclopentyl D-ring, whereas the cyclohexyl D-ring is further formed in the squalene cyclization. The cyclohexyl secondary anti-Markovnikov cation has proven to be stabilized by the Phe-605 in the SHC. As expected, the truncated five-membered ring products were isolated from the SHC<sup>F605A</sup> mutant.<sup>140</sup> The replacement with small



cysteine residue was observed in the OSC homology model. The cyclopentyl D-ring but not cyclohexyl D-ring formation was thus formed due to the loss of the stabilizing interaction (Figure 5.8). Interestingly, the fifth E-ring has never been observed from the natural mammalian oxidosqualene-lanosterol cyclase-mediated cyclization. The dramatic difference in the amino acid arrangement, especially in those stabilizing aromatic residues, was observed between OSC and SHC enzymatic active site. The careful illustration of the spatial position of these functionally important amino acids might provide clues for understanding the mechanistic difference in the E-ring formation.

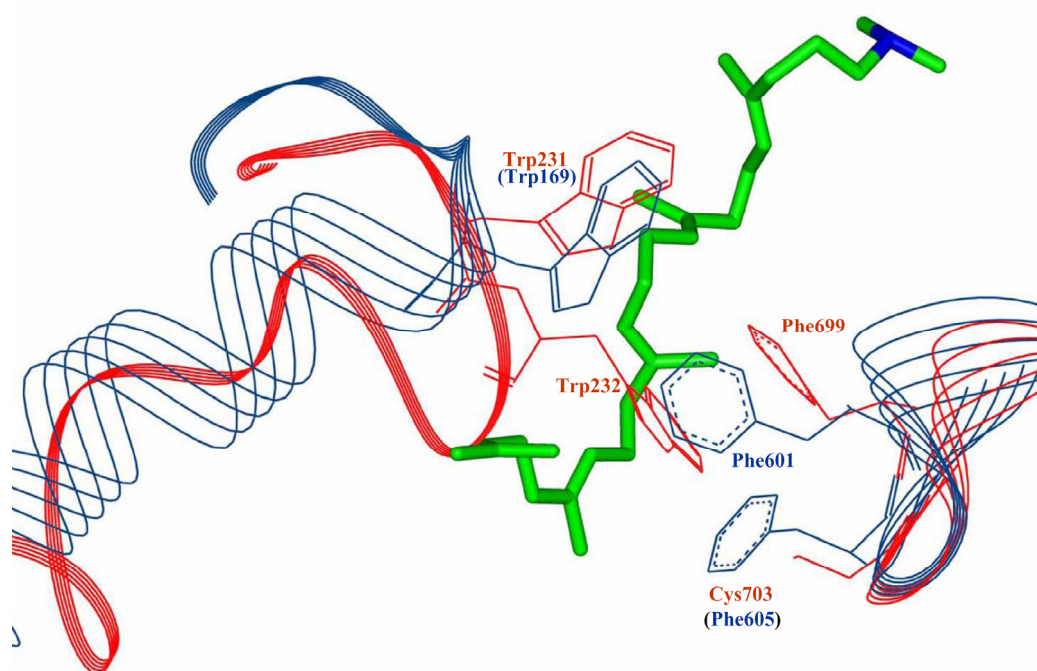


Figure 5.8 The different orientation of Trp-232 from *S. cerevisiae* ERG7 and Phe-601 from *A. acidocaldarius* SHC, which are involved in the formation of C ring. The Phe-605 from *A. acidocaldarius* SHC was proposed for stabilizing the cyclohexyl secondary anti-Markovnikov cation during the formation of D ring. The corresponding Cyc-703 from *S. cerevisiae* ERG7 is also shown.

The functional role of OSC assisting in the skeletal rearrangement has been proposed through the  $\pi$ -electron density surrounding with the intermediate carbocations. The high  $\pi$ -electron density from the aromatic residues were abundantly observed to occupy the position around the intermediate C-6 or C-10 cation, whereas the density was substantially reduced near the position of the C-20 or C-17 cation. Although, it was believed that the protosteryl cation would progress the spontaneous skeletal rearrangement imperceptibly even in the nonenzymatic condition, the highly conserved aromatic residues around the lanosteryl cation indicated the specific importance of the cyclase enzyme in shifting the equilibrium from the protosteryl cation toward the C8/C9 lanosteryl cation.<sup>10, 133</sup> In contrast, due to the proper position of Glu-45 in the SHC, the direct proton abstraction or water quenching termination, rather than skeletal rearrangement was observed for the hopene or hopenol synthesis.<sup>141</sup> Moreover, the critical functional residue for the deprotonation step in the oxidosqualene cyclization was proposed based on the crystal structure of human oxidosqualene cyclase.<sup>105</sup> In parallel, the His-234 (corresponding to His-232 in human OSC) is the only basic residue to occupy the proper position for the role of the deprotonation reaction. However, the hydroxyl group of Tyr-510 (corresponding to Tyr-503 in human OSC) might also serve as the terminal step, due to its closer proximity to the C-8 or C-9 lanosteryl cation (Figure 5.5 and Figure 5.9).<sup>105</sup> Moreover, because of the highly hydrophobic properties in the active site, the water quenching product was never isolated in the natural oxidosqualene cyclase catalyzed-enzymatic reaction.

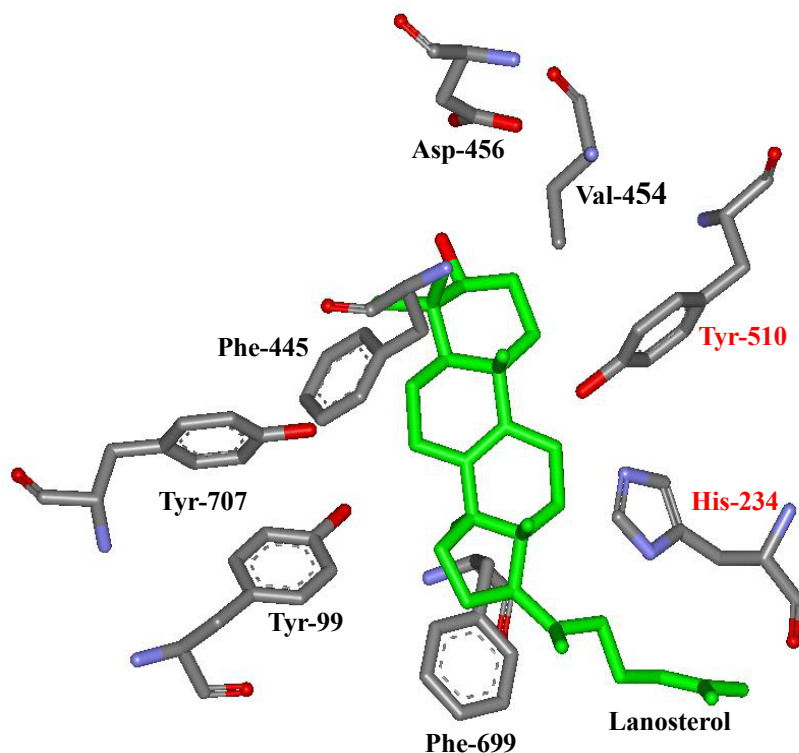


Figure 5.9 Local view of the homology modeling structure of *S. cerevisiae* ERG7. The functional amino acids are shown in the stick representation. The Tyr-510 and His-234 are the proposed base-dyad.

#### 5.2.4 Correlation of the OSC model to the experimental mutagenesis studies

Several functional residues, including Tyr-99, Trp-232, His-234, Trp-390, Trp-443, Phe-445, Lys-448, Tyr-510, Phe-699 and Tyr-707 in *S. cerevisiae* oxidosqualene-lanosterol cyclase (*S. cerevisiae* ERG7) have been respectively elucidated as the catalytically important residues for the cyclase-mediated reaction using site-directed/saturated mutagenesis coupled with bioorganic characterization in our laboratory.<sup>106, 115, 117-119, 125, 142</sup> Diverse products including monocyclic, bicyclic, tricyclic, tetracyclic, truncated rearranged tetracyclic, altered deprotonation tetracyclic, as well as the normal biosynthetic tetracyclic product, lanosterol, were isolated from respective single amino acid substituted *S. cerevisiae* ERG7 mutants. These isolated premature products not only imply the plasticity of the cyclases and the conceivable evolution from the common ancestral cyclase, but also provided the direct evidence

for the existence of the long-sought mechanistic intermediate states during the biosynthesis of lanosterol.

Due to the absence of *S. cerevisiae* oxidosqualene-lanosterol cyclase (*S. cerevisiae* ERG7) crystal structure, the elucidation of the relationship between mutated enzyme and its product diversity is hampered. Fortunately, bioinformatic tools and quantum mechanics simulation might provide some opportunities for rationally examine the role of these functional residues in the complex cyclization/rearrangement reaction. For example, the previously created plausible *S. cerevisiae* ERG7 homology model based on the crystal structure of bacterial squalene-hopene cyclase (*A. acidocaldarius* SHC) or human oxidosqualene-lanosterol cyclase (human OSC) could provide a suitable template for generating the individual homology modeling structures of *S. cerevisiae* ERG7 mutations. In addition, these individual cationic reactive intermediates or the respective annulated products including monocyclic, bicyclic, tricyclic, tetracyclic, truncated rearranged, and altered deprotonation tetracyclic structures could be simulated and docked into the active site cavity of the wild-type *S. cerevisiae* ERG7 and various mutated *S. cerevisiae* ERG7 homology models. Moreover, in consideration of the spatial orientation or electrostatic interaction between the enzymatic active site residues and the various annulated cationic intermediates, the generated homology modeled enzyme/ligand complexed structures were further applied for the energetic minimization module by Sybyl 7.0 software. The chemical structure of these cationic reactive intermediates and various annulated products is shown in Figure 5.10. Moreover, the respective premature products isolated from different mutations are listed in Figure 5.11. The examination or the local view of individual *S. cerevisiae* ERG7 mutation's homology modeling structure complexed with various cationic reactive intermediates or annulated products will be discussed in the following text.

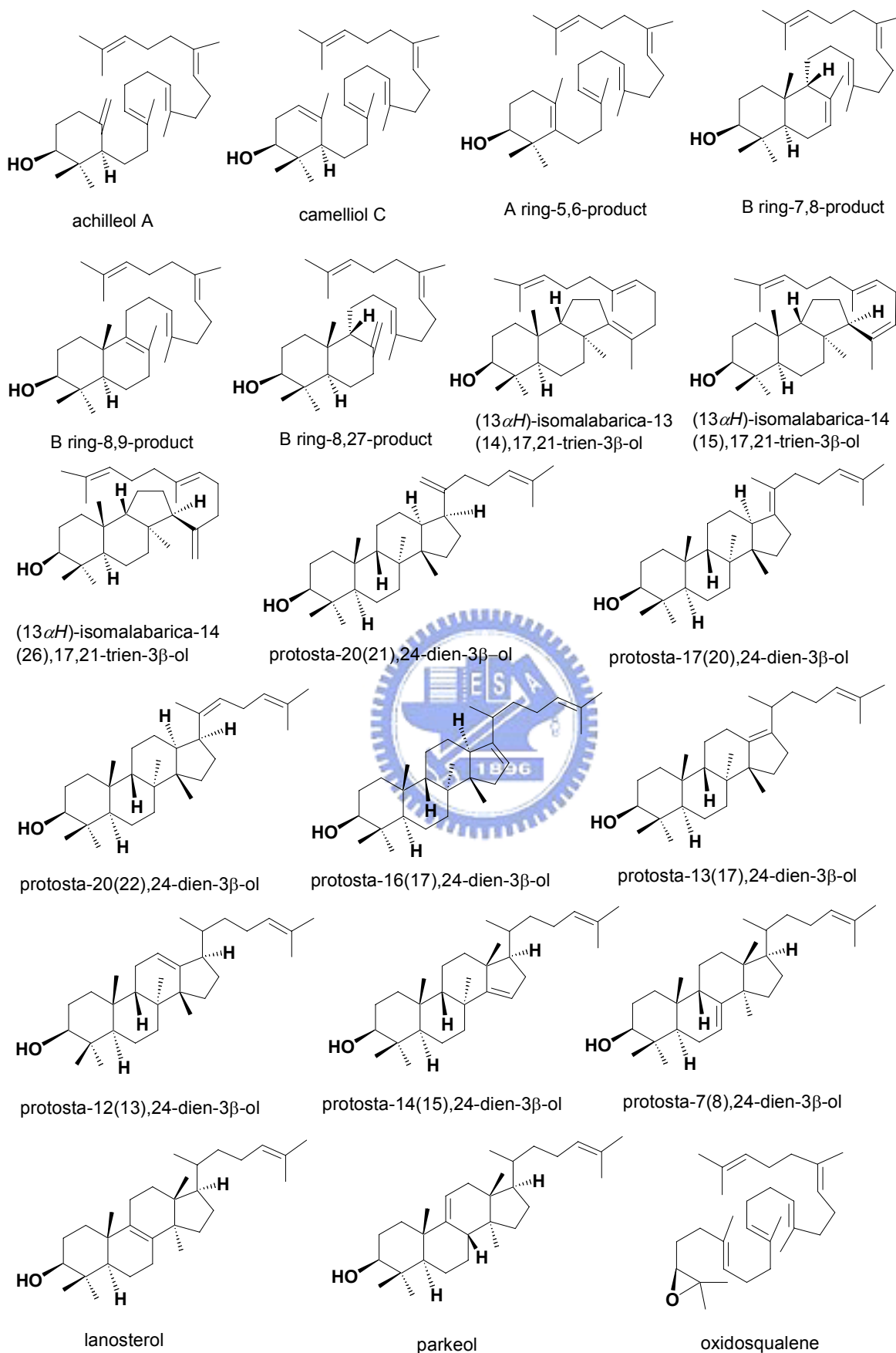


Figure 5.10 The chemical structure of the possible cationic reactive intermediates and various annulated products.



A detailed investigation of the homology modeling structures suggested that a  $\pi$ -electron rich pocket, which was created by previously illustrated catalytically important residues, including Tyr-99, Trp-232, His-234, Trp-390, Trp-443, Phe-445, Tyr-510, Phe-699, and Tyr-707 coupled with other aromatic amino acid residues, including Trp-587, Tyr-239, Trp-194, Phe-528, and Phe-104, existed in the *S. cerevisiae* ERG7 active site. Among these functional residues, the hydrogen-bonding network between His-234 and Tyr-510 was located on the top side of the molecular plane of lanosteryl cation, from which the Tyr-510 was spatially near the C8/C9 position, whereas His-234 was proximal to the C/D ring junction of lanosterol. In addition, the Trp-232, the neighbor residue of His-234, was thought to affect the orientation of His-234 as well as the His-234:Tyr-510 hydrogen-bond dyad.

As illustrated in the previous chapter and in other published data, the single amino acid substitution of *S. cerevisiae* ERG7<sup>Y510X</sup> generated various products from monocyclic achilleol A or camelliol C, and altered deprotonation tetracyclic lanosterol and parkeol.<sup>125, 143</sup> Together with the monocyclic and deprotonated structures from different *S. cerevisiae* ERG7<sup>Y510X</sup> mutants, the possible dual functions of this residue involved in cyclization and deprotonation steps of the oxidosqualene cyclization/rearrangement cascade has been proposed.<sup>106, 142</sup> Moreover, from the observation of the *S. cerevisiae* ERG7 homology model, the hypothesized model of the stabilization of Tyr-510 residue for the monocyclic cyclization or deprotonation reaction was also discussed in the previous chapter.

The isolation of tricyclic (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien-3 $\beta$ -ol, tetracyclic protosta-20,24-dien-3 $\beta$ -ol, and protosta-12,24-dien-3 $\beta$ -ol from various *S. cerevisiae* ERG7<sup>H234X</sup> mutants indicated the catalytic importance of His-234 for the complex cyclization/rearrangement cascade, especially near the C/D ring junction.<sup>115,</sup>

<sup>117</sup> Notably, these incompletely cyclized products further provided the direct



mechanistic evidence for the formation of the chair-boat 6.6.5-fused tricyclic Markovnikov cationic intermediate and chair-boat-chair 6.6.6.5-fused tetracyclic protosteryl cation, that were assigned provisionally to the ERG7-catalyzed biosynthetic pathway. Moreover, the protosta-12,24-dien-3 $\beta$ -ol is the first isolated and characterized truncated rearrangement product, suggesting the important role of His-234 in stabilizing the cationic intermediate generated in the rearrangement process.<sup>115</sup> From the observation of homology models, the position of the N $\epsilon$ 2 of the His-234 imidazole group was found at a distance of about 4.2 Å and 4.0 Å to the C-13 and C-20 of protosteryl cations, respectively (Figure 5.12). Accordingly, the  $\pi$ -electron rich characteristic of His-234 was suitable and optimal for stabilizing the electron deficient cationic intermediates, especially for the chair-boat 6-6-5 tricyclic Markovnikov C-13 cation as well as the protosteryl C-20 cation. The substitution of histidine residue for other amino acids would affect the steric or electrostatic interaction between the respective cationic intermediates and enzymatic active site, and thus resulted in the generation of different ratio of achilleol A, parkeol, (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien-3 $\beta$ -ol, protosta-12,24-dien-3 $\beta$ -ol and protosta-20,24-dien-3 $\beta$ -ol.<sup>115, 117</sup>

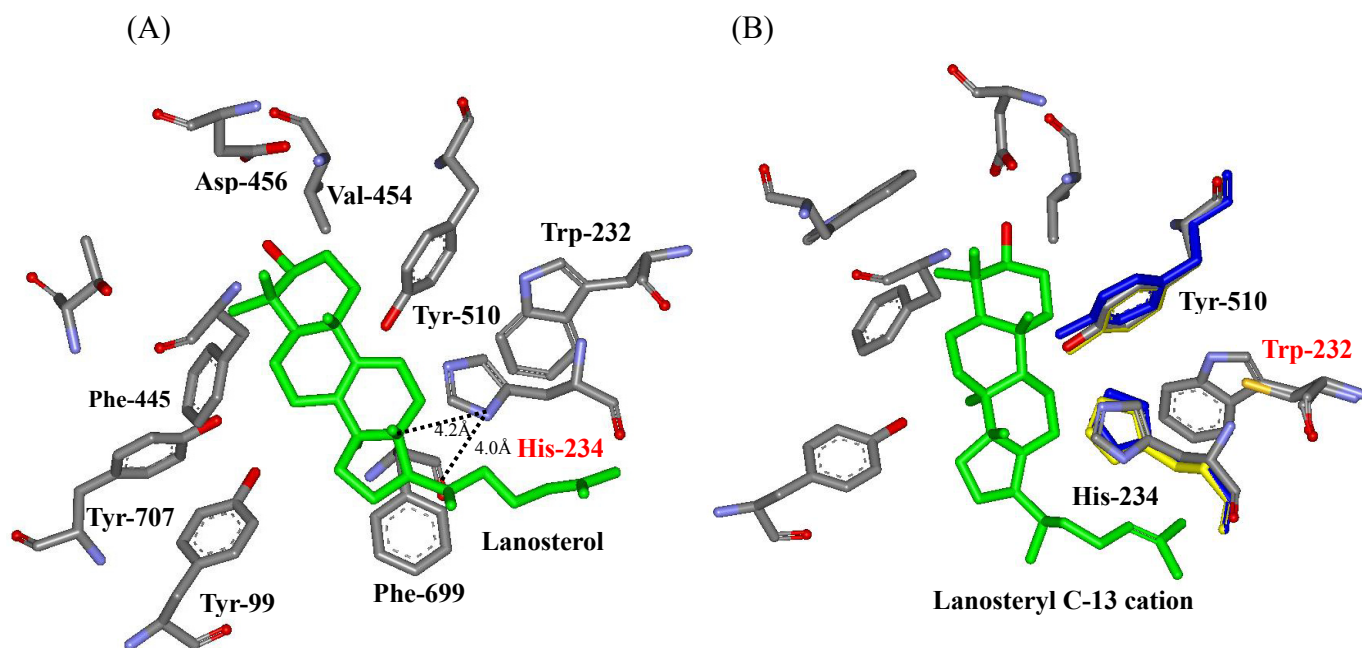


Figure 5.12 (A) Stereo representations of the wild-type *S. cerevisiae* ERG7 homology modeling structure. The putative active-site residues (stick representation) in the active-site cavity are included. Lanosterol is shown in green; while the black dotted lines show the distance between C-13 and C-20 atom of lanosterol with the Nε2 atom of His-234, respectively. (B) Partially superimposed homology modeling structures of wild-type ERG7, ERG7<sup>W232C</sup> and ERG7<sup>W232A</sup> proteins. The original His-234:Tyr-510 catalytic base dyad is shown with stick representation in gray, whereas the yellow color represents the Trp232Cys mutation and blue color represents the Ala substitution, respectively. Moreover, the protosteryl C-13 cation is also included.

On the other hand, the isolation of the same tricyclic (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien-3 $\beta$ -ol from most of *S. cerevisiae* ERG7<sup>Y510X</sup> mutants also implied the altered orientation of His-234:Tyr-510 dyad or spatially changed enzymatic cavity. The impaired catalytic hydrogen-bonding network was further supported from the various homology models of *S. cerevisiae* ERG7<sup>H234X/Y510X</sup> double mutations. First, the hydrogen-bonding interaction between Tyr-510 and His-234 was exchanged for the electrostatic repulsion in the *S. cerevisiae* ERG7<sup>H234Y</sup> mutant, from which the orientation of Tyr-510 was away from the original position and pushed the Tyr-234 residue toward the bottom side of the enzymatic cavity, resulted in the formation of

altered products. In contrast, the *S. cerevisiae* ERG7<sup>H234Y/Y510A</sup> mutant released the electrostatic repulsion and restored the partial active site cavity, which resulted in the production of lanosterol as its sole product (Figure 5.13). Consistently, from the recent experimental results of the *S. cerevisiae* ERG7<sup>H234W/Y510V</sup> mutant that produced the polycyclic products like lanosterol and (13 $\alpha$ H)-isomalabarica-14(26), 17,21-trien-3 $\beta$ -ol; whereas the ERG7<sup>H234W/Y510W</sup> mutant that produced only monocyclic achilleol A, further indicated the spatial importance of Tyr-510 residue in the cyclization reaction. The polycyclic products isolated from the smaller substitution of Tyr-510 in either the *S. cerevisiae* ERG7<sup>H234W/Y510V</sup> mutant or previous *S. cerevisiae* ERG7<sup>H234Y/Y510A</sup> indicated the catalytical importance of His-234 for stabilizing the C-13 and C-20 protosteryl cations respectively. The partial disturbance of the active site residues especially near the C-13 or C-20 of protosteryl cations was observed when the His-234 was changed into other amino acid. The destabilization of the corresponding cationic intermediates resulted in generation of the alternative deprotonated products. On the contrast, the isolation of monocyclic achilleol A from the *S. cerevisiae* ERG7<sup>H234W/Y510W</sup> mutant or *S. cerevisiae* ERG7<sup>Y510W</sup> mutant suggested that the bulky substitution of Tyr-510 dramatically pushed the His-234 away from the original position, blocked the polycyclization progress after the A-ring formation, and resulted in the truncated monocyclic product production. The steric effect of the bulky tryptophan residue, the affected cavity conformation, as well as the coordinative interaction between Tyr-510 and His-234 can be apparently observed from the homology modeling structure of *S. cerevisiae* ERG7 (Figure 5.13).

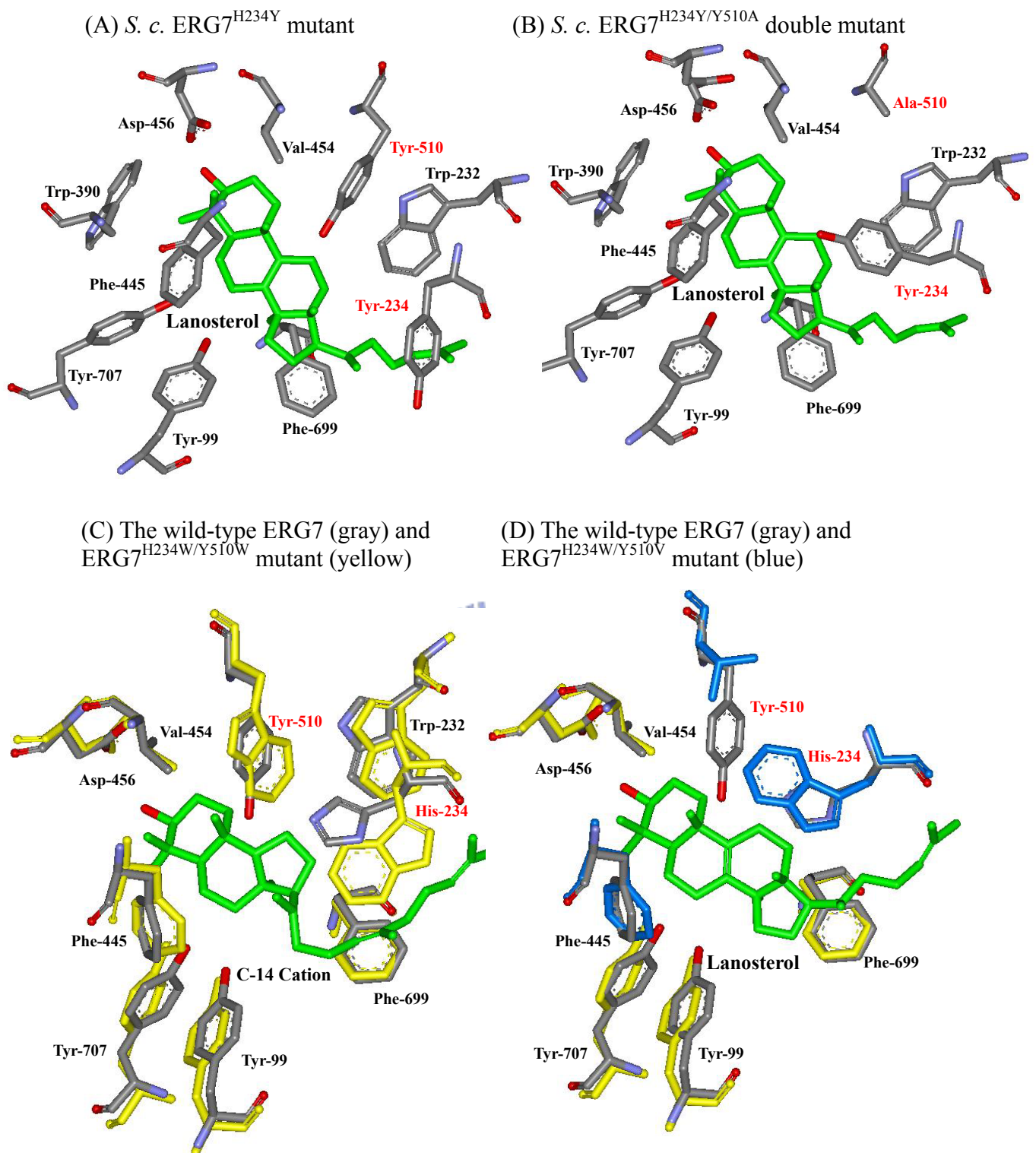


Figure 5.13 The Stereo representations of various *S. cerevisiae* ERG7 mutations modeling structures. (A) ERG7<sup>H234Y</sup> mutant (B) ERG7<sup>H234Y/Y510A</sup> double mutant (C) The partially superimposed ERG7 and ERG7<sup>H234W/Y510W</sup> mutant models (D) The partially superimposed ERG7 and ERG7<sup>H234W/Y510V</sup> mutant models.

On the other hand, the catalytic importance of the Trp-232 might be implied via the indirect interactions from the spatially affected active site residues. From the homology models illustration, the orientation of the hydrogen-bonding network of the His-234:Tyr-510 catalytic base dyad was slightly moved after substituting the bulky Trp-232 residues to other amino acids. Obviously, the distance from His-234 to the protosteryl C-13 cation was changed and slightly close toward the C-11 or C-12 position of lanosteryl cation. Thus, the alternatively deprotonated products, either parkeol or protosta-12,24-dien-3 $\beta$ -ol, were respectively generated in the replacement of the bulky Trp-232 residue to the smaller residues (Figure 5.12).<sup>118</sup>

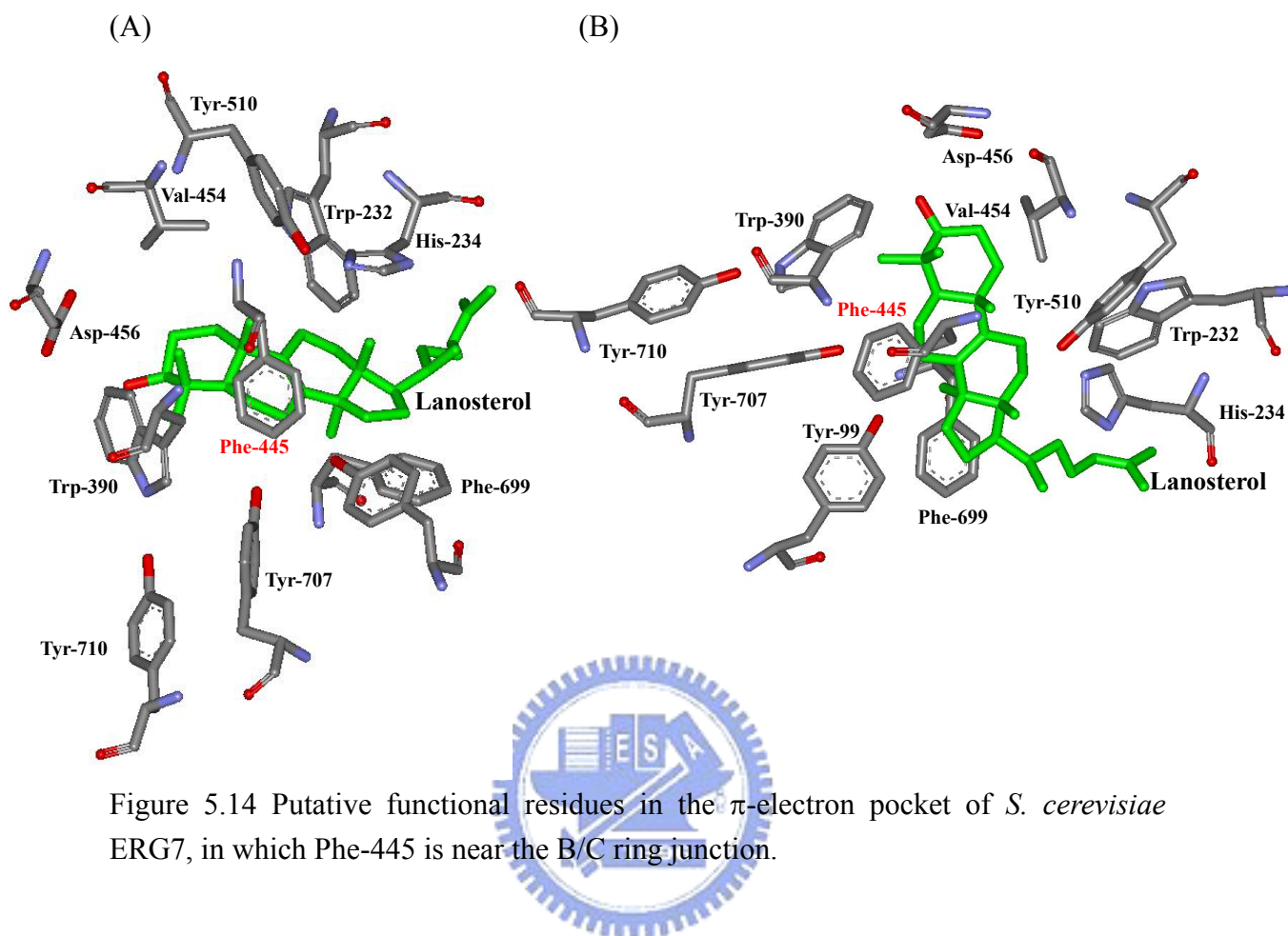
The tricyclic (13 $\alpha$ H)-isomalabarica-14(26),17,21-trien-3 $\beta$ -ol, and three altered deprotonation products: lanosterol, parkeol, and 9 $\beta$ -lanosta-7,24-dien-3 $\beta$ -ol were also observed from the products profile in the *S. cerevisiae* ERG7<sup>F445X</sup> mutants.<sup>119</sup> The careful examination of the mutants' homology models suggested that the substitution of the *S. cerevisiae* ERG7<sup>F445X</sup> significantly influenced the cavity size and steric orientation of the  $\pi$ -electron rich active site pocket. The homology models of *S. cerevisiae* ERG7 revealed that Phe-445 was located spatially proximal to the B/C ring junction and exhibited the near distance from Phe-445 to the C-8/C-14 positions of the lanosteryl cation (Figure 5.14).<sup>119</sup> The Phe-445 also exhibited the definitive interaction with other individual  $\pi$ -electron-rich aromatic residues within the cyclase active site. The distance from the terminal phenyl group to respective amino acid residues are shown in the following: Tyr-99 (3.7 Å), Trp-390 (4.4 Å), Tyr-510 (6.5 Å), Tyr-707 (4.0 Å), and Tyr-710 (6.2 Å). Because these aromatic residues were proposed to provide the high electron density gradient for the equilibrium shift toward the final deprotonation reaction at the C8/C9 cation, substitution of the Phe-445 might partially disturb the transient cation- $\pi$  interaction during the rearrangement step, affect the equilibrium tendency for the final C-8/C-9 proton position, and consequently result in

the formation of three possibly altered deprotonation products.<sup>105, 127</sup> On the other hand, although the relative distances from Phe-445 or His-234 to the C-8 cation of the lanosteryl cation were both approximately 5 Å, the difference of product profiles between Phe-445 and His-234 substituted mutants could also be reasonably examined from the elucidation of the homology models. His-234 exhibited the obvious hydrogen-bonding interaction for the Tyr-510 residue, whereas the hydrogen bond was absent between Phe-445 and Tyr-510 residues due to the long distance about 6.5 Å.<sup>119</sup> Notably, the *S. cerevisiae* ERG7<sup>H234X</sup> site-saturated mutations generated diverse product profiles from monocyclic, tricyclic to tetracyclic products, whereas the only truncated tricyclic and altered deprotonation products were isolated from the *S. cerevisiae* ERG7<sup>F445X</sup> site-saturated mutations. The impacts caused by the Phe-445 substitution on either steric orientation of Tyr-510 position or on the intrinsic Tyr-510:His-234 hydrogen-bonding interaction were not observed in *S. cerevisiae* ERG7<sup>F445X</sup> mutants. The spatially altered active-site cavity structure and the affected electrostatic interaction between the His-234:Tyr-510 hydrogen-bonding network was only observed in the His-234 substitutions.<sup>115, 117</sup>

From the recent site-saturated mutagenesis studies coupled with products characterization in our laboratory, several newly truncated cyclization bicyclic and tricyclic products as well as the truncated rearrangement tetracyclic products, including bicyclic (9*R*,10*S*)-polypoda-8(26),13*E*,17*E*,21-tetraen-3β-ol, tricyclic (13α*H*)-isomalabarica-14*Z*,17*E*,21-triene-3β-ol, (13α*H*)-isomalabarica-14*E*,17*E*,21-triene-3β-ol, tetracyclic protosta-17(20),24-dien-3β-ol, and protosta-13(17),24-dien-3β-ol were respectively isolated from either Tyr-99, Phe-699, or Tyr-707 mutants.<sup>144-146</sup> Although the exact reasons for the production of diverse products in different *S. cerevisiae* ERG7 mutant are complicated, the illustration of the structural basis obtained through the homology modeling might provide suitable insights for



examining the experimental results.



First, the Tyr-99, Phe-699, or Tyr-707 are all located in the bottom half portion of the enzymatic active site cavity, and in the opposite position relative to the previously mentioned His-234, Phe-445, or Tyr-510 amino acid residues (Figure 5.5). According to the observation from homology modeling structure, the Tyr-99 occupied the middle site in the cavity, and interacted with the substrate via its phenolic oxygen side chain. The distance between the phenolic oxygen of Tyr-99 and the tricyclic Markonikov C-14 cation was approximately 4.5 Å, while the Phe-699 residue showed the very similar distance (4.1 Å) to the tetracyclic lanosertyl C-17 cation (Figure 5.15). The short distance allows these two amino acids to stabilize the electron deficient C-14 or C-17 cationic intermediates during the annulation of the C/D ring, respectively.



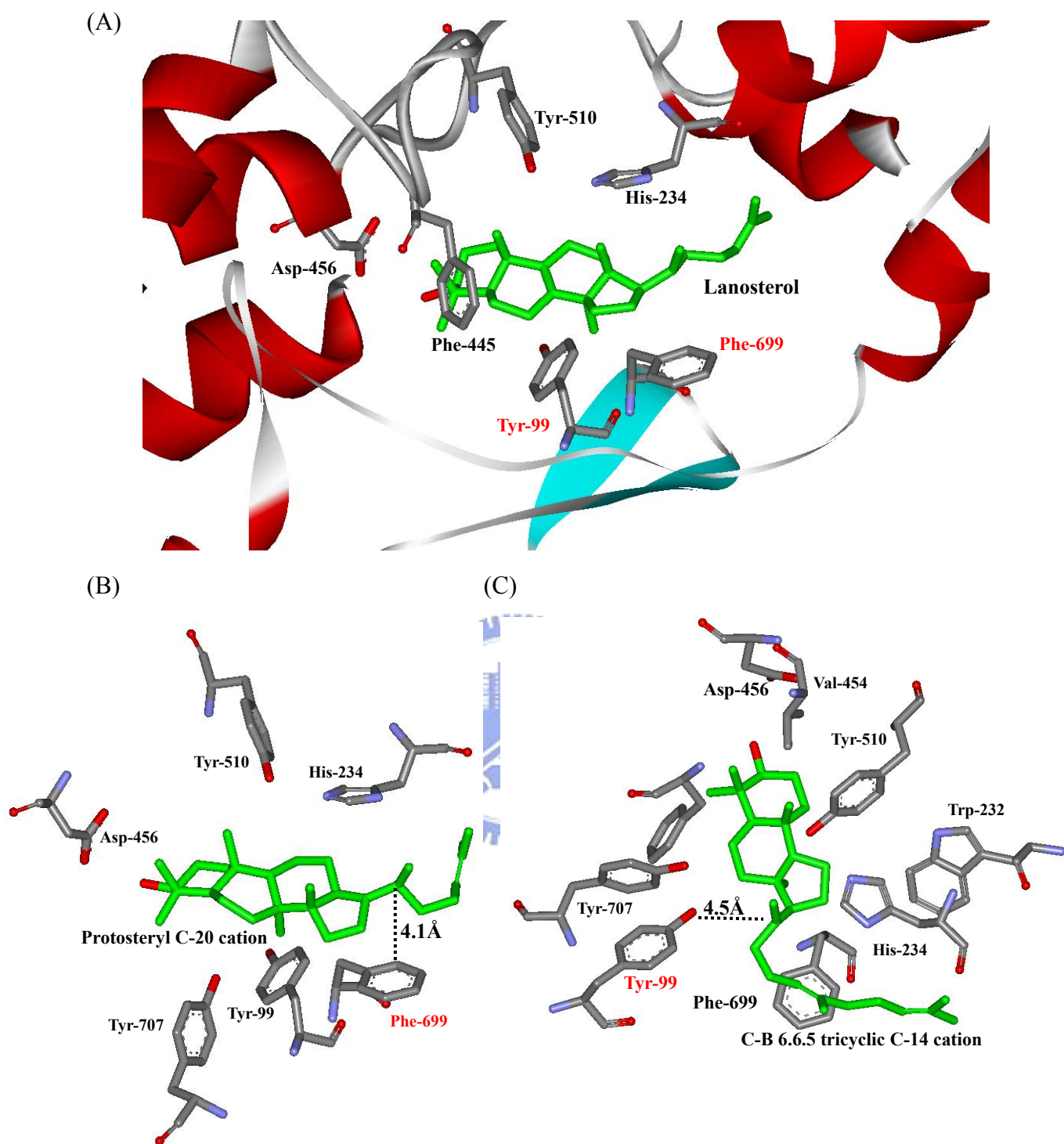
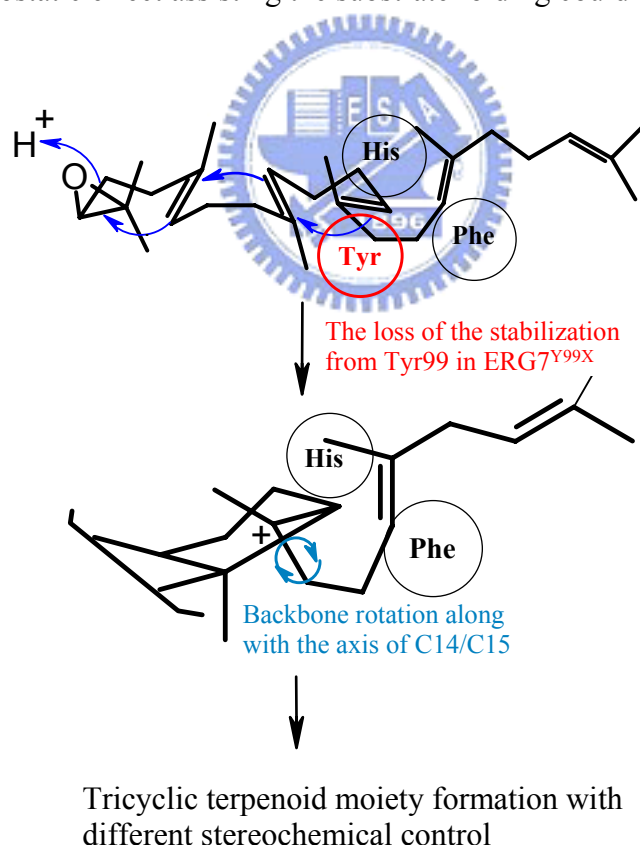


Figure 5.15 (A) Local view of the homology model of *S. cerevisiae* ERG7. The respective catalytically important residues and lanosterol are shown. (B) The distance from C-20 cation to Phe-699 (C) The distance from C-14 cation to Tyr-99

Thus, the simultaneous isolation of different stereochemically (13 $\alpha$ H)-isomalabarica-14Z,17E,21-triene-3 $\beta$ -ol, or (13 $\alpha$ H)-isomalabarica-14E,17E,21-triene-3 $\beta$ -ol from the *S. cerevisiae* ERG7<sup>Y99X</sup> mutants indicated that the affected enzymatic active site allows for the slight rotation of the long carbon side chain along the axis of C-14/C-15 backbone, and resulted in the incorrect deprotonation process (Scheme 5.2). Interestingly, the totally different product profiles between the experimental results of the *S. cerevisiae* ERG7<sup>Y99X</sup> mutants and the theoretical speculation for the corresponding Tyr-98 from human OSC crystal structure might indicate the imperceptible discrepancy or the different spatial neighboring assistance in these two eukaryotic lanosterol cyclases. However, other factors including the steric or electrostatic effect assisting the substrate folding could not be excluded.



Scheme 5.2 The proposed model of Tyr-99 for the stereochemical control.

In addition, the accumulation of tetracyclic protosta-13(17),24-dien-3 $\beta$ -ol and protosta-17(20),24-dien-3 $\beta$ -ol from Phe-699 mutants of *S. cerevisiae* ERG7 clearly indicated the functional role of Phe-699 in stabilizing the lanosteryl C-17 cation intermediate.<sup>146</sup> These two newly discovered tetracyclic structures also demonstrated the occurrence of putative truncated rearranged intermediate states among the backbone rearrangement cascade, from protosteryl cation to the formation of lanosterol. Interestingly, we obtained a Phe-699 mutant that showed a complete product change to generate a sole protosta-13,24-dien-3 $\beta$ -ol.<sup>146</sup> The experimental result of site-saturated mutation in the Phe-699 position clearly exemplified the hypothesis observed from the human OSC crystal structure, which assumed that the Phe-696 occupied the below plane of the substrate molecule and provided the stabilizing interaction for the lanosteryl C-17 cation intermediate.<sup>105</sup>

In other experiments, a long-awaited bicyclic intermediate structure, (9*R*,10*S*)-polypoda-8(26),13*E*,17*E*,21-tetraen-3 $\beta$ -ol, that generated after the ERG7-mediated concerted epoxide ring opening/A-ring cyclization coupled with the arrest at the bicyclic step, was isolated from our recent site-saturated mutants on Tyr-707 residue.<sup>145</sup> The newly trapped bicyclic product together with three altered deprotonation products including lanosterol, parkeol, and 9 $\beta$ -lanosta-7,24-dien-3 $\beta$ -ol, were identified from various *S. cerevisiae* ERG7<sup>Y707X</sup> mutants. The result indicated that the Tyr-707 may play an important role in stabilizing C-8 cation during the formation of second cyclohexyl ring and the final lanosteryl C-8/C-9 cation. Substitutions at this position interrupted the cyclization cascade to generate the bicyclic triterpene alcohol, and interfered with the deprotonation step to generate other three altered deprotonation products. The homology model of *S. cerevisiae* ERG7 model showed that the distance from Tyr-707 to the C-8 cation of bicyclic carbocationic intermediate was 6.2 Å. This long distance causes the elucidation of the

importance of Tyr-707 for the formation of bicyclic structure to be hampered. The crucial importance of Tyr-707 in the formation of the truncated bicyclic chair-boat triterpene alcohol and the contradictory role between Tyr-707 of yeast ERG7 and Tyr-98 of human OSC should be further clarified in the future. However, from the observation of yeast ERG7 homology structure, or human lanosterol synthase structure, the positions of Tyr-707 and/or Tyr-99 are the only functional residues that occupied the suitable orientation to enforce the conformation of secondary cyclohexyl ring into a boat form (Figure 5.15).

### 5.2.5 Quantum mechanism calculation for illustration of the lanosterol synthesis

According to chemical or computational studies, the oxirane ring opening and A-ring formation process were suggested to be concerted.<sup>31, 147</sup> Corey first established that the anchimeric assistance from the nucleophilic C-6/C-7 olefin, as well as the well-positioned acidic group, were necessary for this rapid catalytic initiation reaction.<sup>147</sup> The *ab initio* molecular and density functional theory also provided the evidence to support the concerted mechanism for the oxirane cleavage and A-ring formation.<sup>31</sup> Moreover, the preorganized substrate in the enzymatic active site was also shown to be obligatory, for effectively reducing the distance between tertiary cation and olefin.<sup>148</sup>

In addition, the highly exothermic process was also proposed for the overall oxidosqualene cyclization reaction.<sup>148, 149</sup> About 20 kcal/mol of reaction energy was released with progression from the oxirane ring opening to the B-ring formation, whereas the following annulations proceeded with an exothermicity of about 10 kcal/mol.<sup>148</sup> Although the ring expansion from the C-ring tertiary cation to the secondary cyclohexyl cation was proven to be an endothermic process with about 12 kcal/mol, the better solvation characteristics as well as the stabilization force from the

enzymatic nucleophilic fragment was capable of stabilizing the high energy state.<sup>148</sup> In addition, the D-ring closure was predicted as an exothermic process with 25 kcal/mol.<sup>148</sup>

The free energy surface of each of the carbocationic intermediates and the electrostatic interactions between the respective functional residues and the cationic intermediates were also examined via the molecular dynamics simulation experiments.<sup>149</sup> The importance of the functional residues for lowering the energy barrier was further illustrated from the observation of averaged interatomic distance change between individual carbocationic intermediates and the atom of specific residues. Thus, the enzyme-mediated delicate balance between thermodynamic and kinetic was suggested for controlling the formation of the final tetracyclic lanosterol.<sup>149</sup> Although, the recent theoretically computational calculations proposed some mechanistic views distinct from the experimental evidence, the controversy between computational science and experimental chemistry provided complementary and comparative insights for fully elucidating the mechanisms among cyclization reactions of squalaene/oxidosqualene.<sup>148-151</sup>

In order to elucidate the relationship between the catalytically important amino acids and the different isolated truncated products, the quantum mechanic simulation for the respective carbocationic intermediates was carried out. The relative stability of various truncated intermediates were constructed and optimized at the B3LYP/6-31G\* level of the Gaussian theory. In addition, the parallel calculation of the dammarenyl cation pathway was also performed. The results showed that in neglecting the enzymatic effects, the cyclization cascade to the dammarenyl cation pathway exhibits a progressive reaction energy release. Figure 5.16 shows the partial relative energy profiles derived from quantum mechanic calculations. The similar calculated result of the dammarenyl cation pathway was also observed from previous literature, from

which the A, B, and C-ring formation released most of the reactive energy, whereas the annulation of D-ring produced less energetic changes.<sup>152</sup> In parallel, the thermodynamic tendency for the lanosteryl C-9 cation also existed among the cyclization cascade of the protosteryl cationic pathway. Among the predicted cyclization sequences from the protonated oxidosqualene to the final lanosteryl C-9 cation, about 12 kcal/mol of endothermic process occurred during the C/D ring annulation from the most stable bicyclic C-8 cationic intermediate stage. This energy unfavorable cyclization could be overcome by the following dramatic exothermicity processes, or by the catalytic amino acid from the protein backbone.<sup>148</sup> Thus, the catalytically important residues could be imagined for stabilizing these unstable intermediate stages and for assisting the exact progression until the final product. The critical enzymatic importance via the delicate regulation for the sole product production might contribute the natural specificity among the protosteryl cationic pathway, compared to the diverse product in dammarenyl enzymatic pathway. The amino acid substitution at different spatial positions may alter the electrostatic interaction or the steric control with the special carbocationic intermediate, and further influence the lifetime or the activation energy barrier for the next cyclization among various carbocationic intermediates. Thus, the isolation of the aborted cyclization products from the mutated enzyme could provide the proofs of that the corresponding carbenium ionic intermediate truly existed among the predicted lanosterol biosynthesis pathway. Moreover, these important residues also supported the significance in rigidly or stereochemically control among the predicted cyclization sequences from substrate, various intermediate, to the final product. However, altered cyclization pathways generated in these mutants still could not be excluded. The complete energy profile is listed in the Appendix 5.

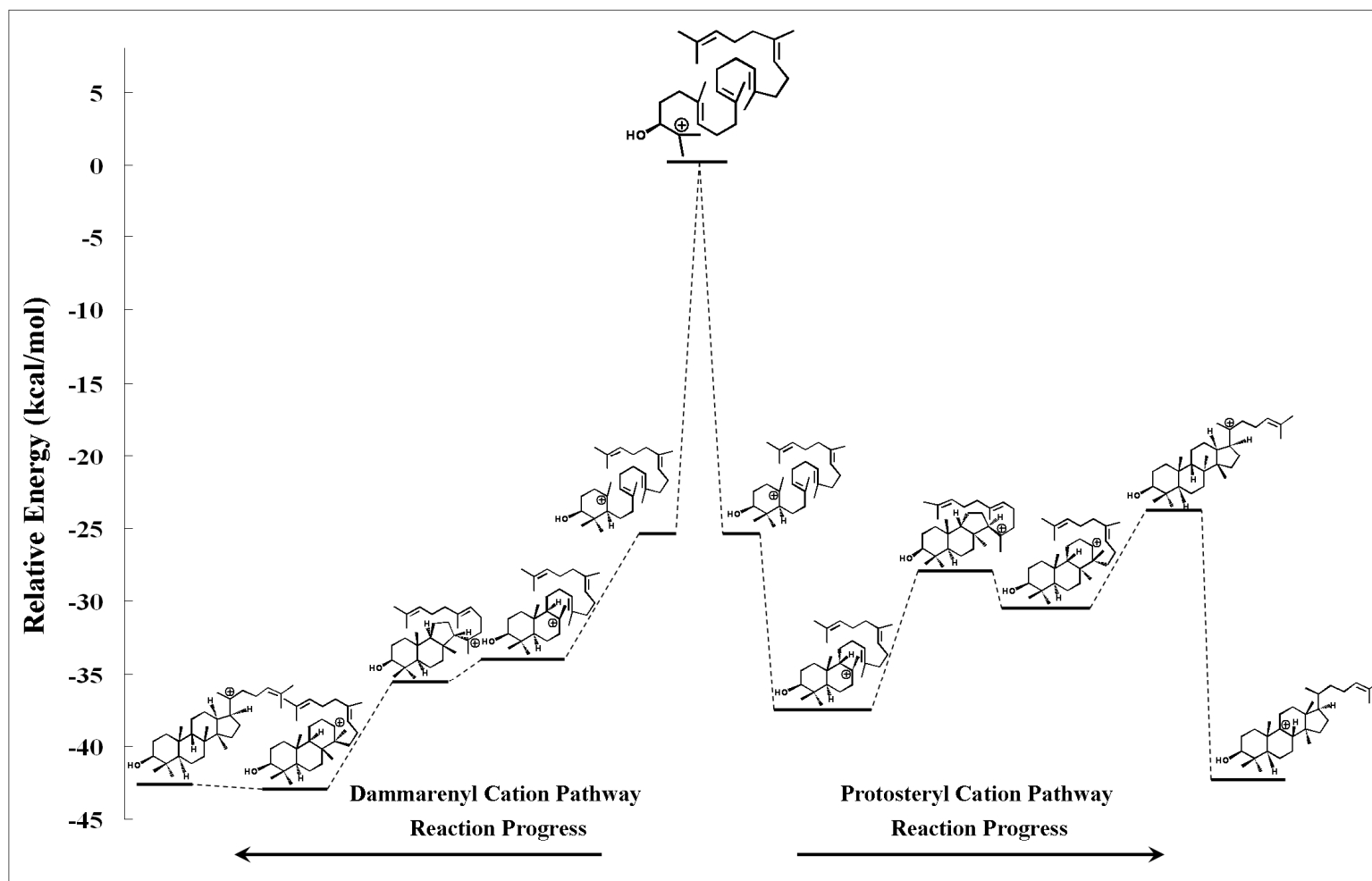


Figure 5.16 The partial relative energy profiles among various intermediate stages, derived from quantum mechanics calculations.



### 5.3 Conclusion and Future Perspective

In conclusion of above results from site-saturated mutations or previous mutagenesis studies on the distinct ERG7 mutations, the functional role of these residues involved in the oxidosqualene cyclization/rearrangement cascade could be elucidated. The proposed enzyme-templated oxidosqualene cyclization/rearrangement cascade as it occurred in the *S. cerevisiae* ERG7 is summarized as follows. First, the Asp-456 coupled with two cysteine residues (Cys-540 and Cys-457) were crucial for the initiation step in ERG7-catalyzed reaction.<sup>30, 56, 105</sup> After the epoxide ring opening and first monocyclic C-10 cation generation, Tyr-510 provided the functional importance for the stabilization of C-6 cationic intermediates.<sup>125</sup> Moreover, Val-454 also contributed the steric effect in assisting the orientation of substrate during the cyclization process.<sup>112</sup> The subsequent cation- $\pi$  annulation generated the chair-boat bicyclic C-8 cation and the chair-boat 6.6.5 Markovnikov tricyclic C-14 cation were stabilized by Tyr-707, Tyr-99, and His-234 residues, respectively.<sup>115, 117, 144, 145</sup> Moreover, His-234 and Phe-699 further provided the highly electronic propriety to stabilize the D-ring closure protosteryl C-20 cation as well as the first backbone rearranged protosta C-17 cation.<sup>115, 117, 146</sup> The Trp-232, the neighboring residue of His-234, showed its unequivocal importance in the rearrangement process, and also exhibits the influence for the coordinative action between His-234 and Tyr-510.<sup>118</sup> In addition to stabilization of the respective carbocationic intermediate during the different stage of cyclization process, His-234 and Tyr-510 further displayed the crucial role to guide the final deprotonation reaction.<sup>115, 117, 125</sup> Moreover, Phe-445 influenced both the C-ring formation and the deprotonation step, distinct from that of His-234 was suggested according the structural insight.<sup>119</sup> Additionally, the product specificity-determining residues, Thr-384 was also investigated previously.<sup>114</sup>

The concomitant isolation of diverse postulated mechanistic products from distinct ERG7 mutants revealed the importance of these residues in controlling the reaction of oxidosqualene cyclization/rearrangement cascade. In addition, these site-directed/saturated mutation experiments help us to elucidate the structure-function relationship between substrate and cyclase, reflect that how minor changes can alter product specificity, and further highlight the potential for increasing the diversity of triterpene skeleton through molecular engineering of cyclase enzyme family. The subtle size or electrostatic changes of active site environment that influence the structure-function relationship of cyclase enzymes could only be thoroughly analyzed by the utilization of site-saturated mutagenesis with all proteinogenic amino acid substitutions.

There are three possible directions that could be considered for the future work in order to better understand the oxidosqualene-lanosterol cyclase-mediated reaction, which is the most remarkable and ingenious biotransformation in the cholesterol biosynthesis pathway.

#### 1. Mechanism of OSC-catalyzed cyclization

Data obtained from site-directed/saturated mutagenesis coupled with random mutagenesis experiments have indicated that some functional residues participate in the complex cyclization reaction and are directly responsible for required critical steps. Moreover, several truncated cyclization/rearrangement structures have been identified from different ERG7 mutants, further implicating that these hypothetically postulated intermediate states may occur among the overall cyclization cascade. However, many questions still remain, particularly concerning the enzymatic steric control involved in the two possible types of pathways for cyclization (*i.e.*, protosteryl cationic or dammarenyl cationic pathway). The steric control for the unique B-ring conformation of oxidosqualene cyclization is the most significant step for the species-dependent

product specificity, because it located at the branch point between the sterol or triterpene biosynthetic pathway. It is fascinating to study how cyclases precisely control the substrate conformation via “chair-boat-chair” conformation for the lanosterol or cycloartenol synthesis, or via the reversed “all-chair” conformation for the production of amyirin or lupenol structure. In spite of the molecular biological investigations that have been undertaken to elucidate amino acid residues critical for the cyclization pathways, the functional residues involved in steric-determination of the cyclohexyl B-ring have never been discovered. Tyr-707 of *S. cerevisiae* ERG7 has showed its influence on the formation of bicyclic structure. The corresponding residues or the spatially neighboring positions in the dammarenyl cation-biosynthetic synthase (e.g.,  $\beta$ -amyirin synthases or lupenol synthases) might provide a chance for altering the structures of the products from the scaffold of “all-chair” form into the scaffold with “chair-boat-chair” form. Moreover, approaches utilizing directed evolution coupled with the oxidosqualene lanosterol cyclase (*erg7*) deficient yeast strain also provides an opportunity to select mutated  $\beta$ -amyirin synthases gene or lupenol synthases gene possessing the lanosterol-synthetic ability. Better understanding the molecular basis of cyclization mechanism, especially in the B-ring conformational distinction, the structural requirements or the natural evolutionary relationship between sterol and triterpene biosynthesis could thus be elucidated clearly.

## 2. Structural Biology

The first crystal structure of human oxidosqualene-lanosterol cyclase has provided valuable structural insights for illustrating the molecular interaction between lanosterol and the enzymatic active-site. Bioinformatics tools also have assisted in creating plausible homology models based on this crystal structure to elucidate the importance of critical amino acids in the *S. cerevisiae* ERG7-catalyzed cyclization.

However, recent site-saturated mutagenesis data and the analysis of diverse products generated from various *S. cerevisiae* ERG7 mutants revealed a contradictory situation between human oxidosqualene-lanosterol cyclase crystal structure and *S. cerevisiae* ERG7 homology models. For example, mutation on Tyr-99 of *S. cerevisiae* ERG7, the corresponding residues of Tyr-98 (hypothesized as sterically important) in human oxidosqualene-lanosterol cyclase, produced the truncated tricyclic structures but not the putative bicyclic structure. On the other hand, the Tyr-707 residue (which corresponds to the Tyr-704 in human oxidosqualene-lanosterol cyclase) unexpectedly isolated the bicyclic intermediates from its mutated cyclase. These data suggested that a slightly altered orientation might exist between the active site cavity of human oxidosqualene-lanosterol cyclase and *S. cerevisiae* ERG7. There were repeated observations of multiple products arising from the different ERG7 mutants. In addition, the mutated cyclase produced the altered products with the deprotonation position, which is unexpectedly distinct from that of the mutated site, implying other influenced amino acids in these single point mutations. In order to understand the role of functional residues assisting in the substrate binding, cationic intermediate stabilization, and the coordinative interaction with other amino acids in an in-depth way, the determination of three-dimensional structures of these mutated cyclases are necessary. The *pichia pastoris* expression system has been used for elucidating the crystal structure of human oxidosqualene lanosterol cyclase. This expression system might be a suitable candidate system for expression and crystallization of different *S. cerevisiae* ERG7 mutants. Moreover, the isolated partially cyclized products might be used for the co-crystallization and detailed understanding of the molecular interaction between the enzymatic active site and the various products.

### 3. The Cyclase Enzyme Plasticity

Mutagenesis studies show that the variation of cyclization reactions can be attained by only small modification of the enzymatic active site. In addition, the sequences comparison and the homology models also suggested that some specific domains in the enzymatic active site might be responsible for the product specificity. Interestingly, various mutated cyclases have been examined for their altered product-specificity from the original cyclase into another cyclase. For example, the *S. cerevisiae* ERG7<sup>H234S</sup> or *S. cerevisiae* ERG7<sup>H234T</sup> completely altered the product-specificity from original lanosterol synthase into either parkeol synthase or protosta-12,24-dien-3 $\beta$ -ol synthase, whereas *A. thaliana* CAS1<sup>H477N/1481V</sup> exhibited the nearly one-hundred percent product conversion from cycloartenol into lanosterol.<sup>117, 153</sup> Moreover, the functional genomic studies on the putative OSC homologous gene from *A. thaliana* genome have characterized many cyclase proteins with novel catalytic functions.<sup>113, 154-160</sup> Recent articles have introduced the detailed reaction mechanism and molecular origin of structural diversity concerning the newly discovered oxidosqualene cyclases in the plant.<sup>157, 159, 161</sup> Careful elucidation of sequence alignment of these cyclases, and comparison with other function-known cyclases might provide valuable information for the catalytic distinctions among the cyclase-catalyzed reactions. Thus, the different artificial cyclases with newly catalytic activity or novel product specificity could theoretically be generated via diverse genetic selections or molecular evolution methods. Moreover, the rationally designed cyclase enzyme could also be obtained by the structure-based alternation after the elegant analysis of the product specificity-determining catalytic residues, which will be revealed from the array of three-dimensional protein structures in the future.

## ***Chapter 6***

# ***Purification, Tandem Mass Characterization, and Molecular Cloning of Oxidosqualene-Lanosterol Cyclase Enzyme from Bovine Liver***

### 6.1 Summary

The oxidosqualene-lanosterol cyclase (OSC) from bovine liver has been successfully purified to homogeneity by using ultracentrifugation, Q-Sepharose, hydroxyapatite, and HiTrap heparin chromatographies. The detergent requirement for its activity in the microsomal fraction was also determined. The purified OSC cyclase displayed a single band with the molecular weight of ~70 and ~140 kDa on the silver-stained SDS-PAGE and the Coomassie-stained Native-PAGE, respectively. Peptide mapping coupled with tandem mass spectrometric determination identified three peptide fragments, ILGVGPDDPDLVR, LSAEEGPLVQSLR, and NPDGGFATYETK, which are highly homologous to human, rat, and mouse OSCs. After successful purification and tandem mass characterization of bovine liver OSC, the encoded gene was determined via PCR based strategy. The deduced amino acid sequence showed >80% identity to the other three mammalian lanosterol cyclases. Genetic complementation of bovine liver OSC gene into a yeast *erg7* knockout strain, CBY57, demonstrated its functional expression and enzymatic activity. The inhibitor binding site of bovine liver OSC was probed with a potent Ro48-8071 inhibitor and determined via LC/MS/MS. However, due to the low peptide coverage, the precise inhibitor binding site still could not be identified at this time. A monoclonal antibody

was also produced and exhibited its activity against the purified OSC via the rational design of an immunogenic oligopeptide antigen based on the homology modeling structure of cyclase.

## 6.2 Introduction

As previously described in the Chapter 1, the purification of oxidosqualene cyclase has always met with difficulties in obtaining the functionally active and soluble enzyme. These difficulties were mainly caused by the detergent- or salt-dependence of cyclase activity. Moreover, due to the instability of the solubilized enzyme, the acquisition of a large amount of purified cyclase protein for the crystal structure determination remains a challenge.<sup>2</sup> Even though the purification of oxidosqualene-lanosterol cyclase has been hampered, some OSCs have been purified to homogeneity. After successful purification of OSCs, the amino acid sequences could be obtained via different methodologies. Recently, numerous cyclase genes were successfully cloned and sequenced, made it is possible to deduce the primary amino acid sequence analysis of OSC. Comparison of amino acids sequence among different eukaryotic oxidosqualene cyclases and prokaryotic squalene cyclases revealed the moderate overall sequence identity and several highly conserved motifs. These results suggested that the cyclase enzymes might maintain the general structure from their common ancestral protein, but alter the critical functional residues in the specific position for developing their diverse products. Thus, from the concept of diverse evolution, understanding the primary amino acid sequence conservation in OSCs is a prerequisite for the future investigation.

On the other hand, photoaffinity labeling is a technique to study the interaction between two different molecules. It has been used broadly in different biological



fields during the past several years. After irradiation reaction occurred, the covalent bond was formed between two molecules. This covalent labeling technology has become a popular tool to investigate the protein-ligand, protein-peptide, or protein-protein interactions in the proteomics studies recently. One of the features in this technique is that the covalent linkage between photoreactive ligand and its binding protein remains even under the denaturing conditions. These covalent binding information could be elucidated by different levels of analysis.<sup>162</sup> In the protein level, SDS-PAGE or Native-PAGE is used to identify the intact proteins, which have been covalent labeled. After enzymatic or chemical degradation, the labeled fragments could be isolated and identified via different liquid chromatographies.<sup>162</sup> The LC/MS/MS instrument provides a very powerful tool to determine the exact photolabeled amino acid.<sup>163</sup>

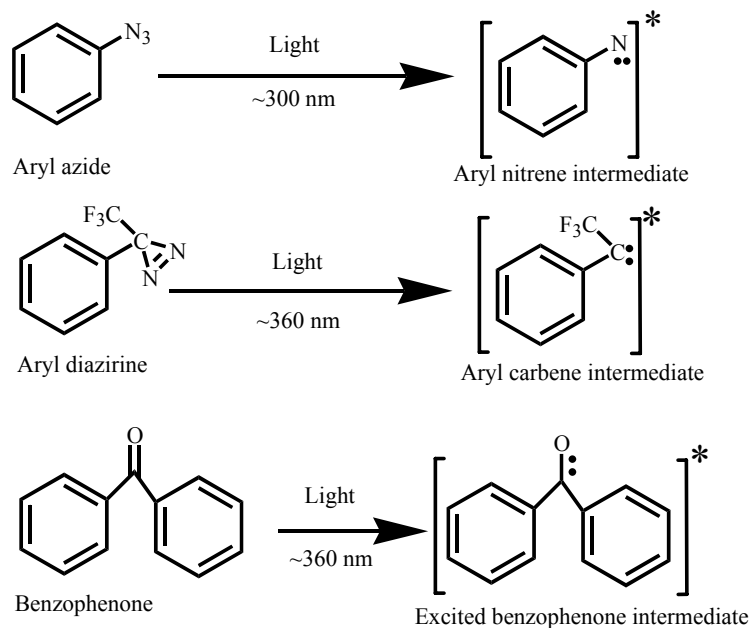
In contrast, although the X-ray crystallography and NMR spectroscopic methods are the most powerful tools to study the structure of the ligand binding, the photoaffinity labeling is thought to be independent from above two approaches and provides another evidence for the ligand binding mode.<sup>164-166</sup> The conformation alteration of the binding ligand might happen during crystallization process from the soluble phase to solid phase; or the binding mode of ligand is absolutely different in these two phases. Consequently, the photoaffinity labeling could fit the requirement complementarily for the structural elucidation of protein-ligand binding events.

Various criteria for photoaffinity probes design were discussed recently.<sup>167</sup> First, it should be stable under the ambient light; second, a photogenerated long-life excitation state should be formed for the covalent linkage reaction; third, a single unambiguous covalent adduct should be generated once; fourth, the covalent linkage products should be avoided for the solvent trapping or releasing. Additionally, the activation wavelength for the probe should be longer than 300 nm to avoid damage to

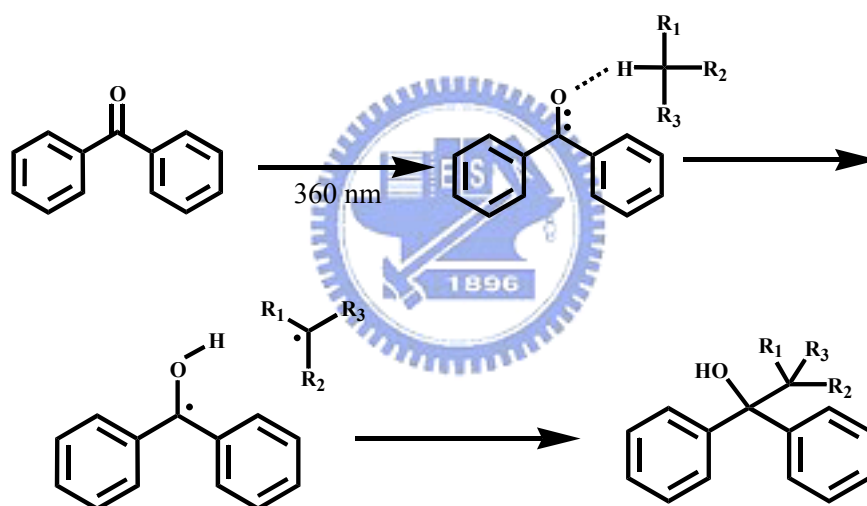
protein. There are three kinds of photoreactive compounds, including aryl azides, aryl diazirine, and benzophenone derivatives extensively used in chemical biology. Their chemical structure and photoinduced mechanism are shown in Scheme 6.1. Among these photoaffinity probes, benzophenone (BP) derivatives are most stable, and the carbonyl oxygen of BPs react preferentially with the neighboring carbon atom within the distance of 3.1Å even in the presence of solvent. This highly efficient covalent modification and the site-specific labeling property of BPs, let it be applied in many biochemical systems.<sup>167, 168</sup> The mechanism of the covalent bond formation between carbon atom and BPs contain three steps: the triple state excitation, the proton abstraction, and the radical recombination (Scheme 6.2).<sup>168</sup>

The Hoffmann-La Roche group has developed a series of BP-containing orally active cholesterol-lowering compounds targeting at OSC. Among them, Ro48-8071 is the most potent OSC (SHC) inhibitor.<sup>169</sup> The tritium-labeled Ro48-8071 was then synthesized and specifically labeled with bacterial SHC or rat liver OSC in either purified or crude extract protein. Competitive displacement experiment also showed that Ro48-8071 might bind in the same location with other two mechanism-based cyclase inhibitors or one nonterpenoid inhibitor.<sup>169, 170</sup> Moreover, the photoaffinity experiment suggested that Ro48-8071 binds at the junction between the central active cavity and substrate entry channel, and then hinders the enzymatic substrate uptaking ability.<sup>171</sup> On the other hands, co-crystallized structure containing Ro48-8071 in SHC or OSC were also established recently.<sup>89, 105</sup> Interestingly, the binding site of Ro48-8071 in these two structures are very different from the previous photoaffinity labeling experiments. The structural conformations were also contradicted with the well-known noncompetitive inhibition mechanism of Ro48-8071. Due to the difference between these experiment results and the overall difference in the solid or liquid phase system, some more extensive experiments are need to clarify these

divergences.



Scheme 6.1 Three major photoreactive groups.



Scheme 6.2 Formation of covalent adducts from photoexcited benzophenone group.

Therefore, the above described importance or puzzlements, involved in the studies of oxidosqualene-lanosterol synthase, provoked us to purify the oxidosqualene-lanosterol synthase from bovine liver and characterize its catalytic activity in the native environment. The peptide mapping coupled with tandem mass spectrometric determination could provide a tool for identifying the *N*-terminal amino acid or internal amino acid sequence. The information obtained from the peptide sequencing and mass characterization will be used to clone the OSC gene from the

bovine liver cDNA library. After cloning and sequencing of the OSC gene, the cyclase activity is confirmed by using the genetic complementation in a yeast *erg7* knockout strain. The sequence comparison will also be performed among the mammalian OSC genes. Moreover, the respective monoclonal antibody will also be developed via the rational designed oligopeptide antigene. In the last part of this chapter, the photoaffinity labeling experiment coupled with the mass spectrometry will be used to further explore the Ro48-8071 binding mode.

## 6.3 Results and Discussion

### 6.3.1 Purification of bovine liver OSC

To date, no more than five oxidosqualene cyclases, mainly cycloartenol synthase and  $\beta$ -amyrin synthase from higher plants, have been purified to homogeneity (Table 6.1). In order to obtain the homologous lanosterol cyclase from the mammalian source, we try to purify the oxidosqualene-lanosterol cyclase from bovine liver. Using knowledge from purification of oxidosqualene-lanosterol cyclase in different species, microsomes from bovine liver homogenates were first treated with various concentrations of Triton X-100. The specific activity of the starting microsomal suspension was apparent in the buffer containing 0.5% (w/v) Triton X-100. Therefore, the concentration of detergent was maintained at this level for subsequent purification steps.

Following recovery of OSC from the microsomal fraction, the crude enzyme solution was subjected to chromatographic separation using Q-Sepharose Fast Flow, hydroxyapatite, and HiTrap heparin columns. In each chromatographic separation, the cyclase enzymatic activity was determined by a standard procedure, and the cyclase-containing fractions were collected for the next column chromatography. The

detailed purification procedure and the cyclase activity assay were described in the experimental section.

In brief, the enzyme adsorbed on the Q-Sepharose Fast Flow column was eluted with IEB buffer containing 20, 50, 70, and 100 mM potassium chloride, and the cyclase activity appeared in fractions containing 70 mM salt. Next, the cyclase containing fractions were applied for hydroxyapatite chromatography fractionation. Interestingly, the cyclase was not retained on hydroxyapatite resin, and exhibited the activity in the flow-through fraction. However, the hydroxyapatite, a mineral material with the formula  $\text{Ca}_5(\text{PO}_4)_3(\text{OH})$ , effectively removed other impurities that occupied abundantly in previous Q-Sepharose fractions. Finally, oxidosqualene-lanosterol cyclase was substantially purified by HiTrap heparin chromatography via elution with HB buffer containing 50 mM potassium chloride. The hydroxyapatite and heparin steps significantly enhanced the fold of purification, yielding 728-fold enrichment in specific activity.

Electrophoresis of the fractions after the different purification steps revealed a single protein band with a molecular weight of about 70 kDa, as visualized by silver stained sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). A single band of about 140 kDa also was revealed on the non-denaturing polyacrylamide gel electrophoresis (PAGE).

The purified homogeneous enzyme was stable for up to 6 months at  $-80^\circ\text{C}$  in the presence of Triton X-100, DTT, and phosphate buffer. Interestingly, the HiTrap heparin chromatography, which was applied to purify oxidosqualene cyclase for the first time, provided an effective method for purifying multi-milligram quantities of apparently homogeneous mammalian oxidosqualene cyclases (Figure 6.1 and Table 6.2).

Table 6.1 Purification and properties of 2,3-oxidosqualene cyclase from vertebrates, higher plants, and yeast.<sup>64</sup>

Source	Molecular weight (kDa)	Purification factor (from microsomes)	Specific activity
<b>Vertebrates</b>			
Pig liver		<b>138</b>	<b>256<sup>a</sup></b>
Pig liver	<b>75</b>	<b>441</b>	<b>2643<sup>b</sup></b>
Rat liver	<b>78</b>		
Dog liver	<b>73</b>		
Human liver	<b>73</b>		
Rat liver	<b>75</b>	<b>1863</b>	<b>436<sup>b</sup></b>
Rat liver	<b>65</b>	<b>128<sup>c</sup></b>	<b>58<sup>d</sup></b>
<b>Yeast</b>			
<i>S.cerevisiae</i>		<b>142</b>	<b>84<sup>a</sup></b>
<i>S.cerevisiae</i>		<b>112</b>	<b>40<sup>c</sup></b>
<i>S.cerevisiae</i>	<b>26</b>	<b>160</b>	<b>4400<sup>f</sup></b>
<b>High Plant</b>			
<b>(<i>R. japonica</i>)</b>			
Cycloartenol synthase	<b>54</b>	<b>139</b>	<b>16<sup>b</sup></b>
$\beta$ -amyrin synthase	<b>28</b>	<b>541</b>	<b>41<sup>b</sup></b>
<b>(Pea)</b>			
Cycloartenol synthase	<b>55</b>	<b>235</b>	<b>167<sup>g</sup></b>
$\beta$ -amyrin synthase	<b>35</b>	<b>1072</b>	<b>28<sup>b</sup></b>

a: nmol/h/mg ; b: mkat/kg; c: from homogenate; d: pmol/min/mg;  
e: mg protein necessary for 20% conversion ratio; f: units/mg; g: pkat/mg

Table 6.2 The purification of oxidosqualene-lanosterol cyclase from bovine liver

Fraction	Total protein (mg)	Total activity (nmol/10min)	Yield (%)	Specific activity (pmol/min/mg)	Purification fold
Microsomal membranes	<b>2417</b>	<b>348</b>	<b>100</b>	<b>2.4</b>	<b>1</b>
Q-sepharose FF	<b>328</b>	<b>228</b>	<b>66</b>	<b>22</b>	<b>9</b>
Hydroxyapatite	<b>105</b>	<b>33.5</b>	<b>10</b>	<b>165</b>	<b>69</b>
HiTrap Heparin	<b>9</b>	<b>16.6</b>	<b>5</b>	<b>1747</b>	<b>728</b>

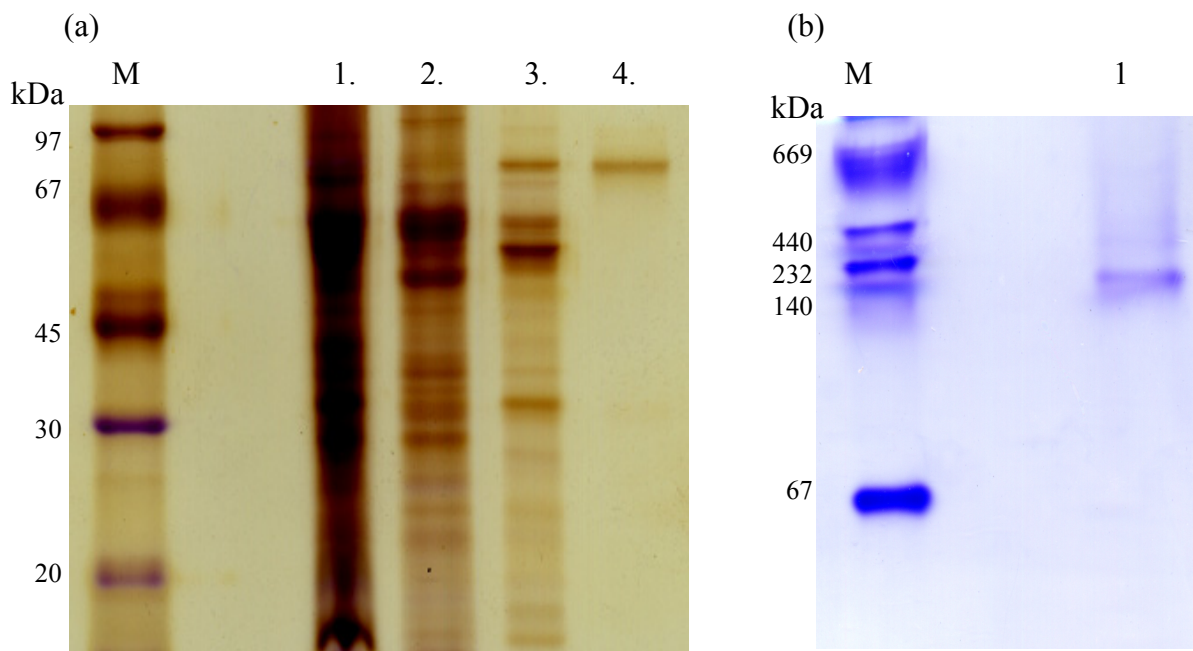


Figure 6.1 (a) Silver-stained SDS-PAGE gel of bovine liver OSC obtained from column chromatography fractions. Lane 1, marker proteins: phosphorylase b (97 kDa), bovine serum albumin (67 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), and trypsin inhibitor (20 kDa); Lane 2, solubilized protein; Lane 3, Q-Sepharose Fast Flow purified protein; Lane 4, hydroxyapatite purified protein; and Lane 5, HiTrap heparin purified protein. (b) Coomassie-stained non-denaturing PAGE gel of bovine liver OSC obtained from HiTrap heparin chromatography. Marker proteins: thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), lactate dehydrogenase (140 kDa), and albumin (67 kDa); Lane 1, HiTrap heparin purified protein.



### 6.3.2 Identification of OSC by tandem mass spectrometry

The purified protein then was subjected to the *N*-terminal sequencing and internal amino acid determination to confirm the identity of the purified protein. Multiple attempts to obtain the *N*-terminal sequence of OSC were unsuccessful, suggesting that the *N*-terminus of the protein may be blocked. For internal amino acid determination, the SDS-PAGE band corresponding to OSC was subjected to in-gel trypsin digestion as described by Rosenfeld *et al.*<sup>172</sup> After digestion, the resulting peptide mixtures were analyzed by nanoscale capillary LC/MS/MS. Among these peptide fragments, one tandem mass spectrum of a doubly charged precursor *m/e* 683.45 is shown in Figure 6.2. Following a database search for the mass spectrum of the peptide and its fragment ions, one unique hit matching the human OSC peptide sequence was identified.<sup>173</sup> The sequence was determined to be <sup>163</sup>ILGVGPDDPDLVR<sup>175</sup> of human OSC. Another peptide, whose sequence was determined to be <sup>495</sup>NPDGGFATYETK<sup>506</sup> from LC/MS/MS analysis, also matched the human OSC protein. Moreover, a “MS-BLAST search” for the homologous proteins identified a third sequence, <sup>248</sup>LSAEEGPLVQSLR<sup>260</sup>. This BLAST-based searching methodology would identify all high-scoring pair (HSP)-regions, the high local sequence similarity regions, between individual peptides in the query and a protein sequence from the database entry.<sup>174</sup> In summary, the results showed that three peptide fragments from in-gel digestion of purified bovine liver OSC were homologous to OSCs from human (EC 5.4.99.7), rat (EC 5.4.99.7), and mouse (EC 5.4.99.8). These data provided evidence that the identity of the purified protein was bovine liver OSC. Individual HSP-aligned sequences of high similarity are shown in Figure 6.3

a	<u>86.10</u>	<u>199.18</u>	256.20	<u>355.27</u>	<u>412.29</u>	509.35	624.37	739.40	836.45	951.48	1064.56	1163.63	1319.73
b	<u>114.09</u>	<u>227.18</u>	<u>284.20</u>	383.27	<u>440.29</u>	<u>537.34</u>	652.37	767.39	864.45	<u>979.47</u>	1092.56	1191.63	1347.73
	Ile	Leu	Gly	Val	Gly	Pro	Asp	Asp	Pro	Asp	Leu	Val	Arg
y	1365.74	<u>1252.65</u>	<u>1139.57</u>	<u>1082.55</u>	<u>983.48</u>	<u>926.46</u>	<u>829.41</u>	<u>714.38</u>	<u>599.35</u>	<u>502.30</u>	<u>387.27</u>	<u>274.19</u>	<u>175.12</u>
z	1348.71	1235.62	1122.54	1065.52	966.45	<u>909.43</u>	812.38	697.35	582.32	485.27	370.24	257.16	158.09

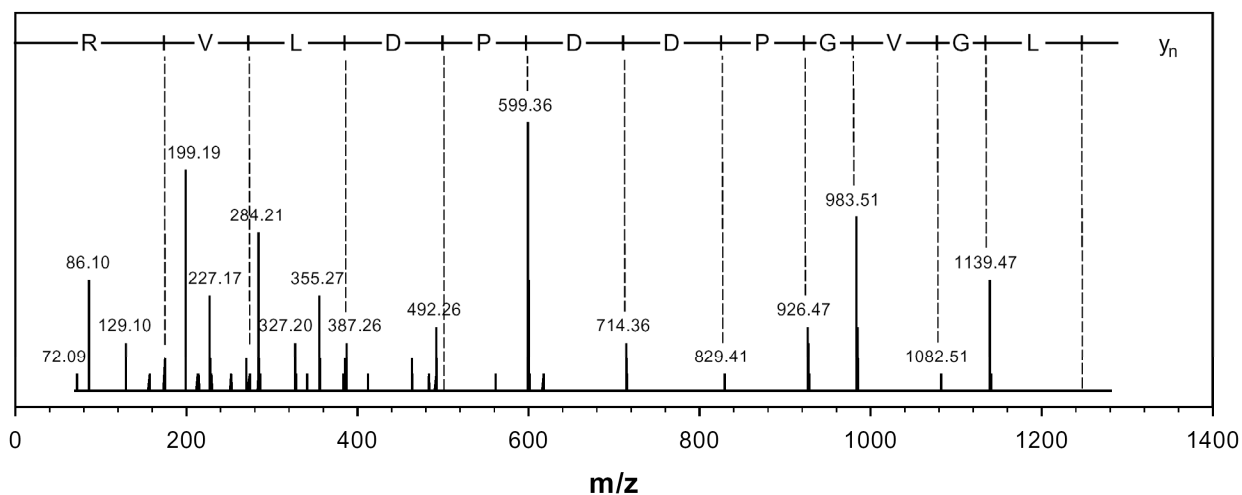


Figure 6.2 Tandem mass spectrum of a doubly charged tryptic peptide at  $m/z$  683.45 from purified bovine liver OSC protein. The complete amino acid sequence, ILGVGPDDPDLVR, was deduced from the mass differences in the y-fragment ion series and is also partly supported by b-ion and a-ion series. The sequence is shown using single letter amino acid symbols. Ions were labeled according to the nomenclature of Roepstorff and Fohlman.<sup>175</sup> Underlined ions indicate the fragment ions observed by the mass spectrometer.



Figure 6.3 Sequence alignments of the trypsin-digested peptides from bovine liver OSC. Three peptide sequences detected by tandem mass spectrometry were homologous to lanosterol synthase (OSC) in human (EC 5.4.99.7), rat (EC 5.4.99.7), and mouse (EC 5.4.99.8). The consensus sequences are boxed in black. These highlighted amino acid residues indicate the similarity of OSCs among different mammalian species.

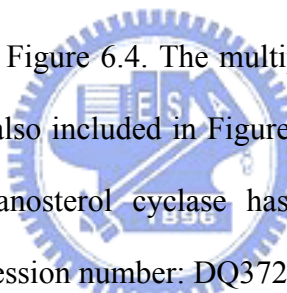
### 6.3.3 Molecular cloning of OSC from bovine liver

Following successfully identified three peptide fragments from purified bovine liver OSC, the gene encoding bovine liver OSC was determined. In addition to three known internal amino acids sequences, we designed the *N*-terminal and *C*-terminal oligonucleotide primers based on the highly identical region of the mammalian sources. One gram of fresh frozen bovine liver was homogenized with mortar and pestle, and the lysed powder was mixed with oligo(dT) cellulose for hybridization. The mRNA library was then eluted from the bead and the concentration of the extracted mRNA was determined as 74 µg/ml. The reverse transcription reaction was then carried out with M-MLV RTase and the transcripts were used as the template for the PCR reaction. After several trails of nested PCR reaction, one of primer pairs (CHC-cDNA-PEP2/CHC-cDNA-PEP3) gave the specific PCR amplicon product with size about 1 kb. The deduced amino acid sequence of this amplicon fragment showed 85% identity to human lanosterol synthase in NCBI protein databank, indicating that it is the portion of cDNA from bovine liver OSC.

The terminal sequence of the bovine liver OSC gene was determined by using the “RACE” method.<sup>176</sup> For the *C*-terminal RACE, two redesigned oligonucleotide primers (CHC-cDNA-Seq1, CHC-cDNA-Seq3), obtained from previous core fragment sequence, and one highly identical *C*-terminal primer, based on mammalian source, were synthesized and applied for the PCR reaction. The deduced amino acid sequence exhibited high sequence homology to the carboxyl terminus of human lanosterol synthase, strongly suggesting that the obtained PCR product is the *C*-terminal region of bovine liver OSC gene.

Moreover, the similar *N*-terminal RACE strategy was used to clone the *N*-terminal fragment of the OSC gene. Several trails of RACE method give no expected PCR product. This unsuccessful amplification might be caused by the lager

size of PCR products or the failed terminal transferase reaction. Another inversed PCR method,<sup>177</sup> which was used for cloning the rat OSC gene, also failed to isolate the expected product. As the cloning of C-terminus region, the highly identical N-terminal primer coupled with two antisense primers were utilized to give the correct product. This PCR product displayed high identity to other known OSCs. Overall, the nucleotide sequence of bovine liver OSC revealed the presence of an ORF with 2,196 bp, encoded a 732 amino acid polypeptide with a molecular mass of 83.2 kDa. The deduced amino acid sequence of OSC exhibited 83%, 82%, and 85% identity to those of the lanosterol synthases from mouse, rat, and human. The theoretical pI of bovine liver OSC was 6.36 from prediction in the *Protein Identification and Analysis Tools in the ExPASy Server*. The nucleotide and deduced amino acid sequences of cDNA from bovine liver OSC is shown in Figure 6.4. The multiple sequence alignment of OSCs among mammalian source is also included in Figure 6.5. The cDNA sequence of the bovine liver oxidosqualene-lanosterol cyclase has successfully submitted to the protein data bank with the accession number: DQ372933.



```

atgaccgagggcacgtgtctgcgtcgaagagggcggaccctataagacagagccagccaca      60
M T E G T C L R R R G G P Y K T E P A T
gacctcagtcgctggcggctcagcaatcaggttgggaggcagacatggacctactctcaa      120
D L S R W R L S N Q V G R Q T W T Y S Q
gaagaagatcctgtccgagagcagctctggcctggaagctcacctcttgggactggataca      180
E E D P V R E Q S G L E A H L L G L D T
aaaagtttctttaaggacttgccgaaagcgcacacagcctgcaggggggctctgaatggg      240
K S F F K D L P K A H T A C R G A L N G
gtgacattttacgctgcactgcagactgaggatgggcattgggagggtgattacggtggc      300
V T F Y A A L Q T E D G H W A G D Y G G
ccactcttctcctgcccaggcctcctgatcacgtgccatgtggcgaacatccctctgccg      360
P L F L L P G L L I T C H V A N I P L P
gccggataccgagaagaaattatacgggtacctgcggtctgtgcagctccccgacggcggc      420
A G Y R E E I I R Y L R S V Q L P D G G
tggggcctgcacattgaggacaagtccacggatatttgggaactgcgcttaactatgtgtct      480
W G L H I E D K S T V F G T A L N Y V S
ttgagaattctgggtgtggggcctgacgatcctgacctggtgagcccggaaccttctt      540
L R I L G V G P D D P D L V R A R N L L
cacaagaaaggtggtgctgtgttcatcccgtcctgggggaagtttggttggctgtcctg      600
H K K G G A V F I P S W G K F W L A V L
aacgtgtacagctgggaaggcctcaatacactgtttccagaaatgtggctgtttcctgac      660
N V Y S W E G L N T L F P E M W L F P D

```

tggatgcccgcacatccctccaccatctggtgccactgtcggcaggtgtacctgcccattg	720
W M P A H P S T I W C H C R Q V Y L P M	
gcctactgctacagcagcggctgagtgccgaggaggcccgctggtccagagcctccgc	780
A Y C Y S T R L S A E E G P L V Q S L R	
caggagctctacctggaggactacagctgcatcgactgggcccgcacaggaacagcgtg	840
Q E L Y L E D Y S C I D W A A H R N S V	
gccccggacgatctgtacacgccgcacagctggctgctccacgtagtatacgccatcctc	900
A P D D L Y T P H S W L L H V V Y A I L	
aacttgtagcagcagccaccacagcaccagcctgcccagtgggccaccagaagctgtat	960
N L Y E R H H S T S L R Q W A T Q K L Y	
gagcacatcgagctgatgaccgttccaccaagtgcacatcgccccgatctcgaaa	1020
E H I A A D D R F T K C I S I G P I S K	
accatcaacatgctcgtgcgctggcatgtagacgggcccagcctctgctgtcttccaggag	1080
T I N M L V R W H V D G P A S A V F Q E	
cacgtgtccaggattcctgactacctctggctgggcccctcgacggcatgaagatgcagggc	1140
H V S R I P D Y L W L G L D G M K M Q G	
accaatggctcacagatctgggacaccgcgtttgccatccaggctttgctggaggcacgt	1200
T N G S Q I W D T A F A I Q A L L E A R	
gcacaacacagggccggagttttggctcctgcctgcggaaggcgcacatgagtatctccggatc	1260
A Q H R P E F W S C L R K A H E Y L R I	
tcacaggtgccagacaatttccccgactaccagaagtactaccgccatgatgagcaagggt	1320
S Q V P D N F P D Y Q K Y Y R H M S K G	
ggcttttagcttcagcagcgtggactgtggctggatcgtggctgactgcacggctgaggcc	1380
G F S F S T L D C G W I V A D C T A E A	
ttgaagtccatcctcctcctgcaggaaaagtgtccctttgtttccaatcatgtcccacga	1440
L K S I L L L Q E K C P F V S N H V P R	
gagcgggcttttgacactgtcgtgtgttggctgagcttgagaaacccccagcggagggttt	1500
E R L F D T V A V L L S L R N P D G G F	
gccacctacgagaccaagcgtgggggacaccttctggagcttctgaaccctccgaggtc	1560
A T Y E T K R G G H L L E L L N P S E V	
ttcggggacatcatgattgactacacgtacgtggagtgcaacctcagcagtgatgcagggc	1620
F G D I M I D Y T Y V E C T S A V M Q A	
ctgaagactttccacaagcagttcccagaccacagggccgggggagatcagggagaccctt	1680
L K T F H K Q F P D H R A G E I R E T L	
gagcagggccttgagttctgtcggcagaagcagagggcctgatggctcctgggaaggtcc	1740
E Q G L Q F C R Q K Q R P D G S W E G S	
tgggggtgtgtgttccagctacggctgcctgggttgggctggaagccttcgctcctgcatgggg	1800
W G V C F T Y G A W F G L E A F A C M G	
cacacctaccataatgggggttgctgtgcagagatctcccggcctgtgacttctctgctg	1860
H T Y H N G V A C A E I S R A C D F L L	
tccccagacatggcagacggaggctggggggaggacttctgagctcctgcaagcagcgtcgc	1920
S R Q M A D G G W G E D F E S C K Q R R	
tatgtgcagagtgcccagctcccagatccacaacacgtgctgggcccctgatggggctgatg	1980
Y V Q S A Q S Q I H N T C W A L M G L M	
gctgtcagggaccctgacgtggcggccctggagagaggagtcagctacctgcttgaagaag	2040
A V R H P D V A A L E R G V S Y L L E K	
cagctgcccattggggactggcctcaggagaacatcagtggggtcttcaacaagtcctgt	2100
Q L P N G D W P Q E N I S G V F N K S C	
gccatcagctacactagctacagaaacgtcttccccatctggaccctcggggcgttctcc	2160
A I S Y T S Y R N V F P I W T L G R F S	
cgtctgcaccctgaccagcccttctgctggccaccctgtga	2199
R L H P D P A L A G H P -	

Figure 6.4 The nucleotide sequence and its deduced amino acid sequence of cDNA from bovine liver OSC





```

R.nLS  QERLYNAVAVLLSMRNSDGGFATYETKRGGYLLELLNPSEVFGDIMIDYTYVECTSAVMQ 540
M.mLS  RERLCDAVDVLLSLRNADGGFATYEKKRGGYLLLELLNPSEVFGDIMIDYTYVECTSAVMQ 540
H.sLS  RERLCDAVAVLLNMRNPDGGFATYETKRGHLLLELLNPSEVFGDIMIDYTYVECTSAVMQ 539
B.tLS  RERLFDTVAVLLSLRNPDGGFATYETKRGHLLLELLNPSEVFGDIMIDYTYVECTSAVMQ 539
      :***  ::*  ***.:** .*****.****:*****
      :***  ::*  ***.:** .*****.****:*****

R.nLS  ALRHFREYFPDHRATESRETlnQGLDFCRKKQRADGSWEGSWGVCFTYGTWFGLEAFACM 600
M.mLS  ALKHFHEHFPDYRAAEVRETLNqGLDFCRRKQRADGSWEGSWGVCFTYGTWFGLEAFACM 600
H.sLS  ALKYFHKRFPPEHRAAEIRETLTQGLEFCRRQQRADGSWEGSWGVCFTYGTWFGLEAFACM 599
B.tLS  ALKTFHKQFPDHRAGEIRETLEQGLQFCRQKQRPDGSWEGSWGVCFTYGAWFGLEAFACM 599
      **:  *::  **::** * ****  ***:***:**:** .*****:*****

R.nLS  GHIYQNRTACA EVAQACHFLLSRQMDGGWGEDFESCEQRRYVQSAGSQVHSTCWALLGL 660
M.mLS  GHTYQDGAACA EVAQACNFLLSQMDGGWGEDFESCEQRRYVQSARSQVHSTCWALMGL 660
H.sLS  GQTYRDGTACA EVSRACDFLLSRQMDGGWGEDFESCEERRYLQSAQSQIHNTCWAMMGL 659
B.tLS  GHTYHNGVACA EISRACDFLLSRQMDGGWGEDFESCKQRRYVQSAQSQIHNTCWALMGL 659
      *:  *::  .****:**:** .****:*****:***:*** **:* .****:***

R.nLS  MAVRHPDISAQERGIRCLLGKQFPNGEWPQENISGVFNKSCAISYTNRYNIFPIWALGRF 720
M.mLS  MAVRHPDITAQERGIRCLLGKQLPNGDWPQENISGVFNKSCAISYTSYRNIFPIWALGRF 720
H.sLS  MAVRHPDIEAQERGVRCLEKQLPNGDWPQENIAGVFNKSCAISYTSYRNIFPIWALGRF 719
B.tLS  MAVRHPDVAALERGVSYLLEKQLPNGDWPQENISGVFNKSCAISYTSYRNVPFIWTLGRF 719
      *****: *  ***:  **  **::***:*****.*****.***:***:***

R.nLS  SSLYPDNTLAGHI 733
M.mLS  SNLYPDNTLAGHI 733
H.sLS  SLYPERALAGHP 732
B.tLS  SRLHPDPALAGHP 732
      *  *:::  :****

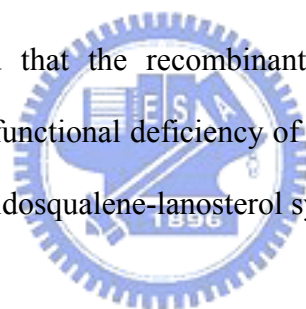
```



Figure 6.5 Multiple sequence alignment of oxidosqualene-lanosterol cyclase (OSC). The sequence shown are *Homo sapiens* OSC (H.sLS), *Mus musculus* OSC (M.mLS), *Rattus norvegicus* OSC (R.nLS), and Bovine liver OSC (B.tLS), respectively. The identical amino acid positions are exhibited above the symbol of “\*”. The highly conserved sequences, at least two of four members, are indicated with the symbol of “.” or “:”.

#### **6.3.4 Genetic complementation of recombinant bovine liver OSC in yeast *erg7* deficient strain**

After the full length of bovine liver OSC cDNA sequence was determined, the functional expression of this gene was subsequently carried out. In order to facilitate the expression of bovine liver OSC in *S. cerevisiae*, the promoter sequence from the *ERG7* gene was fused to the coding sequence of the bovine liver OSC gene, and a complete hybrid *ERG7/b.IOSC* expression plasmid, pCHCRS314BLOSC, was obtained. This hybrid plasmid was checked by restriction enzyme mapping and DNA sequencing analysis. “Plasmid shuffle” method utilized in previous site-directed mutagenesis studies was used to examine the cyclase activity. This pRS314-derived plasmid was next transformed into the TKW14 strain to further confirm the cyclase activity. The results showed that the recombinant gene in pCHCRS314BLOSC plasmid can complement the functional deficiency of lanosterol synthesis, confirming the identity of bovine liver oxidosqualene-lanosterol synthase.



#### **6.3.5 Mass spectrometry determination of the photoaffinity labeling bovine liver OSC with Ro48-8071**

Ro48-8071, a photoactive benzophenone-containing compound, was developed as an orally pharmaceutical agent for lowering serum cholesterol levels in Syrian hamsters, squirrel monkeys, and Göttingen minipigs. Ro48-8071 showed the *in vitro* inhibition for human liver microsomes, rat liver homogenous OSC, and the *in vivo* selective inhibition for the HepG2 cells.<sup>169</sup> In our laboratory, Ro48-8071 also displayed inhibition for bovine liver OSC.<sup>142</sup> The experimental data suggested that the Ro48-8071 might block the OSC activity via a noncompetitive inhibition behavior. However, the controversial results was obtained from the recent crystal structure of SHC.<sup>89</sup> As previously illustrated, the photoactive benzophenone group held the

potential for the cross-linking reaction with its neighboring amino acids in the protein active site.<sup>168</sup> Moreover, the tandem mass spectrum provided a valuable tool to examine the internal amino acid sequence information from protein. In order to understand the inhibition mechanism and to determine the binding site of this potent inhibitor in the mammal OSC, the photoaffinity experiment was carried out.

In order to have sufficient labeled peptide fragments for sequence analysis, the photoaffinity reaction was performed with a threefold molar excess of Ro48-8071. After exposure to UV light, the cyclase inactivation assay was used to confirm the inhibitory ability of Ro48-8071. As expected, there is no observable lanosterol spot in the TLC analysis, whereas the unlabeled protein produced clear product spot. Moreover, the protein bands could be visualized by the coomassie R-250 brilliant blue staining SDS-PAGE, indicating that the photo-labeling process did not dramatically disrupt the enzyme and the inactivated cyclase primarily was caused by the Ro48-8071 inhibition (Figure 6.6). These protein bands were picked for the in-gel digestion and the following LC/MS/MS identification, as previously stated. However, no expected molecular weight change between the unlabeled protein fragments and the photocrossing protein/Ro48-8071 complex was observed in all mass spectra. As shown in Figure 6.7, the total peptide coverage of bovine liver OSC for LC-MS/MS is about 28.5%. The very low coverage is a common challenged phenomenon for analysis of membrane proteins. Because they are not readily soluble in the polar solvents and often undergo aggregation, it is not surprising that analyzing the well known membrane-bound property of OSC makes analysis very difficult. Moreover, the comparison with the result of the previous SHC photoaffinity labeling experiments showed that the predicted binding site of Ro48-8071 is not in the present OSC mass coverage. Hence, the exact binding site of this potent inhibitor in the bovine liver OSC can be further investigated by either increasing the sample concentration to

enhance its peptide coverage or using different combination of chemical and/or enzymatic digestion. For example, a sequential CNBr/trypsin in-gel digestion method coupled with mass spectrometry for analysis of membrane protein analysis was recently published.<sup>178</sup> The critical role of CNBr utilization prior to trypsin cleavage is that CNBr can access the inner hydrophobic sites and cleave the polypeptide chain at the position of methionine residues, generating the smaller peptides for the following photolytic attack of trypsin. This CNBr/trypsin digestion method might provide additional peptides for the purpose of peptide identification.

On the other hand, if the pulses of ionization laser fragmented the Ro48-8071 molecule at different positions, the molecular weight change is hard to predict. This causes the difficulty in judging the crosslinking additive products. Hence, in the future, the more detectable probes either from fluorescent modifications or radioisotopes labeling should be integrated into the moiety of Ro48-8071 molecule for easy analysis.

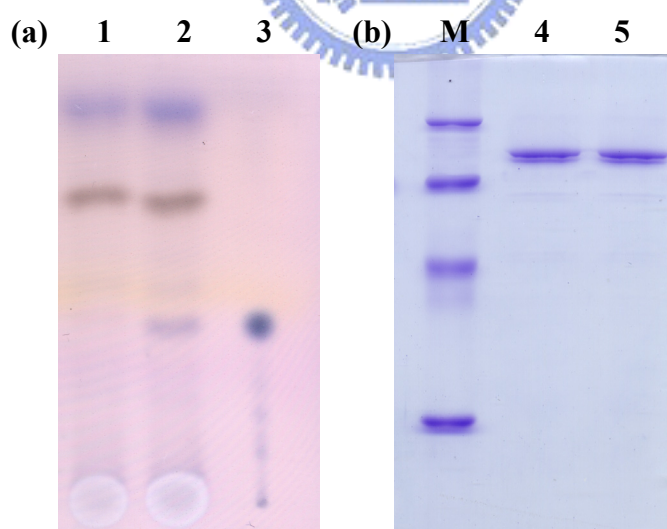


Figure 6.6 (a) TLC analysis of cyclase activity with/without incubation of Ro48-8071 after exposure to UV light. (Lane 1: Activity assay of photoaffinity labeled OSC with Ro 48-8071, Lane 2: Activity assay of wildtype OSC, Lane 3: Standard Lanosterol) (b) SDS-PAGE analysis of photoaffinity labeled OSC with/without Ro48-8071 (Lane M: Protein Markers; Lane 4: photoaffinity labeled OSC with Ro48-8071; Lane 5: wildtype OSC)

1	MTEGTCLRRRGGPYKTEPATDLSRWRLSNQVGR	<b>Q</b> TW <b>T</b> YS <b>Q</b> E <b>E</b> D <b>P</b> V <b>R</b> E <b>Q</b> S <b>G</b> L <b>E</b> A <b>H</b> L <b>L</b> G <b>L</b> D <b>T</b>	60
61	<b>K</b> S <b>F</b> F <b>K</b> D <b>L</b> P <b>K</b> AHTACRGALNGVTFYAALQTEDGHWAGDYG	<b>G</b> P <b>L</b> F <b>L</b> L <b>P</b> G <b>L</b> L <b>I</b> T	120
121	AGYREEIIRYL	<b>R</b> S <b>V</b> Q <b>L</b> P <b>D</b> G <b>G</b> W <b>L</b> H <b>I</b> E <b>D</b> K <b>S</b> T <b>V</b> F <b>G</b> T <b>A</b> L <b>N</b> Y <b>V</b> S <b>L</b> R <b>I</b> L <b>G</b> V <b>G</b> P <b>D</b> D <b>P</b> D <b>L</b> V <b>R</b>	180
181	HKK	<b>G</b> G <b>A</b> V <b>F</b> I <b>P</b> S <b>W</b> G <b>K</b> F	240
241	AYCYSTR	<b>L</b> S <b>A</b> E <b>E</b> G <b>P</b> L <b>V</b> Q <b>S</b> L <b>R</b> Q	300
301	NLYERHHSTSLRQWATQK	<b>L</b> Y <b>E</b> H <b>I</b> A <b>A</b> D <b>D</b> R <b>F</b> T <b>K</b> C	360
361	<b>H</b> V <b>S</b> R <b>I</b> P <b>D</b> Y <b>L</b> W <b>L</b> G <b>L</b> D <b>G</b> M <b>K</b> M	Q <b>G</b> T <b>N</b> G <b>S</b> Q <b>I</b> W <b>D</b> T <b>A</b> F <b>A</b> I <b>Q</b> A <b>L</b> L <b>E</b> A <b>R</b> A <b>Q</b> H <b>R</b> P <b>E</b> F <b>W</b> S <b>C</b> L <b>R</b> K <b>A</b> H <b>E</b> Y <b>L</b> R <b>I</b>	420
421	<b>S</b> Q <b>V</b> P <b>D</b> N <b>F</b> P <b>D</b> Y <b>Q</b> K	Y <b>Y</b> R <b>H</b> M <b>S</b> K <b>G</b> G <b>F</b> S <b>F</b> S <b>T</b> L <b>D</b> C <b>G</b> W <b>I</b> V <b>A</b> D <b>C</b> T <b>A</b> E <b>A</b> L <b>K</b> <b>S</b> I <b>L</b> L <b>L</b> Q <b>E</b> K	480
481	ER	<b>L</b> F <b>D</b> T <b>V</b> A <b>V</b> L <b>L</b> S <b>L</b> R	540
541	LKTFHKQFPDHRAGEIRETLEQGLQFCRQKQRPDGSWEGSWGVCFTYGAWFGLEAFACMG		600
601	HTYHNGVACAEISRACDFLLSRQMADGGWGEDFESCKQRRYVQSAQSQIHNTCWALMGLM		660
661	AVRHPDVAALERGVSYLLEKQLPNGDWPQENISGVFNKSCAISYTSYRNVFPIWTLGRFS		720
721	RLHPDPALAGHP-		732

Figure 6.7 Peptide coverage of bovine liver OSC from LC-MS/MS analysis. The predicted binding sites of Ro48-8071, via the sequence alignment comparison of the previous SHC experiment, are marked with a square box. The shaded amino acids indicated the determined peptide sequences from the LC/MS/MS mass spectrometer.

### 6.3.6 Monoclonal antibody preparation and western blotting analysis

After successfully purified oxidosqualene-lanosterol cyclase from bovine liver, attempt to obtain its respective polyclonal antibody through direct injection of entire purified OSC protein in mice met with failure. The unique biochemical membrane-bound characteristic and the larger molecular size might cause the difficulty to stimulate the immunoreaction in injected mice. Thus, careful consideration for the oligopeptide antigen design was carried out. According to the crystal structure of human OSC and the sequences alignment between human OSC and bovine liver OSC, a homology modeling structure of bovine liver OSC was generated. Figure 6.8 shows the comparison between bovine liver OSC homology model and human OSC crystal structure. Accordingly, one oligopeptide sequence, DGGWGLHIED, was highly conserved among all mammalian OSCs and only changed two residues in that of the respective cyclases from *Saccharomyces*

*cerevisiae* or *Candida albicans*. Moreover, this oligopeptide region was also found to occupy the position on the surface of oxidosqualene-lanosterol cyclase structure (Figure 6.8). Therefore, this oligopeptide sequence might be appropriate for generating the immunogenic antigen. Octameric branches of this oligopeptide on a dendrimer core were thus synthesized from the GeneMed Synthesis Inc. by using the multiple antigenic peptide technology (MAP). The synthesized oligopeptides MAP core was dissolved in PBS solution and added with equivalent volume of adjuvant (Sigma) for injection in BALB/c mice. After three rounds of injection, the immunoassay based on ELISA analysis was subsequently carried out to measure the immunoreactive titer. The monoclonal antibody against the synthesized oligopeptides MAP core was then performed as the following steps: cell fusion of myeloma cell with the spleen cells from the immunized mouse, hybridoma cell culture, ELISA immunoassay determination, mouse ascetic fluid preparation, and the monoclonal antibody purification. The detailed technological processes for the generation of monoclonal/polyclonal antibody were established and executed in our laboratory.<sup>179,</sup>  
<sup>180</sup> After successfully obtaining the monoclonal antibody against the synthesized oligopeptides MAP core, a western blotting experiment of the purified bovine liver OSC was subsequently carried out to confirm the immunogenicity. Figure 6.9 shows the SDS-PAGE and the resultant western blotting analysis with the anti-OSC mAbs. Obvious bands corresponding to the predicted molecular weight of oxidosqualene-lanosterol cyclase were observed from both the SDS-PAGE and the immunoreactive chemiluminescence's reaction on the Kodak autoradiography film.



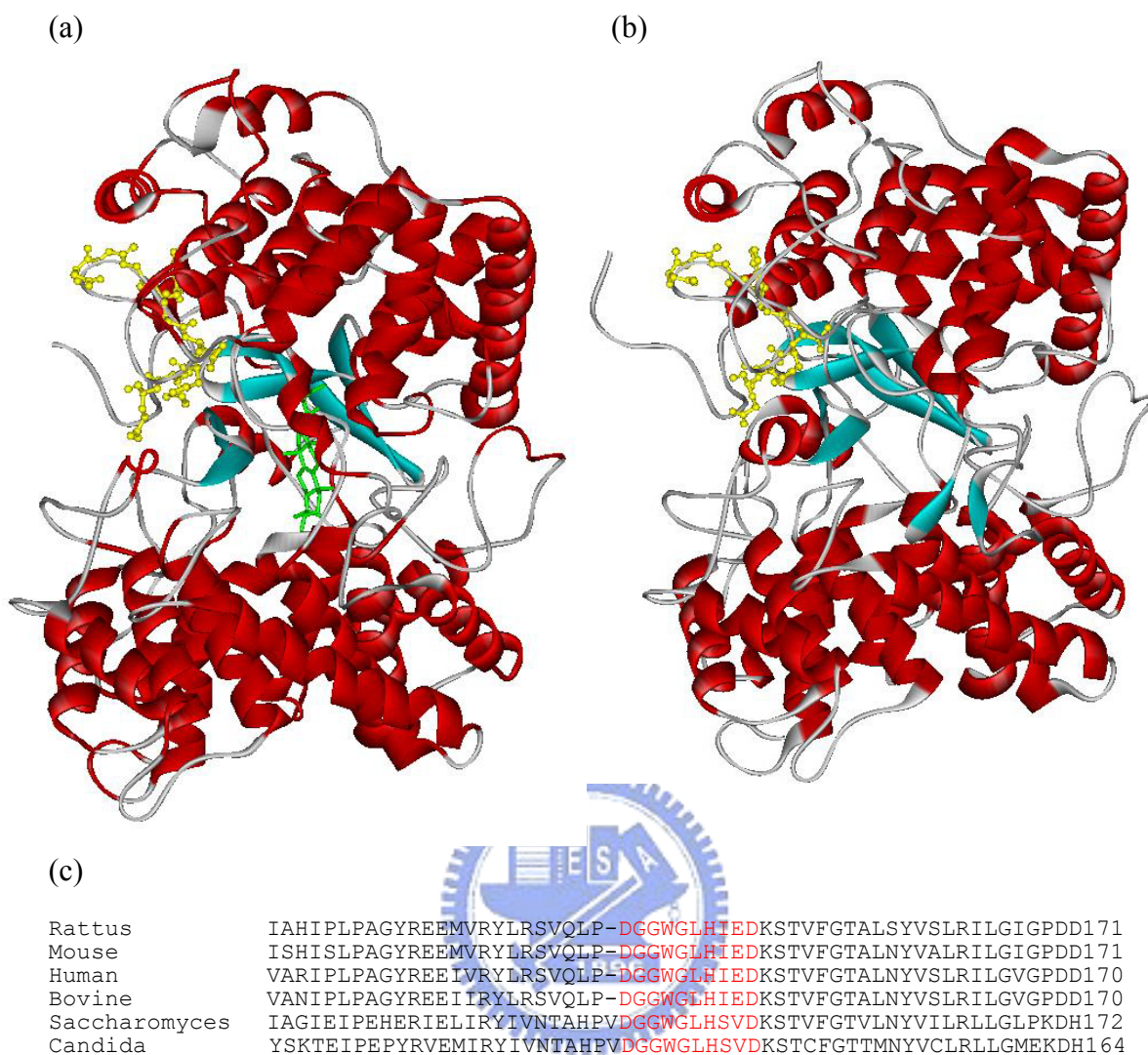


Figure 6.8 (a) Crystal structure of human OSC in complex with lanosterol (green). The oligopeptides region, DGGWGLHIED, is shown in yellow stick representation. (b) Homology modeling structure of bovine liver OSC. The oligopeptides region, DGGWGLHIED, is shown in yellow stick representation. (c) Partial multiple sequence alignment among different OSCs, from which the region of the oligopeptides DGGWGLHIED was colored in red.



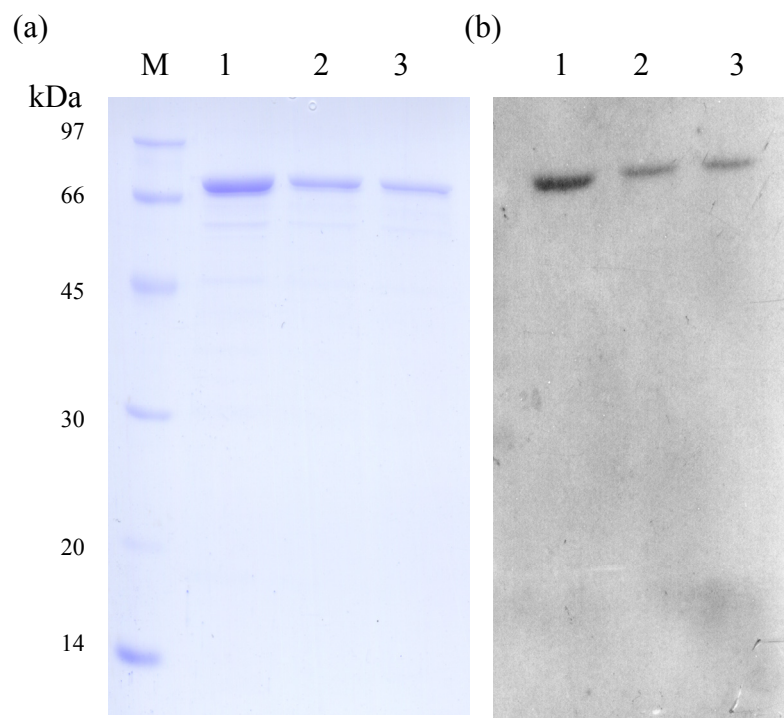
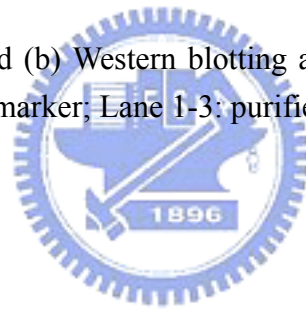


Figure 6.9 (a) SDS-PAGE and (b) Western blotting analysis of purified bovine liver OSC. Lane M: LMW protein marker; Lane 1-3: purified bovine liver OSC.



## *Chapter 7*

### *Development and Application of Fluorescent*

#### *Ro48-8071 Derivative Probes to Study*

#### *Oxidosqualene-Lanosterol Cyclase*

### 7.1 Summary

As previously described results from the Chapter 6, bovine liver OSC had been successfully purified and effectively inhibited by a potent inhibitor, Ro48-8071. But the photoaffinity labeling experiments failed to solve the exact inhibitor-binding site. In this chapter, several Ro48-8071-based fluorescent probes were developed. These fluorescent derivatives, including 4-(4,5-diphenyl-1H-imidazol-2-yl)-phenylboronic acid (DPA), biphenyl-3-boronic acid (BP-3), biphenyl-4-boronic acid (BP-4), naphthalene-1-boronic acid (NA-1), and naphthalene-2-boronic acid (NA-2) moiety, were successfully synthesized from Ro48-8071 using the palladium-catalyzed Suzuki coupling reactions and confirmed via GC/MS and nuclear magnetic resonance (NMR) spectroscopy analysis.

According to the inhibition experiments of all newly synthesized fluorescent probes, it was found that Ro4-NA1, Ro4-NA2 and Ro4-BP4 showed apparent inhibitory activity at the concentration of 100  $\mu$ M. Ro4-BP3 showed worse inhibitory activity than that of Ro4-NA1 at the concentration up to 100  $\mu$ M. Ro4-DPA showed the worst inhibition. The results of inhibition experiment indicated that the fluorescent modification of Ro48-8071 dramatically reduced the inhibitory activity. Results from molecular docking experiments showed that both the interactions and orientation

between the Ro48-8071-derived probes and the active site of bovine liver OSC have dramatically changed. On the other hand, from the fluorescence spectrometric examination, no obvious difference was observed between the bovine liver OSC/Ro4-DPA complex and the free Ro4-DPA molecule. However, after Hi-Trap desalting column purification, the OSC-fraction exhibited slight fluorescent intensity change under picosecond scale spectrography studies. These results suggested that the fluorescent moiety group might interact but not integrate within the protein active site.

## 7.2 Introduction

Fluorescence techniques have been utilized for determination of trace levels of molecules. This non-radioisotope, highly sensitive characteristics have let various fluorescence labeling reagents developed for detection of a wide variety of bioactive compounds. Currently, many labeling reagents have been utilized for conjugation of different functional groups, including amino acids, alcohol, thiols, aldehydes, ketones, and carboxyl groups.<sup>181</sup> However, many pharmaceuticals and agricultural chemicals could not be labeled are due to the lacking of the reactive functional groups. Thus, the new conjugation approaches for specifically labeling are still required. For example, because of the extensive exist of aryl halide in the natural products, the fluorescence labeling reagent for aryl halide is believed as a powerful method in the pharmaceutical applications. Recently, several fluorescent arylboronic acid derivatives have been developed for effectively labeling aryl halides by using the Suzuki coupling reaction, a palladium-catalyzed coupling reaction of aryl halides with aryl boronic acids.<sup>182, 183</sup> Suzuki coupling reaction is one of the most versatile and powerful methods for carbon-carbon bond formation. Among these arylboronic acid derivatives, 4-(4,5-diphenyl-1*H*-imidazol-2-yl) phenylboronic acid (DPA) was first successfully

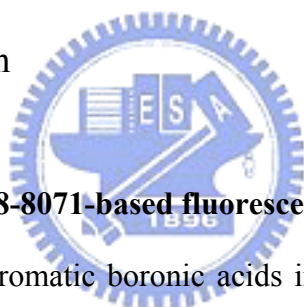
used in the labeling reaction for the aryl halide, and acted as an enhancer for the horseradish peroxidase chemiluminescence reaction.<sup>184</sup>

Aromatic boronic acids have also been illustrated as a glucose sensing indicator via quenching their fluorescence intensity when they interacted with the saccharide molecules. The fluorescence change has been utilized to investigate the relationship between the aromatic ring and saccharide-binding events.<sup>185, 186</sup> On the other hand, many different fluorescent probes also provided applications in the several biochemical fields. For example, the fluorescent probe, bis(8-anilino-1-naphthalenesulfonate), was the first example utilized to study protein conformation change in 1976.<sup>187</sup> The magnesium-dependent conformational change of 30S subunit from *Escherichia coli* ribosome was observed via reversible binding of dye molecule to the specific protein sites at the ribosomal surface. Another rhodamine-labeled geranylgeranylated/methylated cysteine derivative was also used to simulate the C-terminal modification in small G-protein. The protein-protein interaction between small G-protein and its regulation inhibitor GDI proteins was first reported by using the quantitative fluorescence assay.<sup>188</sup> In addition, *p*-Aminobenzamidine, a competitive inhibitor of serine protease, exhibited high fluorescent intensity change when it bound to trypsin or thrombin, whereas the weak fluorescence was observed in the neutral aqueous buffer. The spectral properties of this probe sensitized to the specific hydrophobic interaction or the structural features in the active center of serine proteases were further investigated in 1982.<sup>189</sup> Moreover, a fluorescent substrate analog of undecaprenyl pyrophosphate synthase (UPPs), 7-(2,6-dimethyl-8-diphospho-2,6-octadienyloxy)-8-methyl-4-trifluoromethyl-chromen-2-one geranyl pyrophosphate, was prepared and utilized to study the ligand interactions with *E. coli* UPPs. The dramatic fluorescent intensity quenching caused by the enzymatic hydrophobic environment was also observed. The kinetic

characteristics were also measured by using stopped-flow apparatus. These results indicated that the fluorescent substrate analogs or inhibitor probes might provide a tool to investigate the ligand interactions for a broad class of proteins in native environment, and to provide the comparable results with that of co-crystallized experiment.<sup>190, 191</sup>

In order to expand the methodology for investigating cyclase, different fluorescence labeling reagent will be conjugated to the inhibitor, Ro48-8071, by the palladium-catalyzed Suzuki coupling reaction. The inhibition mechanism will be elucidated using various fluorescence spectrum studies, TLC activity assay, as well as the molecular docking experiments.

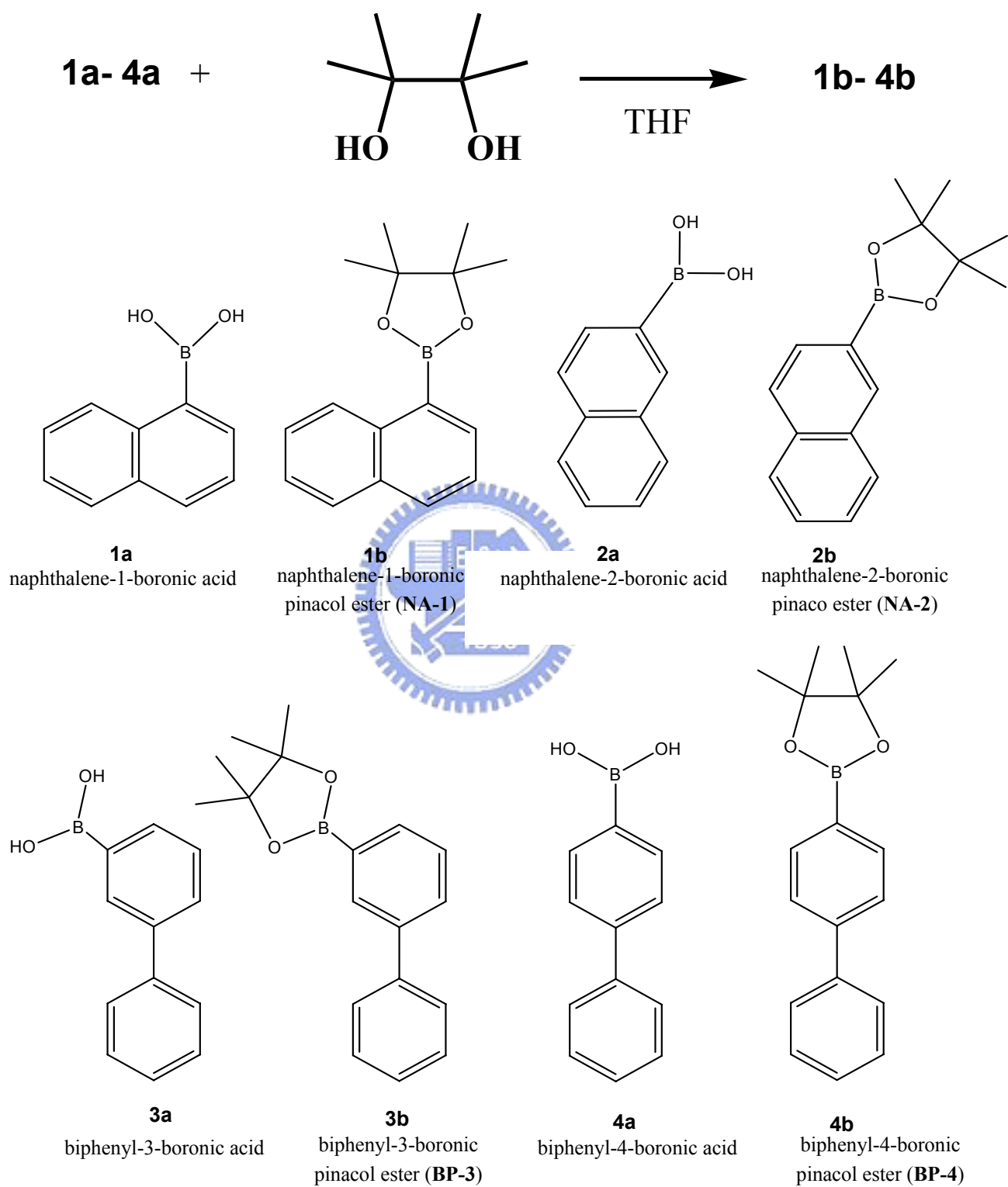
## 7.3 Results and Discussion



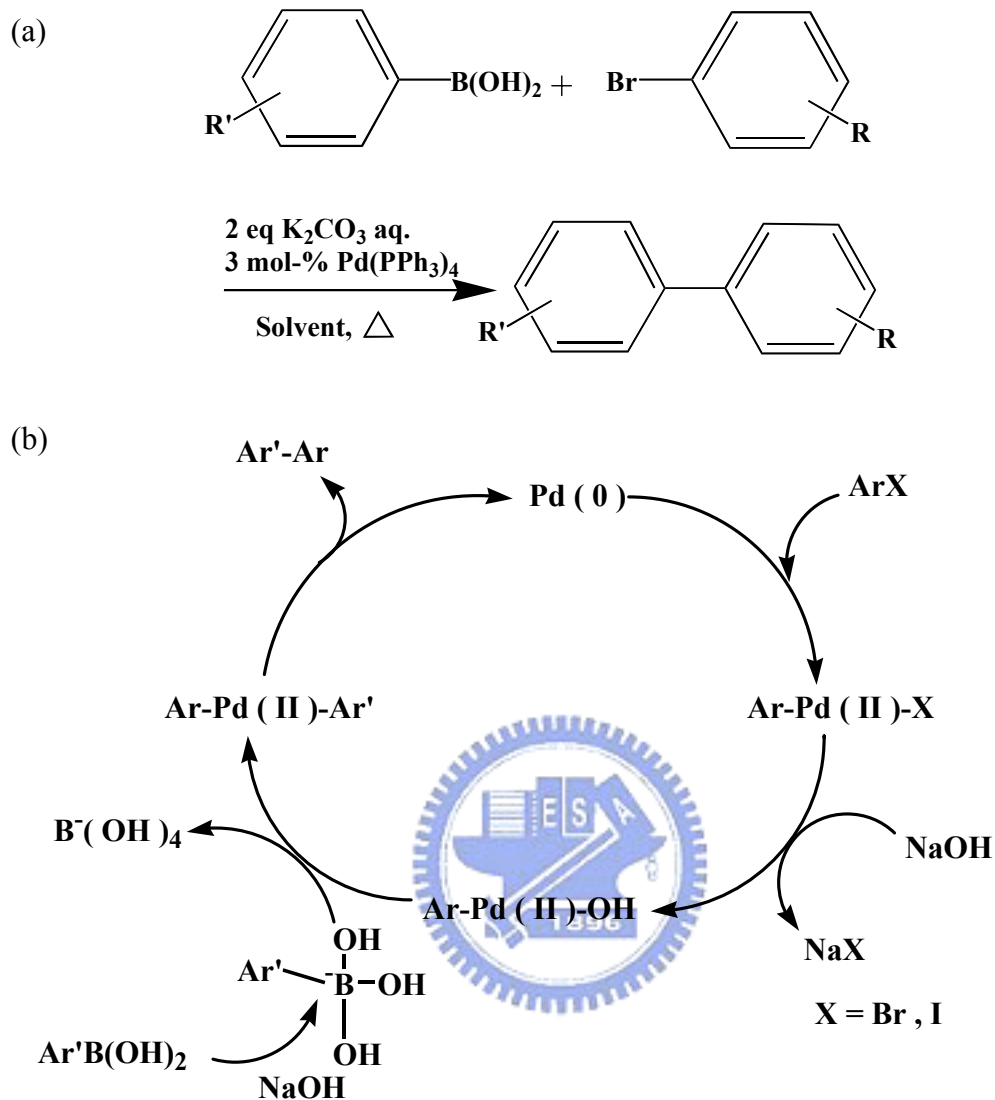
### 7.3.1 Synthesis of newly Ro48-8071-based fluorescent probes.

A series of fluorescent aromatic boronic acids including naphthalene-1-boronic acid, naphthalene-2-boronic acid, biphenyl-3-boronic acid, biphenyl-4-boronic acid were converted into their boronic pinacol ester derivatives via the condensation reaction with the pinacol diol (Scheme 7.1).<sup>192</sup> Four boronic pinacol ester compounds, including naphthalene-1-boronic pinacol ester (**NA-1**), naphthalene-2-boronic pinacol ester (**NA-2**), biphenyl-3-boronic pinacol ester (**BP-3**), biphenyl-4-boronic pinacol ester (**BP-4**) were then used for modification of Ro48-8071 molecule via the Suzuki coupling reaction.<sup>193</sup> The novel C-C single bond is formed via exchanging the aromatic group next to boron atom with alkoxy group from aryl halides. The proposed mechanism of Suzuki coupling is shown in Scheme 7.2. In addition, one fluorescent labeling reagent, 4-(4,5-diphenyl-1H-imidazol-2-yl)phenylboronic acid (DPA) has also been successfully synthesized and coupled with Ro48-8071 by using

the same Suzuki coupling reaction.<sup>184</sup> The structure and coupling reactions are listed in Scheme 7.3 and Scheme 7.4.



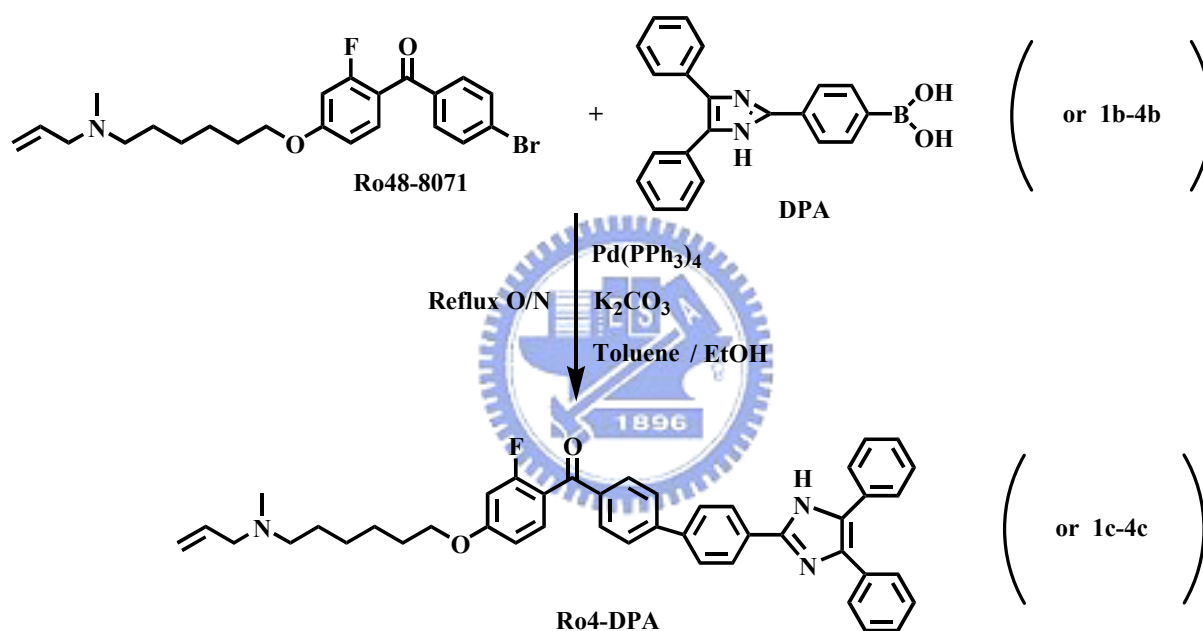
Scheme 7.1 Synthesis of the aromatic boronic pinacol ester



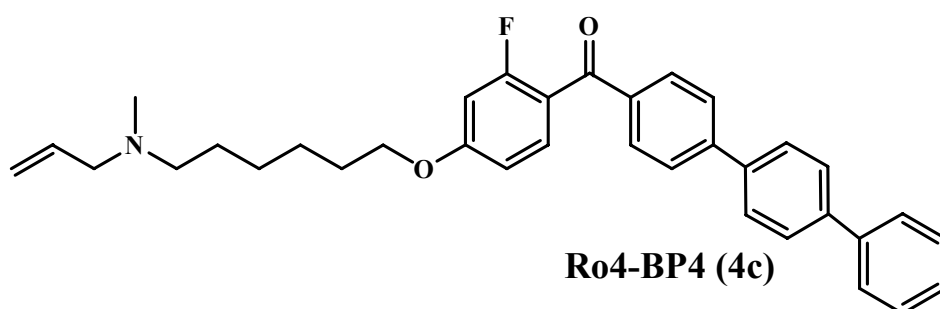
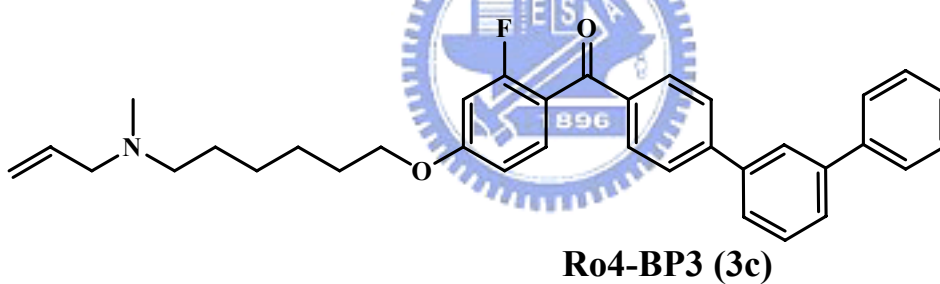
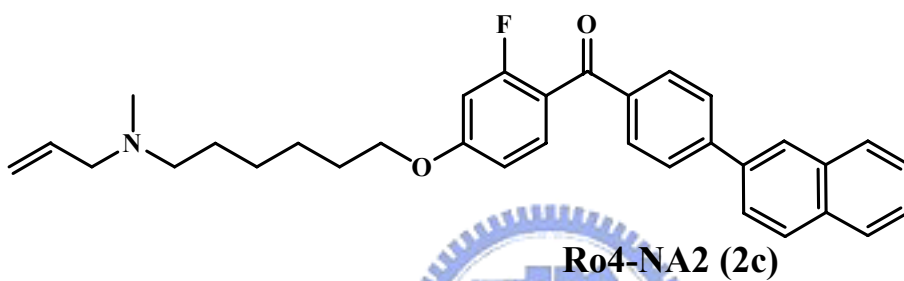
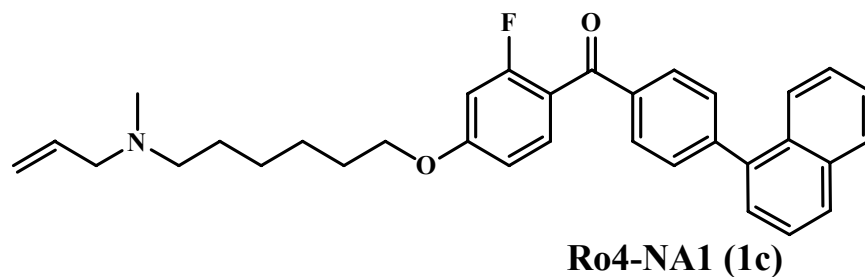
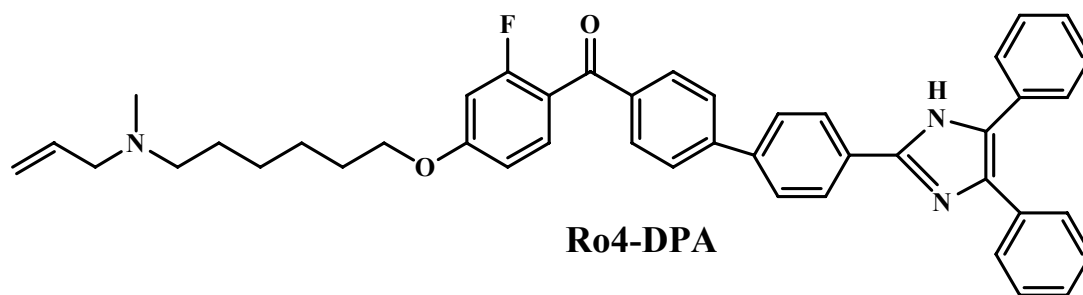
Scheme 7.2 (a) The first published Suzuki coupling reaction.<sup>193</sup> (b) Proposed mechanism of Suzuki coupling reaction, from which the alkali base is sodium hydride.



The crude products were extracted and purified on silica gel using the mobile phase constituted of 0%, 10%, and 20% methanol/CH<sub>2</sub>Cl<sub>2</sub> solution. The spectroscopic data corresponding to the additive modification were identified from the integration values in the <sup>1</sup>H-NMR and from the judgments in the <sup>13</sup>C-NMR as well as the DEPT-NMR spectra. The MS signals also provided the evidence of successful modification for Suzuki coupling reaction. The chemical structures of these newly Ro48-8071-based fluorescent probes are shown in Scheme 7.4.



Scheme 7.3 Derivation reaction of Ro48-8071 with DPA or other four boronic pinacol ester compounds. The Ro48-8071 was synthesized according to the previous literature.<sup>169</sup>

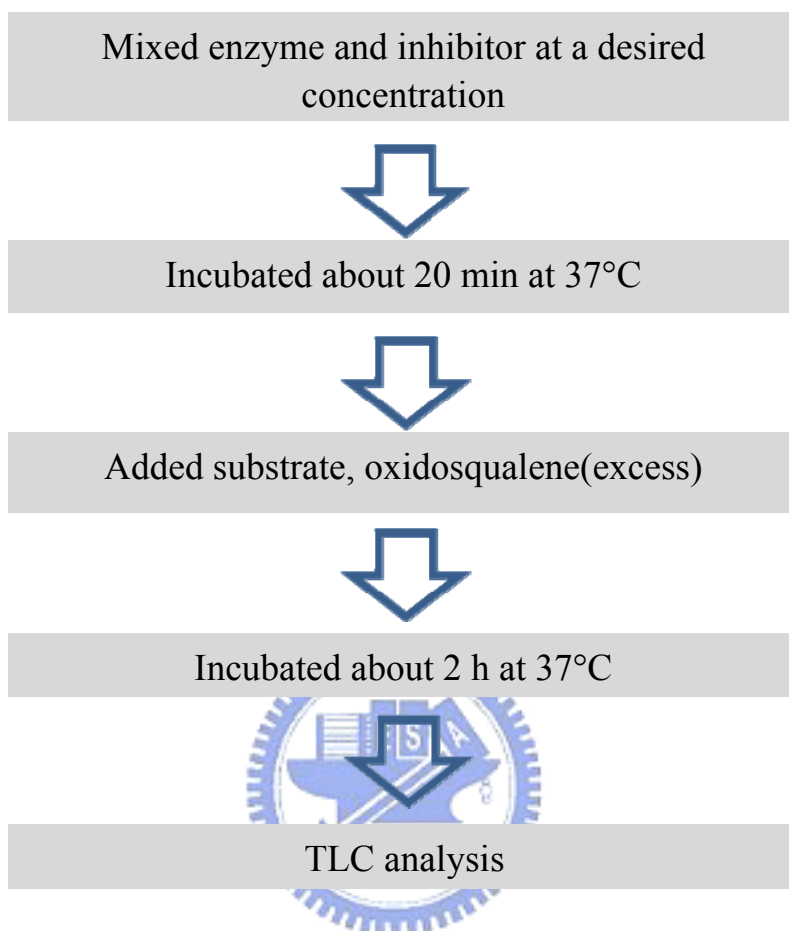


Scheme 7.4 The chemical structure of five newly Ro48-8071-based fluorescent probes.

### 7.3.2 The inhibitory activity of the Ro48-8071-based fluorescent probes

In the present study, the fluorescent modifications were introduced on the benzophenone group of Ro48-8071. In order to understand the inhibitory effect of these fluorescent derivatives and compared with the previous Ro48-8071 data, the inhibition assays were performed by mixing of the substrate, OS, and different concentrations of the inhibitor derivatives with purified bovine liver OSC. After incubation for 2 h, the reactions were extracted and spotted on the TLC. The inhibitory effect was observed and compared to the control experiment that contained no inhibitor in the reaction. The TLC analysis demonstrated that the Ro4-NA1, Ro4-NA2, and Ro4-BP4 showed obvious inhibition for the bovine liver OSC at the concentration of 100  $\mu\text{M}$ , whereas the Ro4-BP3 exhibited less inhibition at the same concentration. The bulky lophine analogue, 4-(4,5-diphenyl-1*H*-imidazol-2-yl) phenylboronic acid (DPA), dramatically abolished the inhibitory activity of Ro48-8071 even at the concentration more than 100  $\mu\text{M}$  (Figure 7.1). Moreover, none of these newly Ro48-8071-based fluorescent probes displayed comparable inhibitory activity with that of parent compound, Ro48-8071. At the concentration of 1  $\mu\text{M}$ , Ro48-8071 exhibited adequate capability to shade the cyclase activity (Figure 7.1). Interestingly, from the illustration of human OSC cyclase crystal structure, the aryl bromide of the Ro48-8071 structure was expected to occupy the substrate entrance channel and not to enter deeply into the enzyme cavity.<sup>105</sup> From the initial consideration, the modification of the aryl bromide should be more reasonable than that of other portions. However, from our present results, the extreme decrement of inhibitory activity was observed from all of the derivatives even in the presence of the less hindrance of naphthalene or biphenyl modification. The flowchart of the inhibitory activity assay of these newly Ro48-8071-based fluorescent probes is shown in Table 7.1

Table 7.1 The inhibitory activity assay of the new Ro48-8071-based fluorescent probes for bovine liver OSC.



(a)

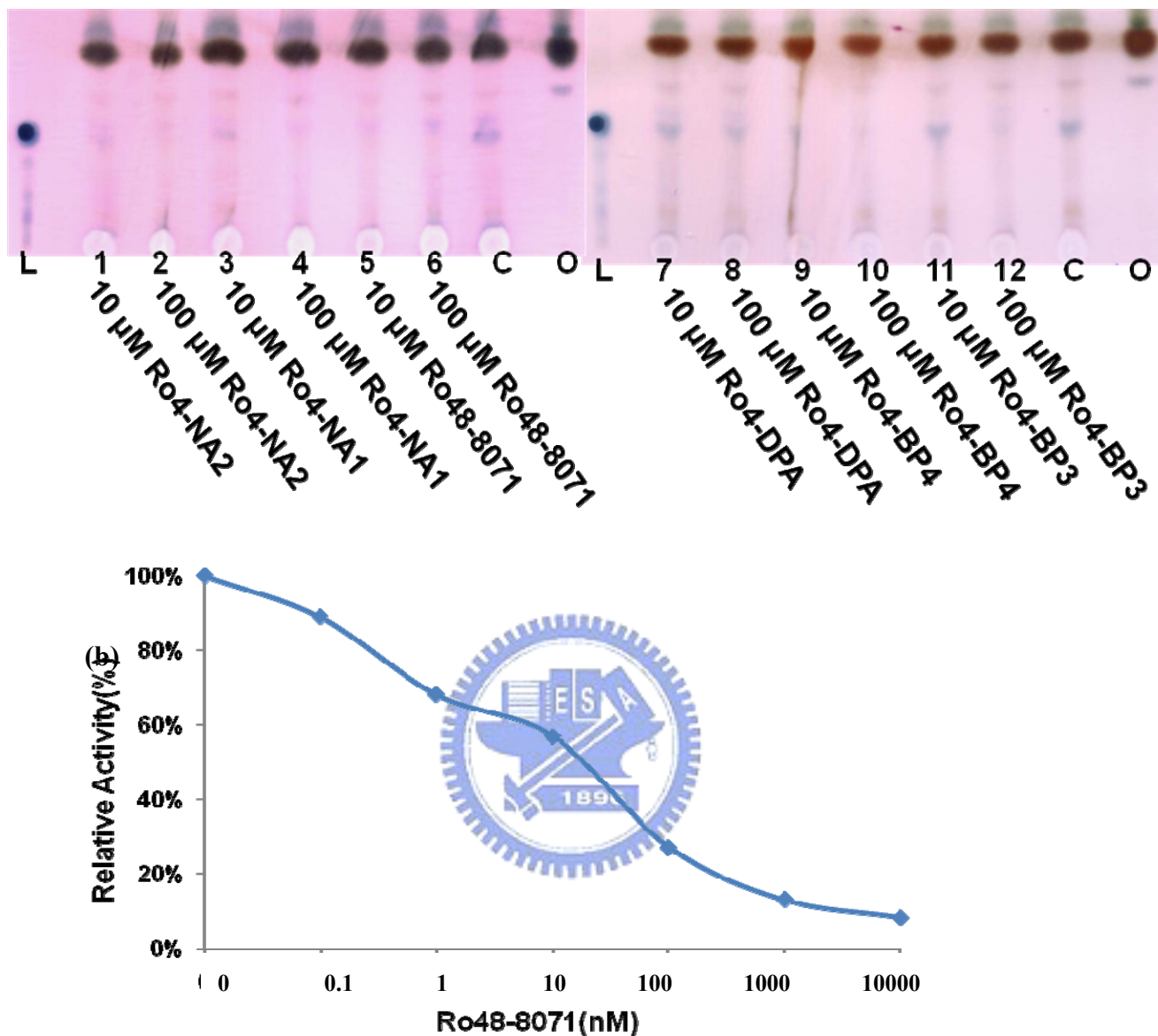
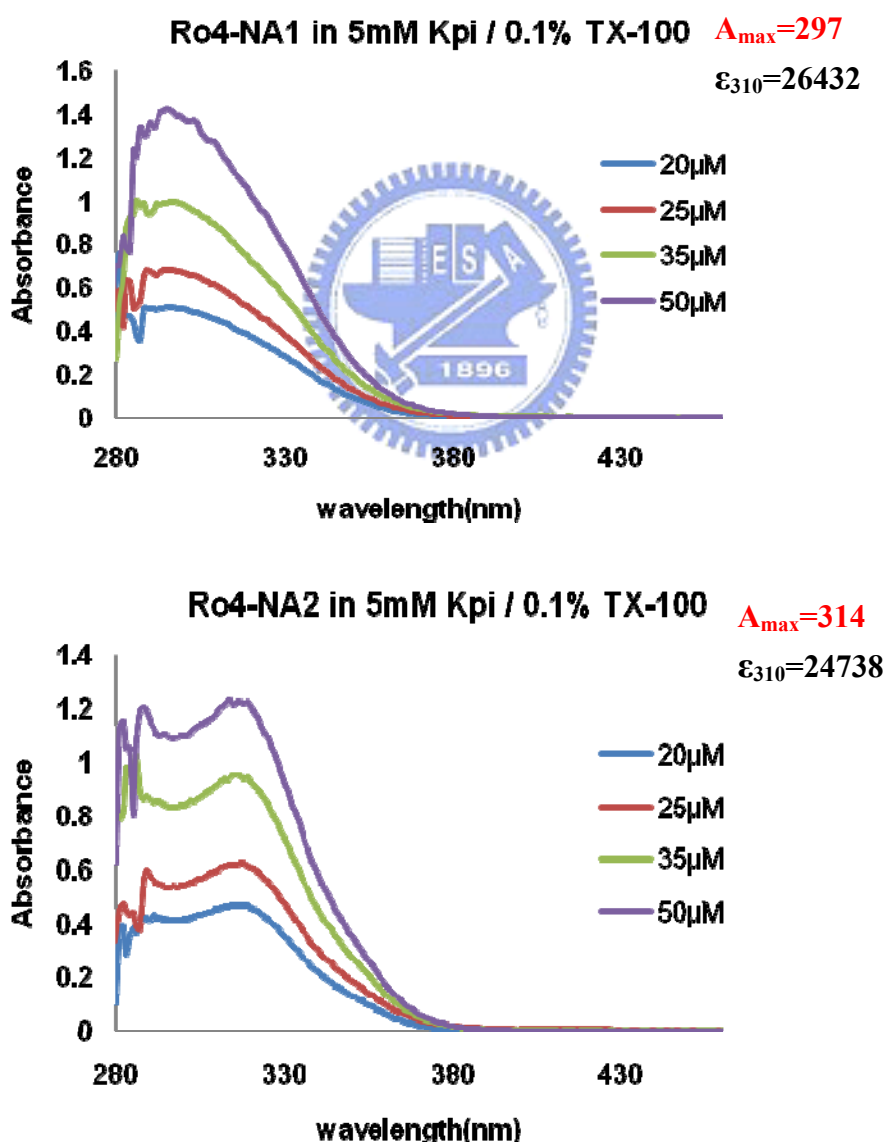


Figure 7.1 (a) TLC analysis of cyclase activity, from which OSC was pre-incubated with different Ro48-8071-based derivatives at either 10 μM or 100 μM concentration (Lane 1~12). In addition, Lane O (negative control): activity assay reaction without cyclase enzyme, Lane C (positive control): activity assay reaction without inhibitor, Lane L: standard lanosterol position. (b) The inhibition of Ro48-8071 for the bovine liver OSC.<sup>142</sup>

### 7.3.3 Spectroscopy of these newly Ro48-8071-based fluorescent probes

The absorption spectra of the newly synthesized Ro48-8071-based fluorescence compounds were measured on a UV-Vis spectrophotometer (DU 7500, Beckman). The molar absorption coefficient was also examined according to the Beer's law. From the UV-Vis spectra, the Ro4-NA1, Ro4-NA2, Ro4-BP3, and Ro4-BP4, respectively displayed the maximal absorption wavelength at 297 nm, 314 nm, 297 nm, and 304 nm, whereas the Ro4-DPA exhibited the maximal absorption wavelength at 365 nm (Figure 7.2).



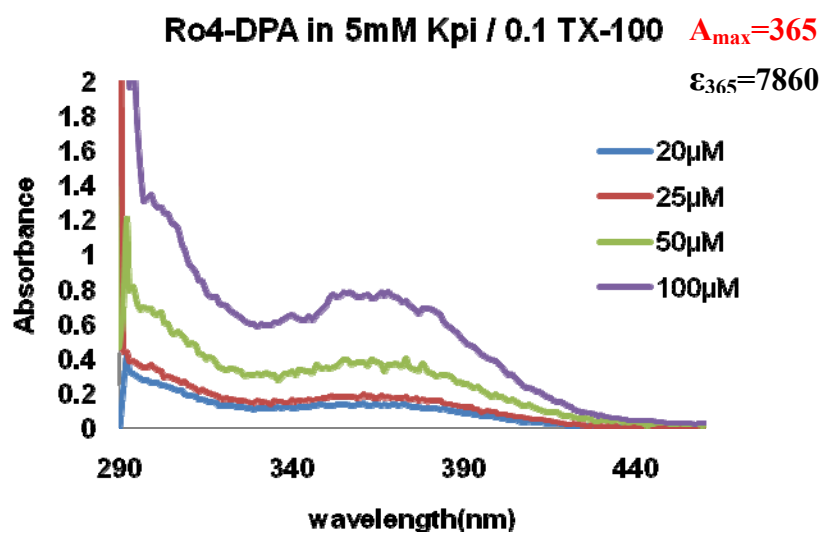
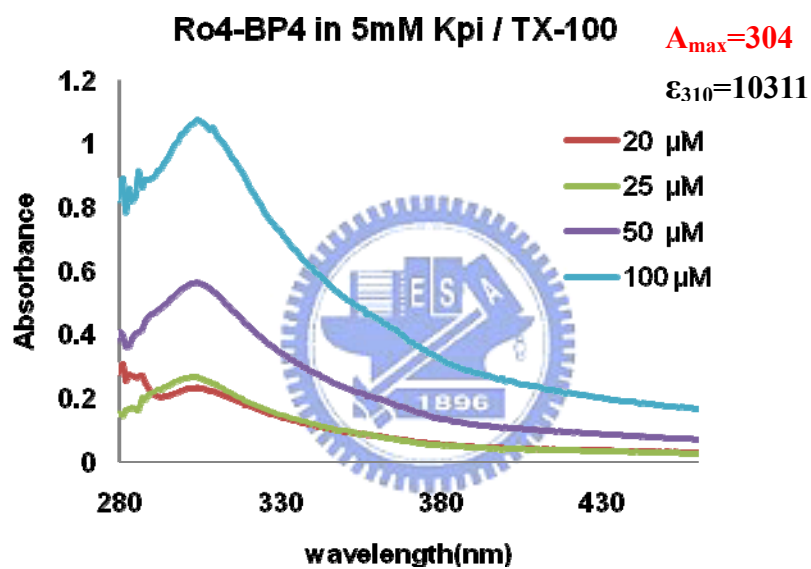
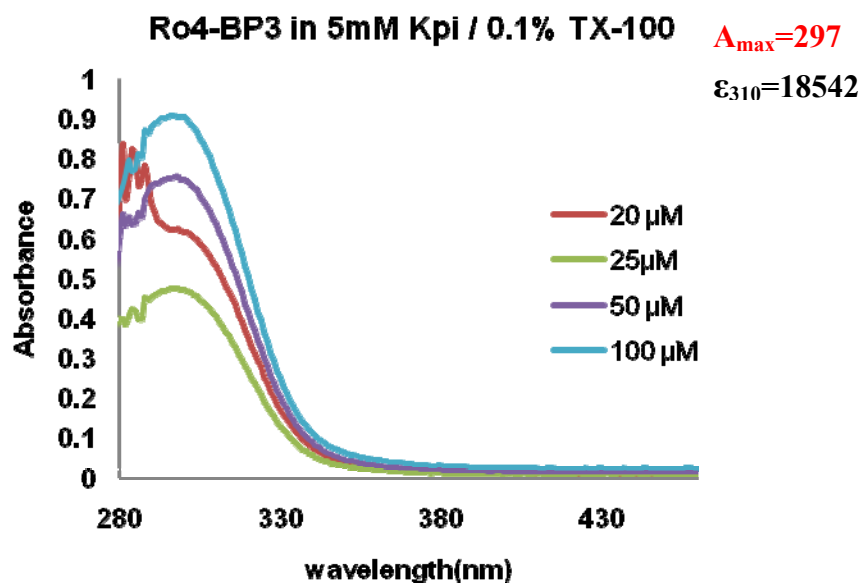


Figure 7.2 The wavelength scanning of fluorescent probes in aqueous solution



According to the UV-Vis spectra of these Ro48-8071-based derivatives, the maximum absorption wavelength was around 300-315 nm except the lophine analogue, Ro4-DPA. The larger red shift might be caused by the intra-molecular conjugated electronic transferring from carbonyl group to the coplanar DPA group. The  $\pi \rightarrow \pi^*$  shift instead of the  $n \rightarrow \pi^*$  shift, the absorption of the carbonyl group, was observed in the UV-Vis spectrum of Ro4-DPA. On the other hand, other derivatives might form a non-planar structure between naphthalene or biphenyl group and the benzophenone moiety, thus led to abrogate the electron transfer phenomenon.

Moreover, the fluorescence spectrums of the Ro48-8071 derivatives in different solution were also taken from 300 nm to 800 nm by using a HitachiF-4500 spectrophotometer. As shown in Figure 7.3, the 25  $\mu$ M Ro4-DPA showed strong fluorescence emission in the aqueous or in the DMSO solution, whereas the weakest emission was observed in the ethanol solution. Interestingly, without addition of detergent TX-100 in the aqueous Kpi buffer, the Ro4-DPA fully eliminates the fluorescence emission. The micellar structure would be formed while the non-ionic detergent TX-100 was solved in the aqueous solution. The hydrophobicity of the inner environment might facilitate the stacking of the fluorescent probe, and lead to enhance the fluorescence emission. In addition, the red shift of fluorescence emission observed in the DMSO solution might also be due to the lower polar property of DMSO. According to the dramatic fluorescent change in different polar solvent, it seems reasonable that when Ro48-8071-based derivatives interacted with the inner hydrophobic OSC active site cavity, the fluorescence intensity change or the maximal emission wavelength shift might be observed. The fluorescence changes might provide the method in measuring the interaction between the fluorescent probes and the OSC protein.

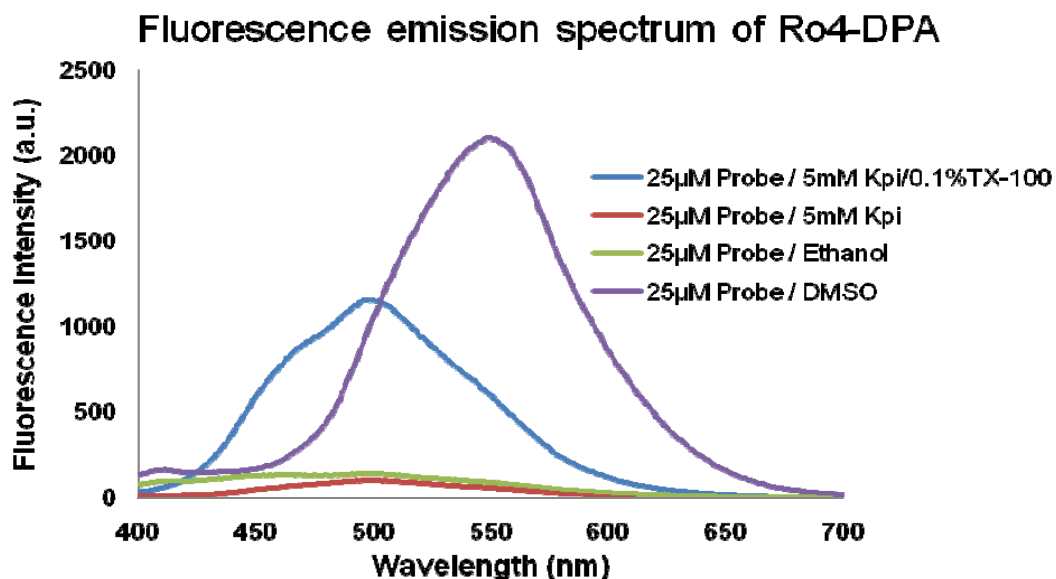


Figure 7.3 The fluorescence spectra of Ro4-DPA in the aqueous solution, ethanol, and DMSO, respectively.

Interestingly, none of other Ro48-8071-based fluorescent probes exhibited the fluorescence characteristic in either aqueous, DMSO solution, or on the TLC plate (Figure 7.4). The fluorescence quenching caused by the solvent mediator effect or the ultrafast fluorescent decay might occurred among these excited Ro48-8071-based fluorescence probes. Due to the lack of the fluorescence emission from naphthalene (Ro4-NA1, or Ro4-NA2) or biphenyl derivatives (Ro4-BP3, or Ro4-BP4), fluorescence titration experiment was only carried out for the Ro4-DPA probe. Fluorescence spectrophotometer was employed to monitor the fluorescence change of Ro4-DPA after adding purified bovine liver OSC to the solution containing Ro4-DPA. As shown in Figure 7.5, there is no obvious change in the fluorescence emission intensity between the free Ro4-DPA control and the Ro4-DPA pre-incubated with OSC samples. The results suggested that the fluorescent moiety of the Ro4-DPA might not either participate in the interaction with OSC or enter in the deeper hydrophobic active site cavity. Moreover, according to the TLC analysis, the fully abrogated inhibitory activity was observed from Ro4-DPA derivatives. The bulky

fluorescent DPA moiety might dramatically influence the binding as well as inhibitory ability of the Ro48-8071 for the OSC, thus we could not observe any fluorescence emission intensity change in the titration studies.

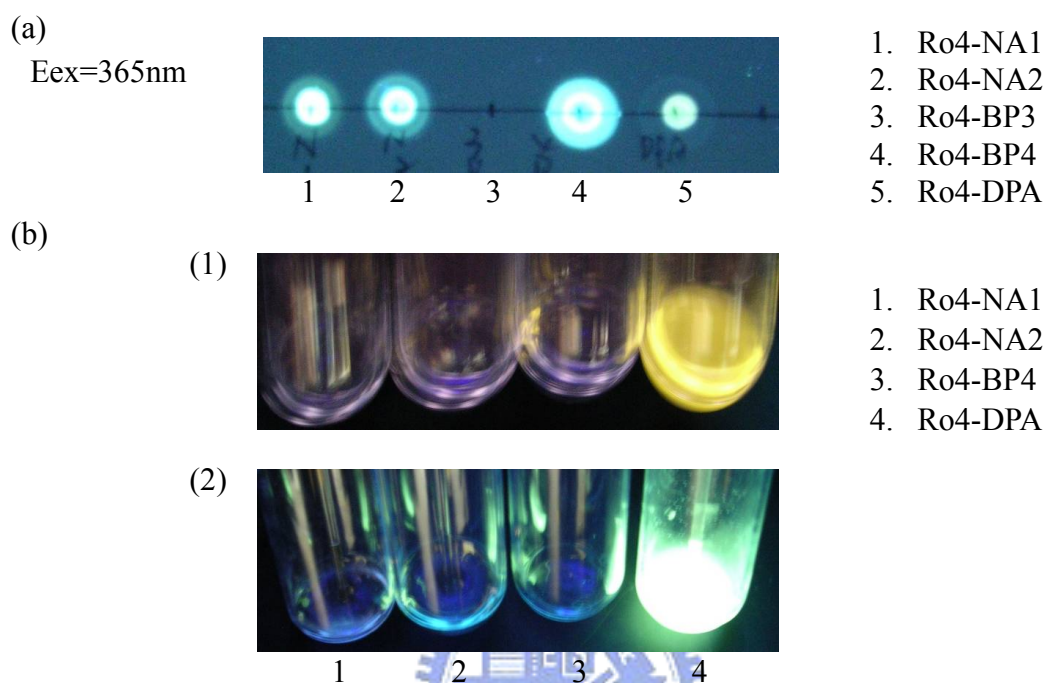


Figure 7.4 Fluorescence characteristic of the various Ro48-8071-based fluorescent probes on either (a) TLC plate or in the solution of (b-1) DMSO and (b-2) the aqueous solution.

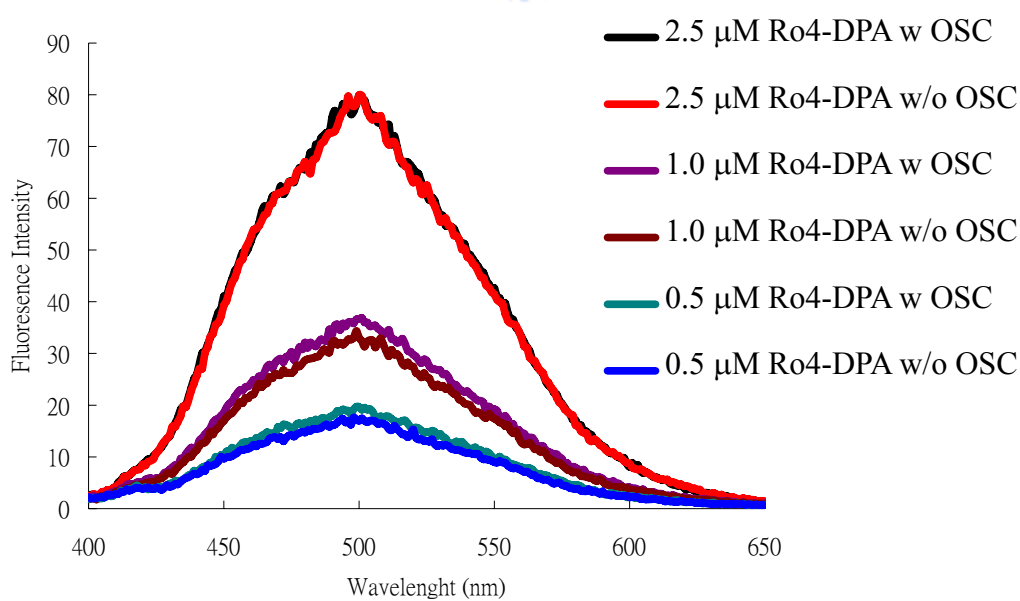


Figure 7.5 Fluorescence spectra of Ro4-DPA (2.5-0.5  $\mu$ M) with/or without addition of bovine liver OSC.

### 7.3.4 Homology modeling and ligand docking studies of newly Ro48-8071-based fluorescent probes

In order to elucidate the influence of fluorescent modification on the Ro48-8071 inhibitory activity as well as to know the interaction between fluorescent probes and OSC cyclase, the homology modeling structures coupled with different ligand docking studies were carried out. The homology models of the bovine liver OSC was created and the different ligands, the Ro48-8071-based fluorescent probes, were individually docked into the bovine liver OSC's modeling structure. The detailed procedure was previously described or in the experimental sections. For each of the Ro48-8071-based fluorescent probes, fifty genetic algorithm runs with standard default parameter settings were performed without any early termination, after which 10 best solutions were chosen for each ligand. Moreover, the scoring functions based on the genetic docking algorithm GOLD 3.0 were used to predict the biological interaction of the ligand. The GOLD 3.0 fitness function consisted of four components: (a) protein–ligand hydrogen bond energy (external H-bond); (b) protein–ligand van der Waals (vdw) energy (external vdw); (c) ligand internal van der Waals energy (internal vdw); and (d) ligand torsional strain energy (internal torsion).<sup>194</sup> The top scoring solution of each probe, predicted by GOLD program, was further used to construct its binding model and the orientation within the enzyme active site cavity. The GOLD fitness values for the Ro4-NA1, Ro4-NA2, Ro4-BP3, and Ro4-BP4 were 43.4767, 70.2055, 63.3292, and 46.6721, individually, whereas all of fifty docking solutions for the Ro4-DPA were negative with averaged fitness values -89.34. This negative fitness represented the loss of binding ability of Ro4-DPA for the OSC enzyme.

The detailed examination of the binding modes and the orientations of these ligand molecules confirmed the predicted GOLD fitness values. According to the

human OSC/Ro48-8071 co-crystallized structure, the quaternary amino group of Ro48-8071 formed the hydrogen-bonding with the putative general acid residue, Asp-455 residue, and subsequently inactivated the cyclase enzyme function. In addition, the benzophenone group of Ro48-8071 stacked with His-232 and Phe-696 also blocked the substrate binding in the enzyme hydrophobic cavity. Among the present docking results, the nitrogen atoms from Ro4-NA1 and Ro4-BP4 exhibited similar distance for the Asp-455 residue (3.9 Å and 4.7 Å relative to 3.1 Å of the Ro48-8071). However, the substituted naphthalene or biphenyl groups of Ro4-NA1 or Ro4-BP4 dramatically twisted the orientation of derivatives and drove the benzophenone group far away from the original position of Ro48-8071 in the enzyme active site channel. This altered binding mode might contribute to the loss of inhibitory activity in the Ro4-NA1 or Ro4-BP4. In contrast, the high GOLD fitness value for the Ro4-NA2 and Ro4-BP3 might be caused by the different orientation of the substituted group in the benzophenone structure. The entire structure of derivatives remained in the original binding channel was observed from the homology modeling structure. However, the distances from nitrogen atom to the aspartic acid are larger than their relative structural isomers (5.0 Å or 5.3 Å relative to the 3.9 Å or 4.7 Å). The loss of hydrogen-bonding might lower the inhibitory activity for these two Ro48-8071 derivatives. Interestingly, all of the docked Ro4-DPA molecules exhibited the reverse orientation compared with that of original Ro48-8071 within the enzyme active site. The negative fitness value and the incorrect binding orientation of the Ro4-DPA might be due to the bulky DPA group or the steric hindrance repulsion. Hence, the loss of binding ability of Ro4-DPA for the cyclase enzyme, based on theoretical homology calculation, provides hits to suggest the abrogated inhibitory activity observed from the experimental TLC activity assay. The binding structures of these Ro48-8071-based fluorescent probes in the bovine liver OSC active site are

shown in Figure 7.6 and Figure 7.7.

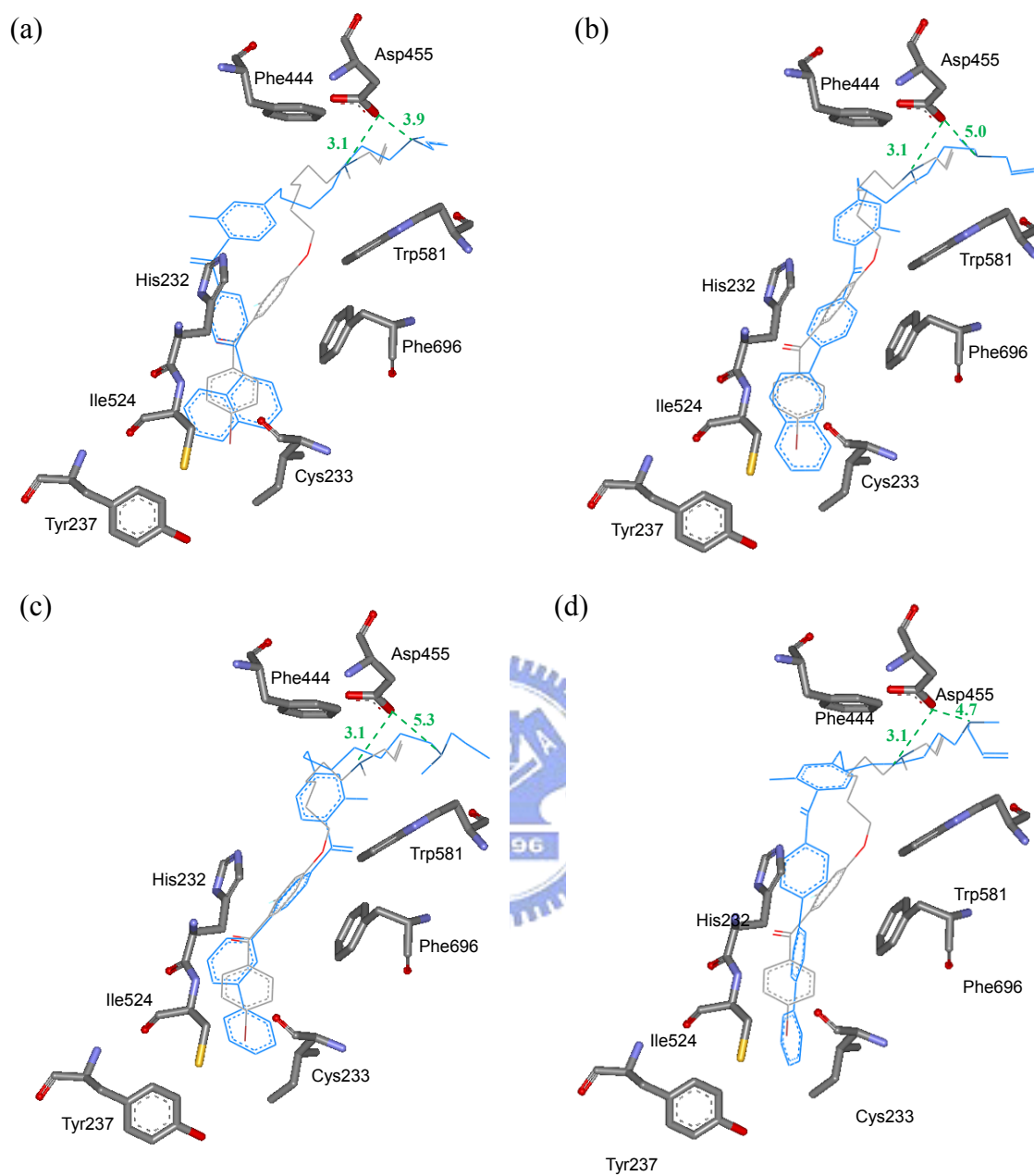


Figure 7.6 The molecular ligand docking results of the (a) Ro4-NA1, (b) Ro4-NA2, (c) Ro4-BP3, and (d) Ro4-BP4 for the bovine liver OSC. The gray molecule represents the original Ro48-8071 molecule, whereas the light blue one shows the corresponding docked Ro48-8071 derivatives. The distances from the nitrogen atom to the aspartic acid residue are also included.

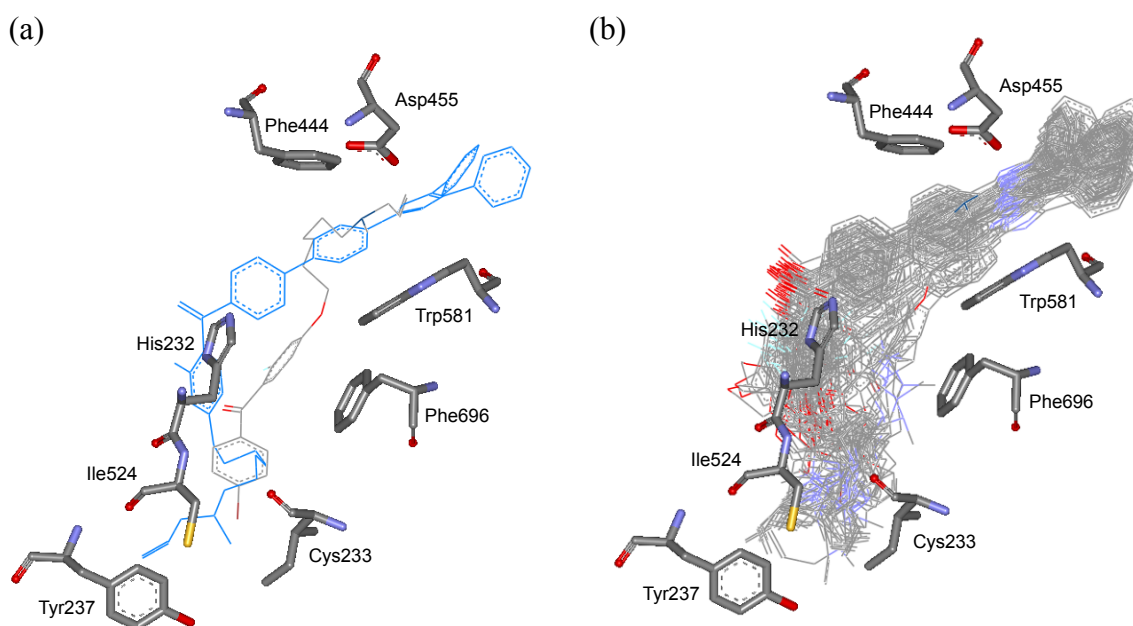


Figure 7.7 (a) The molecular ligand docking results of the Ro4-DPA for the bovine liver OSC. The orientation of Ro4-DPA exhibited the reverse direction, relative to the original position of Ro48-8071. (b) The total fifty docking solutions of the Ro4-DPA.

### 7.3.5 Time-resolved fluorescence studies of Ro48-8071-based fluorescent probes

The laser induced fluorescence decay or the anisotropic analysis of these Ro48-8071 based fluorescent probes were subsequently carried out. Results from these picosecond-scale spectrographic studies might provide some useful clues to understand the fluorescent characteristics or the interaction behaviors of these fluorescent probes. One of the Ro48-8071 based fluorescent probes, Ro4-NA1, was pre-incubated with the purified bovine liver OSC at 37°C for 2 h, and subsequently applied for a PD-10 desalting columns to remove the free probes. The fluorescent lifetimes detection or the anisotropy analysis were performed by using a time-correlated single-photon counting spectrometer (TCSPC, Fluo-Time 200, PicoQuant) coupled with an excitation system from a picosecond laser system (LDH-P-C-440 and PDL-800B, PicoQuant) at 375 nm. The detailed experiment to measure the fluorescent lifetime or the anisotropic change were described and



executed elsewhere.<sup>195</sup> From the results of fluorescence decay of the sole Ro4-NA1 in HB buffer, the major component of decay rates was 0.4 ns (76%) whereas others two were 1.4 ns and 2.9 ns (Figure 7.8). Moreover, for the Ro4-NA1/OSC mixture, the fluorescence decays are 0.1 ns (80%), 1.2 ns (15%), and 3.3 ns (5%), respectively. Accordingly, the major difference between the Ro4-NA1/OSC and free Ro4-NA1 molecule is the first lifetime ( $\tau_1$ ) whereas the second and third lifetimes are similar. Without the theoretical longer lifetime appearing in the experiment of Ro4-NA1/OSC indicated that the fluorescent probe might not be tightly trapped by cyclase. However, results from the difference of first decay rate implied that some energy transfer or other non-radioactive process between ligand and protein might occur and thus further influence the cyclase activity.

Moreover, from the results of anisotropy analysis, the time-resolved fluorescence anisotropic change of two conditions (Ro4-NA1/OSC mixture and the free Ro4-NA1 molecule) was almost the same and also fitted for the double exponential decay. As shown in Figure 7.8, the first depolarized lifetime was pulse limited and another was  $\sim 9.3$  ns. This result suggested that Ro4-NA1 dye molecule might not bind tightly in the OSC's cavity or Ro4-NA1 dye molecule might leave away from the cyclase rapidly. Moreover, the longer linker design of these Ro48-8071-based fluorescent probes might also cause the facile rotation of phenyl rings in the enzyme active site and lead to the loss of fluorescent change in the anisotropic analysis.

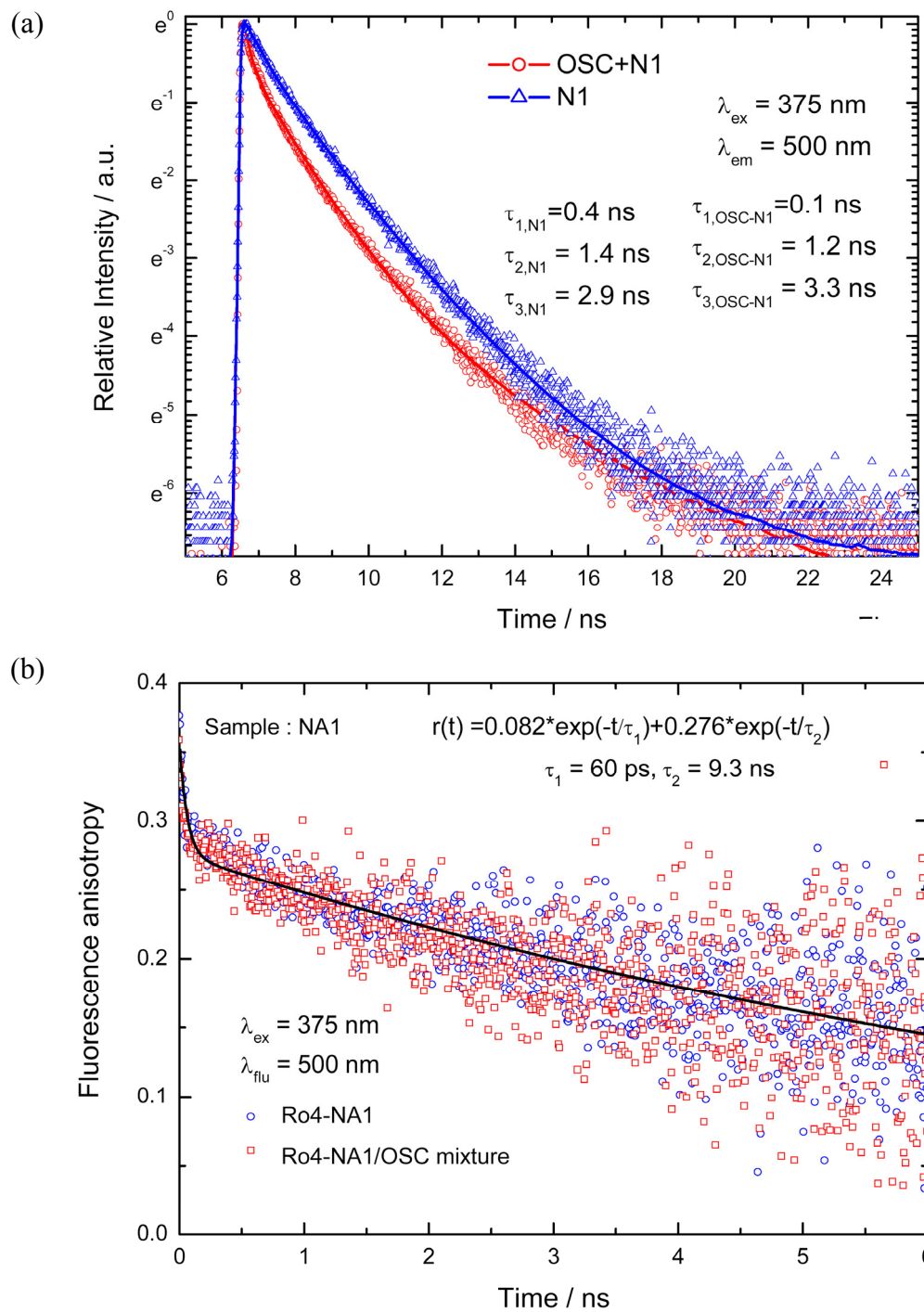


Figure 7.10 (a) The laser induced fluorescence decay, or (b) the anisotropy analysis of the Ro4-NA1 fluorescent probes. The transient fluorescence intensity decay from either sole Ro4-NA1 in HB buffer (red line) or the Ro4-NA1 pre-incubated with OSC (blue one) are measured and compared, respectively.

## Chapter 8

### *Identification of DNA Aptamers for Bovine Liver*

### *Oxidosqualene-Lanosterol Cyclase via Systematic Evolution of Ligands by Exponential Enrichment*

#### 8.1 Introduction

The functional nucleic acids that specifically bind and regulate the cellular protein were first discovered in 1989.<sup>196</sup> Interestingly, Turek and Gold then established a standard generation procedure for identification of RNA ligand with high affinity for T4 DNA polymerase as well as various organic dyes.<sup>197</sup> Such functional nucleic acid molecules were termed as “aptamer” by Ellington and Szostak.<sup>198</sup> Aptamer molecules were derived from an *in vitro* SELEX (Systematic Evolution of Ligands by EXponential enrichment) selection.<sup>197, 198</sup> In general, SELEX consists of several iterative steps. The first step of SELEX process is to construct a single-stranded nucleic acid library. The chemically synthesized oligonucleotide containing a centralized random sequence segment (typical about 30 ~ 40 nucleotides) flanked with two fixed terminal sequences was directly used as the starting material library for the first round of DNA-SELEX or subjected for the reverse transcription to generate the material for the first round of RNA-SELEX. Due to the restricted chemical synthesis scale, the diversity of random library was commonly limited to  $10^{15}$  distinct molecules even though the theoretical value could be composed of  $10^{24}$  molecules ( $4^{40} = 1.2 \times 10^{24}$ ). However, this kind of initial library was big enough to generate a high probability for obtaining a desired aptamer. Second, this random

sequence library was incubated with the target molecule and the different washing strategies were subsequently chosen for separating the unbound oligonucleotides from the target/oligonucleotide complex. For example, for the small organic molecules which were pre-immobilized on the solid supporting matrix, the buffer solution was enough to remove the unbound oligonucleotides. For the protein targets, the filter binding assay with nitrocellulose membrane or the gel mobility shift assay were usually selected to analyze the binding behavior between oligonucleotides and proteins. In addition, the affinity chromatographic column or the ELISA microtiter plates were also been utilized for the protein immobilization during the SELEX process. Subsequently, after dissociating the bound oligonucleotides from the target molecules via harsh or denaturing extraction, the standard ethanol precipitation was carried out to concentrate and purify the residual oligonucleotides. The next step is to amplify the resultant nucleic acids pool via the PCR-based reaction. Alternatively, the single-stranded DNA separation or the reverse transcription was performed to obtain the starting materials for the next round of SELEX selection. Theoretically, the next SELEX cycle utilized the newly generated nucleic acids library will enrich those oligonucleotides which have higher affinity for the same target molecule than that of previous SELEX round. The affinity of bound oligonucleotides exponentially increases for each round of SELEX cycles. After several repeating rounds of these stages, the nucleic acids with high affinity for the target molecules were isolated via general molecular cloning technology and followed with DNA sequencing analysis.

Interestingly, a derivative of SELEX technology, photo-SELEX, from which the substrate of either DNA polymerase or RNA polymerase was exchanged for the 5'-iodo or 5'-bromo dUTP or UTP in the PCR reaction has been included in the normal SELEX procedure.<sup>199-201</sup> These evolved aptamer molecules allow the photo-crosslinking reaction occurred between target protein and the photo-aptamer

molecules after exposing them under long wavelength UV light. This kind of laser induced photo-crosslinking process has been mentioned recently to further facilitate the design of aptamer-assisting diagnostic array system.<sup>201, 202</sup> In parallel, the random double-stranded DNA library has also been used in the SELEX procedure to identify the dsDNA consensus for the interest protein.<sup>203, 204</sup> Moreover, in order to isolate nucleic acid aptamers with the extremely high specificity, the counter selection strategies or the negative selection strategies were broadly carried out during the SELEX procedure.<sup>205</sup> On the other hand, the stringent selection conditions might also be applied to adjust the binding characteristic of the resultant nucleic acids. Theoretically, the nucleic acid molecule with highest specificity and the highest affinity for the target molecule could be isolated and confirmed for its binding property via different biophysical studies. The “winner” of nucleic acid isolated from the *in vitro* section was named “aptamer”, which was originally derived from Latin word “aptus”, meaning “to fit”. Moreover, after completely sequencing of the full-length of aptamer, the essential region for the binding could be determined via design of the different truncate forms of aptamer. Usually, the nucleotide sequences that are not important for direct binding with target or for facilitating the correct structure folding should be eliminated from the full-length. The scheme for the selection of DNA aptamer to the specific target molecule is shown in Figure 8.1. Interestingly, the specificity of existing aptamer could also be further changed for another target molecule via the second selection of SELEX for the cognate targets by using a biased library which was produced from the existing aptamer.<sup>206</sup> For example, the specificity change from L-arginine to L-citrulline of definite L-arginine-aptamer molecule could be used to demonstrate the power of the *in vitro* selection for optimizing or modifying the specific aptamer on demand.<sup>206, 207</sup>

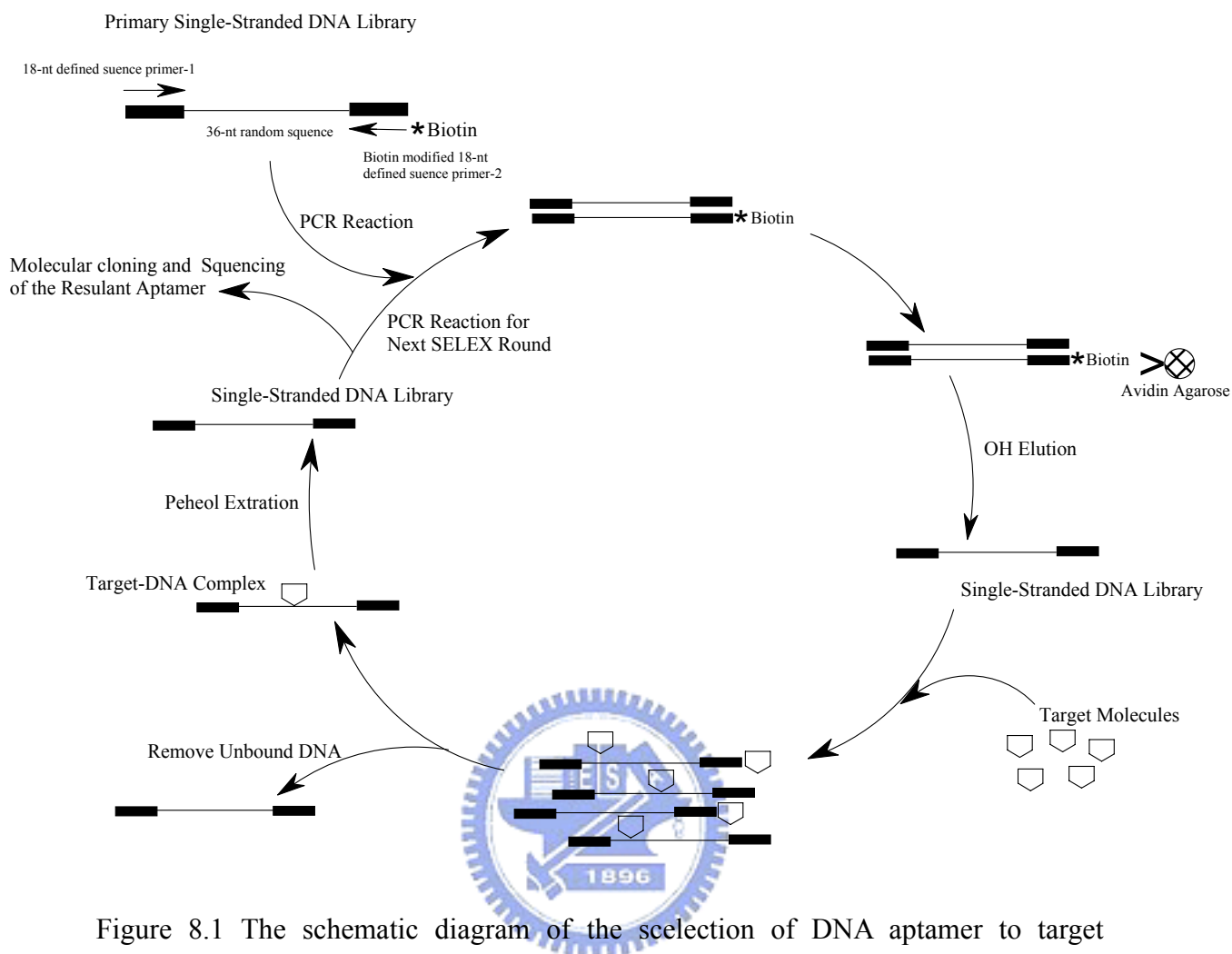


Figure 8.1 The schematic diagram of the selection of DNA aptamer to target molecule, *i.e.* the general SELEX procedure.

As we know, for a given target molecule, both antibody and aptamer exhibiting comparable binding affinity as well as high specificity could be theoretically isolated. However, the production of antibody was derived from the *in vivo* immunoreactions, whereas the aptamer was derived from the *in vitro* SELEX procedures. Moreover, aptamers could also be obtained via chemical synthesis. Thus, different reporter molecules such as biotin, radioisotope, fluorophores, could be easily modified on the skeleton of aptamer molecules. Aptamer molecules could also specifically targeted at the non-immunogenic toxic molecules and be applied for the non-physiological selection condition. Hence, the target range of antibody was narrower than that of the

aptamer molecule. Moreover, the batch to batch variation of definite antibody caused by the *in vivo* immunoreactions process was significantly bigger than that of relative distinct production of aptamer. Further, the thermal denaturation was also reversible for an aptamer molecule. Importantly, the specificity and discrimination for an aptamer could further be regulated easily during the aptamer screening procedure. Thus, the application and stability of aptamer is broader than that of antibody molecules. The advantages of aptamer over antibody and other unique characteristics of both molecules are listed in Table 8.1.

Table 8.1 The Comparison of functional characteristics between aptamers and antibodies.<sup>208</sup>

	Aptamers	Antibodies
Affinity	Low nM - pM	Low nM - pM
Specificity	High	High
Production	In vitro process	In vivo process
Target range	Wide: ions, small organic molecules, protein, whole cells, toxic molecules, etc	Narrow: only immunogenic compounds
Batch to batch variation	Little	Significant
Chemical modification	Easy and straightforward	Limited
Thermal denaturation	Reversible	Irreversible
Shelf-life	Unlimited	Limited

Up to date, various targets including organic dye<sup>209</sup>, amino acids<sup>210</sup>, biological cofactors<sup>211</sup>, antibiotics<sup>212</sup>, peptides<sup>213</sup>, proteins<sup>214</sup>, even complex mixtures such as whole virus particle<sup>215, 216</sup>, cell membrane<sup>217</sup>, live trypanosomes<sup>218</sup>, and whole mammalian endothelial cell<sup>219</sup> have been selected to isolate their own aptamers from diverse library, further indicating that the aptamer molecules have the property of broad target range. The online aptamer database listing all of the selected aptamers was publicly available on the internet (<http://aptamer.icmb.utexas.edu/>). The high target affinity and high discrimination specificity of aptamer molecules allowed them



to be used as the recognition elements in the diagnostic application. The first fluorescent-tag modified aptamer was utilized in the flow cytometry system in 1996.<sup>220</sup> After then, different recognition elements involving the functional aptamer molecules were dramatically developed. These engineered aptamer molecules have been employed in the diverse array of diagnostic systems containing the quartz crystal microbalance<sup>221</sup>, surface plasmon resonance technique<sup>222</sup>, electrochemical sensor<sup>223-228</sup>, and different FETs fabricated with single-walled carbon nanotubes (SWNTs)<sup>229, 230</sup>. Among these, the electrochemical sensors coupled with functional aptamers, compared to that of antibody-form biosensors, have another obvious advantage during the regeneration process. The bound protein could be easily removed from the electrode surface without damaging the sensing aptamer molecule itself, whereas the harsh washing step would denature the bound antigen as well as the antibody in the antibody-form biosensors. Hence, the repeatable utilization and reversible stability of aptamer molecules provide the superiority over the diagnostic systems based on the antibody.

Conversely, many aptamer molecules not only display high specificity and high affinity for the particular target but also disrupt the respective biological activity of this target. Thus, the aptamer molecules are also been referred as one of the prime potential drug candidates. For the purpose of application of aptamers in pharmaceutical medicine, the stability of nucleic acids is the critical issue for prior consideration. The endonuclease-resistant aptamer was developed after changing the 2'-OH group of ribose to 2'-NH<sub>2</sub> or 2'-F modifying sugars on the pyrimidines moiety. This kind of chemical modification provided not only the stability of aptamer but also the compatibility for enzymatic reaction in the SELEX protocol.<sup>231</sup> Moreover, the phosphorothiolate linkage modification<sup>232</sup> or the chemically synthesized enantiomer aptamer based on the parent-SELEX against the enantiomer of target provided

additional clues to produce the nuclease-insensitive ligand.<sup>233</sup> On the other hand, the PEG conjugation<sup>234</sup> or the attachment of aptamers on the surface of liposome<sup>235</sup> have also been illustrated to enhance the circulation time, to avoid the renal excretion, as well as to increase the efficacy of aptamers in body fluid. The “lead optimization strategy” which was broadly used in the medical chemistry field, to modify the “lead compounds” for obtaining a more effective drug could also be utilized in the aptamer procedures. Conspicuously, after a decade of preclinical development for optimizing its biological effects, the Pegaptanib, an anti-VEGF aptamer for ocular vascular disease, was notably shown in the clinical trials for being the potential first aptamer therapeutic agent for use in humans.<sup>236</sup>

After the determination of sequence, the structures of aptamer-ligands complex could further be solved by using different methods of NMR spectroscopy. Thus, these structures could be utilized to elucidate clearer insights about the crucial interaction involved in the binding between the definite aptamer and its ligand molecule. Compared with the antisense nucleic acids, aptamer molecules can form the characteristic three-dimensional structure via their intrinsic intramolecular interaction or the intermolecular interaction with the target molecules.<sup>202</sup> Usually, the continued mismatch of nucleotide sequences often contributed the characteristic binding pocket for small ligand molecules. Notably, these flexible and unstructured regions of aptamer will become the well-organized structure when aptamer molecule interacted with its respective ligands.<sup>237</sup> Moreover, the base-base stacking, base-triples formation, hydrophobic interactions, hydrogen-bonds interactions, electrostatic interactions, and different sets of edgewise interactions from the base of nucleotides represented the combination of different stabilizing forces within the aptamer-ligand complex and also contributed the discrimination among the relative ligand analogs.<sup>238-240</sup> In parallel, the complexed structures between aptamers and relatively large protein have also been

resolved via the X-ray crystallography.<sup>241, 242</sup> Compared with the previously illustrated aptamer-small molecule interactions via the formation of specific binding pocket to incorporate its respective ligands, the protein-bound aptamers always formed unique folds for integrating into the structure of larger protein.<sup>243</sup> Interestingly, from the results of the well-resolved aptamer-ligand complex, the solved structures derived from the *in vitro* SELEX selection were highly similar to the binding structures found in nature.<sup>239, 244, 245</sup> Hence, these *in vitro* selected aptamers not only could provide the high affinity and specificity for the diagnostic or therapeutics application but also facilitate the understanding about structural characterization in natural occurring systems. The building architecture, interaction patterns, or regulation principle involved in either protein-nucleic acid interactions or small molecule-nucleic acids interaction could thus be revealed particularly.

As previous chapters described, the great biological and medicinal significance of oxidosqualene cyclase fulfills a unique role in designing new antifungal, hypocholesteremic, and phytotoxic drugs. The preliminary information obtained from sequence comparisons, site-directed/saturated mutagenesis, and mechanistic inhibition studies have provided a basis for rational designing of drugs for therapeutic purpose. On the other hand, the utilization of random selection approaches to gain the novel type of inhibitors was rarely applied to the oxidosqualene cyclase. The combinatorial chemistry methodology, a technique for screening very large combinatorial libraries of biopolymers against a specific target via an iterative *in vitro* selection process, offers the convenience of repeatable amplification of their members and also makes the screening process rate faster and easier. These combinatorial libraries either based on replicable nucleic acids or polypeptides have been utilized for screening the definite protein binder to act as their inhibitors. The combinatorial polypeptide-based phage display technology has also been used to find the

OSC-binding peptide in our laboratory.<sup>246</sup> In order to extend the research field of oxidosqualene cyclase, we try to utilize another *in vitro* random selection methodology, SELEX, to isolate the high-affinity binding ligands for oxidosqualene cyclase. As previously, SELEX begins with a large population of single-stranded nucleic acid molecules and usually challenges with the specific task. The task herein is to find the aptamer molecule specifically bind to the purified bovine liver OSC protein. Theoretically, single-stranded oligonucleotides with the high affinity for the bovine liver OSC could be isolated from the random sequence library. Especially, if the binding domain of this high-affinity aptamer located at or near the active site of the OSC protein, it might be regarded as a potential drug-candidate to inhibit the OSC activity. After the iterative SELEX selection, the nucleotide sequence of the corresponding aptamer molecule is determined by using the automatic DNA sequencing. Subsequently, the radioisotope-labeled ssDNA from the individual clones are used to confirm the binding behavior between the aptamers and OSC via the gel mobility shift assay. The dissociation constant of individual aptamer are also determined using the surface plasmon resonance (SPR) in a BIACORE instrument. Finally, the sequence analysis and the secondary structure predication of individual aptamer molecules are determined by the ClustalW and the mfold program, respectively.

## 8.2 Results and Discussion

### 8.2.1 *In vitro* selection of DNA aptamer for bovine liver OSC

The synthetic single-stranded DNA library was first subjected to the PCR amplification with two defined primers, from which one of terminal primers was pre-modified with biotin molecule. This PCR amplicon was subsequently applied to

the Avidin-agarose bead for immobilizing the biotinylated double-stranded DNA. After washing with binding buffer, the single-stranded DNA library was generated and collected via an elution of 0.2 N NaOH solutions, and then purified by using the ethanol precipitation. The polystyrene microtiter plates which possess the hydrophobic surface for immobilization of various macromolecules were chosen for the first six rounds of SELEX selections.<sup>247, 248</sup> Accordingly, the DNA concentration used in every round was kept constant at 50 pmol/well, whereas the amount of coated bovine liver OSC was decreased in the last two rounds of SELEX selection in order to enhance the stringency of the selection. The total amount of DNA and bovine liver OSC utilized in the SELEX experiment are described in Table 8.2. In order to eliminate the non-specific DNA binder for the polystyrene as well as the blocking solution, the pre-selection of DNA library for the microtiter plates solely coated with blocking solution was first carried out. This “panned” polynucleotide library was then loaded into the protein-coated wells for the SELEX selection. After adequate washing with binding buffer, the guanidine thiocyanate solution was then used to disrupt the absorption of protein on the surface of microtiter plates as well as the DNA molecules which interacted with the coated protein. The residual DNA molecules via guanidine thiocyanate elution from the surface of the microtiter plate were collected, precipitated, and amplified with the conserved primers. The recovery of DNA amplicon was analyzed on the agarose gel with SYBR green I staining. The reactions without immobilizing OSC protein were used as background control in all rounds of SELEX selection. The PCR amplicon recovered from these “negative wells” were also compared with the PCR products obtained from the wells coating with the target protein. The staining gel from SELEX-3<sup>rd</sup>-Cycle selection exhibited a clearer DNA band which corresponds to the correct PCR amplicon, indicating that the negative-selection was successful to remove the non-specific DNA binder of the

supporting matrix. After the sixth round of SELEX cycles, the binding ability of resultant DNA to the OSC protein was further determined by gel mobility shift assay. The representative patterns of DNA recovery from the OSC-immobilizing wells are shown in Figure 8.2.

Table 8.2 The concentration of OSC or DNA library load into each SELEX round. The total volume of either protein or DNA is fixed at 200  $\mu$ l. These values just donate the amount loaded into each wells. It does not represent the amount of absorbed macromolecules.

Round	OSC / $\mu$ g per well	DNA/ pmol per well
1	30	50
2	30	50
3	30	50
4	30	50
5	15	50
6	3	50

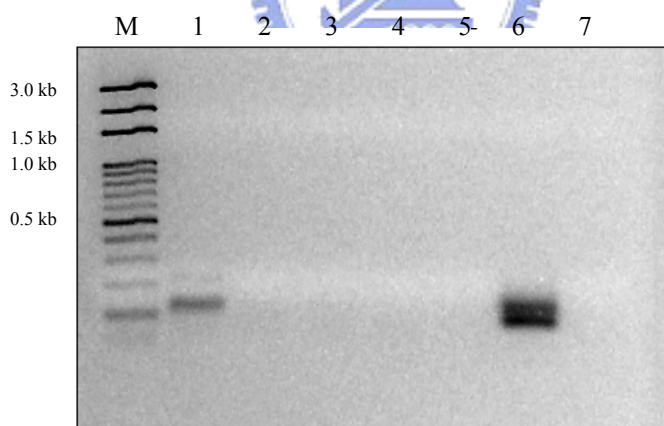


Figure 8.2 PCR amplicon of the resultant DNA, which was eluted from the surface of the microtiter plates, during SELEX-3<sup>rd</sup>-cycle. Lane-M is the commercial 3.0 kb marker consists of 100 bp to 3,000 bp DNA fragment mixtures. Lane-1 to Lane-5 represent the PCR products of eluted DNA from the linearly diluted OSC (from 30  $\mu$ g to 3 ng) coating wells. Lane-6 is the PCR product of original single-stranded DNA library. Lane-7 represents the PCR product of eluted DNA from the BSA-coated well, indicating that the SELEX experiment with “pre-panning procedure” effectively removed the DNA binder for other than cyclase protein.

### 8.2.2 Gel mobility shift assay of the binding between DNA aptamer and OSC

After finishing six rounds of SELEX selection on microtiter plates, the guanidine thiocyanate eluted DNA was amplified and purified according to the method described above. In order to examine the binding ability of these DNA and bovine liver OSC, the gel mobility shift assay was carried out. The single-stranded DNA library was obtained and labeled with the adenosine 5'-[ $\gamma$ - $^{32}$ P] triphosphate. Different concentrations of DNA probes were incubated with constant concentration of OSC in the total volume of 40  $\mu$ l at 37°C for 30 min. These reaction mixtures were then fractionated on the non-denatured polyacrylamide gel to analyze their binding ability for the bovine liver OSC. The salmon sperm DNA and the BSA protein were used as non-competitive controls, respectively. According to the result of autoradiography obtained from the Image Scanner, the retardant radioisotope bands were clearly observed in the reaction mixture containing the OSC protein. On the other hand, non-binding pattern appeared on the detected gel in the reaction with sole labeled DNA (Lane-2 and Lane-4, Figure 8.3). Moreover, under the fixed concentration of cyclase protein, the clearer retardant radioisotope band was observed in the reactions with the larger amount of DNA probe (Lane-2 and Lane-3, Figure 8.3). In addition, various concentrations of OSC were reacted with the fixed amount of [ $\gamma$ - $^{32}$ P] labeled DNA. As shown in Figure 8.4, the retardant radioisotope bands become darker and correlated with the increasing concentrations of OSC. It suggests that the interaction between aptamer and OSC is dose-dependent.



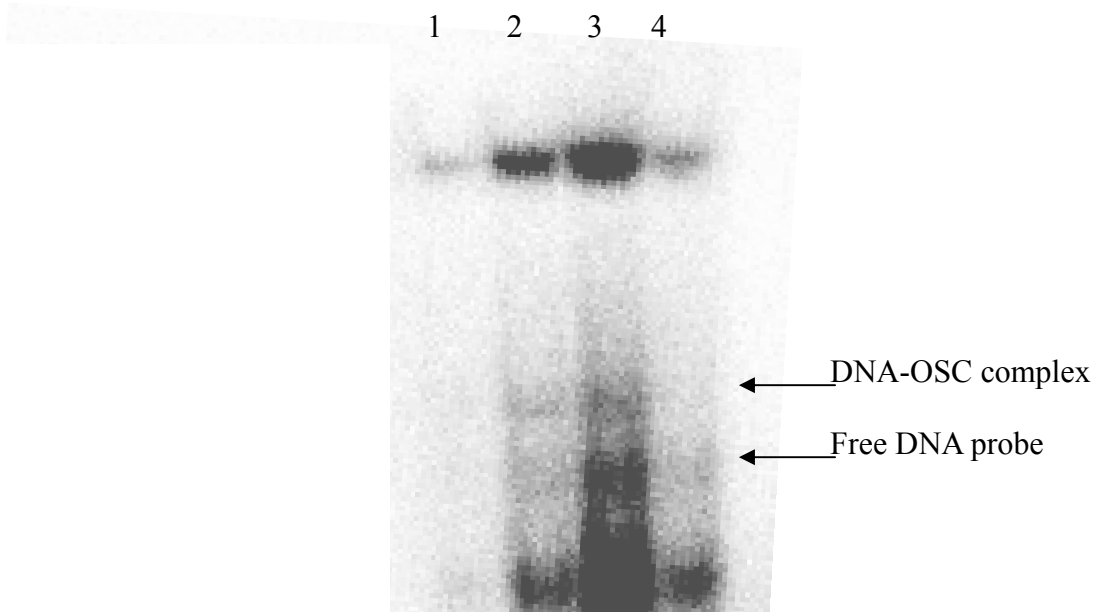


Figure 8.3 Gel mobility shift assay of different amounts of DNA probe interacted with the fixed bovine liver OSC. Lane-1: 2.25 ng DNA probe incubated with OSC. Lane-2: 11.25 ng probe incubated with OSC. Lane-3: 22.5 ng probe incubated with OSC. Lane-4: sole 11.25 ng DNA probe. The retardant radioisotope DNA band is observed in the binding reaction with cyclase protein

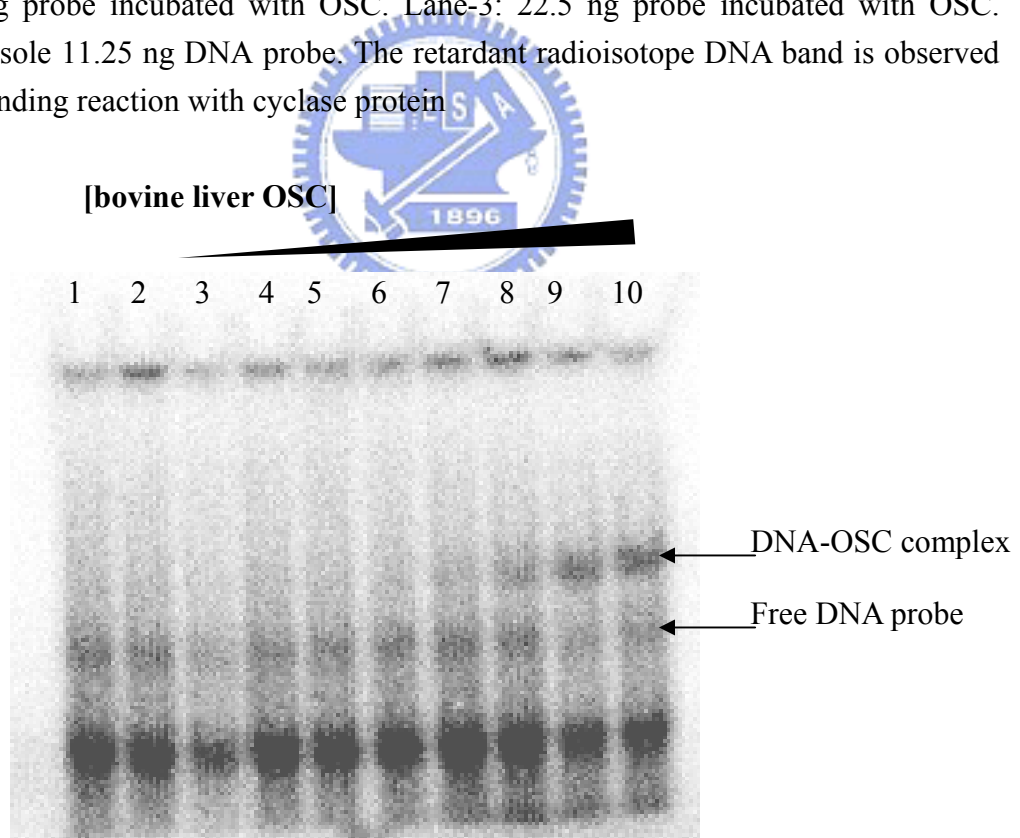
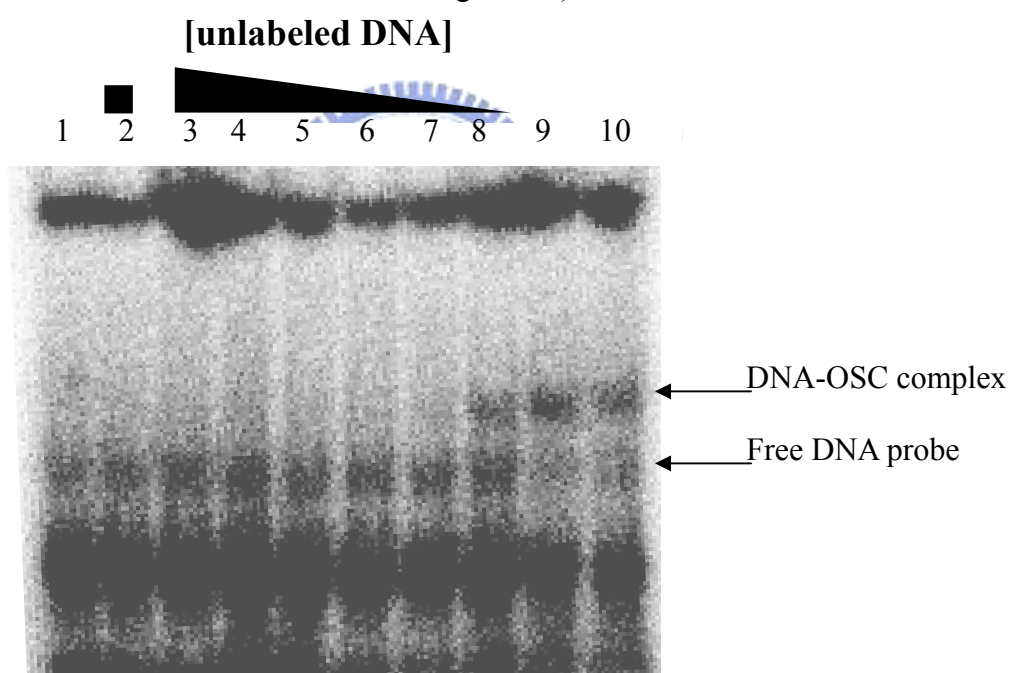


Figure 8.4 Gel mobility shift assay of fixed amount of DNA probe (11.25 ng) interacted with the various amounts of bovine liver OSC. Lane 1-2: Binding reaction without cyclase. From Lane 3-10: The binding reactions were added with increasing amount of the cyclase (from 0.1 ng to 10  $\mu$ g).

In order to confirm the specificity of interaction between bovine liver OSC and the DNA probes, the competition experiment with the unlabeled probe was also carried out. As shown in Figure 8.5, the addition of excessive amount of unlabeled DNA will interfere the binding between  $[\gamma\text{-}^{32}\text{P}]$  labeled DNA probes and the bovine liver OSC. Hence, the decrement of intensity of the retardant radioisotope bands correlated with the addition of unlabeled DNA (From the radioisotope bands of Lane -3 to Lane-10 in Figure 8.5). Without addition of bovine liver OSC, the excess unlabeled DNA could not produce any retardant radioisotope bands, indicating the binding specificity between DNA probe and the cyclase protein (Comparison of radioisotope bands in Lane-1 and Lane-2 in Figure 8.5).



**Figure 8.5** Gel mobility shift assay of competition experiment by using the unlabeled DNA probe. The excess unlabeled competitive probes were added into the reaction mixture. The amounts of the unlabeled competitive probes in Lane-3 to Lane 8 are more concentrated than that of  $[\gamma\text{-}^{32}\text{P}]$  labeled DNA probes (from 40-fold to 1-fold), respectively. Lane-9 and Lane-10 is the positive control without addition of any unlabeled competitive probes. On the other hand, no OSC was added in the Lane-1 and Lane-2 well, except 20-fold concentration of unlabeled competitive probes was exclusively added in the reaction of Lane-2, which was used to indicate the binding specificity between cyclase and DNA probe.

### 8.2.3 Gel mobility shift assay provide alternative selection strategy for SELEX

Subsequently, the observed retardant radioisotope bands from the gel mobility shift assay were excised and purified.<sup>249</sup> The obtained PCR amplicon exhibited similar size to that of primary single-stranded DNA library, indicating that the retardant radioisotope bands were the reliable complex of cyclase and the enriched DNA library (Figure 8.6). The electrophoretic gel mobility shift assay (EMSA) could provide another selection strategy to screen the aptamer molecules. Thus, the PCR amplicon from the retardant radioisotope bands was subjected to produce the novel single-stranded library for next round of SELEX cycles. Another three rounds of SELEX (SELEX-7<sup>th</sup>-cycle to SELEX-9<sup>th</sup>-cycle) were carried out through the EMSA strategy. The retardant radioisotope bands were observed among the each round of SELEX and provided the new materials for the next SELEX round.

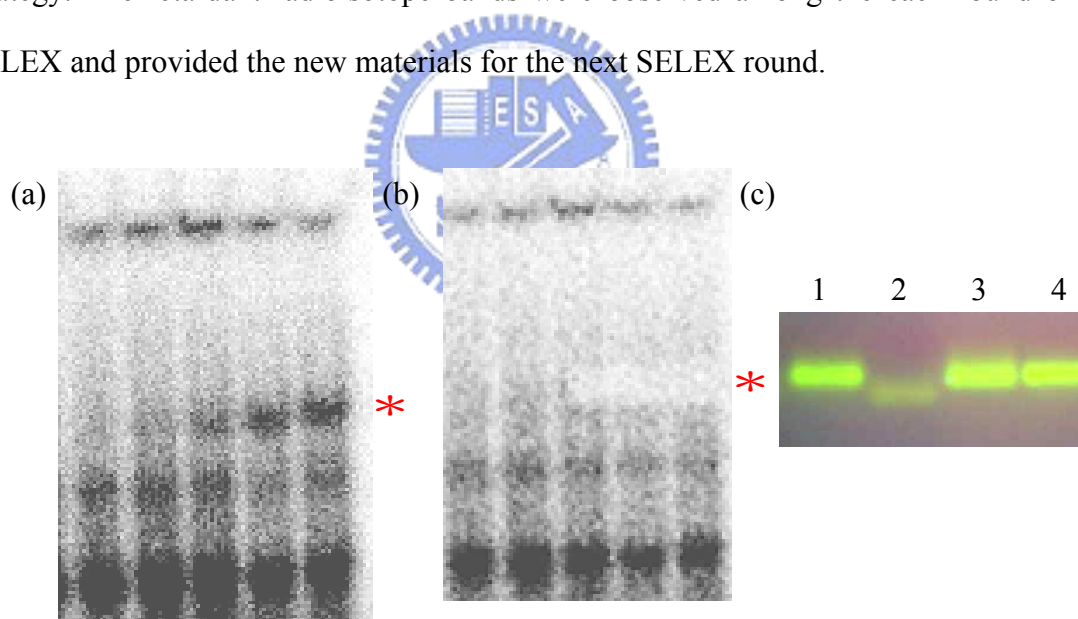


Figure 8.6. Gel mobility shift assay of isolated DNA molecules for the bovine liver OSC. The symbol ” \* ” on the (a) and (b) represent the region of retardant radioisotope bands for excision and purification. (c) The PCR amplicon of recovered DNA fragment from the excised region, indicating the reliability of EMSA. Lane-1 is the PCR product of primary single-stranded DNA library. Lane-2 is the negative control without addition of templates. Lane-3 and Lane-4 are the PCR products of excised fragment of retardant radioisotope bands.

#### 8.2.4 Molecular cloning and DNA sequencing analysis of individual aptamer

After completing the ninth round of SELEX selection, the autoradiography band of DNA retardation was isolated and cloned into the pCR2.1-TOPO vector. The full lengths of sequence in the respective colonies were subsequently determined. We sequenced twenty-six colonies isolated from the retardant radioisotope bands of the SELEX-9<sup>th</sup>-cycle, and another fifteen colonies isolated from the SELEX-6<sup>th</sup>-cycle. No consensus sequence was detected from the sequenced clones. Theoretically, the distribution of nucleotide diversity should be narrow down via iterative screening of DNA library population against the same target molecule. The conserved motif, which is involved in the binding between target and DNA nucleotide, should be “panned” from the vast ocean-like single-stranded DNA library. However, the non-conserved sequences among the entire population of the selected aptamer clones have previously been discovered from other independent investigations.<sup>250-252</sup> Smolke, *et al.* isolated approximately 60 kinds of aptamer sequences for an unique small molecule target, codeine, and none of the selected clones exhibited the predominant predicted structure.<sup>250</sup> Similarly, Ikebukuro’s group also founded a diverse array of DNA sequences targeting at two kinds of proteins, Vascular endothelial growth factor (VEGF) and quinoprotein glucose dehydrogenase (PQQGDH).<sup>251, 252</sup> The sequences of aptamer clones from our SELEX-6<sup>th</sup>-cycle and SELEX-9<sup>th</sup>-cycle experiments respectively, are listed in Figure 8.7 and some representative secondary structure predicted from mfold server are also included in Figure 8.8.<sup>253</sup> According to the results of secondary structure prediction, almost all of the selected aptamer molecules adopt a unique folding structure. However, the secondary structure could not completely reveal the binding characteristics between target protein and the respective nucleic acids. The fairly critical motifs for protein binding should further be determined via other structure-based analyses or experiments with truncated mutation

construction.

(a) 5'-GGATCCGAGCTCCACGTG-36N-GCTAGCGTCGACCGTACG-3'  
variation regions

(b)

125-6-14	-----CCATACAGACGGTTATTACACCTGCCTACCTCTCC--	36
126-6-10	-----CCGGAACGTTTCATAGACCCATCACTAACTAGGCGG	36
125-6-08	-----CGGAGGCTACGACGCAAAGTATCTAACAGCTCGCC---	36
126-6-12	-----AGGATGCTAAGCTGCAACAGGGTGGAACTAAGCCCC---	36
126-6-03	-----ACACTAA-CATTATCTGCACCCTCGCCTCCCC-----	31
126-6-07	-----CCATGCTAAACACCAAGGACCCCAACGTTACTGCCC--	36
125-6-04	-----ACAACCTCACCCGCAAACAATTCCCTGACCCCTGTC---	36
126-6-05	-----CACACCGCTACCCGCCAATGACCTTCACCCACCG-----	35
125-6-10	GCAGAGGGTCTCGTTAAGCGATAATCATCCTCACTC-----	36
125-6-06	-----ACATCATCCTCCGAGGCAGTTATCGTCCCATGCCCC--	36
126-6-06	----CTGCCCATTTCCCTTCTACCTCATGCTACCCAGGGC-----	37
125-6-09	-----CCCACCAGCCATCCGTGTCATAACGCCCTCACTGC-	36
126-6-04	-----GGCACCTCGTTCCCGTGTAACCAGTGTGAGGGCGC---	36
126-6-09	----CCACTCACACAAGCTTCTATTTGCGCTTCCC-----	31
121-6-01	-----CCCCACCCTCACATTGAAGAAACCGTCATACATGCC---	37

(c)

121-9-01	-----CCACTACACGCCTTAACCACAGTGCCTCGCCGCTC-----	35
125-9-13	-----CCACT-CACGCTTCGAAC-CCGCGCCTACCCCCATC----	34
125-9-11	-----CAAG-CACCCAGCAATC-CCCAACCTCGCCCCGTGTACC	37
125-9-12	----GGCATAACACACCATGAACTCTCGAATAACATATGC-----	36
125-9-04	---CAACCCTAACCTCAACGACAAATCTAACCCCTCCCC-----	36
126-9-10	-----CCAGCATGCGTACGACAA-TCCGCCACTCCTCCGGCC--	36
125-9-06	-----AGCAGCTGCAAAAACCTCTCCATCCCCCCCCCGGGCC-	37
125-9-15	-----CAAACCTCC--CAGAGCATGTGGATCACACTACCCCTGC-	36
126-9-01	-----ACACTCCAGCCGAGATTGTGTACTG-AGTGCC-----	31
126-9-05	-----CCACTAACTCCCCTGCCTAACTACTTTCTCAGCC----	36
126-9-09	-----GGCTAACAAGTAGGTCCCACCCGCTAGCCCACACCC---	36
125-9-02	----CCCTCCATTTCCCCCGTGCTACACCTGCACCGGGCC-----	36
125-9-03	-----GCCATGTCACCC-CGCTATCCCAACACCGTTGCTCGC-	36
126-9-07	-----CGTCCATCCCTGGTGGAAACATACGCTCCCCTATGCC----	36
125-9-07	CCCCCCTGCACGAAAGACACAG-CCTGTGCCTACCC-----	35
125-9-10	---CCGCACACGACCAATGCA--CGTACGCACACTATCGCCC---	37
125-9-09	----CCCACAAACCCCAAGCA--CTAATGCCTCCTCCCGGCC---	36

```

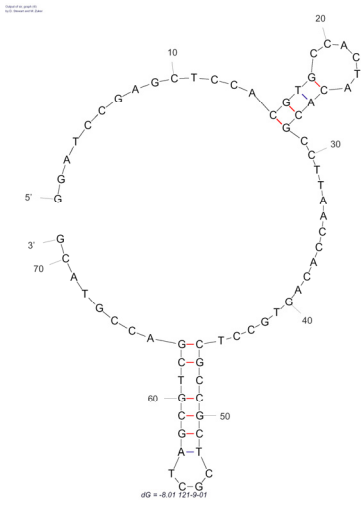
126-9-04      -----CCACAAACCCGTTTCCTCCTTATGCCTTCTCACTGCC--- 37
125-9-01      -----TACAGAC-CAAGGAGTTACAGGC-CCACCCCTGCCCTCG 37
125-9-08      -----CCAAAAACACTAACAGCCACATACGTTACCGCCGCC---- 36
126-9-11      ----GGCACAATC---AACAGATAAAGAC-TCTCCCATGCTGCC- 36
126-9-12      --GGGTCCAA-ATCACAGCACAGCTT--AACGACAGACTCC---- 36
126-9-02      -----CCCGGCACAGCATCACC-CGAGAAACCACAGGCCCCC--- 36
126-9-06      -----CCCA--ACATCATCACCGCATGAAACCGCCTAATGCC--- 35
125-9-05      -----CCACCGTGATGATAGCGTCGGGGACGCTGTCCGGCC--- 36
126-9-08      -----CGCAGCCATCATCATTTGAGCATAATCCCAACCAGCC--- 36

```

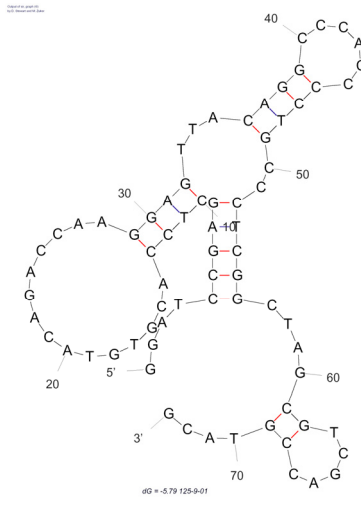
Figure 8.7 Sequence alignment of the respective aptamer clones. (a) single-stranded DNA template for generating primary DNA pool, from which the variation regions were chosen for the following sequence alignment; (b) 15 sequences of the aptamer clones from the SELEX-6<sup>th</sup>-cycle; (c) 26 sequences of the aptamer clones from SELEX-9<sup>th</sup>-cycle



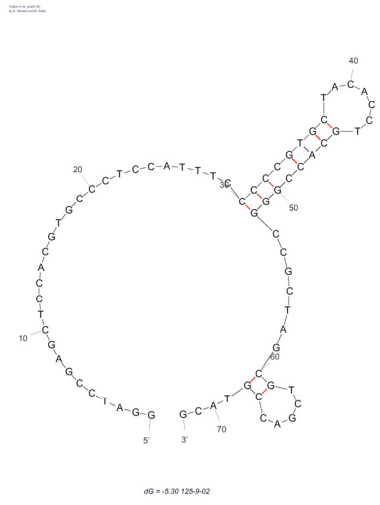
(121-9-01)



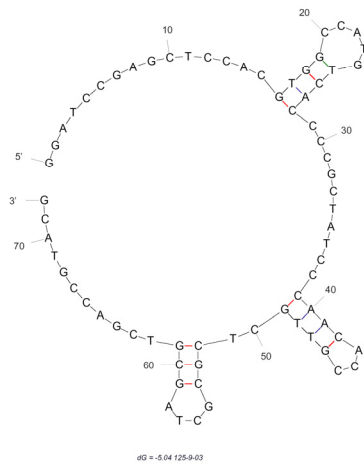
(125-9-01)



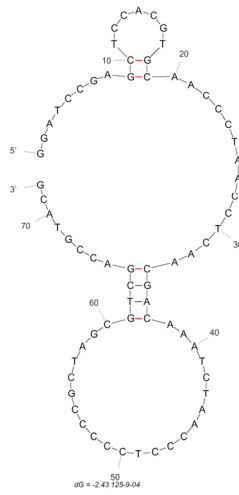
(125-9-02)



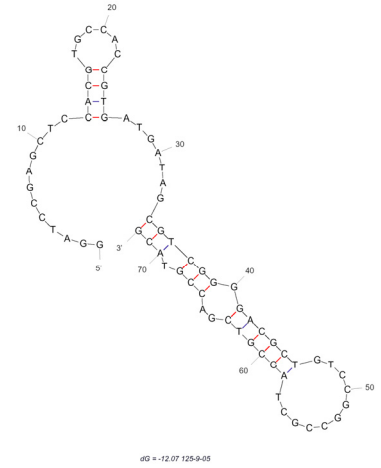
(125-9-03)



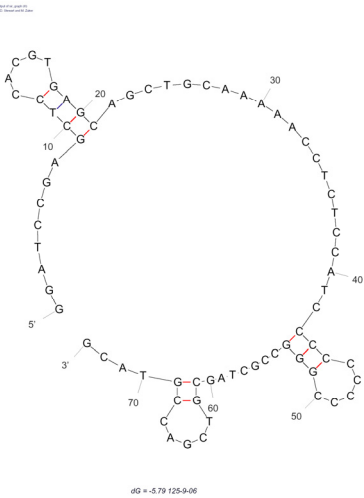
(125-9-04)



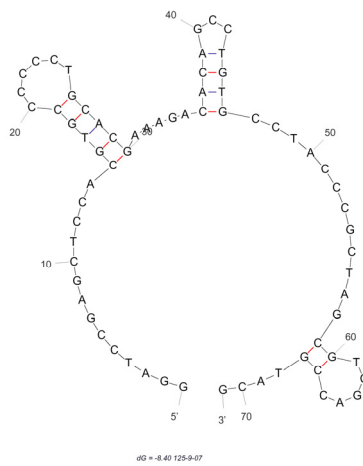
(125-9-05)



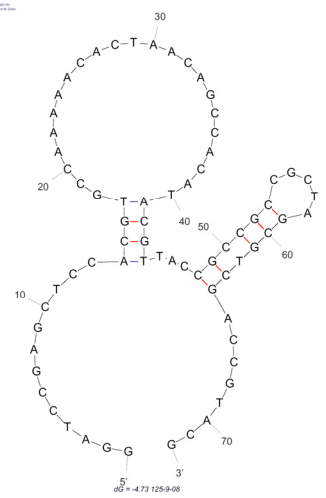
(125-9-06)



(125-9-07)

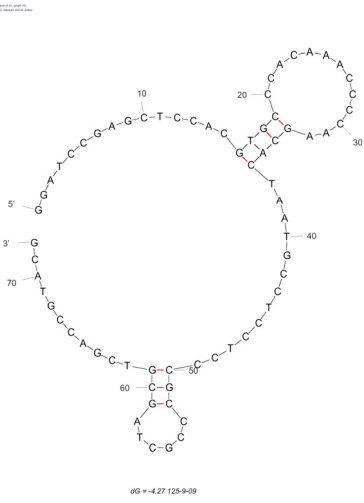


(125-9-08)

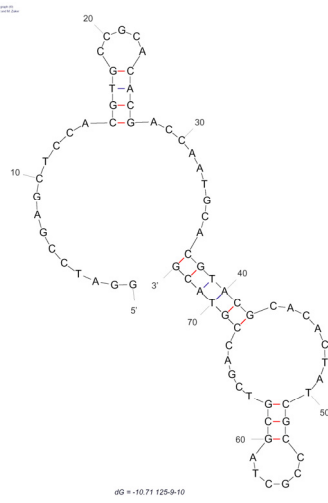




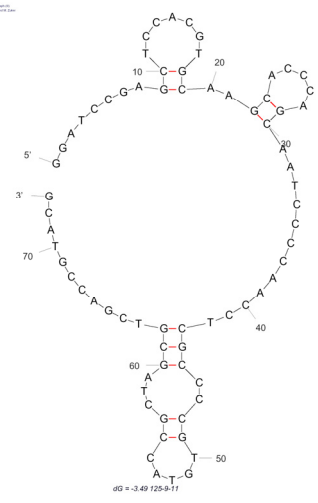
(125-9-09)



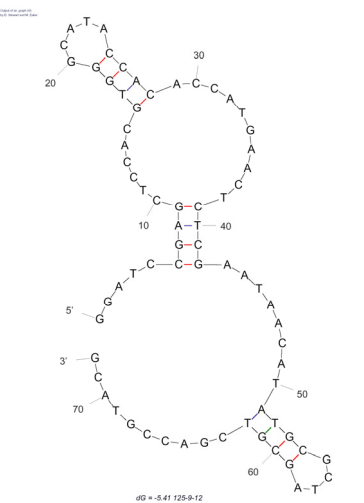
(125-9-10)



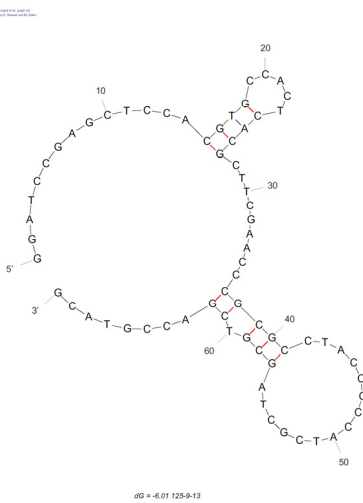
(125-9-11)



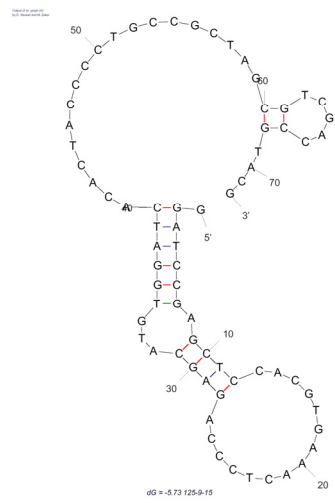
(125-9-12)



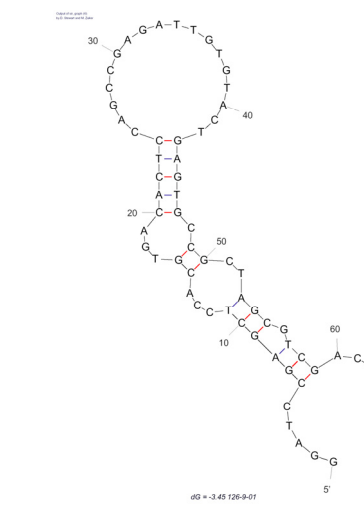
(125-9-13)



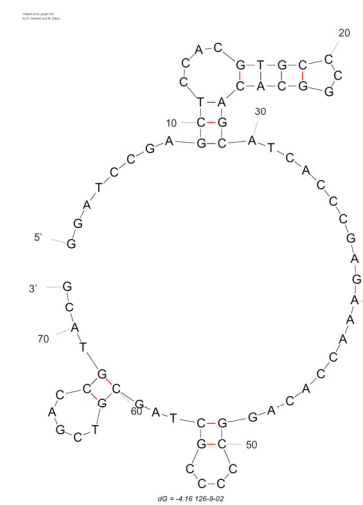
(125-9-15)



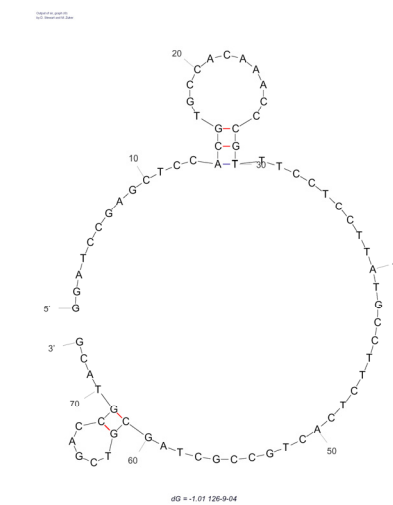
(126-9-01)



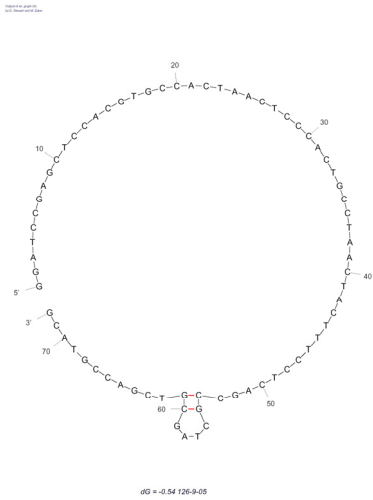
(126-9-02)



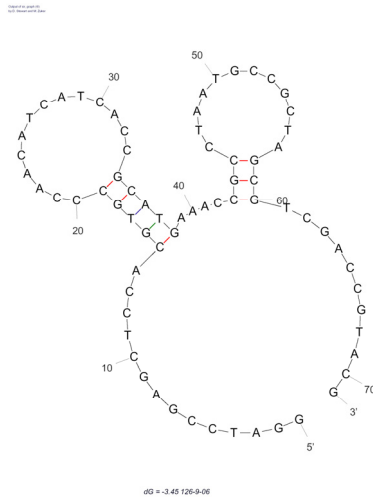
(126-9-04)



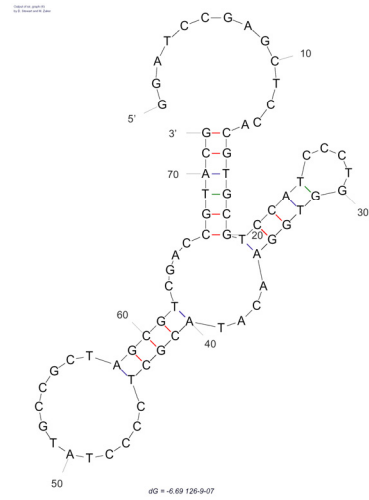
(126-9-05)



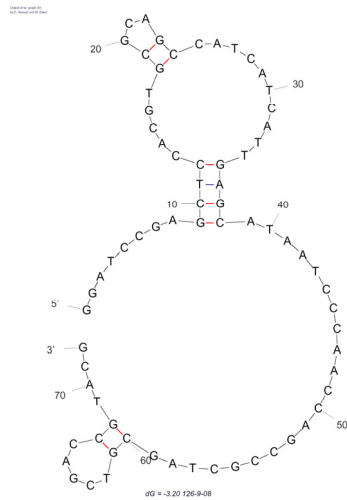
(126-9-06)



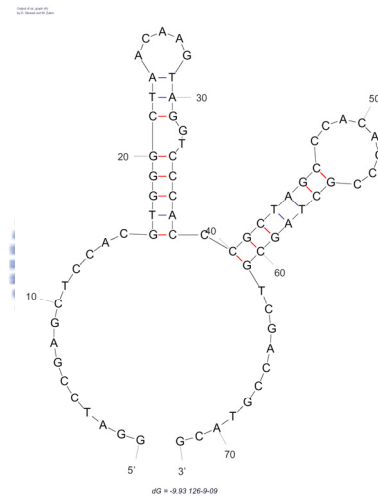
(126-9-07)



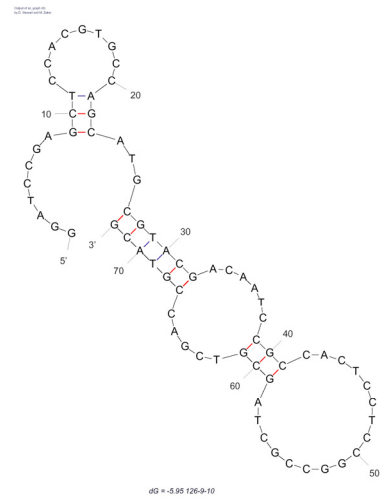
(126-9-08)



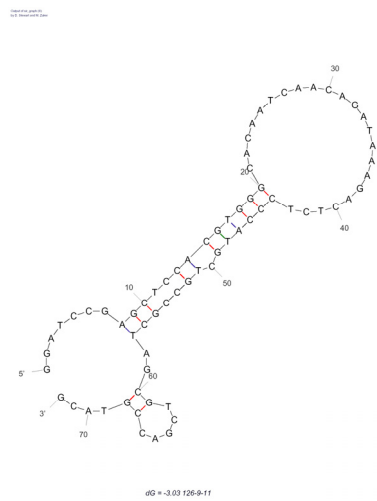
(126-9-09)



(126-9-10)



(126-9-11)



(126-9-12)

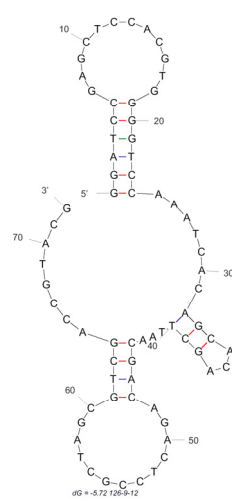


Figure 8.8 The representative predicted secondary structures of aptamers from the SELEX-9<sup>th</sup>-cycle, via the mfold sever prediction.<sup>253</sup>

### 8.2.5 The binding characterization of the individual aptamer molecules

After successfully cloning the respective aptamers from the enriched library, the individual single-stranded DNA aptamer with the uniquely characterized sequence was utilized to illustrate their binding property for bovine liver OSC. As shown in Figure 8.9, the individual aptamers still interacted with the purified OSC and resulted in the obvious retardant radioisotope band on the EMSA gel. Compared with the previous parent SELEX library, the retardant band was clearer among these unique aptamer molecules. The composition of unique-sequence DNA molecules might eliminate the contamination of other weak-binding DNA molecules from the distribution of the previously enriched parent SELEX library. The representative EMSA gel is shown in Figure 8.9. On the other hands, the divergent retardant pattern produced from some individual aptamers revealed the different binding characteristics of aptamers for the bovine liver OSC protein. Further investigation should be carried out for better understanding about the interaction between individual aptamers and the cyclase protein.

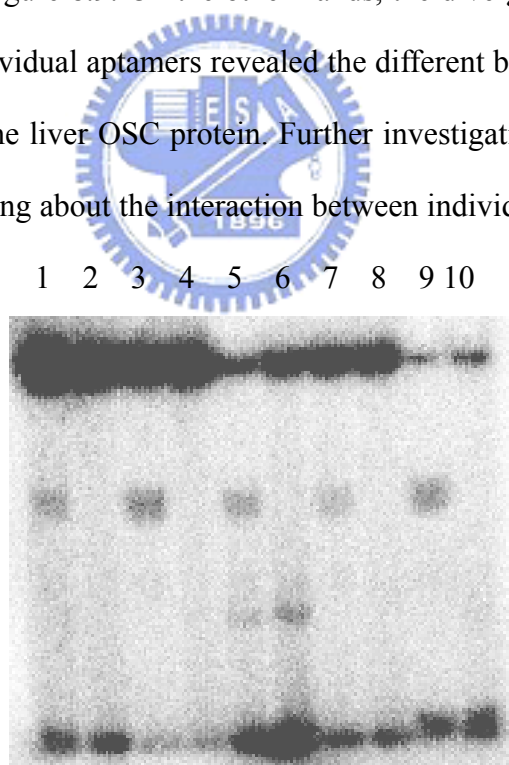


Figure 8.9 Gel mobility shift assay of individual DNA aptamers for the bovine liver OSC. Among these reactions, Lane-1/Lane-2, Lane-3/Lane-4, Lane-5/Lane-6, Lane-7/Lane-8, and Lane-9/Lane-10 represent the different sets of EMSA experiment, from which the purified OSC protein was only included in the odd numbers of binding reaction, respectively. Accordingly, the clear retardant band is only observed in the reaction containing cyclase protein.

Among these single-stranded aptamer molecules, aptamer-125-9-13 exhibited the apparent interaction with bovine liver OSC from the result of gel mobility shift assay. Next, in order to quantify the equilibrium dissociation constant of the interaction between bovine liver OSC and this aptamer molecule (aptamer-125-9-13), the surface plasmon resonance assay (SPR) was carried out. Notably, the biotinylated position of the primer pair described herein is distinct from the primer pairs used in the screening procedure. The end of original biotinylated primer used in the selection is changed to the native “hydroxyl group”, whereas the end of another primer is exchanged from the native “hydroxyl group” for the “biotin group”. This reversed biotin modification is used for immobilizing the sense-stranded of aptamer on the surface of flow cell. Following the immobilization of denatured biotin-labeled aptamer molecules on the surface of the flow cell of the sensor chip, serial dilution of cyclase protein sample were injected into these flow cells. The free biotin molecules were also immobilized on the surface of the other flow cell. The detected response from the biotin-immobilized flow cell was used as the blank subtraction for evaluating the binding constant. The corresponding  $K_d$  value was thus generated from the observation of the concentration-dependent binding response data. Consistently, this aptamer (125-9-13) also displayed the obvious binding behaviors with cyclase protein and has the dissociation constant of approximately 2 nM, from the measurement of surface plasmon resonance (SPR) in a BIACORE instrument. The sensorgram profile of aptamer-125-9-13 for different concentrations of bovine liver OSC was shown in Figure 8.10. The measured binding kinetic of this aptamer was also included.

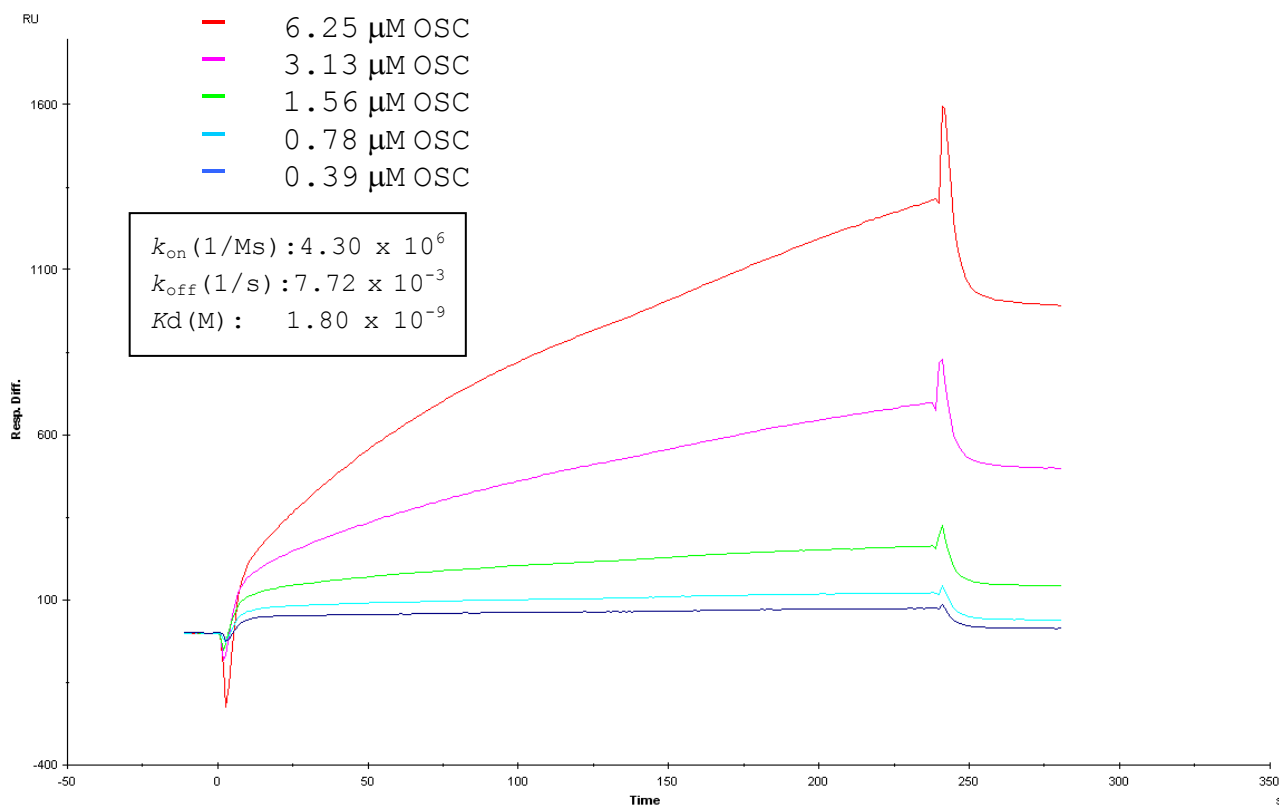
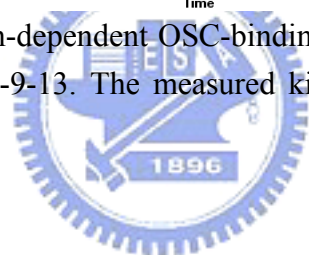


Figure 8.10 The concentration-dependent OSC-binding response was revealed on the SPR analysis of aptamer-125-9-13. The measured kinetic data of this aptamer was also included in the inset.



### 8.3 Conclusion

In this section of thesis, we have used an *in vitro* selection methodology, SELEX, to find the single-stranded DNA binder for the bovine liver OSC. The microtiter plates and EMSA assay were alternatively utilized in our SELEX experiments. After nine rounds of SELEX cycles, a diverse array of aptamer molecules was respectively isolated from the original single-stranded library. However, this kind of sequence divergence has also been observed in other reported research. From the results of gel mobility shift assay, most of the selected aptamers exhibited the expected interaction with the target protein. The non-labeled competitive probe was also used to illustrate the specific binding between target and DNA molecule. One of the aptamer molecules was further used to characterize its binding properties by using the SPR instrument, and its kinetic values for bovine liver OSC was also obtained. In the future, structure-based investigations, including the mutagenesis studies or the construction of the truncated form aptamers will be carried out to examine the interaction between individual aptamers and the bovine liver OSC. These aptamer molecules might provide clues for developing different applications for oxidosqualene-lanosterol cyclase in either diagnostics or the pharmaceutical fields in the future.

## Chapter 9

### Experimental Section

#### 9.1 Materials

##### 9.1.1 Bacterial strains and molecular cloning/expression vectors

Except stated elsewhere, bacterial growth condition, agar plate preparation, recombinant DNA purification, DNA sequence determination, and agarose gel electrophoresis were performed according to standard procedures or the commercial kit.<sup>249</sup> Phagemid pBluescript IISK(+), *Escherichia coli* XL1-Blue were purchased from Stratagene. Plasmid vector pRS314 was kindly provided by Dr. Botstein (Stanford University). The *ERG7* gene including the promoter sequence was cloned into pRS314 vector to create the pRS314ERG7WT plasmid. In addition, in order to facilitate the expression of extrogeneous cyclase in yeast, the promoter sequence from *ERG7* gene was fused in front of the recombinant *CAS1* gene to create the pTKP5 plasmid.<sup>106</sup> The haploid yeast strain CBY57[pZS11] (*ERG7* $\Delta$ ::*LEU2 ade2-101 lys2-801 his3- $\Delta$ 200 leu2- $\Delta$ 1 trp1- $\Delta$ 63 ura3-52* [pZS11]) was derived from the diploid strain CBY1 (MATa/ $\alpha$  *ERG7* $\Delta$ ::*LEU2 ade2-101 his3- $\Delta$ 200 lys2-801 trp1- $\Delta$ 63 ura3-52*).<sup>254</sup>

##### 9.1.2 Enzymes, chemicals, equipments, and reagents

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

(2) Chemicals: The following section lists the chemicals utilized in this thesis. The manufacturers were included in the square bracket.

Acetic acid [Merck], Acetic anhydride [Sigma], Acetone [Merck], 40% Acrylamide solution [GE Healthcare], Adenine [Sigma], Adenosine 5'-[ $\gamma$ -<sup>32</sup>P] triphosphate [GE Healthcare], Agarose-LE [USB], Aluminum chloride [Merck], Ammonium persulfate (APS) [Gibco BRL], Ampicillin sulfate [Sigma], Anisaldehyde [Merck], Bacto<sup>TM</sup> Agar [DIFCO], Biphenyl-3-boronic acid [Alfa], Biphenyl-4-boronic acid [Alfa], Benzamidine [Sigma], 4-Bromobenzoyl chloride [Fluka], Bromophenol blue [USB], Coomassie brilliant R250 [Merck], Cyclohexane [Merck], Deoxyribonucleotide triphosphate (dNTP) 100mM Solutions [GE



Healthcare], 1,6-dibromohexane [Fluka], Dichloromethane [Merck], Dimethyl sulfoxide [Merck], 4-(4,5-diphenyl-1H-imidazol-2-yl)phenylboronic acid (DPA) [ACROS], Di-potassium hydrogen phosphate (KPi) [Merck], Dithiothreitol (DTT) [GE Healthcare], DNA 10kb Ladder [BioBasic Inc.], D-Sorbitol [Sigma], Ethanol (95% and 99%) [Merck], Ether [Merck], Ethyl acetate [Merck], Ethylenediamine-tetraacetic acid (EDTA) [Merck], Ergosterol [Sigma], 3-Fluoroanisole [Aldrich], Fumaric acid [Sigma], Geneticine (G418) [Gibco], Glucose [Sigma], Glycerol [Merck], Hemin chloride [Merck], Hexane [Merck], Histidine [Sigma], Hybond ECL nitrocellulose membrane [GE Healthcare], Hydrobromic acid 48% [Riedel-de Haen], Hydrochloric acid 37% [Merck], Lanosterol [Sigma], LB Broth (Miller) [DIFCO], Lysine [Sigma], Methanol [Merck], Methioine [Sigma], N-allylmethylamine [Fluka], Naphthalene-1-boronic acid [Alfa], Naphthalene-2-boronic acid [Alfa], N-bromosuccinimide (NBS) [Sigma], Nitrobenzene [Merck], N,N-dimethylacetamide [Merck], N,N'-Methylene-bis-acrylamide [GE Healthcare], Oligonucleotide Primers [BioBasic Inc.], Pinacol [ACROS], Phenylmethylsulfonyl fluoride (PMSF) [Sigma], Potassium carbonate [Merck], Potassium chloride [Merck], Potassium hydroxide [Merck], Polyvinylidene difluoride (PVDF) membrane [GE Healthcare], Pyridine [Sigma], Pyrogallol [Merck], Sea sand [Merck], Silica gel [Merck], Silver nitrate [Merck], Sodium bicarbonate [Merck], Sodium carbonate [Merck], Sodium dodecylsulfate (SDS) [Gibco], Sodium sulfate [Merck], SYBR<sup>®</sup> Green I [Roche], Squalene 99% [ACROS], Tetrahydrofuran (THF) [Merck], Trichloroacetic acid (TCA) [Merck], TLC plate [Merck], Tris-base [USB], Triton X-100 [Sigma], Tryptone [DIFCO], Tryptophan [Sigma], Tween 80 [Merck], Yeast Extra [DIFCO], Yeast Nitrogen Base w/o amino acid [DIFCO], Uracil [Sigma]

These chromatographic columns were obtained from GE Healthcare: Q-Sepharose Fast Flow, HiTrap Heparin Columns, and disposable PD-10 Desalting Columns, whereas the Hydroxyapatite HT Gel was purchased from Bio-Rad. Chromatography Silica Gel 200~450 mesh was obtained from Fisher. Agarose-binding Avidin D column was prepared by Vector Labs.

(3) Equipments: The following instruments were used in this thesis: ABI PRISM<sup>®</sup> 3100 Genetic Analyzer [Applied Biosystems], Allegra<sup>™</sup> 21R Centrifuge [Beckman Coulter], Avanti<sup>®</sup> J0E Centrifuge [Beckman Coulter], Colling Circulator Bath Model B401L [Firstek Scientific], Centrifuges 5415R [Eppendorf], DU 7500 Spectrophotometer [Beckman Coulter], Electrophoresis Power Supply EPS 301 [GE healthcare], GeneAmp<sup>®</sup> PCR System 9700 Thermal Cycler [Applied Biosystems], Hoefer<sup>®</sup> HE 33 Mini Horizontal Submarine Unit [GE Healthcare], Kodak Electrophoresis Documentation and Analysis System 120 [Kodak], Orbital Shaking

Incubator Model-S302R [Firstek Scientific], Pulse Controller [BioRad], Rotary Vacuum Evaporator N-N Series [EYELA], Steritop™ 0.22µm Filter Unit [Millipore], 32 Karat™ System Gold® HPLC [Beckman Coulter], FL-4500 Fluorescence Spectrophotometer [Hitachi], Agilent 6890N gas-chromatography equipped with a DB-5 column [Agilent], Agilent Technologies model 5973 MSD mass spectrometer [Agilent], Instant Imager [Packard Bioscience]

(4) Reagents:

**30% Acrylamide/1% Bisacrylamide solution**

Dissolve 5 g N,N'-Methylene-bis-acrylamide into 375 mL 40% Acrylamide solution. Adjust the volume into 500 mL with ddH<sub>2</sub>O. Store it at 4°C.

**50 X ALTHMU solution**

0.2% Adenine, 0.3% Lysine, 0.2% Tryptophan, 0.2% Histidine, 0.2% Methonine, and 0.2% Uracil in ddH<sub>2</sub>O. Sterilize and store it at 4°C.

**50 X ALHMU solution**

0.2% Adenine, 0.3% Lysine, 0.2% Histidine, 0.2% Methonine, and 0.2% Uracil in ddH<sub>2</sub>O. Sterilize and store it at 4°C.

**ALHMU/Heme/Ergosterol plate**

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

**Ampicillin stock solution (100mg/mL)**

Dissolve 1 g Ampicillin sulfate in 10 ml ddH<sub>2</sub>O. Filter it through a 0.22 µm pore size filter and stock it at -20°C.

**10% APS solution**

Dissolve 1 g Ammonium persulfate into 10 ml ddH<sub>2</sub>O.

**Blocking buffer**

1 X PBS buffer + 5% skim milk.

**Development solution**

7.5 g Na<sub>2</sub>CO<sub>3</sub>, 125 µL formaldehyde in 250 ml H<sub>2</sub>O.

**6 X DNA loading dye**

0.25% Bromophenol blue and 30% glycerol in ddH<sub>2</sub>O.

**20% EA developing solution**

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

**Ergosterol supplement solution**

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

**G418 stock solution (1 g/mL)**

Dissolve 500 mg G418 in 500 µl sterile ddH<sub>2</sub>O. Store it at 4°C under darkness.

### **50% Glucose solution**

500 g Glucose was dissolved in 1 L ddH<sub>2</sub>O, and sterilized it.

### **80% Glycerol solution**

80 ml Glycerol was added in 20 ml ddH<sub>2</sub>O and sterilized. Store it at 4°C.

### **Heme solution**

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

### **Homogenization Buffer I (HB I), pH 7.4**

100 mM Tris-base, 1 mM EDTA-Na<sub>2</sub>, 1 mM DTT, 1 mM Benzamidine, and 40 µg/mL PMSF solution. Store it at 4°C. Adjuste the pH value to 7.4 before utilization.

### **Homogenization Buffer II (HB II), pH 7.4**

20 mM Tris-base, 1 mM EDTA-Na<sub>2</sub>, 1 mM DTT, 1 mM Benzamidine, and 40 µg/mL PMSF solution. Store it at 4°C. Adjuste the pH value to 7.4 before utilization.

### **Hydroxyapatite Buffer (HAB) and Heparin Buffer (HB), pH 7.4**

5 mM KPi, 1 mM DTT, and 0.5% Triton X-100 solution, pH 7.4. Store it at 4°C.

### **Ion Exchange Buffer (IEB), pH 7.4**

20 mM Tris-base, 1 mM EDTA-Na<sub>2</sub>, 1 mM DTT, 1 mM Benzamidine, 40 µg/mL PMSF, and 0.5% Triton X-100 solution. Store it at 4°C. Adjuste the pH value to 7.4 before utilization.

### **LB medium**

25 g LB Broth was dissolved in 1 L ddH<sub>2</sub>O and sterilized.

### **LB plate**

25 g LB Broth and 20 g Bacto™ Agar was dissolved in 1L ddH<sub>2</sub>O and sterilized. The sterile solution was poured and dispersed in Petri dishes before coagulation.

### **10 X LiOAc**

1 M Lithium acetate, adjusted to pH 7.5 with diluted acetic acid.

### **LiOAc/TE solution**

Add 1 ml 10 X LiOAc and 1 ml 10 X TE buffer into an 8 ml ddH<sub>2</sub>O solution.

### **PBS solution**

Mix 1.37 g Na<sub>2</sub>HPO<sub>4</sub>, 0.35 g NaH<sub>2</sub>PO<sub>4</sub>, and 8.77 g NaCl into 1 L ddH<sub>2</sub>O. Adjuste the pH value to 7.4

### **5 X Sample buffer**

Add 0.5 mL 1M Tris (pH 6.8), 0.8 mL Glycerol, 0.8 mL 20% SDS solution, 0.4 mL β-Mercaptoethanol, and 0.2 mL 0.05% Bromophenol blue into 10 ml ddH<sub>2</sub>O.

### **20% SDS solution**

Dissolve 10 g Sodium dodecylsulfate(SDS) into 50 ml ddH<sub>2</sub>O.

### **SDS Running Buffer**

Dissolve 14.4 g Glycine, 3 g Tris, and 1 g SDS into 1 L ddH<sub>2</sub>O.

### **SDS-PAGE staining solution**

1 g Coomassie brilliant blue R-250, 400 mL Methanol, and 100 mL Acetic acid were mixed. Adjust the volume into 1L with ddH<sub>2</sub>O.

### **SDS-PAGE destaining solution I**

400 mL Methanol and 100 mL Acetic acid were mixed. Adjust the final volume into 1 L with ddH<sub>2</sub>O.

**SDS-PAGE destaining solution II**

50 mL Methanol and 70 mL Acetic acid were mixed. Adjust the final volume into 1 L with ddH<sub>2</sub>O.

**SD medium**

0.17% Yeast Nitrogen Base w/o amino acid was dissolved in ddH<sub>2</sub>O and sterilized.

**5 X Sequencing buffer**

Dissolve 4.85 g Tris-base and 0.203 g MgCl<sub>2</sub> into 100 ml ddH<sub>2</sub>O, and adjust it to pH9.0

**1 M Sorbitol solution**

182.2 g D-sorbitol was dissolved in 500 ml ddH<sub>2</sub>O and sterilized. Store it at 4°C.

**10 X SYBR Green solution**

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

**50 X TAE buffer**

Add 242 g Tris-base, 57.1 ml Acetic acid, and 0.5 M EDTA into 800 ml ddH<sub>2</sub>O. Adjust the total volume into 1 L and pH value into pH 8.5. Store it at room temperature. Dilute the concentration to 1 X with ddH<sub>2</sub>O. Adjust the pH value to 7.5~7.8 before utilization.

**10 X TE buffer**

0.1 M of Tris-HCl and 0.01 M EDTA, pH 7.5

**TLC staining solution**

40 ml concentrated H<sub>2</sub>SO<sub>4</sub> was slowly added into 200 ml ethanol, followed by a addition of 12 ml Acetic acid and 16 ml Anisaldehyde. The solution was finally added with another 600 mL Ethanol with careful and slight agitation.

**Transfer buffer**

Add 3 g Tris, 14.4 g Glycine, 200 ml Methanol in to 500 ml ddH<sub>2</sub>O. Adjust final volume into 1 L with ddH<sub>2</sub>O.

**1.5 M Tris solution (pH 8.8)**

Dissolve 91 g Tris into 500 ml ddH<sub>2</sub>O. Adjuste the pH value to 8.8.

**1 M Tris solution (pH 6.8)**

Dissolve 61 g Tris into 500 ml ddH<sub>2</sub>O. Adjuste the pH value to 6.8.

**Washing buffer (1 X PBST)**

100 ml 10 X PBS, 900 ml ddH<sub>2</sub>O, and 1 ml Tween-20.

## 9.2 General experimental procedures

### Chapter 3

#### 9.2.3.1 Construction of alanine-scanning and site directed/saturated mutagenic plasmids.

For the construction of different alanine-scanning or site-directed/saturated mutagenic plasmids, pRS314ERG7WT plasmid was used as the template for PCR reaction. Mutated plasmids were constructed by using the QuikChang Site-Directed Mutagenesis kit (Stratagene Inc., La Jolla, CA) according to the manufacturer's protocols. After PCR amplification with designed primers, recombinant plasmids were subjected for *Dpn* I digestion to remove the parent plasmid, and transformed it into the XL1-Blue cells for DNA purification. The exact substitutions at corresponding position were determined by using an ABI PRISM 3100 auto-sequencer.

#### 9.2.3.2 Preparation of competent yeast cell (CBY57 and TKW14)

Pick the yeast TKW14 stock from refrigerator into a 3 ml SD medium containing 60  $\mu$ l 50 X ALTHMU solution, 120  $\mu$ l 50% Glucose solution, 60  $\mu$ l Heme solution and 60  $\mu$ l Ergosterol supplement solution, and then incubate it at 30°C for three days until turbidity. Transfer the yeast cells into a 100 ml SD medium containing the same nutrition additive condition, and incubate it at 30°C for another 12-18 hours. After OD<sub>600</sub> of yeasts arrived at 1.0~1.5, the yeast cells were collected by centrifugation at 3,000 rpm, 10 min at 4°C. Add 35 mL aseptic ddH<sub>2</sub>O to suspend the cell pellet, and centrifuge it again at 3,000 rpm, 10 min at 4°C. Repeat above step twice. Add 25 ml 1M D-sorbitol solution to suspend the cell pellet, and wash it by centrifugation. Finally, add  $n \times 50 \mu$ L ( $n$  represent the number of the transformed plasmids) 1 M D-sorbitol into the cell pellet, and suspend it gently on ice for 5 min. Before electroporation, 50  $\mu$ l competent cells were added into each of 1.5 ml microtube with 5  $\mu$ l recombinant plasmids, respectively. The preparation protocol for the yeast CBY57 strain is almost the same, except the yeast cultural growth medium are changed into ALHU solution and Glucose solution.

#### 9.2.3.3 Cyclase activity assay by using plasmid shuffle method in CBY57 strain

The pRS314-derived ERG7 mutated plasmids were electroporated into a yeast cyclase-deficient haploid strain CBY57[pZS11]. The pRS314 and pRS314ERG7WT plasmids, *TRP1* centromeric plasmids with no insert and the wild-type *S. cerevisiae* ERG7 gene, were used as the negative and positive control, respectively. The yeast transformants were plated on SD+Ade+Lys+His+1 M Sorbitol medium plates for three days to determine the presence of both pZS11 and pRS314-derived plasmids. Individual colonies were picked and grown in a 10 ml Ade+Lys+His+Ura liquid

culture. Aliquots of 100  $\mu$ l of each culture were simultaneously plated on the SD+Ade+Lys+His+Ura and SD+Ade+Lys+His+Ura+1 mg/ml 5'-Fluoroorotic acid medium plates, and grown them for another three days to elucidate the complementation effect. Colonies that grew on the non-5'-FOA plates, but not on 5'-FOA plates were separately picked and grown in SD+Ade+Lys+His+Ura liquid medium for mutated cyclase characterization. The presence of both pZS11 and pRS314-derived plasmids were verified by using yeast plasmid minipreps and restriction endonuclease analysis.

#### **9.2.3.4 Cyclase activity assay by using Ergosterol complementation in TKW14**

The mutated plasmids were transformed into TKW14 strain, a yeast *hem1/erg7* double deficient strain, by using the electroporation. Aliquots of 120  $\mu$ L culture were plated onto SD+Ade+Lys+His+Met+Ura+Hemin+Erg+G418 plates, and replated it on the same plates except without addition of Ergosterol, for selection of the cyclase complement ability. The pRS314 and pRS314ERG7WT plasmids were also transformed as negative and positive control, respectively. Different TKW14C2 [pERG7<sup>mutations</sup>] transformants were then separately grown in SD+Ade+Lys+His+Met+Ura+Hemin+Erg medium at 30°C with shaking (220 rpm) for seven to ten days for the products characterization.

#### **9.2.3.5 Lipids extraction and column chromatography.**

The yeast transformants with different cyclase mutated plasmids were grown in the 500 ml SD liquid medium containing Ade+Lys+His+Met+Ura+Hemin+Erg at 30°C with 220 rpm for several days. Yeast cells were harvested by centrifugation at 6,000 rpm for 10 min. The washed cell pellets were then suspended in a solution containing 50 ml Ethanol, 50 ml 30% KOH, and 0.1 g Pyrogallol.<sup>255</sup> This reaction was refluxed at 110°C for 3 h. The hydrolysate was extracted three times with total 500 ml Petroleum ether. The organic phase was collected and concentrated by using a rotary evaporator. The nonsaponifiable lipids were then dissolved in CH<sub>2</sub>Cl<sub>2</sub>, spotted on the TLC plates, and developed with 4:1 Hexane/Ethyl acetate solution. TLC plates were subjected to the staining buffer and heated until the products pattern appeared. In order to fractionate the different products pattern in these mutants, large scale of cultural mediums were harvested, washed, and saponified by refluxing as previously. The nonsaponifiable lipids (NSLs) were extracted and fractionated by using silica gel column chromatography with 19:1 Hexane/Ethyl acetate solution. The separated fraction was then concentrated and analyzed on the TLC plate. The same R<sub>f</sub> value spots were collected and separated in either native form or the acetylated form on silica gel or on the HPLC. The resultant products were analyzed by using GC, GC-MS, or different NMR spectra.



### 9.3.3.6 Acetylation modification and the alkaline hydrolysis reaction.

The acetylation modification of the triterpene alcohol products was performed according to the previous literature.<sup>256</sup> The dry triterpene alcohols were first dissolved in 2 ml pyridine. Then the excess amount of 10 mM Acetic anhydride was added. The solution was stirred overnight at room temperature. The acetylation reaction was monitored by TLC analysis. After 16 h, 5 ml of water was added to stop the reaction, and three times extraction with 10 ml CH<sub>2</sub>Cl<sub>2</sub> were carried out. The total organic phase was collected, dried over with Sodium sulfate, and then evaporated by using a rotary evaporator. The acetylated products were then applied for Silica chromatography, and analyzed by using GC-MS and NMR spectra. The acetylated derivatives could also be hydrolyzed to yield the native nonsaponifiable lipid.<sup>124</sup> The acetylated derivatives were refluxed in the 5% KOH/Methanol solution for 3 h. After extraction with Petroleum ether, the evaporated crude extracts were separated and analyzed, as previously described.

### 9.3.3.7 GC and GC-MS column chromatography

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

### 9.3.3.8 The novel gene disruption strain TKW14 construction

Haploid yeast strain TKW14 (MATa or MATα *ERG7Δ::LEU2 ade2-101 lys2-801 his3-Δ200 leu2-Δ1 trp1-Δ63 ura3-52 HEM1Δ::kan<sup>R</sup>*), a *hem1/erg7* double knockout strain, was derived from the haploid strain CBY57[pZS11]. A 1.6 kb DNA fragment containing the hybrid *HEM1/kanMX6* gene was obtained from PCR amplification with pFA6-KanMX6 plasmid as template. The amplicon from pFA6-KanMX6 plasmid encoded a G418-resistant gene.<sup>257, 258</sup> Two designed primers, Hem1-F and Hem1-R, not only process the 18 nucleotide sequences identical to G418-resistant gene but also have the flanking 60-70 nucleotide sequences, which is homologous to the upstream region of the start codon and the downstream region of the stop codon from the *HEM1* gene, respectively. Thus the homologous remobilization reaction was carried out in the yeast cell by using the lithium acetate transformation protocol.<sup>259</sup> First, CBY57[pZS11] cells were inoculated in a 3 mL YPD+Adenine medium, and



grew it at 30°C over night. The overnight cultural medium was transferred into a 25 mL fresh YPD+Ade medium and re-incubated at 30 °C with shaking until value of OD<sub>600</sub> to 1.0. After centrifugation, the cells were washed with a 4 mL LiOAc/TE solution. After twice washing steps, the cell pellet was suspended in a 1 mL LiOAc/TE solution. 100 µL of the concentrated cells were mixed with 3 µL salmon sperm DNA (10 mg/mL, boiled it for 1 min and then immediately put it on ice) and 3 µg transformed DNA. Next, 400 µL of 50% PEG<sub>3350</sub>/LiOAc/TE was added into the mixture. This cell suspension was mixed gently, incubated at 30°C with shaking for 30 min, heat shocked at 42°C for 20 min, and finally held on ice for 5 min. The transformants were plated onto SD+Ade+Lys+His+Met+hemin+750 mg/L G418 plates and grown at 30°C for three to five days. The resulting transformants were then re-grown overnight in a 3 mL SD+Ade+Lys+His+Met+hemin+750 mg/L G418 medium. Aliquots of 100 µL each culture were then plated on SD+Ade+Lys+His+Met+hemin+ G418+5'-FOA+Erg and SD+Ade+Lys+His+Met+hemin+G418+5-FOA plates for another three to five days to confirm the complementation effect, respectively. Correct targeting of the PCR-made KanMX6 module into the genomic locus was verified by using the PCR reaction with two specific primers (Heme1-N, and Heme1-C). The PCR amplicon and an *NcoI* fractionized fragment with expected size were observed from the Agarose gel, indicating the accurate homologous recombination. The four primers utilized in the disruption experiment are listed in the following.

Oligonucleotide Primers	Sequence
Heme1-F	5'-ATGCAACGCTCCATTTTTGCGAGGTTTCGGTAACTCCTCTGCCGC TGTTCCACACTGAATCGTACGCTGCAGGTCGAC-3'
Heme1-R	5'-T TACTGCTTGATACCACTAGAAACCTCTAGTTGTTTAAACGATGGG GTCTCTAACATTAGGGTTCAAGATCGATGAATTCGAGCTCG-3'
Heme1- N	5'-ATTTTTGCGAGGTTTCGG-3'
Heme1- C	5'-ACCACTAGAAACCTCTAG-3'

## Chapter 4

### 9.2.4.1 Construction of chimeric enzyme library between oxidosqualene-lanosterol cyclase and oxidosqualene-cycloartenol synthase

To set up the domain swapping study, we constructed 10 chimeric clones that were divided in two parts. In the first parts (Chimera-1~4), plasmid pRS314ERG7WT containing the wild-type *ERG7* gene and pTKP5 containing the wild-type *CASI* gene were individually digested with the restriction enzymes *NotI* and *PstI*. The smaller digested fragment of pRS314ERG7WT was ligated into the larger fragment of pTKP5 digested fragment to produce the Chimera-1, which included *ERG7* gene in the *N*-terminus and *CASI* gene in the *C*-terminus; The smaller digested fragment of pRS314ERG7WT was ligated into the digested native pRS314 plasmid to produce the Chimera-2, which only included the truncated *N*-terminal *ERG7* gene. Similarly, the Chimera-3 bearing the *N*-terminal *CASI* gene coupled with *C*-terminal *ERG7* gene, or the Chimera-4 including the truncated *N*-terminal *CASI* were also constructed.

The Chimeras-5~ -10 was constructed by using two stage PCR reactions, respectively. Primary PCR step consisted of two separating PCR reactions for two truncated regions. Each PCR reactions were performed with a hybrid primer, based on two different cyclases sequence in the joint region, and an outer primer, respectively. Another PCR reaction was also performed by using the complementary hybrid primer and another opposite direction outer primer. These two PCR reactions are carried out with different plasmids as templates, respectively. The primary PCR products were purified. The equal amount of primary PCR product were mixed and used as the template for the secondary PCR reaction with previous two outer primers. The amplified chimeric DNA fragments from secondary PCR were digested with *NotI/XhoI*, and ligated into the same enzymes digested pRS314 plasmid to obtain the Chimera-5 ~ -10 plasmids. The hybrid primer pairs used in the construction of chimeric enzyme library are listed in the following table.

Table 9.2.4.1 Primers sequence in the chimerical library construction.

Oligonucleotide Primers	Sequence
Chimeric-5	
O390C416-C1	5'-CAAATGGTGTGCAAACCTGGGATACAGGTTTTGCTATTCA-3'
O390C416-C2	5'-TGAATAGCAAAACCTGTATCCCAGGTTTGCACACCATTG-3'
Chimeric-6	
C416O390-C1	5'-ATAACGGAAGCCAGCTATGGGATTGTGCGTTTGCCATTCA-3'
C416O390-C2	5'-TGAATGGCAAACGCACAATCCCATAGCTGGCTTCCGTTAT-3'
Chimeric7	
O169C191-C1	5'-TACGTTTATTGGGTCTACCCAACGATGGAGATGGAGATAT-3'
O169C191-C2	5'-ATATCTCCATCTCCATCGTTGGGTAGACCCAATAAACGTA-3'
Chimeric8	
C191O169-C1	5'-TAAGGTTGCTTGGAGAAGGACCCAAGGACCACCCGGTTTG-3'
C191O169-C2	5'-CAAACCGGGTGGTCCTTGGGTCTTCTCCAAGCAACCTTA-3'
Chimeric9	
C546O524-C1	5'-CTTGTTAGAGCTAATCAACCCTGCTGAAGTTTTTGGTAA-3'
C546O524-C2	5'-TTACCAAAAACCTCAGCAGGGTTGATTAGCTCTAACCAAG-3'
Chimeric10	
O524C546-C1	5'-CAATGGAAACCTTGAATCCTGCAGAAACCTTTGGCGATAT-3'
O524C546-C2	5'-ATATCGCCAAAGGTTTCTGCAGGATTCAAGTTTCCATTG-3'

## Chapter 5

### 9.2.5.1 General Materials

Sequence information of human OSC or *S. cerevisiae* Erg7 were extracted from the NCBI database (*H. sapiens* OSC: p48449<sup>78</sup>, *S. cerevisiae* Erg7: p38604<sup>52</sup>). The crystal structure of human OSC was obtained from the Protein Data Bank (PDB entry 1W6K; structure of human OSC in complex with lanosterol<sup>105</sup>). Ligand and the detergent molecules in the 1W6K structure were removed before homology model building. Mutated residue was obtained by altering the wild-type gene. Different ligand molecules for the molecular docking experiment were generated by using Chem3D ultra software.

### 9.2.5.2 Sequence alignment, secondary structure prediction, and molecular modeling.

Multiple sequence alignments was generated by using the program CLUSTAL W in the websites (<http://www.ebi.ac.uk/clustalw/>).<sup>260</sup> The result of alignment was verified by considering the location of the previously reported highly conserved

residues. The secondary structure elements predictions were carried out via bidirectional recurrent neural networks (SSpro, and PHDsec program).<sup>131, 132</sup> The homology modeling structures were performed by using the Insight II module software (Accelrys software package interface). The plausible conformation of the ERG7 protein was generated based on the sequence alignment with the reference proteins, human OSC structure in PDB format file 1W6K. The ERG7 protein was firstly aligned with human OSC and then used to search the matching segments from the amino acid sequences and the conformational similarities. The resulting structure was built according to the default values of parameters. This homology modeling structure was then refined by energy minimization module in the SYBYL 7.0 program (Tripos application software interface). The hydrogen atom was first added on the protein, as well as on the ligand molecule. The charge of the atoms was calculated according to the default value of “Geister Huckel charge”, except for the atom of the various cationic intermediates, which was separately modified as formal charge at +1.00. The minimization algorithm was selected as “Powell”, and the termination method normally was set up to "Gradient" with a 0.05 kcal/mol. The total numbers of minimization steps depend on the energy convergence.

#### **9.2.5.3 Ligand docking within the ERG7 homology modeling**

The genetic algorithm (GA) for protein-ligand docking software “Gold” was chosen for calculating the docking modes of different cationic intermediates into the binding site of the newly built ERG7 homology modeling structure.<sup>194</sup> First, the ERG7 modeled structure file and the ligand molecule files were load into the “Gold”, respectively. Second, the “Gold active site residue” was defined at “atom number-3684”, which corresponds to the Cys-457 in the ERG7 protein. This atom is theoretically close to the center of the active site. The “active site cavity” was then defined as a diameter of 10 Å. After submitting the ligands into the Gold software, four to ten docking modes with slightly different orientation were generated in a “Gold output files”. The final model was chosen according to the detailed verification and comparison with the lanosterol molecule in the human OSC crystalized structure. After the determination of the docking mode, the ligand molecule was then added into the homology model file by using the “WebLab ViewerPro” Software (Molecular Simulations Inc). The coordination section of docked ligand was altered according to the HETATM records in the PDB format. The connecting bond and series number in the ligand were also carefully verified. The integrated file with the protein-ligand mixture was then uploaded into the “SYBYL” program, and saved as a “mol 2” file. Finally, the energy minimization with “SYBYL” was carried out as previously described, including the hydrogen atom addition, charge alteration, and energy minimization calculation.

#### 9.2.5.4 Quantum mechanical calculations protocol

The following three software packages were used for the quantum mechanical calculations, including Chemical Office Ultra 8.0, GaussView 3.0 and Gaussian 03. Chemical Office and GaussView calculations were carried out in the local computer, whereas the Gaussian 03 was executed via the PC cluster in the National Center for High-Performance Computing in Taiwan (NCHC, Taiwan). The preliminary chemical structure was built with Chemical Office and the respective Gaussian input files were also created by using Chemical 3D. The GaussView was then used to check the stereochemistry or the geometry of these compounds. Geometry optimization and single-point energies minimum was then calculated by using Gaussian 03 at the RHF approach (quantum mechanical method). The calculation bases used for optimization are from hf/3-21g, b3lyp/3-21g, to b3lyp/6-31g. In each basis setting, the calculation was completed until the stationary point was found, and the force items were converged. The positive value of individual vibration frequencies were got with the final basis setting: b3lyp/6-31g\* geom=check guess=read scf(maxcycle=999) freq(noraman) opt(maxcycle=100). These value represented the optimized structure are under the single-point energies minimum. The relative energy of these chemical molecules to oxidosqualene state was measured under their energy minimum states.

## Chapter 6

#### 9.2.6.1 Microsomal membrane preparation and purification of oxidosqualene-lanosterol cyclase from bovine liver

Homogenization and all subsequent purification procedures were performed at 4°C. Fresh bovine liver (300 g) were first blended with 600 mL HB I buffer at 4°C and then centrifuged for 30 min at 10,000 rpm to remove the cellular debris. The supernatant was subsequently centrifuged at 40,000 rpm for 1 h. The resulting microsomes were resuspended in HB I buffer (200 mL) and recentrifuged for 1 hr at 40,000 rpm. Washed microsomes were then suspended in 100 mL HB II buffer, supplemented with 0.5% Triton X-100, and gently stirred for 1 h. Crude enzyme extract was separated from the microsomal membrane fraction by centrifugation at 40,000 rpm for 1 h. The crude enzyme extract was first applied on a Q-Sepharose Fast Flow column (about 80 mL resin) which has been equilibrated with IEB buffer. The resin was washed with 320 mL IEB buffer and then eluted with 2 column volumes of IEB buffer containing 20, 50, 70, and 100 mM KCl. The functionally active fractions were pooled, dialyzed against HAB buffer, and concentrated using YM30 filter membrane. The cyclase-containing sample was then loaded on the hydroxyapatite column (50 mL) equilibrated with the HAB buffer. The column was washed with 400

mL HAB buffer. Since OSC does not adsorb on the gel, fractions of preceding 10 mL of flowthrough were collected and assayed for cyclase activity. The functionally active fractions were next applied on HiTrap heparin columns (4 × 5 mL) equilibrated with HB buffer, and then eluted with 3.5 volumes of stepped 50 mM, 100 mM, and 1 M KCl. Cyclase enzyme was eluted and collected. Portions (200 µL) of each fraction were assayed for OSC activity. Protein concentrations were determined by the BCA assay kit (PIERCE) with bovine serum albumin as the standard.

#### **9.2.6.2 Oxidosqualene-lanosterol cyclase activity assay and polyacrylamide gel electrophoresis**

Enzymatic activity was determined by a standard procedure established in our laboratory. The reaction mixture contains 5 mM phosphate buffer (pH 7.4), 0.5% Triton X-100, and 50 µM [*R,S*] 2,3-oxidosqualene in a final volume of 200 µL cyclase-containing protein fraction. The reaction was incubated at 37°C for 2 h with gently shaking. After the termination of reaction with addition of equal volume of 15% ethanolic KOH, the sample was extracted twice with 400 µL CH<sub>2</sub>Cl<sub>2</sub>. The pooled organic layers were evaporated to dryness and redissolved in 50 µL CH<sub>2</sub>Cl<sub>2</sub>. The product extract was then applied onto a Silica gel Thin-layer chromatography (TLC) plate (Merck No. 5714-3), developed with Hexane/Ethyl acetate (4:1, v/v), and visualized with sprayed anisaldehyde reagent with the heat incubation.

Polyacrylamide gel electrophoresis was carried out to determine the purity of the cyclase-containing protein and the probably molecular weight according to the proposed method of Laemmli. All of the denatured samples were subjected to a 10% Polyacrylamide-SDS gel. After electrophoresis was performed in fixed 120 voltages, the Coomassie brilliant blue R250 staining solution or the silver nitrate staining solution were used to visualize the gel, respectively. Coomassie brilliant blue R250 staining solution was used to stain the gel for 1 h. For destaining process, the gel was firstly left in Destaining buffer I for 1 h and then put in Destaining buffer II until the protein band was obviously visualized. In the silver staining, the gel was put in the 50% methanol with gently shaking for 10 min. After another washing in the 5% methanol, the gel was left into the 40 µM DTT solution for 10 min. After three washing with water to remove all residual solution, the gel was soaked in the 0.1% silver nitrate solution for 10 min. This silver nitrate solution should be carefully removed by three short washings with water. Then 1/3 amount of development solution was added to wash the gel until solution color turn to yellow. Pour off the solution and add another 2/3 development solution in the gel-containing box for gently shaking. After protein band become visualized, stop the staining with 1% Acetic acid. The non-denatured polyacrylamide gel electrophoresis (Native PAGE) was performed on 7.5% polyacrylamide gel with 100 volts at 4°C. The sample buffer



and the gels do not contain any SDS and mercaptoethanol, and the running condition should be remained at 4°C to avoid the protein denature.

### **9.2.6.3 Peptide sequencing by tandem mass spectrometry**

(This portion was cooperated with Prof. Yu-Ju Chen and Hsin-Kai Liao from Institute of Chemistry, Academia Sinica, Taipei, Taiwan)

The OSC-containing fractions were collected, concentrated, desalted by using a Centriprep concentrator and then applied to the SDS-PAGE. After the electrophoresis separation, the gel was stained with Coomassie blue R250. The corresponding OSC band was excised and applied for in-gel digestion with trypsin in 25 mM, pH 8.0 NH<sub>4</sub>HCO<sub>3</sub> solution.<sup>261</sup> Small aliquot of the generated peptide mixture was analyzed by a nanoflow capillary liquid chromatography tandem mass spectrometry (Ultimate, LC Packings, Amsterdam, Netherlands). The column of nanoscale capillary liquid chromatography is a reversed-phase C18 column (15 cm x 75 mm I.D., 3 mm) with a flow rate of 200 nL/min. These tryptic peptides were separated by a 30 min gradient of 5–50% solvent B (solvent A is 2% acetonitrile/0.1% formic acid, whereas solvent B is 98% acetonitrile/0.08% formic acid). The eluted peptides were then introduced into a Q-TOF mass spectrometer (QSTAR Pulsar, Applied Biosystems, Foster City, CA, USA). Tandem mass spectra were automatically collected under the Information-Dependent Acquisition (IDA) during the 60 min LC/MS/MS run. Peptide product ion spectra generated by LC/MS/MS were searched against the SWISS-PROT protein databases using the Mascot sequence database search engine.<sup>173</sup> The MS BLAST was also carried out for searching the consensus aligned protein sequence from database by using previously produced peptide fragment by mass spectrometry.<sup>174</sup>

### **9.2.6.4 Molecular cloning of bovine liver oxidosqualene-lanosterol cyclase.**

Frozen bovine liver was grinded and subjected to the cDNA preparation using the Poly(A) Pure<sup>TM</sup> Kit with slight modification (QIAGEN, Hilden, Germany). In brief, mRNA (0.4 µg) was reverse-transcribed to produce the cDNA mixture by using the M-MLV RTase (Invitrogen, Carlsbad, CA, USA) with 10 mM each dNTP and 0.5 µg oligo (dT)<sub>12-18</sub> primer in a total volume of 20 µL. This mixture was incubated at 65°C for 5 min before centrifugation at 14,000 rpm for 1 min. The supernatant was added with 4 µL 5X first strand buffer, 2 µL 0.1M DTT, and 1 µL RNaseOUT<sup>TM</sup> ribonuclease inhibitor. This solution was incubated at 37°C for 2 min. Subsequently, 200 units M-MLV RTase was added in the solution mixture and reincubated at 37°C for 1 h before termination reaction by heating at 70°C for 15 min. The resulting cDNA library was directly used as the template for the following PCR reactions.

Three degenerated oligonucleotide primers (CHC-cDNA-PEP1, CHC-cDNA-PEP2, CHC-cDNA-PEP3) corresponding to internal polypeptide sequences which



were obtained from previous tandem mass spectrometry were synthesized. (Peptide1: <sup>165</sup>GVGPPDDPD<sup>172</sup>; Peptide2: <sup>249</sup>SAEEGPLVQSLR<sup>260</sup>; Peptide3: <sup>506</sup>KTEYTAFGGD<sup>497</sup>)

The nucleotide sequences of these three primers are listed as follow: CHC-cDNA-PEP1:5'-GGNGTNGGNCCNGAYGAYCCTGAY-3'; CHC-cDNA-PEP2:5'-AGTGCCGARGARGGNCCNCTGGTYCARAGYCTCCGC-3'; CHC-cDNA-PEP3:5'-YTTNGTYTCRTANGTNGCRAANCCNCCRTC-3'. PCR reaction was first carried out with CHC-cDNA-PEP1 and CHC-cDNA-PEP3 primers by using *Taq* DNA polymerase (BioBasic Inc., Toronto, Canada) with 15 mM MgCl<sub>2</sub>, 10 mM each dNTPs, 1 µg each primer, 1.5 µL Glycerol, 2 µL cDNA template, 5 µL 10 X *Taq* buffer, and DEPC water in the final volume of 50 µL. Nested PCR was also carried out with CHC-cDNA-PEP2 and CHC-cDNA-PEP3 primers by using the first PCR product solution as the template. The amplified 0.9 kb DNA fragments were purified and sequenced by using an ABI PRISM 3100 auto-sequencer. Computer assisting analysis of DNA sequences were also performed at the NCBI website with BLAST programs against GeneBank databases. Subsequently, the RACE method was applied for 5'- or 3'-end amplifications. For the 3'-end amplification, one specific primer (CHC-cDNA-Seq1) was designed in the sense direction, based on the obtained core fragment sequence. Two anchor primers (CHC-cDNA-XC1: based on the highly conserved C-terminal sequences from homologous mammalian lanosterol synthases; and CHC-OSC-RACE: the adaptor primer coupled with poly(dA)<sub>17</sub> primer) were designed and synthesized in the antisense direction, respectively. The nucleotide sequences of these two primers are listed as follows: CHC-cDNA-Seq1: 5'-GGCTCACAGATCTGGGA-3'; CHC-cDNA-XC1:5'-CCGCTCGAGTCANNNGTGGCCAGCAAGGGYNNNATCAGG-3'; OSC-RACE:5'-GACTCGAGTCGACATCGATTTTTTTTTTTTTTTTTTTT-3'. For the C-terminal RACE reaction, first PCR reaction was carried out with CHC-cDNA-PEP2 and OSC-RACE primers using bovine liver cDNA library as the template. Nested PCR reaction was subsequently carried out with CHC-cDNA-Seq1 and CHC-cDNA-XC1 primers as previously described. For the 5'-RACE amplification, reverse transcription of bovine liver mRNA using CHC-cDNA-PEP3 as a primer was prepared. The resultant cDNA mixture was then used for the first PCR reaction with CHC-cDNA-EN1 and CHC-cDNA-PEP3 primers, and for the second PCR reaction with CHC-cDNA-EN1 and CHC-cDNA-Seq2 primers by using the first PCR product as the template. The nucleotide sequences of these primers are listed as follows: CHC-cDNA-EN1:5'-TCTAGAATTCCATATGACNGAGGGCACGTGTCTGCG-3'; CHC-cDNA-Seq2: 5'-GCCGCAGGCTGGTGCTG-3'. Accordingly, the PCR amplicons of N- and C-terminal DNA fragments with CHC-cDNA-EN1/CHC-cDNA-Seq2 and CHC-cDNA-pep1/CHC-cDNA-XC1 primer

produced 1,698 and 936 bp fragments containing 441 bp overlapping sequence, respectively. In addition, the promoter sequence from the yeast *ERG7* gene should be fused in front of the bovine liver OSC gene to facilitate the expression of bovine liver OSC in the yeast *S. cerevisiae*. The *Not I/Pst I* fractionized fragment of the pTKERG7SK14 which contains *S. cerevisiae* *Erg7* promoter and partial *ERG7* gene was first ligated into the yeast-*E. coli* shuttle vector pRS314 to construct plasmid pCHC-RSC2. The amplified N-terminal DNA fragment was then digested with *NdeI* and *NcoI* restriction endonucleases and subcloned into a pCHC-RSC2 plasmid, which was digested previously with the same enzymes to produce pCHC-cDNA-N. Finally, the *NcoI/XhoI* restricted fragment from the amplified C-terminal cDNA fragment was excised and ligated into *NcoI-XhoI* digested pCHC-cDNA-N, to generate completed pRS314BLOSC plasmid. The insert region of the bovine liver OSC coding sequence was sequenced and verified by using an ABI PRISM 3100 auto-sequencer.

#### **9.2.6.5 Genetic complementation of recombinant bovine liver OSC in a yeast *erg7* deficient strain.**

Plasmid pRS314BLOSC plasmid DNA was introduced into CBY57[pZS11] by electroporation as previously described. The yeast transformant, CBY57([pZS11][pRS314BLOSC]), was plated and grown on SD+Ade+Lys+His+1M Sorbitol medium to select the presence of both pZS11 and pRS314BLOSC plasmids. Several colonies of CBY57([pZS11][pRS314BLOSC]) were picked and grown overnight in SD+Ade+Lys+His+Ura medium at 30°C. Aliquots (50 µL) from each culture were plated on SD+Ade+Lys+His+Ura and SD+Ade+Lys+His+Ura+1mg/mL 5'-fluoroorotic acid (5'-FOA) medium and grown for another three days at 30°C to elucidate the complementation effects. Colonies that grew on 5'-FOA-containing plates were grown separately in SD+Ade+Lys+His+Ura medium and subjected to yeast minipreps. The presence of pRS314BLOSC plasmid, but not pZS11, was verified using restriction endonuclease digestion and DNA sequencing.

#### **9.2.6.6 Nucleotide Sequence Accession Number and DNA Sequence Analysis**

Nucleotide sequence analysis and protein sequence comparisons were performed with GCG and Clustal W program. Database searching was also performed using the BLAST network services at the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>).

#### **9.2.6.7 Photoaffinity labeling of Ro48-8071 in bovine liver OSC and mass spectrometry determined**

Laboratory synthesized Ro48-8071 was dissolved in DMSO for generating the 10 mM stock solution. A desired concentration about threefold molar excess of Ro48-8071 was used in the photoaffinity experiment. After the organic solvent evaporating, 200 µL bovine livers OSC (0.8 mg/mL) in HB buffer was added into the

2 mL micro-centrifuge tubes containing Ro48-8071. Without Ro48-8071 addition cyclase-containing solution was used as the positive control for enzyme inactivation assay. Enzyme-inhibitor mixture was then incubated at 30°C for 2 h in the shaker incubator. After pre-incubation, the micro-centrifuge tubes were transferred into the quartz cuvette for exposure under UV light (365 nm at 1,200  $\mu\text{J}/\text{cm}^2$ ) for 30 min at 25 °C. The small amount of solution (50  $\mu\text{L}$ ) from each of the photoaffinity labeled enzyme-Ro48-8071 mixture, or from the wild-type OSC stock solution was added with the synthetic substrate, oxidosqualene, to check the enzyme activity. Samples of each incubation mixture were subjected for the Speed-vac concentrator to enhance the sample concentration for trypsin digestion. The same amounts (about 64  $\mu\text{g}$ ) of each samples (wild-type OSC and OSC/Ro48-8071 complex) were subjected to the 12% SDS polyacrylamide gel electrophoresis (PAGE) and visualized by coomassie brilliant blue R250 staining. The protein bands were then picked up for the in-gel digestion coupled with the LC/MS/MS analysis.

#### **9.2.6.8 Monoclonal antibody preparation and Western blotting analysis**

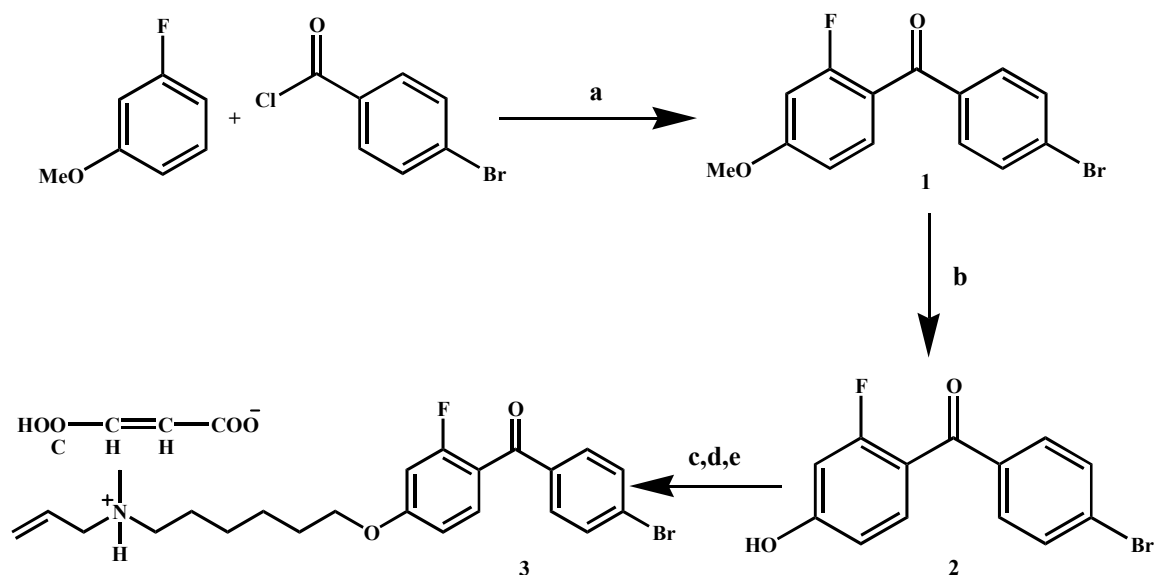
The anti-OSC polyclonal antibody (mAb) was prepared according to the established procedure in our laboratory by injecting BALB/c mice with an immunogenic bovine liver OSC oligopeptides.<sup>179, 180</sup> After immunoreactive titer determination with ELISA assay, the myeloma cell line (FO cell) was fused with the spleen cells of the immunized mouse for preparation of the monoclonal antibody. After the completion of preparation, the ELISA assay was further used to measure the immunoreaction by limiting dilutions. The entailed procedure for antibody preparation was listed in our previous literature.<sup>179, 180</sup> Moreover, for the western blotting analysis, after gel electrophoresis the resultant SDS-PAGE and a nitrocellulose (NC) membrane were soaked immediately in the transfer buffer. The protein transferring process from SDS-PAGE into the NC membrane was carried out with a blotting apparatus at 90 Volt for 1 h at 4°C. The nitrocellulose membrane was then blocked with blocking buffer for 1 h at RT. Subsequently, the membrane was washed three times with 1 X PBS buffer briefly and incubated with the our primary anti-OSC monoclonal antibody (mAb) at 1:500 dilutions in 1 X PBS buffer for 1 h at RT with gentle shaking. After five times of 10 min washing with 1 X PBST buffer, the membrane was incubated with the goat anti-mouse IgG conjugated HRP at 1:5,000 dilutions in 1 X PBS buffer for another 1 h at RT with gentle shaking. Finally, the membrane was washed with 1X PBST buffer for 5 times. The immunoreactive bands were visualized in a autoradiography film (Kodak Digital Sciences) with the ECL chemiluminescence's reagent (GE Healthcare).

## Chapter 7

### 9.2.7.1 Chemical synthesis of substrates or Ro48-8071 inhibitor

The (*R,S*)-2,3-oxidosqualene was synthesized according to the well established protocol in our laboratory. In brief, 10 ml squalene (20.94 mmole) was added in a solution consisted of 160 mL THF and 40 mL of water at 0°C. After the turbid suspended solid was appeared, another 50 mL THF dissolved with 3.75 g NBS was added slowly. The mixture was allowed to stir at 4°C for 3 h. Two times of 500 mL Hexane extractions were carried out. The combined organic layer was dried and concentrated by evaporation. The residue was subjected to silica gel chromatography with 0% ~ 10% Ethyl acetate/hexane to collect the bromohydrin, a colorless oil product. The bromohydrin was then dissolved in 100 mL Methanol and be added with 345 mg potassium carbonate. The resulting mixture was stirred at room temperature for 15 h. After Methanol was removed by evaporation, water and Hexane were added to the residue. The aqueous layer was separated and extracted twice with Hexane. The combined organic phase was dried by evaporation to give final (*R,S*)-2,3-oxidosqualene as colorless oil.

Ro48-8071 is an inhibitor of OSC that has low-density lipoprotein (LDL) cholesterol lowering activity similar to the 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) inhibitor Simvastatin. The chemical synthesis of Ro48-8071 was according to the published protocol from the F. Hoffmann-La Roche Ltd in Switzerland.<sup>169</sup> The simplified synthesized routine of Ro48-8071 was described in Scheme 9.2.7.1.



Scheme 9.2.7.1. The synthesis route of Ro48-8071<sup>169</sup>

(a) Nitrobenzene/ $\text{AlCl}_3$ , (b)  $\text{HBr}/\text{CH}_3\text{COOH}$ , (c) 1,6-dibromohexane/  $\text{K}_2\text{CO}_3$  in acetone, (d) *N*-allylmethylamine in *N,N*-dimethylacetamide, (e) fumaric acid in ethanol.

### 9.2.7.2 Chemical synthesis of boronic pinacol ester derivatives.

Aromatic boronic acids, including naphthalene-1-boronic acid, naphthalene-2-boronic acid, biphenyl-3-boronic acid, biphenyl-4-boronic acid were converted into their boronic pinacol ester derivatives via the condensation reaction with the pinacol diol.<sup>192</sup> 10 mmol respective aromatic boronic acid chemicals and 12 mmol pinacol were dissolved in 50 mL anhydrous THF. The reaction mixture was refluxed for 12h at 70°C under nitrogen. Molecular sieves were used to remove water. Evaporating the solvent under vacuum, the residue was separated on a silica gel by using 10% Ethyl acetate/Hexane solution to collect the final pinacol ester derivatives. These four boronic pinacol ester compounds, including naphthalene-1-boronic pinacol ester (NA-1), naphthalene-2-boronic pinacol ester (NA-2), biphenyl-3-boronic pinacol ester (BP-3), biphenyl-4-boronic pinacol ester (BP-4) were then used for modification of Ro48-8071 molecule via the Suzuki coupling reaction.

### 9.2.7.3 Chemical synthesis of 4-(4,5-diphenyl-1H-imidazol-2-yl)phenylboronic acid (DPA)

Phenylboronic acid derivate, 4-(4,5-diphenyl-1H-imidazol-2-yl)phenylboronic acid (DPA) was synthesized according to a method described in previous literature.<sup>182</sup> To a mixture of benzil (1.0 mmol) and ammonium acetate (8.0 mmol) in 5 ml of acetic acid was added with 4-formylphenylboronic acid (1.0 mmol) and refluxed for 24 h. After the solution was cooled, the resulting crystals were filtered off, washed with water, and dried. The NMR characterizations of this compound were identical to the authentic standard spectrum.

### 9.2.7.4 Synthesis of novel Ro48-8071 based fluorescence probes via Suzuki coupling reaction

The Suzuki coupling reaction, a palladium-catalyzed cross-coupling reaction of aryl halides with aryl boronic acids, was carried out for connection of DPA with Ro48-8071.<sup>193</sup> Firstly, the equal amount (1 mmol) of Ro48-8071 and DPA were dissolved in 50 mL toluene. To this solution were added with Pd(PPh<sub>3</sub>)<sub>4</sub> (0.003 mmol), K<sub>2</sub>CO<sub>3</sub> (25 mL, 2 M), and Ethanol (25 mL), and then be purged with N<sub>2</sub>. After heating at 100°C for overnight, the solution was filtered, extracted with Ethyl acetate, and concentrated to dryness by rotary evaporator. The resulting crude was purified by silica chromatography using CH<sub>2</sub>Cl<sub>2</sub>-Methanol (90:10) to give Ro48-8071 fluorescence derivative. Four boronic pinacol ester compounds, including NA-1, NA-2, BP-3, and BP-4 were also used to modify the Ro48-8071 molecule via the similar protocol.

### 9.2.7.5 Spectroscopic studies of novel Ro48-8071-based fluorescence probes

The absorption spectra of novel Ro48-8071-based fluorescence probes in ethanol, water, and DMSO were taken on a UV-vis spectrophotometer (DU 7500,



Beckman). The molar absorption coefficient was examined at maximum absorption wavelength. The Fluorescence emission spectrums were also recorded on a HitachiF-4500 spectrophotomer from 365 nm to 800 nm.

#### **9.2.7.6 The inhibition activity of the newly Ro48-8071-based fluorescent probes.**

Different concentrations of Ro48-8071-based fluorescent probes from 100 mM to 1 nM were incubated with 0.05 mg bovine liver OSC at 37°C for 20 min, respectively. After pre-incubation reaction, the enzymatic substrate, OS, was added into the enzyme mixture solution and incubated for another 1.5 h at 37°C. The OSC activity assay was carried out according to the previous mentioned protocol in section 9.2.6.2.

## **Chapter 8**

### **9.2.8.1 Generation of a degenerate single strand DNA library**

72-mer DNA templates was prepared by using solid-phase phosphoramidite synthetic chemistry (Bio Basic Inc), containing 36 random nucleotide sequences flanked by two defined terminal nucleotide regions which are allowed for the PCR amplification reaction. The full sequence of this polynucleotide is listed in the following: 5'-biotin-GGATCCGAGCTCCACGTG-(N)<sub>36</sub>-GCTAGCGTCGACCGTAC G-3'. The synthetic DNA was first purified by PAGE and amplified by PCR reaction with the defined primers to generate a population of polynucleotide library. Accordingly, one of the primer pair was modified with biotin, whose sequence is shown below: 5'-biotin-GGATCCGAGCTCCACGTG. In parallel, the other primer is complementary to the 3'-end of fixed terminus with sequence as follow: 5'-CGTACGGTCGACGCTAGC. The resultant biotinylated double-stranded DNA amplicon was subjected to an Avidin-agarose column (Vector Labs) which was equilibrated with buffer-A (10mM Tris-HCl pH7.5, 50mM NaCl, and 1mM EDTA). After incubation for 30 min at 25°C, the unbound double-stranded DNA was removed with five times of column volume of buffer-A. The matrix-bound double-stranded DNA was then denatured with 500 µl of 0.2 N NaOH solutions for 15 min at 37°C. The resultant free single-stranded DNA was collected and precipitated with ethanol. After purification via centrifugation, the single-stranded DNA library was then dissolved in the selection buffer (10mM HEPES pH 7.4, 75mM NaCl, 5mM MgCl<sub>2</sub>, and 1mM CaCl<sub>2</sub>) and stored at -20°C before selection.

### **9.2.8.2 *In vitro* selection by using immobilized protein on a microtiter plates**

Polystyrene microtiter plates are known to bind a wide range of proteins through hydrophobic interaction. Theoretically, the generated single-stranded DNA library could be “panned” against the target protein which was immobilized on the surface of

a microtiter plate, and the captured DNA species could be eluted and amplified. Different concentrations of purified oxidosqualene cyclase were loaded into individual wells of a microtiter plate (Nunc Co, Denmark) and incubated overnight at room temperature (“OSC only”). At the same time, an aliquot of HB buffer (5 mM Kpi buffer, pH 7.4) were added into the other wells to make a negative control (“BSA only”). After overnight incubation, the non-derivatized surface were blocked by 3% (w/v) BSA solution (100  $\mu$ l) and incubated at 37°C for another 3 h. After three times of washing with PBST solution (1 X PBS solution plus 0.1% (v/v) Tween-20), the derivatized plates are carefully drained and stored at 4°C. The single-stranded DNA library dissolved in the binding buffer was denatured at 95°C for 5 min and then chilled on ice for another 5 min. After the performance of negative “panning” selection with “BSA only” wells for 1 h, the single-stranded DNA library was then transferred from “BSA only” wells into the “OSC only” wells. The mixtures were incubated for another 2 h at room temperature, and then be washed via three times with selection buffer to discard the unbound DNA. After the selection, the guanidine thiocyanate solution (200  $\mu$ l) was added into the “OSC only” wells, and the plate was heated at 80°C for 15 min on a heat block. Finally, the mixture was removed and precipitated. The PCR reaction was carried out to amplify the resultant DNA library for the next round of *in vitro* selection. The guanidine thiocyanate solution contains 20 mM Tris-acetate pH 8.3, 4 M guanidine thiocyanate, and 1mM DTT.

#### **9.2.8.3 *In vitro* Selection by using immobilized protein on filters**

UltraBind membrane (Pall) possesses the functional aldehyde groups for the covalent bond formation with amino groups of proteins. The binding efficiency and stability of membrane are theoretically stronger than that of microtiter plates. The high specificity with amino group also led UltraBind membrane suitable for the *in vitro* selection process. After cutting the membrane into pieces with suitable diameter, the immersion of these membranes with the protein solution for 2 h at 37°C was carried out. After drying of the membrane, 1% nonfat milk was added for blocking. Subsequently, the membranes were rinsed thrice with PBST buffer. The denatured single-stranded DNA were first added into the reaction tube with membranes, which were coated with sole blocking buffer, for 2 h at 37°C (Negative selection). The solution was then transferred into the reaction tube with membrane coated with OSC for another 2 h. After washing with binding buffer for three times, the membranes were transferred into the PCR reaction to generate the novel library for next round of *in vitro* selection.



#### **9.2.8.4 Electrophoretic mobility shift assay for the analysis of DNA-protein interactions**

The binding ability of single-stranded DNA to OSC was determined by using gel mobility shift assay (EMSA). The radiolabeling reaction of DNA with adenosine 5'-[ $\gamma$ -<sup>32</sup>P] triphosphate was carried out by using T4 polynucleotide kinase at 37°C for 1 h. These radiolabeled DNA probes were denatured at 95°C for 5 min, and then chilled in the ice before use. At first, different concentrations of DNA probes were incubated with the constant concentration of OSC in total volume of 40  $\mu$ l reaction at 37°C for 30 min. Then 8  $\mu$ l 5 X DNA loading dye was added into the reaction mixture. These mixtures were separated with 5% non-denaturing polyacrylamide gel electrophoresis at 4°C for 1 h at 110 V. After completion of the electrophoresis, lift the gel with the Whatman® paper and cover it with the plastic wrap. The autoradiograph was detected by using an Image Scanner (Packard Inc). Mobility shift assay of different concentrations of OSC with the constant concentration of DNA probe were also performed in the same way.

#### **9.2.8.5 Molecular cloning and DNA sequencing analysis**

The autoradiographic band of the DNA retardation corresponding to the DNA-protein complex were excised with the razor blade, and the DNA extraction from the polyacrylamide gel were performed according to the procedure described in the Molecular Cloning.<sup>249</sup> The amplified dsDNA from the excised band were cloned into pCR2.1-TOPO vector (Invitrogen Inc), and sequenced by using ABI automated DNA sequencer 3100. The software ClustalW program was used to analyze the sequence alignment of these resultant DNAs. Moreover, the online mfold program was utilized to predict the secondary structure of respective aptamer molecule.

#### **9.2.8.6 Binding analysis of individual aptamer with Biacore**

The respective clone with the unique aptamer was first used for PCR reaction with two terminal primers. Notably, the biotinylated position of the primer pair described herein is distinct from the primer pairs used in the screening procedure. The original biotinylated primer used in the selection is changed for the native “hydroxyl group”, whereas the end of another primer is changed from the native “hydroxyl group” for the “biotin group”. This reversed biotin modification is used for immobilizing the sense-stranded of aptamer on the surface of flow cell. After PCR amplification of desired aptamer clone, the PCR amplicon was denatured under heating for 5 min, and then be chilled on the ice for another 5 min. Before immobilization, the streptavidin-coated sensor chip SA (Biacore) was first conditioned with three times washing of the consecutive 1-minute injections of 1 M NaCl in 50 mM NaOH. The 100  $\mu$ l denatured DNA molecules was then immobilized on the surface of one flow cell at a flow rate of 5  $\mu$ l/min with the HBS buffer (GE

healthcare). The excess unbound biotinylated DNA as well as the possibly producing double-stranded DNA was washed by using three times of washing with 1 M NaCl in 50 mM NaOH buffer. In parallel, the other flow cell was coated with free biotin molecule, used as the background control. After finishing the immobilization, the streptavidin-coated sensor chip SA was conditioned again. Subsequently, the different concentration of OSC protein was load into the flow cells with the flow rate of 10  $\mu$ l/ml. After injection, bound proteins were washed with the 100  $\mu$ l regeneration buffer (1 M NaCl in 50 mM NaOH). The injections and regenerations were repeated until obtainment of the desired sensorgrams profiles for analysis with BioEvolution software (Biacore, GE healthcare).



## References

1. Segura, M. J.; Jackson, B. E.; Matsuda, S. P. *Nat. Prod. Rep.* **2003**, 20, (3),304-17.
2. Abe, I.; Rohmer, M.; Prestwich, G. D. *Chem. Rev.* **1993**, 93, 2189-2206.
3. Robinson, R. *J. Soc. Chem. Ind. (London)* **1934**, 53, 1062-1063.
4. Bloch, K.; Rittenberg, D. *J. Biol. Chem.* **1942**, 145, 625-36.
5. Woodward, R. B.; Bloch, K. *J. Am. Chem. Soc.* **1953**, 75, 2023-4.
6. Tchen, T. T.; Bloch, K. *J. Biol. Chem.* **1957**, 226, (2), 931-9.
7. Langdon, R. G.; Bloch, K. *J. Biol. Chem.* **1953**, 200, (1), 135-44.
8. Tchen, T. T.; Bloch, K. *J. Biol. Chem.* **1957**, 226, (2), 921-30.
9. Corey, E. J.; Russey, W. E.; Ortiz de Montellano, P. R. *J. Am. Chem. Soc.* **1966**, 88, 4750-1.
10. Van Tamelen, E. E.; Willett, J. D.; Clayton, R. B.; Lord, K. E. *J. Am. Chem. Soc.* **1966**, 88, 4752-4.
11. Barton, D. H.; Jarman, T. R.; Watson, K. C.; Widdowson, D. A.; Boar, R. B.; Damps, K. *J. Chem. Soc. [Perkin I]* **1975**, (12), 1134-8.
12. Boutaud, O.; Dolis, D.; Schuber, F. *Biochem. Biophys. Res. Commun.* **1992**, 188, (2), 898-904.
13. Corey, E. J.; Gross, S. K. *J. Am. Chem. Soc.* **1967**, 89, 4661-2.
14. Anding, C.; Heintz, R.; Ourisson, G. *C R Acad Sci Hebd Seances Acad Sci D* **1973**, 276, (2), 205-7.
15. Bujons, J.; Guajardo, R.; Kyler, K. S. *J. Am. Chem. Soc.* **1988**, 110, 604-6.
16. Xiao, X. Y.; Prestwich, G. D. *Tetrahedron Lett.* **1991**, 32, 6843-6.
17. Corey, E. J.; Russey, W. E. *J. Am. Chem. Soc.* **1966**, 88, 4751-2.
18. Corey, E. J.; Virgil, S. C.; Liu, D. R.; Sarshar, S. *J. Am. Chem. Soc.* **1992**, 114, 1524-5.
19. Corey, E. J.; Ortiz de Montellano, P. R.; Lin, K.; Dean, P. D. *J. Am. Chem. Soc.* **1967**, 89, (11), 2797-8.
20. Clayton, R. B.; Van Tamelen, E. E.; Nadeau, R. G. *J. Am. Chem. Soc.* **1968**, 90, 820-1.
21. Corey, E. J.; Lin, K.; Jautelat, M. *J. Am. Chem. Soc.* **1968**, 90, 2724-6.
22. Ceruti, M.; Rocco, F.; Viola, F.; Balliano, G.; Grosa, G.; Dosio, F.; Cattel, L. *Eur. J. Med. Chem.* **1993**, 28, 675-82.
23. Taton, M.; Benveniste, P.; Rahier, A. *Biochem. Biophys. Res. Commun.* **1986**, 138, (2), 764-70.
24. Gerst, N.; Duriatti, A.; Schuber, F.; Taton, M.; Benveniste, P.; Rahier, A. *Biochem. Pharmacol.* **1988**, 37, (10), 1955-64.
25. Rahier, A.; Taton, M.; Benveniste, P. *Biochem. Soc. Trans.* **1990**, 18, (1), 48-52.
26. Cattel, L.; Ceruti, M.; Viola, F.; Delprino, L.; Balliano, G.; Duriatti, A.; Bouvier-Nave, P. *Lipids* **1986**, 21, (1), 31-8.
27. Ceruti, M.; Balliano, G.; Viola, F.; Grosa, G.; Rocco, F.; Cattel, L. *J. Med. Chem.* **1992**, 35, (16), 3050-8.
28. Balliano, G.; Milla, P.; Ceruti, M.; Viola, F.; Carrano, L.; Cattel, L. *FEBS Lett.* **1993**, 320, (3), 203-6.
29. Ceruti, M.; Rocco, F.; Viola, F.; Balliano, G.; Milla, P.; Arpicco, S.; Cattel, L. *J. Med. Chem.* **1998**, 41, 540-554.
30. Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc.* **1997**, 119, 1277-88.
31. Gao, D.; Pan, Y. K. *J. Am. Chem. Soc.* **1998**, 120, 4045-46.
32. Van Tamelen, E. E.; James, D. R. *J. Am. Chem. Soc.* **1977**, 99, 950-2.
33. Van Tamelen, E. E.; Hopla, R. E. *J. Am. Chem. Soc.* **1979**, 101, 6112-4.
34. Boar, R. B.; Couchman, L. A.; Jaques, A. J.; Perkins, M. J. *J. Am. Chem. Soc.* **1984**, 106, 2476-7.
35. Renoux, J. M.; Rohmer, M. *Eur. J. Biochem.* **1986**, 155, 125-33.
36. Corey, E. J.; Virgil, S. C.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Sarshar, S. *J. Am. Chem.*

- Soc.* **1995**, 117, 11819-20.
37. Hoshino, T.; Hoshino, T.; Sakai, Y. *Chem. Commun.* **1998**, 1591-2.
  38. Corey, E. J.; Cheng, H. *Tetrahedron Lett.* **1996**, 37, 2709-12.
  39. Van Tamelen, E. E.; Sharpless, K. B.; Hanzlik, R.; Clayton, R. B.; Burlingame, A. L.; Wszolek, P. *C. J. Am. Chem. Soc.* **1976**, 89, 7150-1.
  40. Corey, E. J.; Virgil, S. C. *J. Am. Chem. Soc.* **1991**, 113, 4025-6.
  41. Corey, E. J.; Virgil, S. C.; Sarshar, S. *J. Am. Chem. Soc.* **1991**, 113, 8171-2.
  42. Ourisson, G. *Pure Appl. Chem.* **1989**, 61, 345-8.
  43. Rohmer, M.; Bouvier, P.; Ourisson, G. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, 76, 847-51.
  44. Ourisson, G.; Rohmer, M.; Poralla, K. *Ann. Rev. Microbiol.* **1987**, 41, 301-33.
  45. Johnson, W. S.; Lindell, S. D.; Steele, J. J. *J. Am. Chem. Soc.* **1987**, 109, 5852-3.
  46. Johnson, W. S.; Telfer, S. J.; Cheng, S.; Schubert, U. *J. Am. Chem. Soc.* **1987**, 109, 2517-8.
  47. Poralla, K. *Bioorg. Med. Chem. Lett.* **1994**, 4, 285-90.
  48. Poralla, K.; Hewelt, A.; Prestwich, G. D.; Abe, I.; Reipen, I.; Sprenger, G. *Trends Biochem. Sci.* **1994**, 19, (4), 157-8.
  49. Xiao, X. Y.; Prestwich, G. D. *J. Am. Chem. Soc.* **1991**, 113, 9673-4.
  50. Abe, I.; Prestwich, G. D. *J. Biol. Chem.* **1994**, 269, (2), 802-4.
  51. Wendt, K. U.; Poralla, K.; Schulz, G. E. *Science* **1997**, 277, 1811-5.
  52. Shi, Z.; Buntel, C. J.; Griffin, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, 91, 7370-4.
  53. Abe, I.; Bai, M.; Xiao, X. Y.; Prestwich, G. D. *Biochem. Biophys. Res. Commun.* **1992**, 187, (1), 32-8.
  54. Feil, C.; Sussmuth, R.; Jung, G.; Poralla, K. *Eur. J. Biochem.* **1996**, 242, (1), 51-5.
  55. Abe, I.; Dang, T.; Zheng, Y. F.; Madden, B. A.; Feil, C.; Poralla, K.; Prestwich, G. D. *J. Am. Chem. Soc.* **1997**, 119, 11333-4.
  56. Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc.* **1997**, 119, 1289-96.
  57. Abe, I.; Liu, W.; Oehlschlager, A. C.; Prestwich, G. D. *J. Am. Chem. Soc.* **1996**, 118, 3180-1.
  58. Seckler, B.; Poralla, K. *Biochimica et biophysica acta* **1986**, 881, 356-363.
  59. Hoshino, T.; Williamsa, H. J.; Chunga, Y.; Scott, A. I. *Tetrahedron* **1991**, 47, 5925-32.
  60. Duriatti, A.; Schuber, F. *Biochem. Biophys. Res. Commun.* **1988**, 151, (3), 1378-85.
  61. Carrano, L.; Noe, M.; Grosa, G.; Milla, P.; Denaro, M.; Islam, K. *J. Med. Vet. Mycol.* **1995**, 33, (1), 53-8.
  62. Abe, I.; Sankawa, U.; Ebizuka, Y. *Chem. Pharm. Bull.* **1992**, 40, 1755-70.
  63. Grosa, G.; Viola, F.; Ceruti, M.; Brusa, P.; Delprino, L.; Dosio, F.; Cattel, L. *Eur. J. Med. Chem.* **1994**, 29, 17-23.
  64. Cattel, L.; Ceruti, M. *Crit. Rev. Biochem. Mol. Biol.* **1998**, 33, (5), 353-73.
  65. Balliano, G.; Grosa, G.; Milla, P.; Viola, F.; Cattel, L. *Lipids* **1993**, 28, (10), 903-6.
  66. Milla, P.; Lenhart, A.; Grosa, G.; Viola, F.; Weihofen, W. A.; Schulz, G. E.; Balliano, G. *Eur. J. Biochem.* **2002**, 269, (8), 2108-16.
  67. Abe, I.; Ebizuka, Y.; Sankawa, U. *Chem. Pharm. Bull.* **1988**, 36, 5031-4.
  68. Abe, I.; Ebizuka, Y.; Seo, S.; Sankawa, U. *FEBS Lett.* **1989**, 249, 100-4.
  69. Abe, I.; Sankawa, U.; Ebizuka, Y. *Chem. Pharm. Bull.* **1989**, 37, 536-8.
  70. Abe, I.; Sankawa, U.; Ebizuka, Y. *Chem. Pharm. Bull.* **1992**, 40, 1755-60.
  71. Abe, I.; Bai, M.; Xiao, X. Y.; Prestwich, G. D. *Biochem. Biophys. Res. Commun.* **1992**, 187, 32-8.
  72. Kusano, M.; Abe, I.; Sankawa, U.; Ebizuka, Y. *Chem. Pharm. Bull.* **1991**, 39, 239-241.
  73. Moore, W. R.; Schatzman, G. L. *J. Biol. Chem.* **1992**, 267, 22003-6.
  74. Corey, E. J.; Matsuda, S. P. *J. Am. Chem. Soc.* **1991**, 113, 8172-4.
  75. Corey, E. J.; Matsuda, S. P. T.; Bartel, B. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, 91, 2211-5.
  76. Abe, I.; Prestwich, G. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, 92, 9274-8.

77. Kusano, M.; Shibuya, M.; Sankawa, U.; Ebizuka, Y. *Biol. Pharm. Bull.* **1995**, *18*, 195-7.
78. Baker, C. H.; Matsuda, S. P. T.; Liu, D. R.; Corey, E. J. *Biochem. Biophys. Res. Commun.* **1995**, *213*, 154-60.
79. Kelly, R.; Miller, S. M.; Lai, M. H.; Kirsch, D. R.; Kelly, R. *Gene* **1990**, *87*, 177-83.
80. Buntel, C. J.; Griffin, J. H. *J. Am. Chem. Soc.* **1992**, *114*, 9711-3.
81. Corey, E. J.; Matsuda, S. P. T.; Bartel, B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11628-32.
82. Sung, C. K.; Shibuya, M.; Sankawa, U.; Ebizuka, Y. *Biol. Pharm. Bull.* **1995**, *18*, 1459-61.
83. Kushiro, T.; Shibuya, M.; Ebizuka, Y. *Eur. J. Biochem.* **1998**, *256*, 238-44.
84. Shibuya, M.; Zhang, H.; Endo, A.; Shishikura, K.; Kushiro, T.; Ebizuka, Y. *Eur. J. Biochem.* **1999**, *266*, 302-7.
85. Herrera, J. B. R.; Bartel, B.; Wilson, W. K.; Matsuda, S. P. T. *Phytochemistry* **1998**, *49*, 1905-11.
86. Haralampidis, K.; Trojanowska, M.; Osbourn, A. E. A. *Adv. Biochem. Eng.* **2002**, *75*, 31-49.
87. Wendt, K. U.; Feil, C.; Lenhart, A.; Poralla, K.; Schulz, G. E. *Protein. Sci.* **1997**, *6*, (3), 722-4.
88. Wendt, K. U.; Lenhart, A.; Schulz, G. E. *J. Mol. Biol.* **1999**, *286*, 175-87.
89. Lenhart, A.; Weihofen, W. A.; Pleschke, A. E.; Schulz, G. E. *Chem Biol.* **2002**, *9*, (5), 639-45.
90. Hoshino, T.; Kouda, M.; Abe, T.; Ohashi, S. *Biosci. Biotechnol. Biochem.* **1999**, *63*, (11), 2038-41.
91. Sato, T.; Abe, T.; Hoshino, T. *Chem. Commun.* **1998**, (23), 2617-8.
92. Pale-Grosdemange, C.; Feil, C.; Rohmer, M.; Poralla, K. *Angew. Chem., Int. Ed.* **1998**, *37*, 2237-40.
93. Sato, T.; Hoshino, T. *Biosci. Biotechnol. Biochem.* **1999**, *63*, 2189-98.
94. Merkofer, T.; Pale-Grosdemange, C.; Wendt, K. U.; Rohmer, M.; Poralla, K. *Tetrahedron Lett.* **1999**, *40*, 2121-4.
95. Pale-Grosdemange, C.; Merkofer, T.; Rohmer, M.; Poralla, K. *Tetrahedron Lett.* **1999**, *40*, 6009-12.
96. Hoshino, T.; Sato, T. *Chem. Commun.* **1999**, (19), 2005-6.
97. Full, C.; Poralla, K. *FEMS Microbiology Letters* **2000**, *183*, 221-4.
98. Hoshino, T.; Shimizu, K.; Sato, T. *Angew. Chem., Int. Ed.* **2004**, *43*, 6700-3.
99. Hoshino, T.; Abe, T.; Kouda, M. *Chem. Commun.* **2000**, 441-2.
100. Sato, T.; Kouda, M.; Hoshino, T. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 728-38.
101. Sato, T.; Sasahara, S.; Yamakami, T.; Hoshino, T. *Biosci. Biotechnol. Biochem.* **2002**, *66*, 1660-70.
102. Hoshino, T.; Sato, T. *Chem. Commun.* **2002**, 291-301.
103. Wendt, K. U.; Schulz, G. E.; Corey, E. J.; Liu, D. R. *Angew. Chem., Int. Ed.* **2000**, *39*, 2812-33.
104. Corey, E. J.; Matsuda, S. P.; Baker, C.; T. H.; Alice, Y.; Cheng, H. *Biochem. Biophys. Res. Commun.* **1996**, *219*, 327-31.
105. Thoma, R.; Schulz-Gasch, T.; D'Arcy, B.; Benz, J.; Aebi, J.; Dehmlow, H.; Hennig, M.; Stihle, M.; Ruf, A. *Nature* **2004**, *432*, 118-22.
106. Wu, T. K.; Griffin, J. H. *Biochemistry* **2002**, *41*, (26), 8238-44.
107. Hart, E. A.; Hua, L.; Darr, L. B.; Wilson, W. K.; Pang, J.; Matsuda, S. P. *J. Am. Chem. Soc.* **1999**, *121*, 9887-8.
108. Meyer, M. M.; Xu, R.; Matsuda, S. P. *Org. Lett.* **2002**, *4*, 1395-8.
109. Herrera, J. B. R.; Wilson, W. K.; Matsuda, S. P. *J. Am. Chem. Soc.* **2000**, *122*, 6765-6.
110. Matsuda, S. P.; Darr, L. B.; Hart, E. A.; Herrera, J. B.; McCann, K. E.; Meyer, M. M.; Pang, J.; Schepmann, H. G. *Org. Lett.* **2000**, *2*, 2261-3.
111. Lodeiro, S.; Segura, M. J.; Stahl, M.; Schulz-Gasch, T.; Matsuda, S. P. *Chembiochem* **2004**, *5*, 1581-5.
112. Joubert, B. M.; Hua, L.; Matsuda, S. P. T. *Org. Lett.* **2000**, *2*, 339-41.
113. Kolesnikova, M. D.; Wilson, W. K.; Lynch, D. A.; Obermeyer, A. C.; Matsuda, S. P. T. *Org. Lett.* **2007**, *9*, 5223-6.



114. Meyer, M. M.; Segura, J. R.; Wilson, W. K.; Matsuda, S. P. T. *Angew. Chem., Int. Ed.* **2000**, 39, 4090-3.
115. Wu, T. K.; Liu, Y. T.; Chang, C. H. *Chembiochem* **2005**, 6, (7), 1177-81.
116. Wu, T. K.; Wen, H. Y. *The Dissertation for the Degree of Master* **2007**.
117. Wu, T. K.; Liu, Y. T.; Chang, C. H.; Yu, M. T.; Wang, H. J. *J. Am. Chem. Soc.* **2006**, 128, (19), 6414-9.
118. Wu, T. K.; Yu, M. T.; Liu, Y. T.; Chang, C. H.; Wang, H. J.; Diau, E. W. *Org. Lett.* **2006**, 8, (7), 1319-22.
119. Wu, T. K.; Liu, Y. T.; Chiu, F. H.; Chang, C. H. *Org. Lett.* **2006**, 8, (21), 4691-4.
120. Wu, T. K.; Li, W. H. *The Dissertation for the Degree of Master* **2007**.
121. Segura, M. J.; Lodeiro, S.; Meyer, M. M.; Patel, A. J.; Matsuda, S. P. *Org. Lett.* **2002**, 4, 4459-62.
122. Lewis, T. A.; Taylor, F. R.; Parks, L. W. *J. Bacteriol.* **1985**, 163, 199-207.
123. Gollub, E. G.; Liu, K. P. D., J.; Adlersberg, M.; Sprinson, D. B. *J. Biol. Chem.* **1977**, 252, 2846-54.
124. Akihisa, T.; Arai, K.; Kimura, Y.; Koike, K.; Kokke, W. C. M. C.; Shibata, T.; Nikaido, T. *J. Nat. Prod.* **1999**, 62, 265-68.
125. Wu, T. K.; Chang, C. H. *Chembiochem* **2004**, 5, (12), 1712-5.
126. Lodeiro, S.; Wilson, W. K.; Shan, H.; Matsuda, S. P. *Org. Lett.* **2006**, 8, 439-42.
127. Schulz-Gasch, T.; Stahl, M. *Journal of Computational Chemistry* **2003**, 24, 741-53.
128. Kushiro, T.; Shibuya, M.; Ebizuka, Y. *J. Am. Chem. Soc.* **1999**, 121, 1208-16.
129. Back, K.; Chappell, J. *Proc. Natl. Acad. Sci. U. S.A.* **1996**, 93, 6841-5.
130. Kushiro, T.; Shibuya, M.; Masuda, K.; Ebizuka, Y. *J. Am. Chem. Soc.* **2000**, 122, 6816-24.
131. Cheng, J.; Randall, A.; Sweredoski, M.; Baldi, P. *Nucleic Acids Research* **2005**, 33, 72-6.
132. Rost, B.; Sander, C. *J. Molecular Biol.* **1993**, 232, 584-99.
133. Van Tamelen, E. E.; Willett, J.; Schwartz, M.; Nadeau, R. *J. Am. Chem. Soc.* **1966**, 88, 5937-8.
134. Higgins, D.; Thompson, J.; Gibson, T. *Nucleic Acids Res.* **1994**, 22, 4673-80.
135. Pollastri, G.; Przybylski, D.; Rost, B.; Baldi, P. *Proteins* **2002**, 47, 228-35.
136. Pontius, J.; Richelle, J.; Wodak, S. J. *J. Molecular Biol.* **1996**, 264., 121-36.
137. Luthy, R.; Bowie, J. U.; Eisenberg, D. *Nature* **1992**, 356, 83-5.
138. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, 26, 283-91.
139. Morris, A. L.; MacArthur, M. W.; Hutchinson, E. G.; Thornton, J. M. *Proteins* **1992**, 12, 345-64.
140. Schmitz, S.; Full, C.; Glaser, T.; Albert, K.; Poralla, K. *Tetrahedron Lett.* **2000**, 42, 883-5.
141. Dang, T.; Prestwich, G. D. *Chem. Biol.* **2000**, 7, 643-9.
142. Wu, T. K.; Huang, C. Y.; Ko, C. Y.; Chang, C. H.; Chen, Y. J.; Liao, H. K. *Arch. Biochem. Biophys.* **2004**, 421, (1), 42-53.
143. Lodeiro, S.; Wilson, W. K.; Shan, H.; Matsuda, S. P. *Org. Lett.* **2006**, 8, (3), 439-42.
144. Wu, T. K.; Li, W. H. *Submitted* **2007**.
145. Wu, T. K.; Wang, T. T. *Manuscript in preparation* **2008**.
146. Wu, T. K.; Wen, H. Y. *Manuscript in preparation* **2008**.
147. Corey, E. J.; Stass, D. J. *J. Am. Chem. Soc.* **1998**, 120, 3526-7.
148. Jenson, C.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1997**, 119, 10846-54.
149. Rajamani, R.; Gao, J. *J. Am. Chem. Soc.* **2003**, 125, 20768-81.
150. Hess B. A., J. *J. Am. Chem. Soc.* **2002**, 124, 10286-7.
151. Hess B. A., J. *Org. Lett.* **2003**, 5, 165-7.
152. Matsuda, S. P.; Wilson, W. K.; Xiong, Q. *Org. Biomol. Chem* **2006**, 4, 530-43.
153. Lodeiro, S.; Schulz-Gasch, T.; Matsuda, S. P. *J. Am. Chem. Soc.* **2005**, 127, 14132-3.
154. Shan, H.; Wilson, W. K.; Phillips, D. R.; Bartel, B.; Matsuda, S. P. *Org. Lett.* **2008**.
155. Lodeiro, S.; Xiong, Q.; Wilson, W. K.; Kolesnikova, M. D.; Onak, C. S.; Matsuda, S. P. *J. Am.*

- Chem. Soc.* **2007**, 129, 11213-22.
156. Kolesnikova, M. D.; Obermeyer, A. C.; Wilson, W. K.; Lynch, D. A.; Xiong, Q.; Matsuda, S. P. *Org. Lett.* **2007**, 24, 2183-6.
157. Phillips, D. R.; Rasbery, J. M.; Bartel, B.; Matsuda, S. P. *Curr. Opin. Plant Biol.* **2006**, 9, 305-14.
158. Xiong, Q.; Wilson, W. K.; Matsuda, S. P. *Angew. Chem., Int. Ed.* **2006**, 13, 1285-8.
159. Shibuya, M.; Xiang, T.; Katsube, Y.; Otsuka, M.; Zhang, H.; Ebizuka, Y. *J. Am. Chem. Soc.* **2007**, 129, 1450-5.
160. Xiang, T.; Shibuya, M.; Katsube, Y.; Tsutsumi, T.; Otsuka, M.; Zhang, H.; Masuda, K.; Ebizuka, Y. *Org. Lett.* **2007**, 8, 2835-8.
161. Abe, I. *Nat. Prod. Rep.* **2007**, 24, 1311-31.
162. Jahn, O.; Eckart, K.; Tezval, H.; Spiess, J. *Anal. Bioanal. Chem.* **2004**, 378, (4), 1031-6.
163. Jahn, O.; Tezval, H.; Spiess, J.; Eckart, K. *International Journal of Mass Spectrometry* **2003**, 228, 528-40.
164. Hashimoto, M.; Hatanaka, Y. *Chem. Pharm. Bull.* **1999**, 47, (5), 667-71.
165. Hatanaka, Y.; Hashimoto, M.; Kanaoka, Y. *J. Am. Chem. Soc.* **1998**, 120, 453-54.
166. Hatanaka, Y.; Ishiguro, M.; Hashimoto, M.; Gastinel, L. N.; Nakagomi, K. *Bioorg. Med. Chem. Lett.* **2001**, 11, (3), 411-3.
167. Dorman, G.; Prestwich, G. D. *Trends Biotechnol.* **2000**, 18, (2), 64-77.
168. Prestwich, G. D.; Dorman, G.; Elliott, J. T.; Marecak, D. M.; Chaudhary, A. *Photochem. Photobiol.* **1997**, 65, (2), 222-34.
169. Morand, O. H.; Aebi, J. D.; Dehmlow, H.; Ji, Y. H.; Gains, N.; Lengsfeld, H.; Himber, J. J. *Lipid Res.* **1997**, 38, (2), 373-90.
170. Abe, I.; Zheng, Y. F.; Prestwich, G. D. *Biochemistry* **1998**, 37, (17), 5779-84.
171. Dang, T.; Abe, I.; Zheng, Y. F.; Prestwich, G. D. *Chem. Biol.* **1999**, 6, (6), 333-41.
172. Rosenfeld, J.; Capdevielle, J.; Guillemot, J. C.; Ferrara, P. *Anal. Biochem.* **1992**, 203, (1), 173-9.
173. Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis.* **1999**, 20, 3551-67.
174. Shevchenko, A.; Sunyaev, S.; Loboda, A.; Shevchenko, A.; Bork, P.; Ens, W.; Standing, K. G. *Anal. Chem.* **2001**, 73, 1917-26.
175. Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, 11, (11), 601.
176. Zhang, Y.; Frohman, M. A. *Methods Mol. Biol.* **1997**, 69, 61-87.
177. Towner, P.; Gärtner, W. *Nucleic Acids Res.* **1992**, 20, 4669-70.
178. Quach, T. T.; Li, N.; Richards, D. P.; Zheng, J.; Keller, B. O.; Li, L. *J. Proteome Res.* **2003**, 2, (5), 543-52.
179. Wu, T. K.; Wang, Y. K. *The Dissertation for the Qualified Examination in the Degree of Doctor of Philosophy* **2006**.
180. Wu, T. K.; Chen, W. H. *The Dissertation for the Degree of Master* **2006**.
181. Yamaguchi, M.; Ishida, J., *Modern Derivatization Methods for Separation Sciences: Reagent for FL Detection.* ed.; Wiley, Chichester: 1999; 'Vol.' p 99.
182. Kuroda, N.; Kawazoe, K.; Nakano, H.; Wada, M.; Nakashima, K. *Luminescence* **1999**, 14, (6), 361-4.
183. Kuroda, N.; Murasaki, N.; Wada, M.; Nakashima, K. *Luminescence* **2001**, 16, (2), 167-72.
184. Kuroda, N.; Sugihara, S.; Sugihara, Y.; Wada, M.; Kishikawa, N.; Ohba, Y.; Nakashima, K. *J. Chromatogr. A* **2005**, 1066, 119-25.
185. Suenaga, H.; Yamamoto, H.; Shinkai, S. *Pure&Appl. Chem.* **1996**, 68, 2179-86.
186. Yoon, J.; Czarnik, A. W. *J. Am. Chem. Soc.* **1992**, 114, 5874-5.
187. Favaudon, V.; Pochon, F. *Biochemistry* **1976**, 15, (18), 3903-12.
188. Mondal, M. S.; Wang, Z.; Seeds, A. M.; Rando, R. R. *Biochemistry* **2000**, 39, (2), 406-12.
189. Evans, S. A.; Olson, S. T.; Shore, J. D. *J. Biol. Chem.* **1982**, 257, (6), 3014-7.
190. Chen, A. P.; Chen, Y. H.; Liu, H. P.; Li, Y. C.; Chen, C. T.; Liang, P. H. *J. Am. Chem. Soc.* **2002**,



- 124, (51), 15217-24.
191. Chen, Y. H.; Chen, A. P.; Chen, C. T.; Wang, A. H.; Liang, P. H. *J. Biol. Chem.* **2002**, *277*, (9), 7369-76.
192. Zhang, J.; Geddes, C. D.; Lakowicz, J. R. *Anal. Biochem.* **2004**, *332*, 253-60.
193. Miyaura, N.; Suzuki, A. *Chem. Commun.* **1979**, 866-7.
194. Jones, G.; Willett, P.; Glen, R. C. *J. Mol. Biol.* **1995**, *245*, (1), 43-53.
195. Luo, L.; Chang, C. H.; Chen, Y. C.; Wu, T. K.; Diau, E. W. *J. Phys. Chem. B.* **2007**, *111*, 7656-64.
196. Cullen, B. R.; Greene, W. C. *Cell* **1989**, *58*, 423-6.
197. Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505-10.
198. Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818-22.
199. Willis, M. C.; LeCuyer, K. A.; Meisenheimer, K. M.; Uhlenbeck, O. C.; Koch, T. H. *Nucleic Acids Res.* **1994**, *22*, 4947-52.
200. Jensen, K. B.; Atkinson, B. L.; Willis, M. C.; Koch, T. H.; Gold, L. *Proc. Natl. Acad. Sci. U. S.A.* **1995**, *92*, 12220-4.
201. Gold, L.; Collins, B.; Willis, M. C.; Koch, T. H. *J. Bacteriol.* **1999**, *81*, 167-78.
202. Brody, E. N.; Gold, L. *Reviews in Molecular Biotechnology* **2000**, *74*, 5-13.
203. Blackwell, T. K.; Weintraub, H. *Science* **1990**, *250*, 1104-10.
204. Wright, W. E.; Binder, M.; Funk, W. *Mol. Cell. Biol.* **1991**, *11*, 4104-10.
205. Jenison, R. D.; S.C., G.; Pardi, A.; Polisky, B. *Science* **1994**, *263*, 1425-1429.
206. Jayasena, S. D. *Clinical Chemistry* **1999**, *45*, 1628-50.
207. Burgstaller, P.; Famulok, M. *Angew. Chem.* **1994**, *33*, 1084-7.
208. Lee, J. O.; So, H. M.; Jeon, E. K.; Chang, H.; Won, K.; Kim, Y. H. *Anal. Bioanal. Chem.* **2007**, *390*, 1023-32.
209. Ellington, A. D.; Szostak, J. W. *Nature* **1992**, *355*, 850-2.
210. Harada, K.; Frankel, A. D. *EMBO J.* **1995**, *14*, 5798-811.
211. Burke, D. H.; Gold, L. *Nucleic Acids Res.* **1997**, *25*, 2020-24.
212. Wallace, S. T.; Schroeder, R. *RNA* **1998**, *4*, 112-23.
213. Williams, K. P.; Liu, X. H.; Schumacher, T. N. M.; Lin, H. Y.; Ausiello, D. A.; Kim, P. S.; Bartel, D. P. *Proc. Natl. Acad. Sci. U. S.A.* **1997**, *94*, 11285-90.
214. Dang, C.; Jayasena, S. D. *J. Mol. Biol.* **1996**, *264*, 268-78.
215. Pan, W.; Craven, R. C.; Qiu, Q.; Wilson, C. B.; Wills, J. W.; Golovine, S.; Wang, J. F. *Proc. Natl. Acad. Sci. U. S.A.* **1995**, *92*, 11509-13.
216. Wang, J.; Jiang, H.; Liu, F. *RNA* **2000**, *6*, 571-83.
217. Morris, K. N.; Jensen, K. B.; Julin, C. M.; Weil, M.; Gold, L. *Proc. Natl. Acad. Sci. U. S.A.* **1998**, *95*, 2902-7.
218. Homann, M.; Goring, H. U. *Nucleic Acids Res.* **1999**, *27*, 2006-14.
219. Blank, M.; Weinschenk, T.; Priemer, M.; Schluesener, H. *J. Biol. Chem.* **2001**, *276*, 16464-8.
220. Davis, K. A.; Abrams, B.; Lin, Y.; Jayasena, S. D. *Nucleic Acids Res.* **1996**, *24*, 702-6.
221. Liss, M.; Petersen, B.; Wolf, H.; Prohaska, E. *Anal. Chem.* **2002**, *74*, 4488-95.
222. Kawakami, J.; Imanaka, H.; Yokota, Y.; Sugimoto, N. *J. Inorg. Biochem.* **2000**, *82*, 197-206.
223. Ikebukuro, K.; Kiyohara, C.; Sode, K. *Anal. Lett.* **2004**, *37*, 2901-9.
224. Ikebukuro, K.; Kiyohara, C.; Sode, K. *Biosens. Bioelectron.* **2005**, *20*, 2168-72.
225. Xiao, Y.; Lubin, A. A.; Heeger, A. J.; Plaxco, K. W. *Angew. Chem., Int. Ed.* **2005**, *117*, 5592-5.
226. Xiao, Y.; Piorek, B. D.; Plaxco, K. W.; Heeger, A. J. *J. Am. Chem. Soc.* **2005**, *127*, 17990-1.
227. Baker, B. R.; Lai, R. Y.; Wood, M. S.; Doctor, E. H.; Heeger, A. J.; Plaxco, K. W. *J. Am. Chem. Soc.* **2006**, *128*, 3138-9.
228. Lai, R. Y.; Plaxco, K. W.; Heeger, A. J. *Anal. Chem.* **2007**, *79*, 229-33.
229. So, H. M.; Won, K.; Kim, Y. H.; Kim, B. K.; Ryu, B. H.; Na, P. S.; Kim, H.; Lee, J. O. *J. Am. Chem. Soc.* **2005**, *127*, 11906-7.

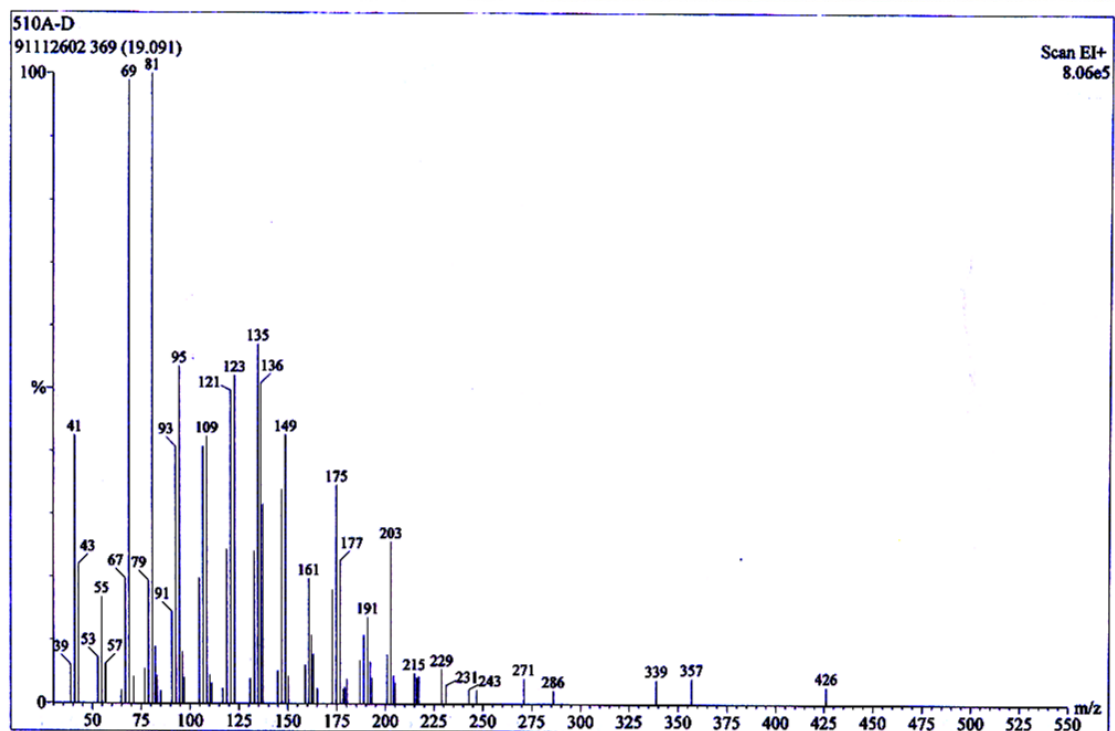
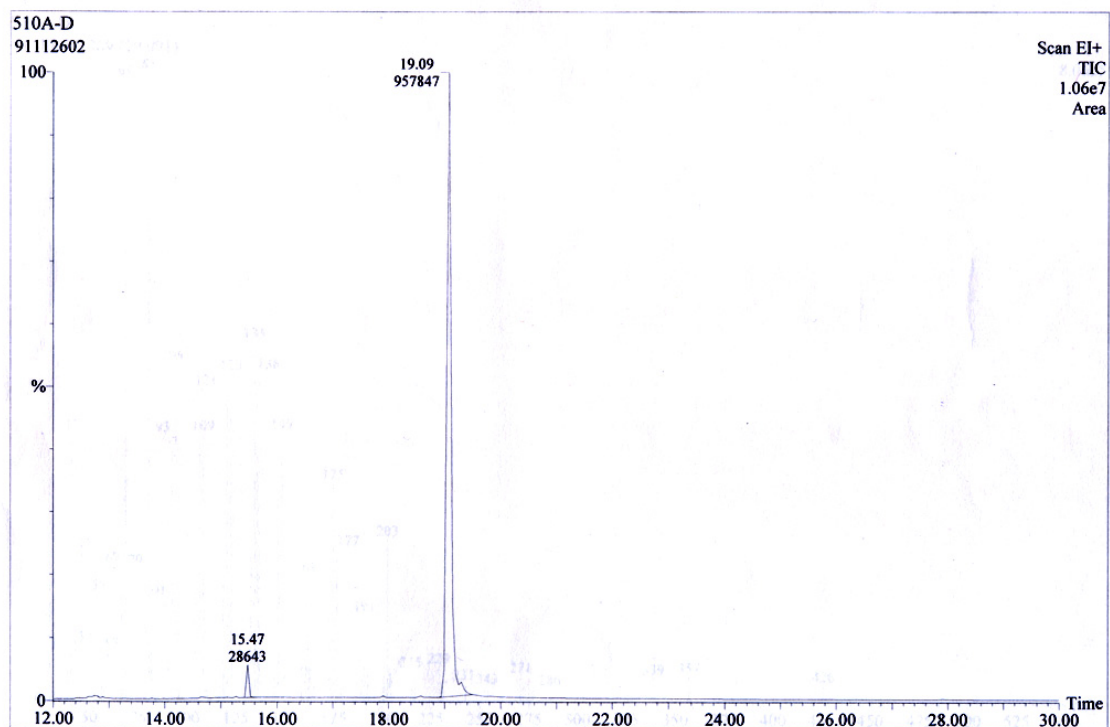
230. Maehashi, K.; Katsura, T.; Kerman, K.; Takamura, Y.; Matsumoto, K.; Tamiya, E. *Anal. Chem.* **2007**, *79*, 782-7.
231. Jellinek, D.; Green, L. S.; Bell, C.; Lynott, C. K.; Gill, N. *Biochemistry* **1995**, *34*, 11363-72.
232. Andreola, M. L.; Calmels, C.; Michel, J.; Toulme, J. J.; Litvak, S. *Eur. J. Biochem.* **2000**, *267*, 5032-40.
233. Klusmann, S.; Nolte, A.; Bald, R.; Erdmann, V. A.; Furste, J. P. *Nat. Biotechnol.* **1996**, *14*, 1112-5.
234. Tucker, C. E.; Chen, L.; Judkins, M. B.; Farmer, J. A.; Gill, S. C.; Drolet, D. W. *J. Chromatography B: Biomedical Sciences and Applications* **1999**, *732*, 203-12.
235. Willis, M. C.; Collins, B.; Zhang, T.; Green, L. S.; Sebesta, D. P.; Bell, C.; Kellogg, E.; Gill, S. C.; Magallanez, A.; Knauer, S.; Bendelle, R. A.; Gill, P. S.; Janjic, N. *Bioconjug. Chem* **1998**, *9*, 633.
236. Ng, E. W.; Shima, D. T.; Calias, P.; E.T.Jr., C.; Guyer, D. R.; Adamis, A. P. *Nature Reviews in Drug Discovery* **2006**, *5*, 123-32.
237. Patel, D. J.; Suri, A. K. *Reviews in Molecular Biotechnology* **2000**, *74*, 39-60.
238. Dieckmann, T.; Suzuki, E.; Nakamura, G. K.; Feigon, J. *RNA* **1996**, *2*, 628-40.
239. Fan, P.; Suri, A. K.; Fiala, R.; Live, D.; Patel, D. J. *J. Mol. Biol.* **1996**, *258*, 480-500.
240. Zimmermann, G. R.; Jenison, R. D.; Wick, C. L.; Simorre, J.; Pardi, A. *Nat. Struct. Biol.* **1997**, *4*, 644-9.
241. Convery, M. A.; Rowsell, S.; Stonehouse, M. J. *Nat. Struct. Biol.* **1998**, *5*, 133-9.
242. Rowsell, S.; Stonehouse, N. J.; Convery, M. A. *Nat. Struct. Biol.* **1998**, *5*, 970-75.
243. Hermann, T.; Patel, D. J. *Science* **2000**, *287*, 820-5.
244. Wimberly, B.; Varani, G.; Jr., T. *Biochemistry* **1993**, *32*, 1078-87.
245. Huang, S.; Wang, Y. X.; Draper, D. E. *J. Mol. Biol.* **1996**, *258*, 308-21.
246. Wu, T. K.; Yeh, C. Y. *The Dissertation for the Degree of Master* **2005**.
247. Zhang, X. M.; Shao, N. S.; Chi, M. G.; Sun, M. J. *Acta. Pharmacol. Sin.* **2003**, *24*, 711-14.
248. Dorlet, D. W.; Jenison, R. D.; Smith, D. E.; Pratt, D.; Hicke, B. J. *Combinatorial Chemistry & High Throughput Screening* **1999**, *2*, 271-8.
249. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular cloning: A laboratory manual 2nd ed* **1989**.
250. Win, M. N.; Klein, J. S.; Smolke, C. D. *Nucleic Acids Res.* **2006**, *34*, 5670--82.
251. Ikebukuro, K.; Takase, M.; Sode, K. *Nucleic Acids Symposium Series* **2007**, *51*, 403-4.
252. Ikebukuro, K.; Hasegawa, H.; Sode, K. *Nucleic Acids Symposium Series* **2007**, *51*, 399-400.
253. Zuker, M. *Nucleic Acids Res.* **2003**, *31*, 3406-15.
254. Buntel, C. *Stanford University, Stanford, CA* **1996**.
255. Ryder, N. S.; Seidl, G.; Troke, P. F. *Antimicrob. Agents. Chemother.* **1984**, *25*, 483-87.
256. Freimund, S.; Kopper, S. *Carbohydrate Research* **1998**, *308*, 195-200.
257. Wach, A.; Brachat, A.; Pohlmann, R.; Philippsen, P. *Yeast* **1994**, *10*, 1793-1808.
258. Wach, A. *Yeast* **1996**, *12*, 259-65.
259. Geitz, R. D.; Woods, R. A. *Methods Enzymol.* **2002**, *350*, 87-96.
260. Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, (22), 4673-80.
261. Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. *Anal. Chem.* **1996**, *68*, 850-8.

# Appendix

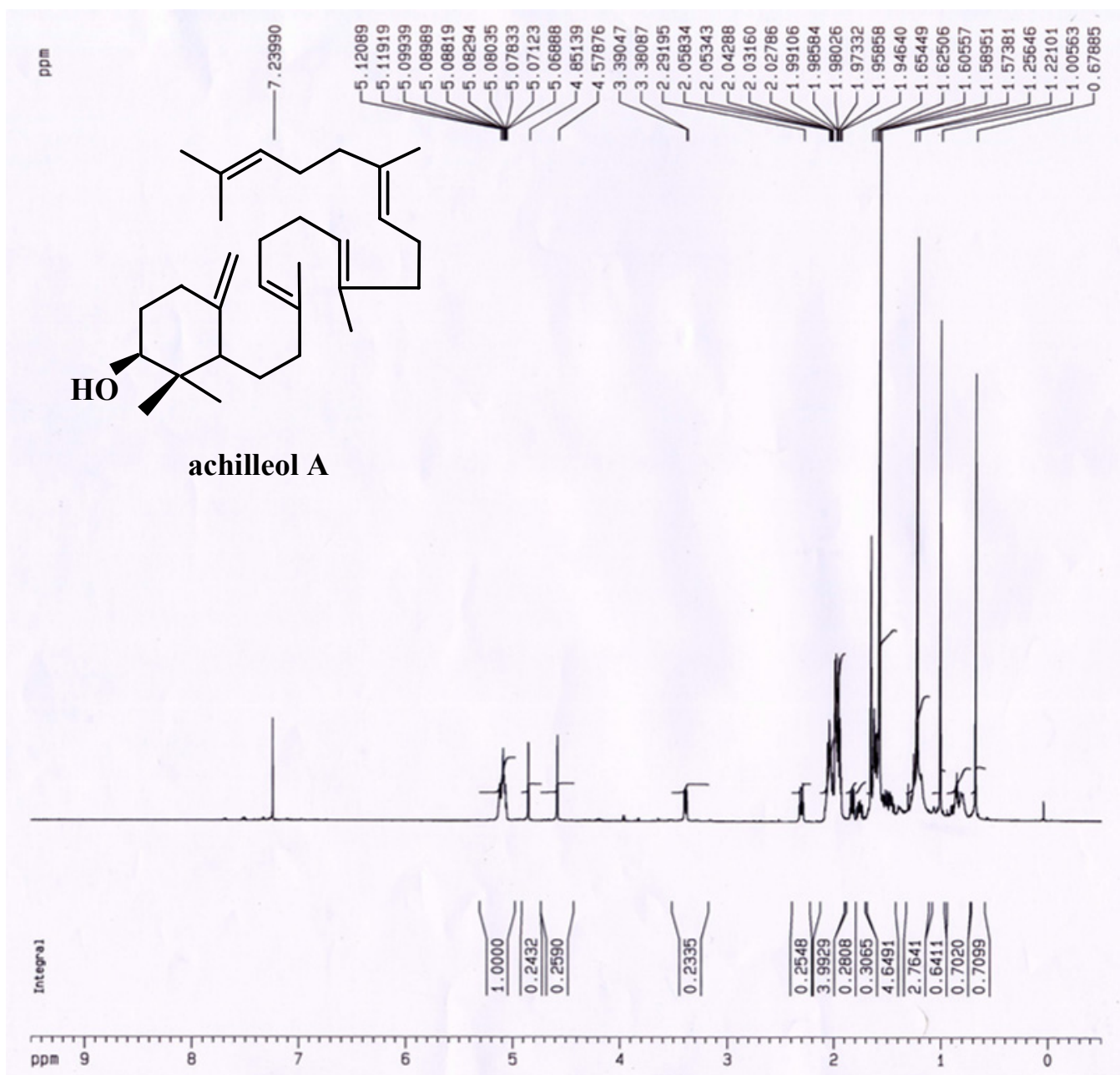
## Appendix 3.1

Appendix 3.1 Monocyclic triterpene achilleol A characterization.

1. The GC relative retention time (relative to ergosterol) is -2.14.
2. MS  $m/z = 426 [M]^+$ , 357, 339, 286, 271, 203, 175, 149, 135, 121, 109, 95, 81, 69, 68, 55, 41.
3. The GC-MS and NMR data were consistent with those published in the literature<sup>1,106</sup>

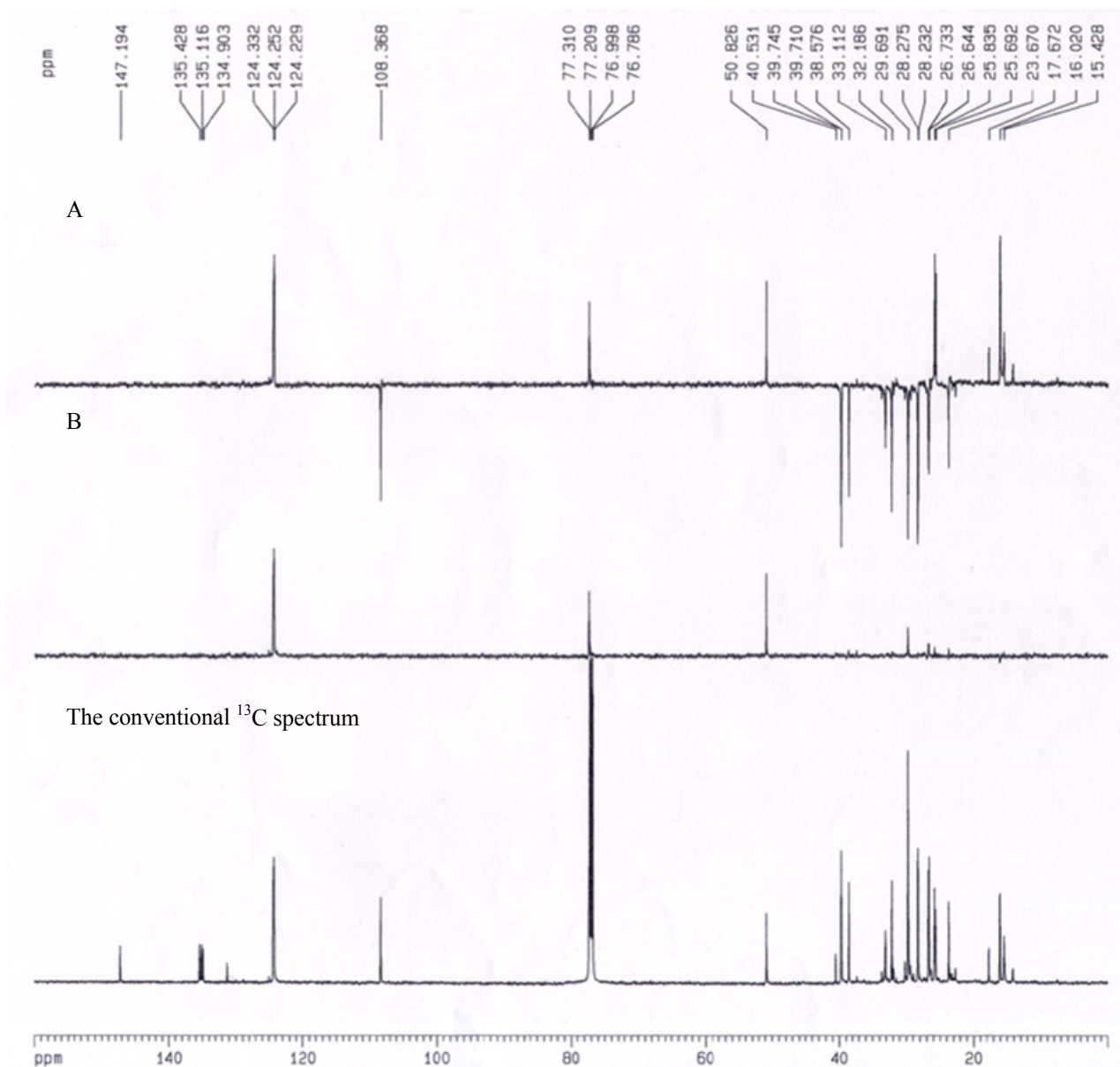
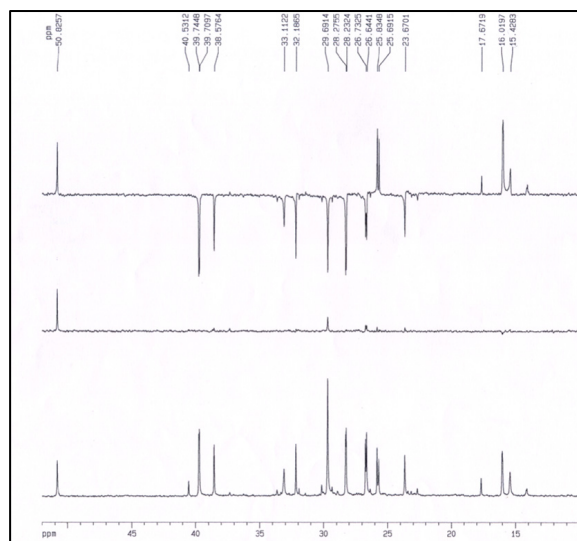


The  $^1\text{H-NMR}$  spectrum of achilleol A



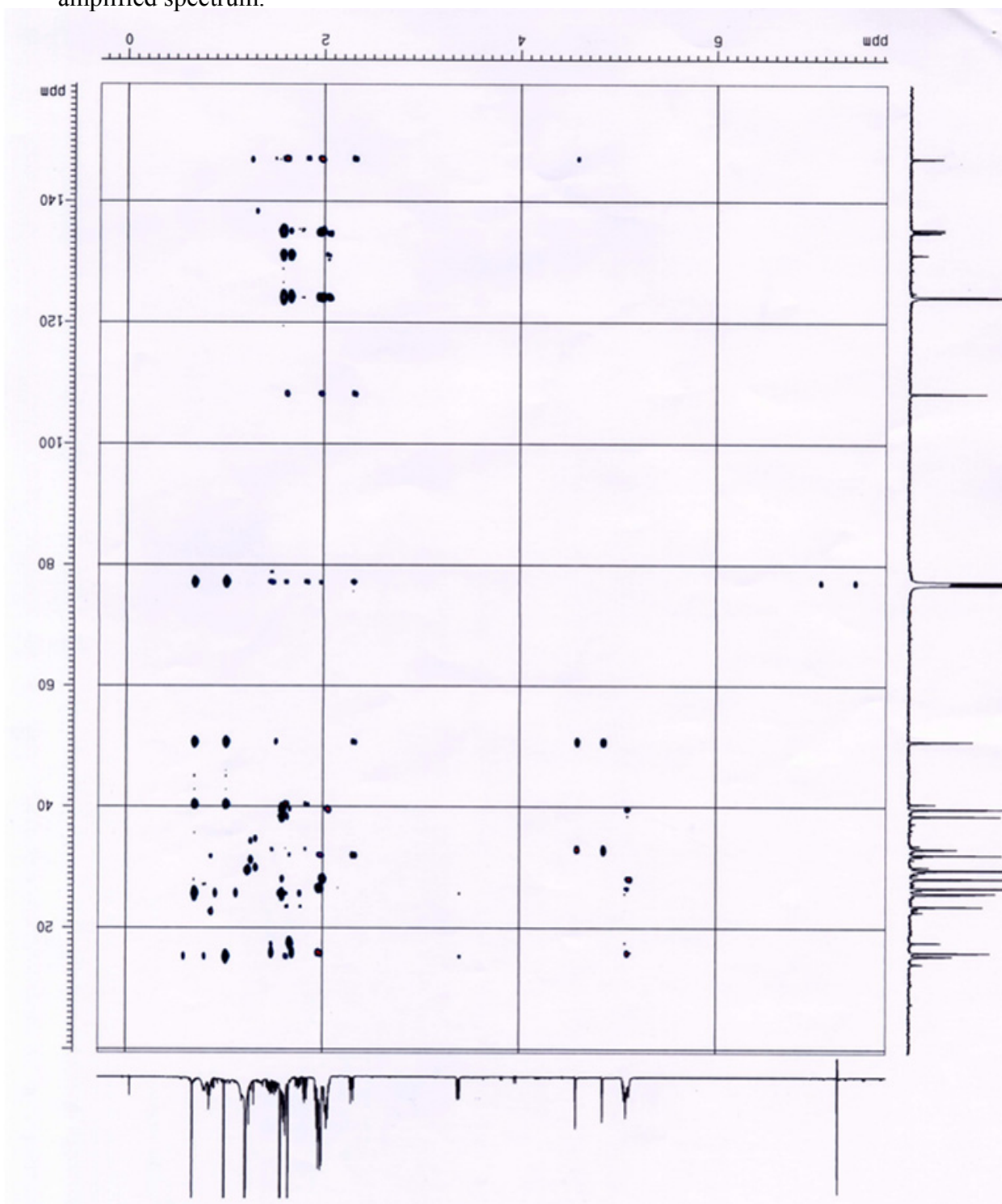
The  $^{13}\text{C}$ -NMR (DEPT) spectrum of the achilleol A. The inset indicates the partial spectrum.

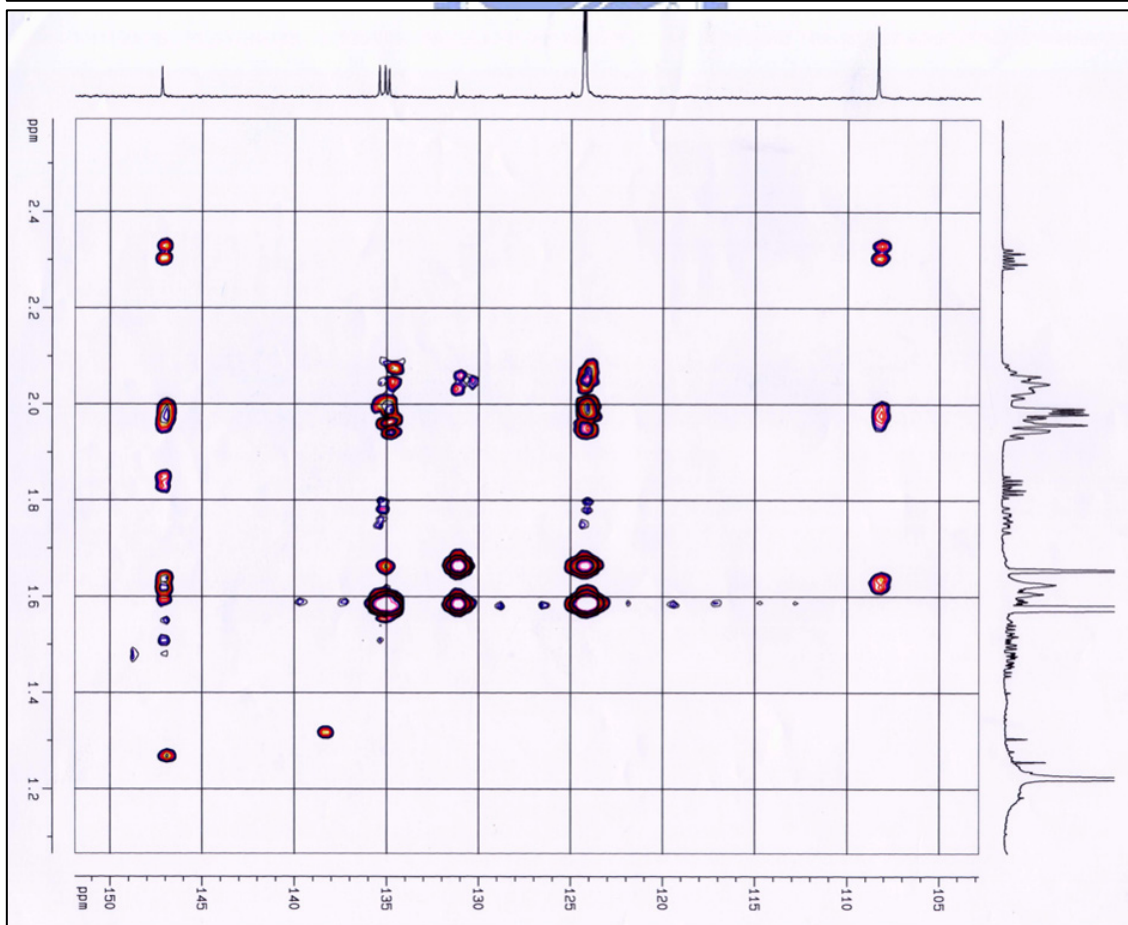
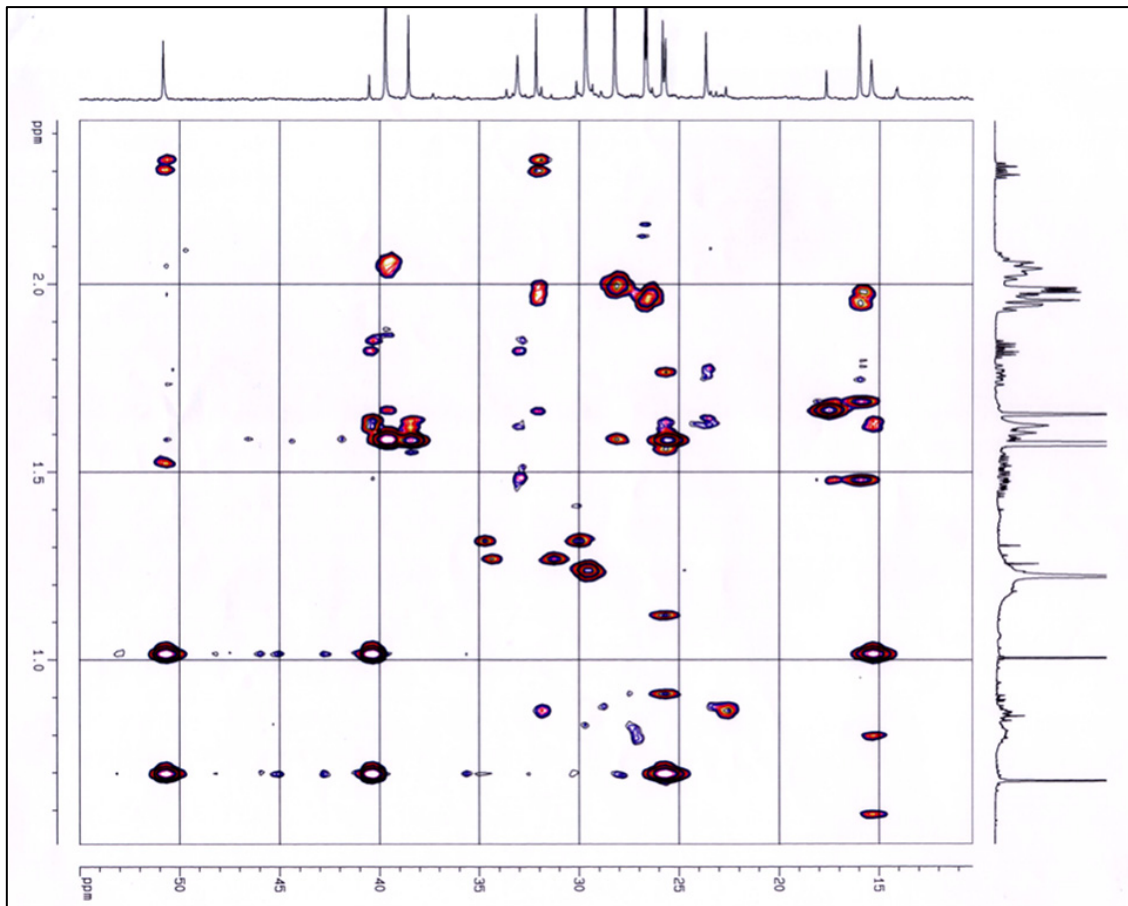
- Subspectrum A,  $\text{CH}_3$  and  $\text{CH}$  up,  $\text{CH}_2$  down.
- Subspectrum B,  $\text{CH}$  up.
- The conventional  $^{13}\text{C}$  spectrum is at the bottom of the figure.





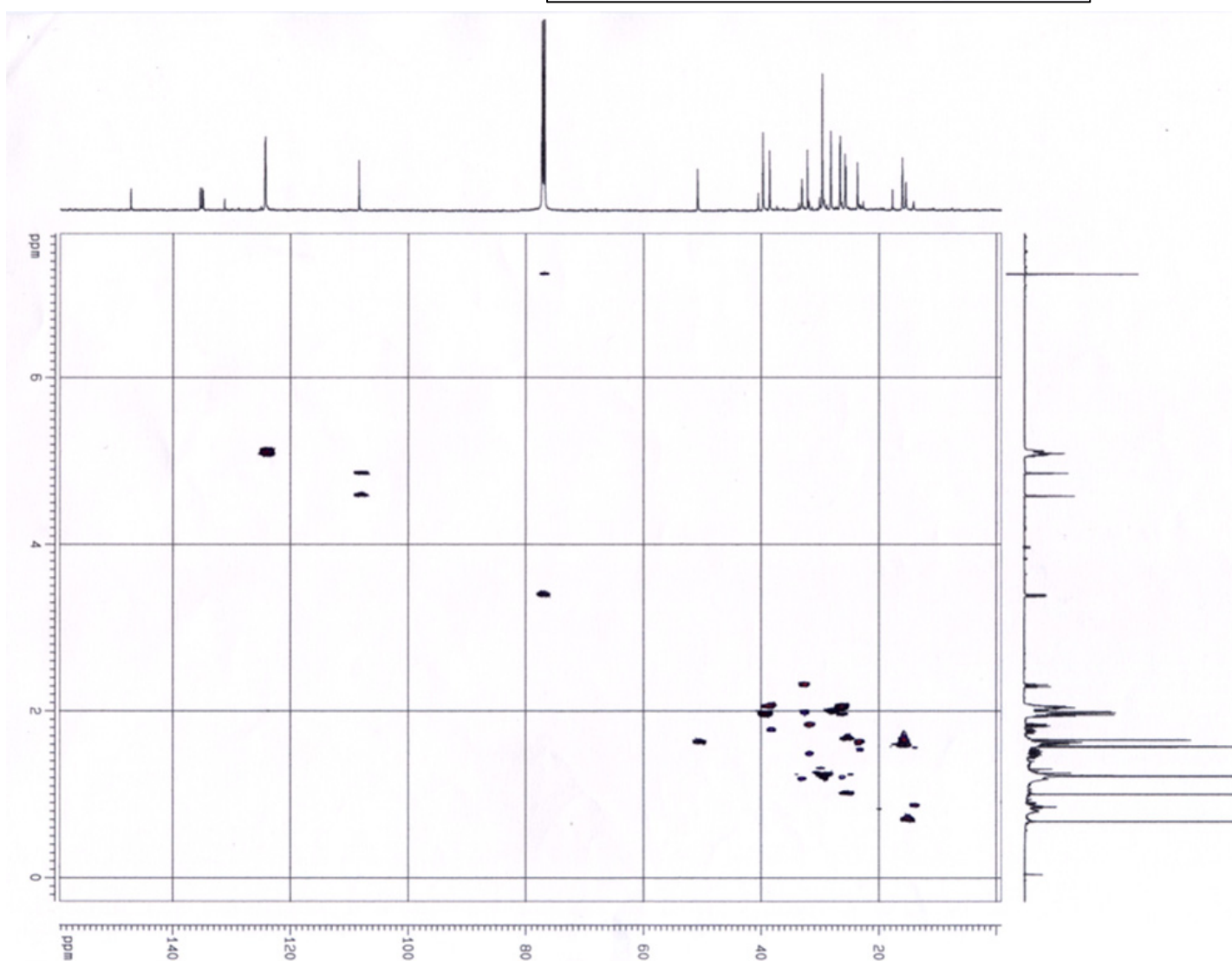
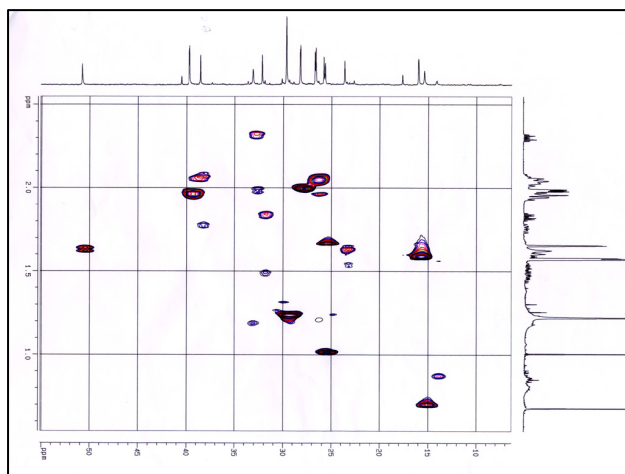
The HMBC spectrum of achilleol A. The inset on next page indicates the partially amplified spectrum.





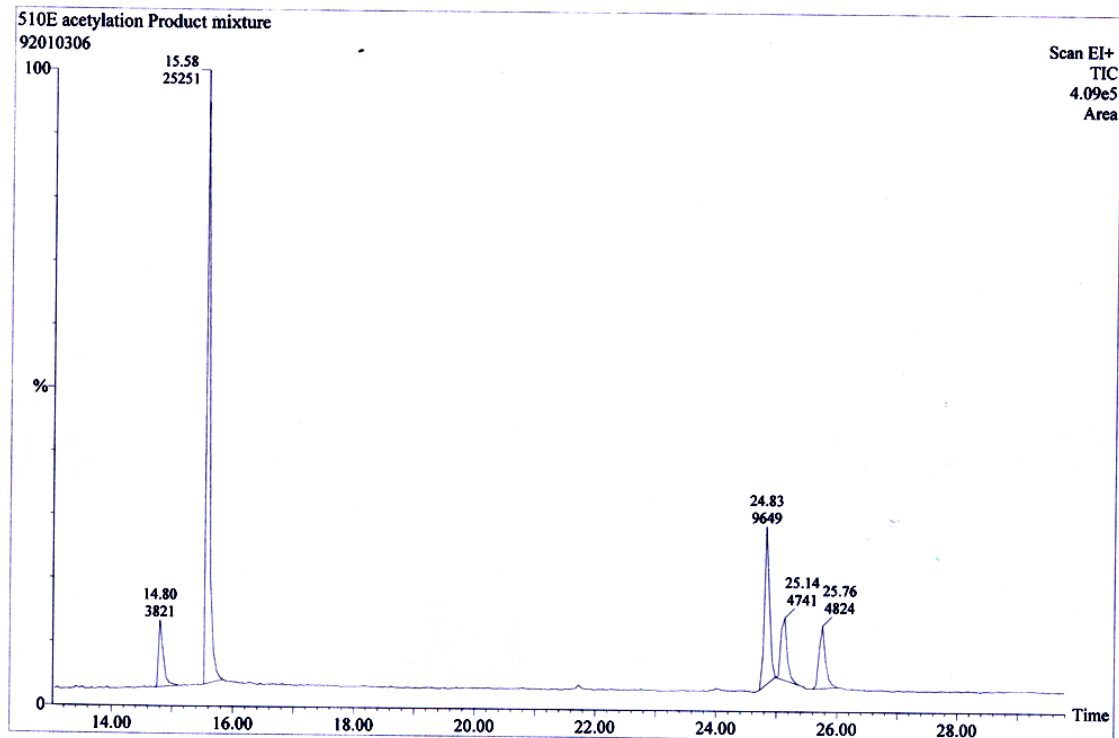


The HMQC spectrum of achilleol A. The inset indicates the partially amplified spectrum.

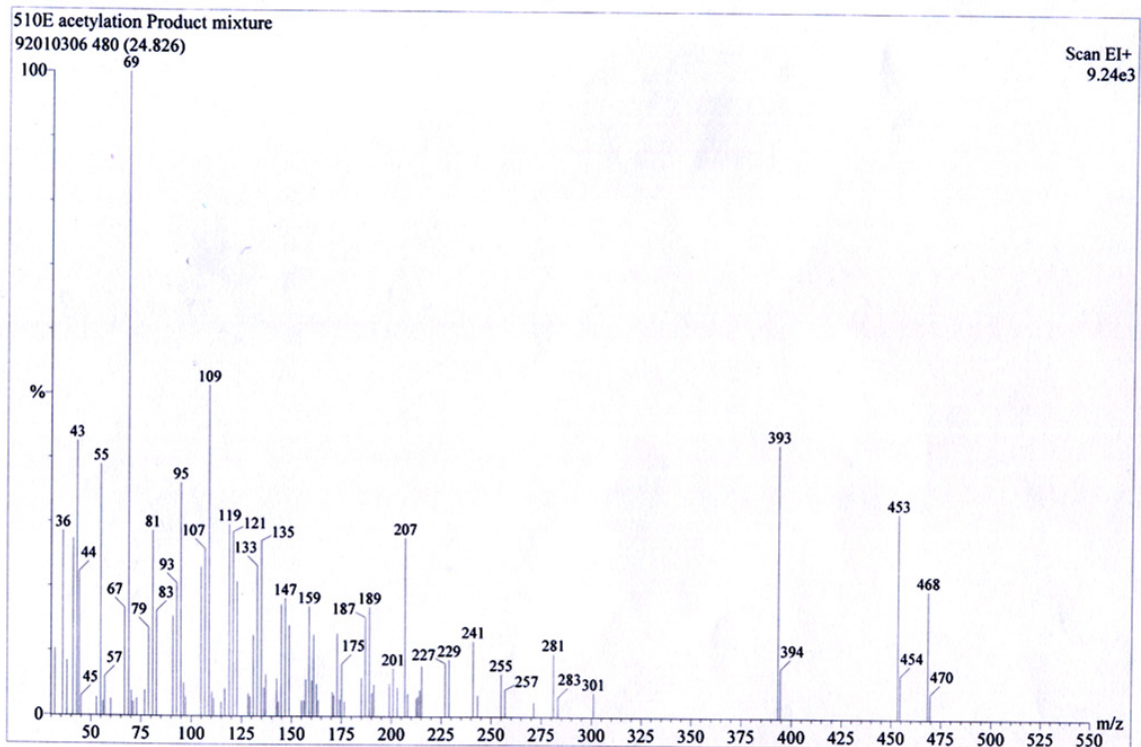


## Appendix 3.2

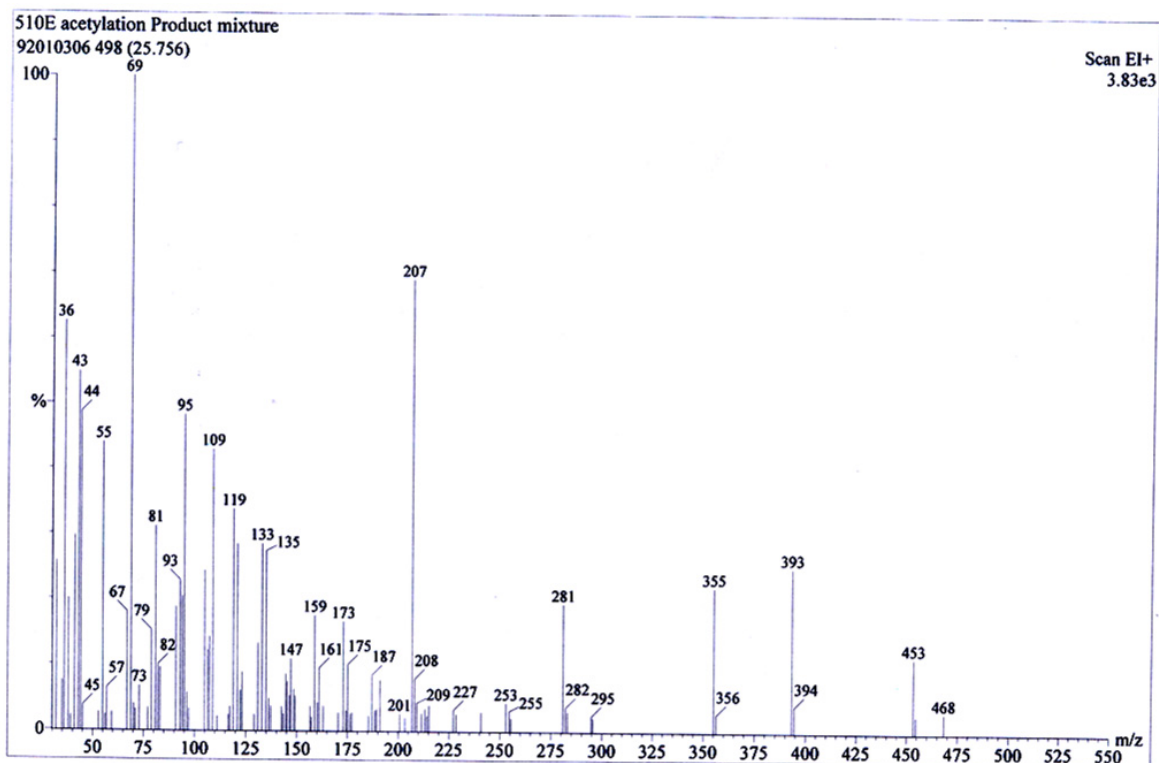
Appendix 3.2 The GC-MS spectrum of lanosterol-positioned product in Tyr510Ala mutant



**Lanosteryl acetate** The GC relative retention time of lanosterol (measured relative to ergosterol) is +1.64. MS for lanosteryl acetate,  $m/z=468 [M]^+$ , 453, 393, 339, 301, 281, 255, 241, 229, 189, 159, 109, 95, 81, 69(100), 55.



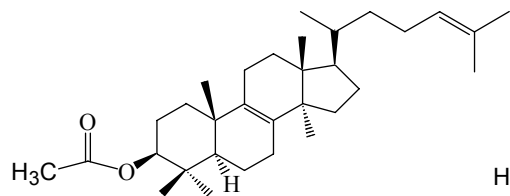
**Parkeyl acetate** The GC relative retention time of parkeol (measured relative to ergosterol) is +2.34. MS for parkeyl acetate,  $m/z = 468[M]^+$ , 453, 393, 355, 295, 207, 173, 159, 133, 109, 95, 81, 69 (100), 55.



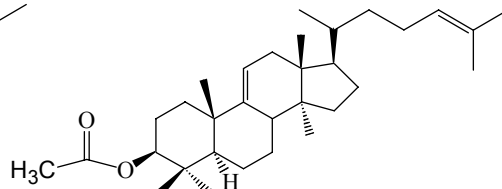
## Appendix 3.3

### Appendix 3.3 Lanosteryl acetate and parkeyl acetate NMR characterization

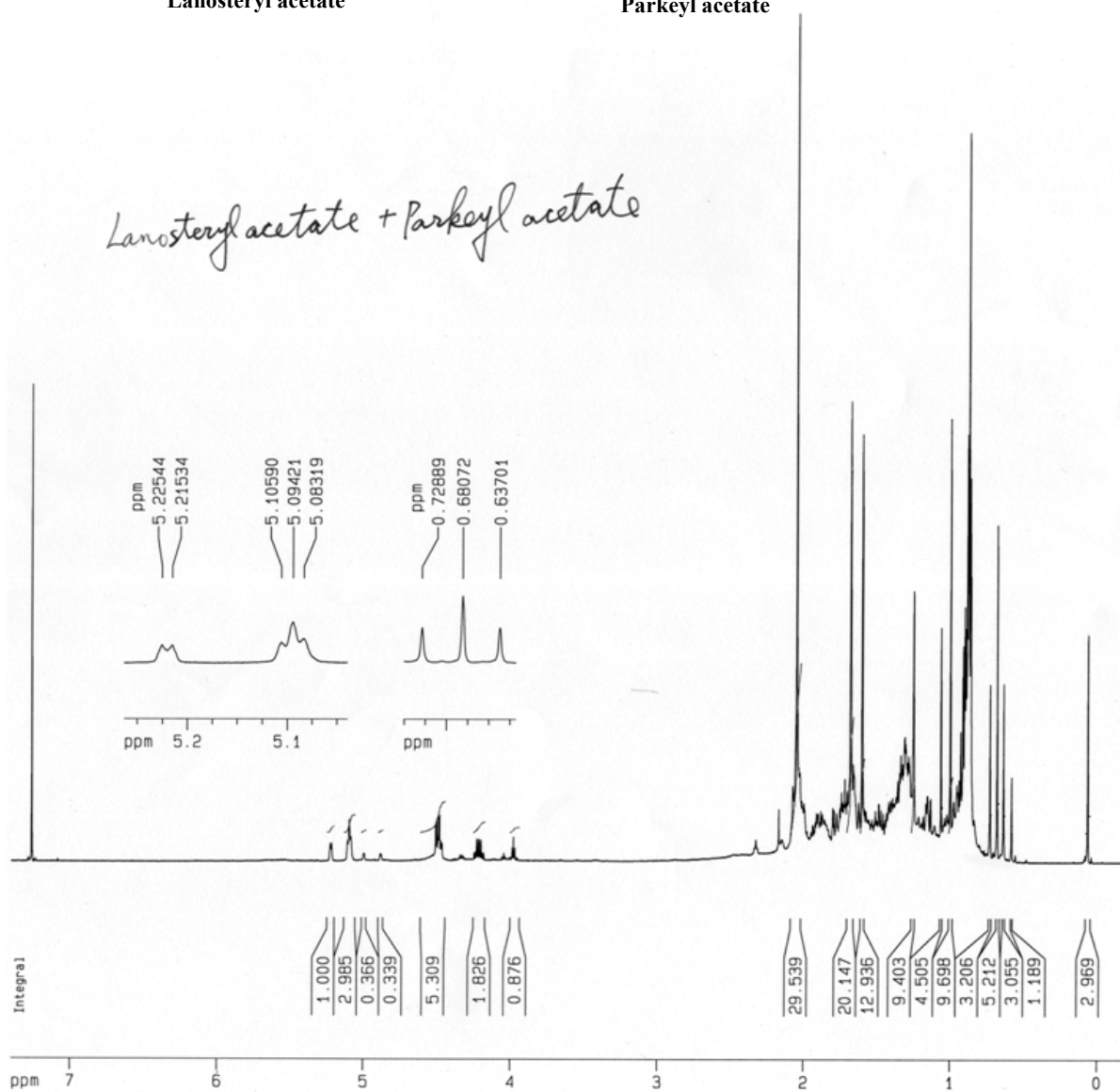
The  $^1\text{H}$ -NMR spectrums of lanosteryl acetate and parkeyl acetate



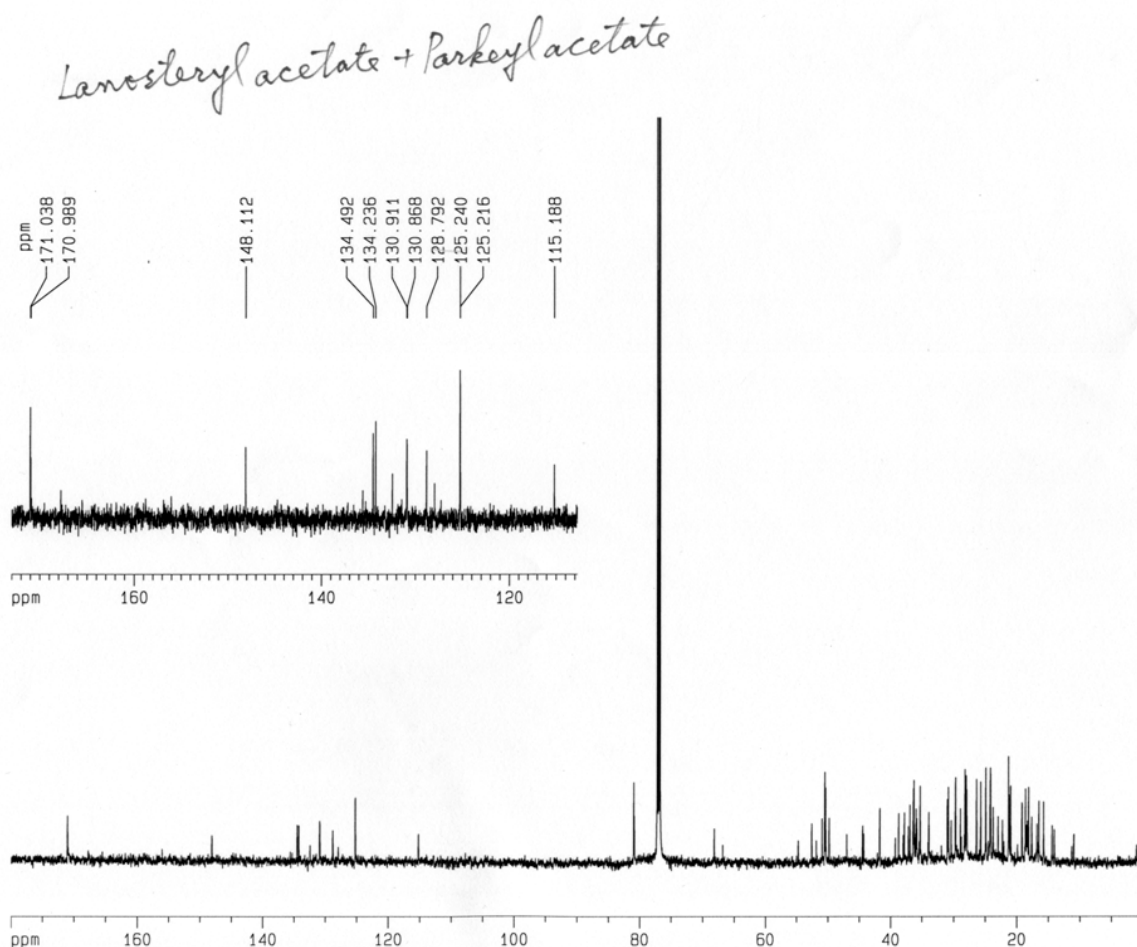
Lanosteryl acetate



Parkeyl acetate



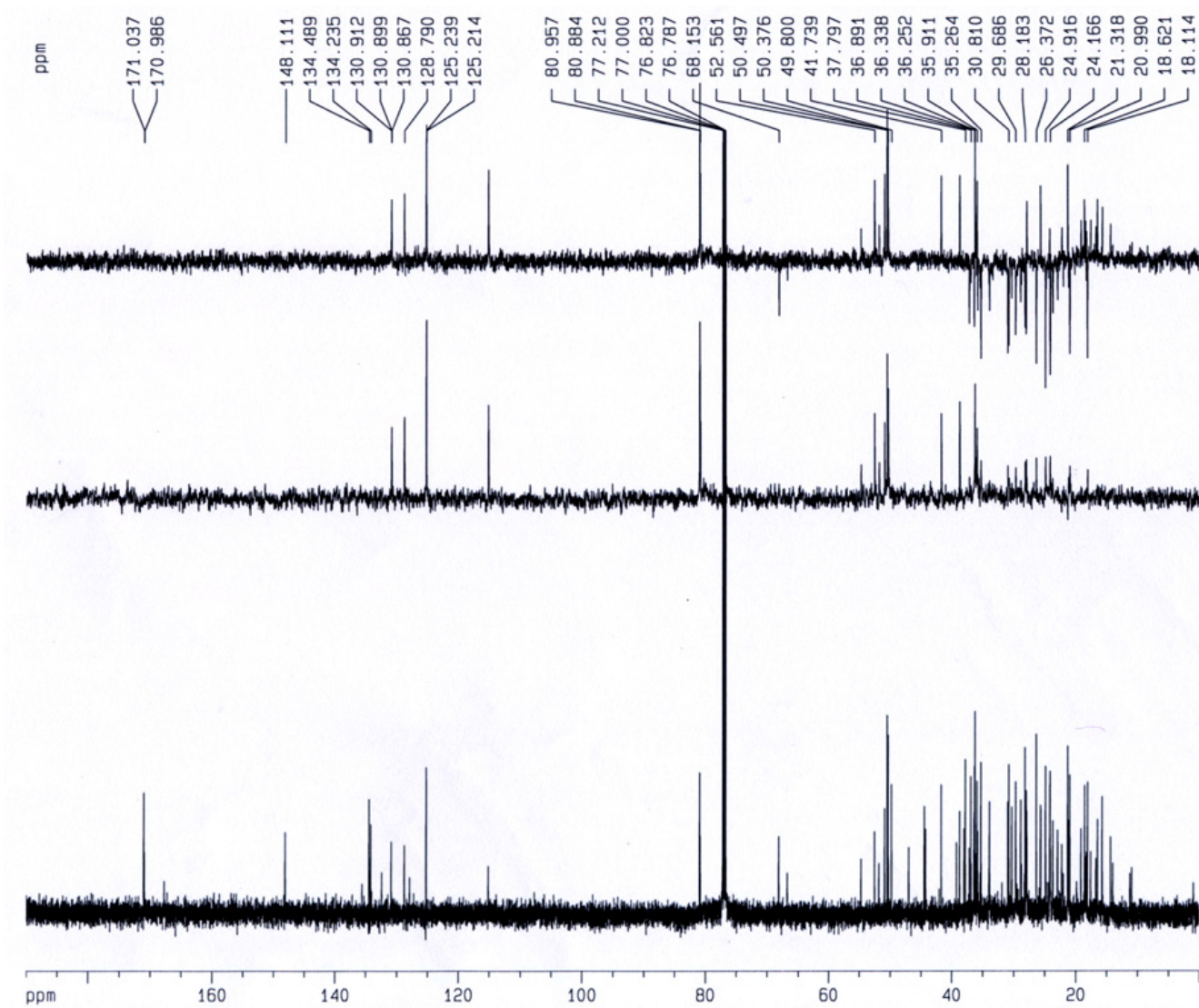
The  $^{13}\text{C}$ -NMR spectrum of the lanosteryl acetate and parkeyl acetate. The inset indicates the partial  $^{13}\text{C}$ -NMR spectrum.



$^{13}\text{C}$ -NMR (150.68MHz, 30mM solution in  $\text{CDCl}_3$ ,  $25^\circ\text{C}$ )

$\delta$  15.74 (C-18), 16.51 (C-29), 17.62 (C-27), 18.11 (C-6), 18.62 (C-21), 19.16 (C-19), 20.97 (C-11), 21.32 ( $\text{CH}_3\text{CO-}$ ), 24.17 (C-2), 24.23 (C-30), 24.92 (C-23), 25.74 (C-26), 26.37 (C-7), 27.90 (C-28), 28.18 (C-16), 30.81 (C-15), 30.96 (C-12), 35.26 (C-1), 36.25 (C-20), 36.34 (C-22), 36.89 (C-10), 37.80 (C-4), 44.47 (C-13), 49.80 (C-14), 50.38 (C-17), 50.50 (C-5), 80.96 (C-3), 125.24 (C-24), 130.87 (C-25), 134.24 (C-9), 134.49 (C-8), 171.02 ( $\text{CH}_3\text{CO-}$ ). Chemical shift ( $\pm 0.02$  ppm) were referenced to the  $\text{CDCl}_3$  signal at  $\delta$  77.0 ppm

DEPT spectrum of the lanosteryl acetate and parkeyl acetate.

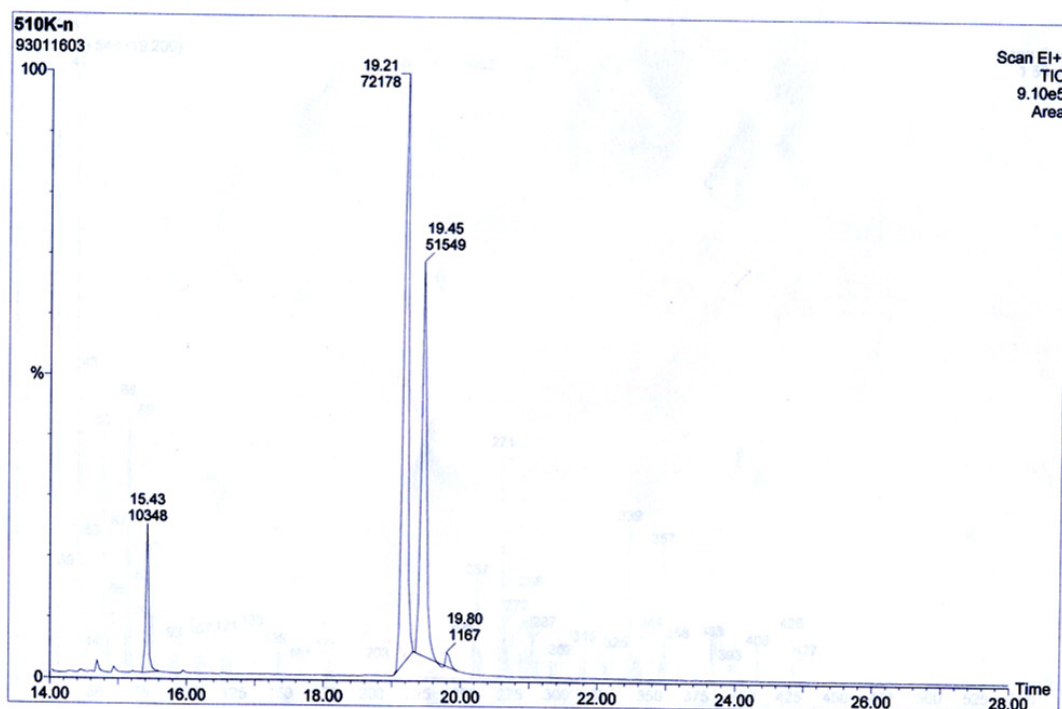




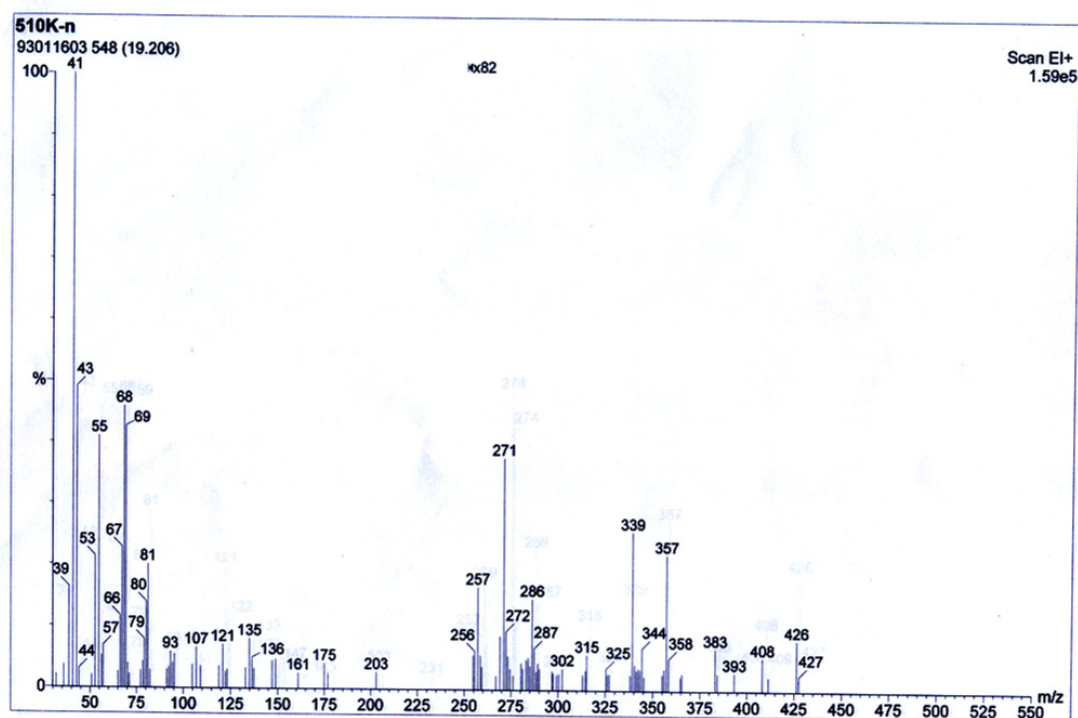
## Appendix 3.4

### Appendix 3.4 Monocyclic triterpenes achilleol A and camelliol C characterization.

1. The GC relative retention time (measured relative to ergosterol ) of achilleol A is  $-2.14$ . MS  $m/z = 426 [M]^+$ , 357, 339, 286, 271, 203, 175, 149, 135, 121, 109, 95, 81, 69, 68, 55, 41
2. The GC relative retention time (measured relative to ergosterol) of camelliol C is  $-1.9$ . MS  $m/z = 426 [M]^+$ , 408, 383, 357, 339, 315, 286, 274, 259, 231, 203, 121, 93, 81, 69 68, 55, 41 (100).

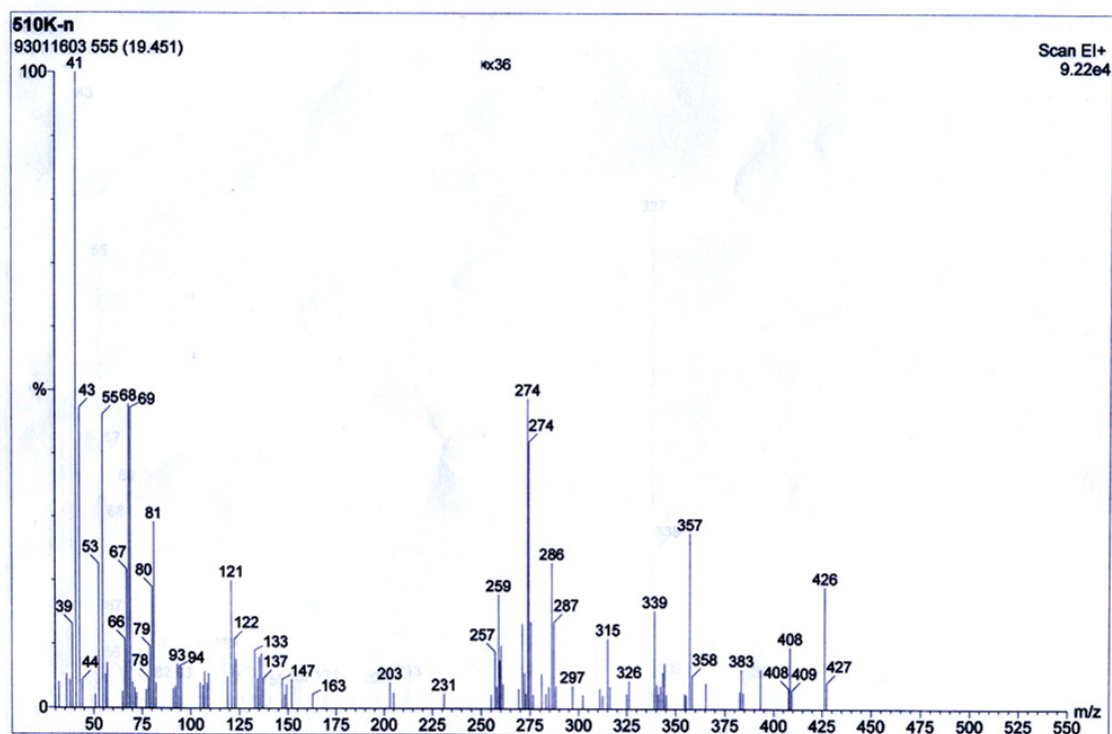


### 3. achilleol A



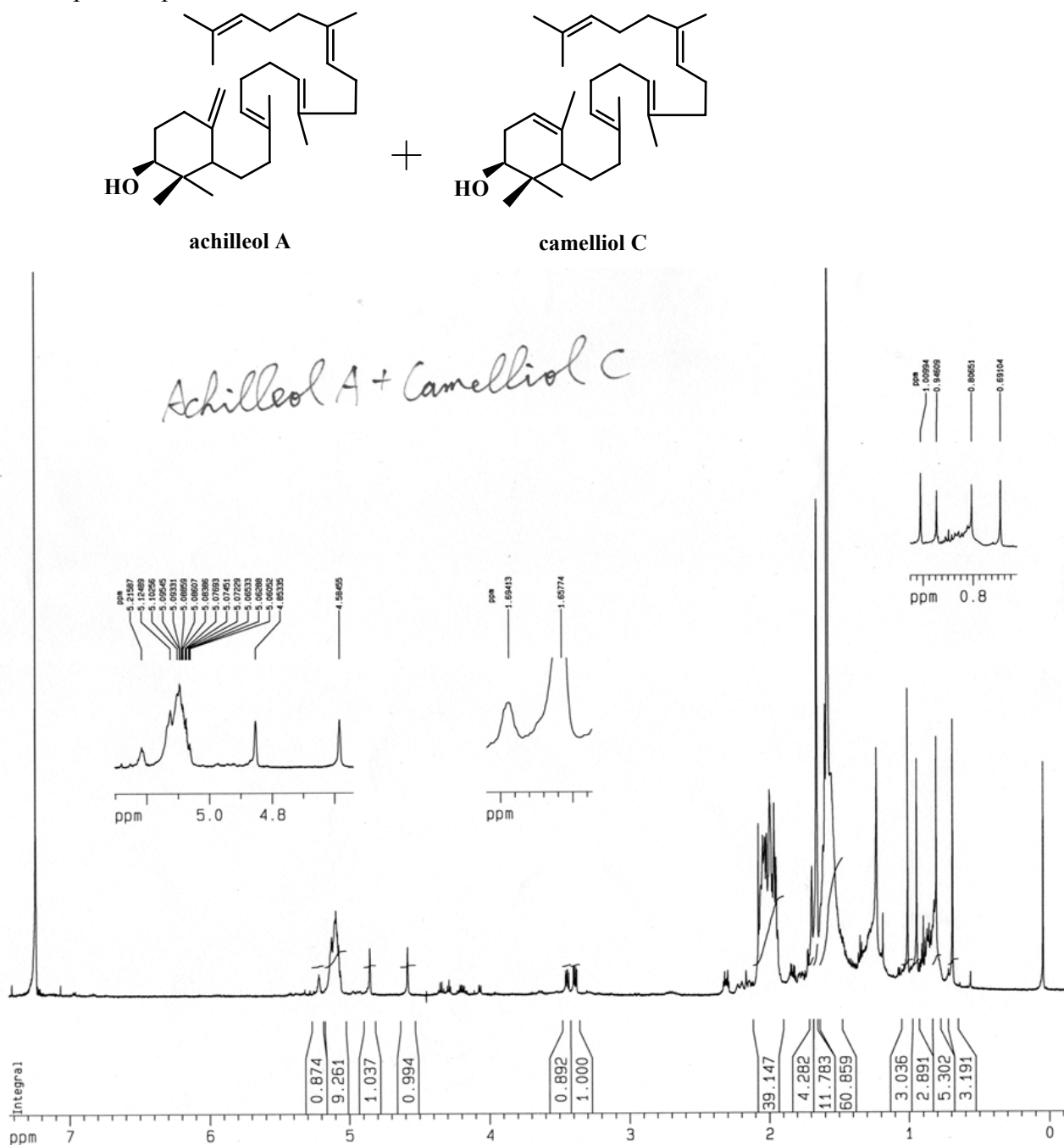


4. camelliol C



## 5. NMR assignments of camelliol C

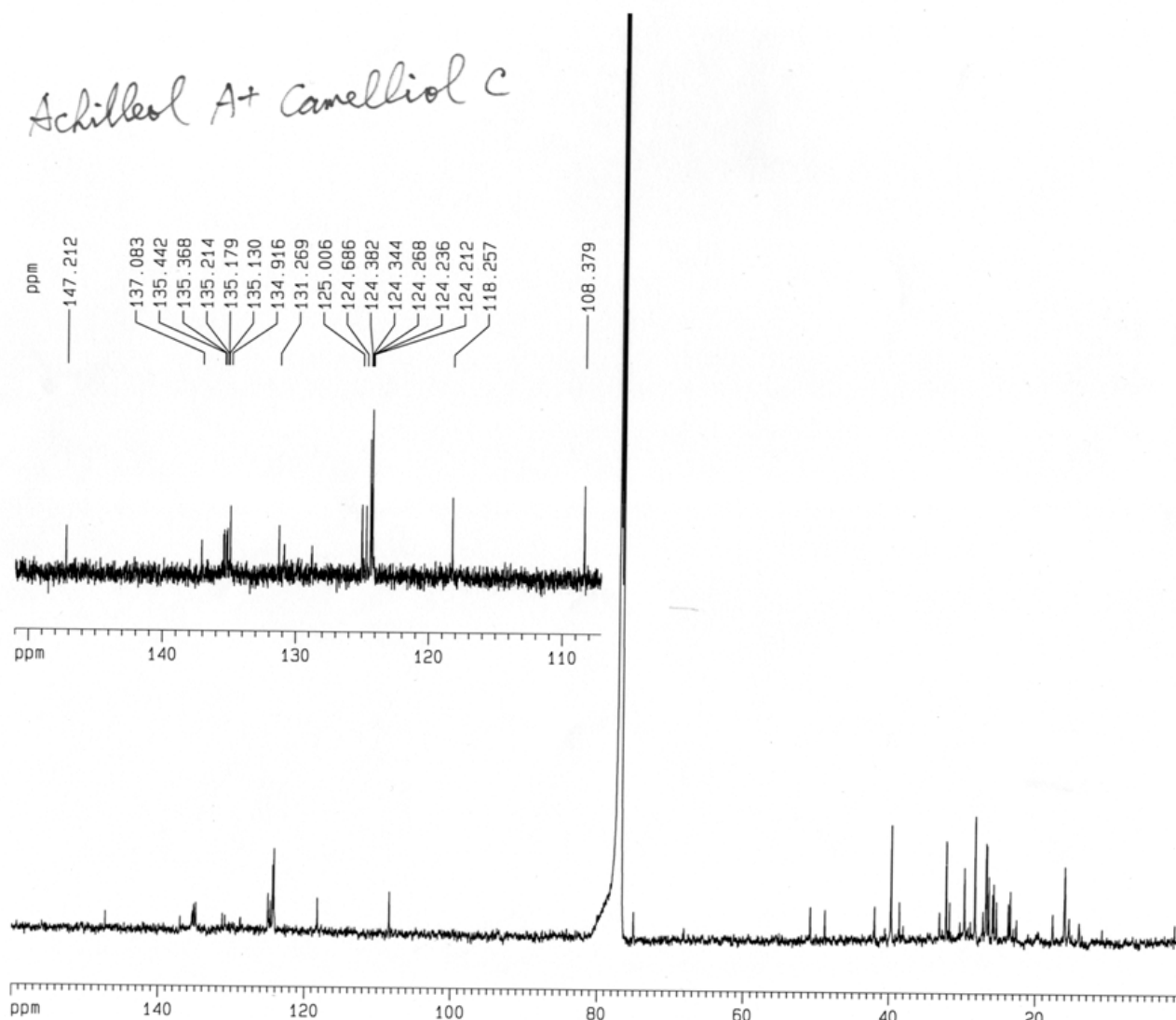
The  $^1\text{H}$ -NMR spectrum of achilleol A and camelliol C. The insets indicate the partial amplified spectrum.



$^1\text{H}$ -NMR of representative spectra (600 MHz, 5mM solution in  $\text{CDCl}_3$ ,  $25^\circ\text{C}$ )  $\delta$  0.807 (s,3H,C-24), 0.943 (s,3H, C-23), 1.348 (M, 1H), 1.560-1.613 (M, 13H), 1.658 (s,3H,C-30), 1.694 (s, 3H, C-25), 1.760 (m, 1H), 1.952-2.078 (m, 1H), 2.300 (M, 1H), 3.433-3.456 (dd,  $J=8.2$  and  $5.5\text{Hz}$ ,1H), 5.061-5.125 (m, 4H), 5.216 (m,1H). Chemical shift were referenced to  $\text{Si}(\text{CH}_3)_4$  and are generally accurate to  $\pm 0.001$  ppm.

The  $^{13}\text{C}$ -NMR spectrum of the achilleol A and camelliol C.

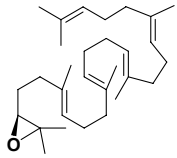
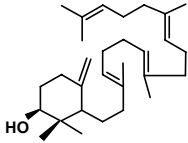
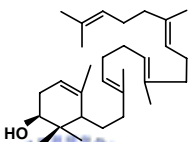
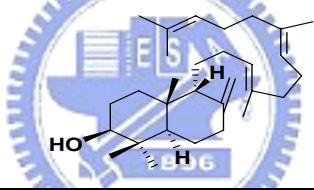
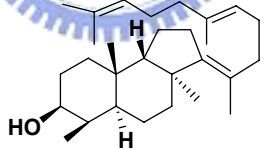
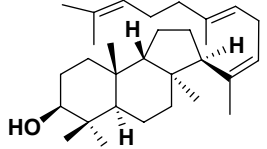
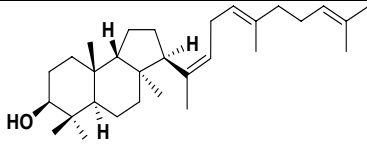
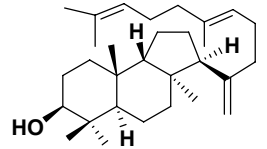
The inset spectrum indicates the partial  $^{13}\text{C}$ -NMR (DEPT) spectrum



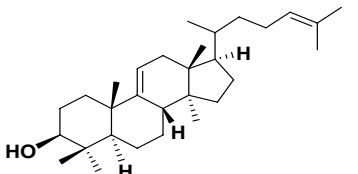
$^{13}\text{C}$ -NMR (150.68MHz, 30 mM solution in  $\text{CDCl}_3$ , 25°C)

$\delta$  16.03(C-26), 16.07(2C, C-27 & C-28), 16.10 (C-24), 17.67 (C-29), 22.57 (C-25), 25.31 (C-23), 25.70 (C-30), 26.66 (C-16), 26.75 (C-20), 27.14 (C-6), 28.29 (2C, C-11 & C-12), 31.75 (C-2), 38.08 (C-4), 39.76 (2C, C-15 & C-19), 41.99 (C-7), 48.83 (C-5), 75.04 (C-3), 118.26 (C-1), 124.21 (C-13), 124.24 (C-17), 124.38 (C-21), 124.69 (C-9), 131.27 (C-22), 134.92 (C-14), 135.18 (C-8), 135.44 (C-18), 137.08 (C-10).

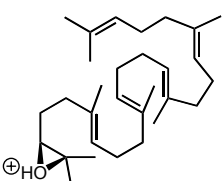
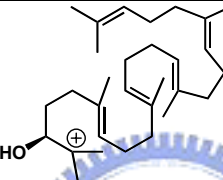
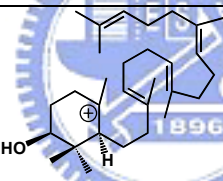
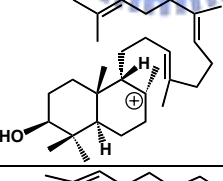
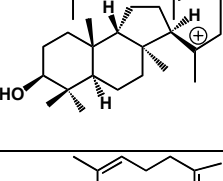
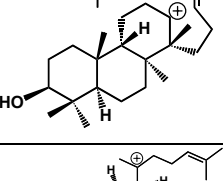
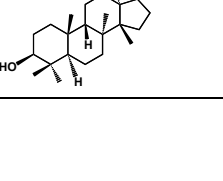
## Appendix 5 Relative energy of respective product in protersterol pathway

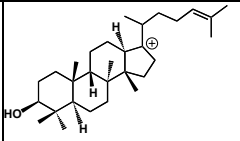
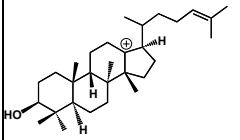
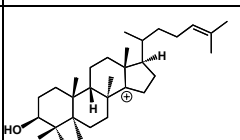
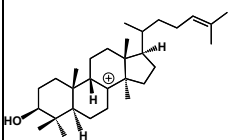
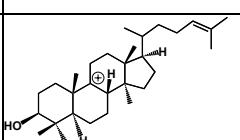
compound		Energy compared to OS (kcal/mol)
Oxidosqualene		0
achilleol A		-16.1208
camelliol C		-15.1568
(9 <i>R</i> ,10 <i>S</i> )-polypoda-8(26),13 <i>E</i> ,17 <i>E</i> ,21-tetraen-3β-ol		-19.1249
isomalabarica-13(14),17,21- trien-3β-ol		-21.7355
(13α <i>H</i> )-isomalabarica- 14 <i>Z</i> ,17,21-trien-3β-ol		-17.8477
(13α <i>H</i> )-isomalabarica- 14 <i>E</i> ,17,21-trien-3β-ol		-20.7868
(13α <i>H</i> )-isomalabarica- 14(26),17,21-trien-3β-ol		-16.347

protosta-20(21),24-dien-3 $\beta$ -ol		-21.792
protosta-17(20),24-dien-3 $\beta$ -ol		-28.9162
protosta-20(22),24-dien-3 $\beta$ -ol		-22.3912
protosta-16(17),24-dien-3 $\beta$ -ol		-27.9756
protosta-13(17),24-dien-3 $\beta$ -ol		-33.5624
protosta-12(13),24-dien-3 $\beta$ -ol		-30.7924
protosta-14(15),24-dien-3 $\beta$ -ol		-28.3661
protosta-7(8),24-dien-3 $\beta$ -ol		-30.6062
lanosterol		-41.7754


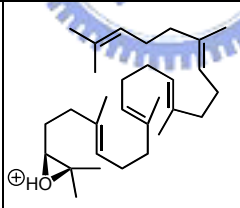
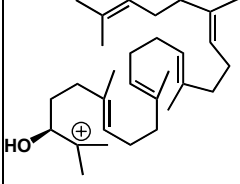
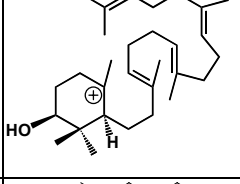
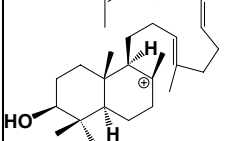
parkeol		-41.2894
---------	--	----------

Relative energy of respective cationic intermediates in protosterol pathway

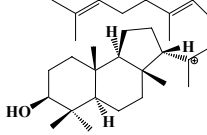
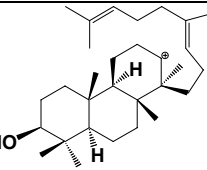
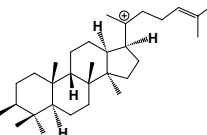
compound		Energy compared to C-2 cation (kcal/mol)		
Protonated OS		8.48		
Linear C-2 cation		Without prefolding	Pre-chair-chair	Pre-chair-boat-chair
		0	-17.107	-16.772
Monocyclic C-10 cation		-25.156		
Bicyclic C-8 cation		-36.912		
6-6-5 Tricyclic C-14 cation		-27.501		
6-6-6 Tricyclic C-13 cation		-29.730		
Protosteryl C-20 cation		-24.794		

Protosteryl C-17 cation		-35.761
Protosteryl C-13 cation		-34.472
Protosteryl C-14 cation		-36.561
Lanosteryl C-8 cation		-38.339
Lanosteryl C-9 cation		-42.106

Relative energy of respective cationic intermediates in dammarenyl pathway

compound		Energy compared to C-2 cation (kcal/mol)
Protonated OS		8.48
Linear C-2 cation		0
Monocyclic C-10 cation		-25.156
Bicyclic C-8 cation		-34.425



6-6-5 Tricyclic C-14 cation		-35.512
6-6-6 Tricyclic C-13 cation		-42.549
Dammarenyl C-20 cation		-42.452

