Achilleol A Synthase Evolved from *Saccharomyces Cerevisiae* Oxidosqualene-Lanosterol Cyclase His234Trp/Tyr510Trp Double Mutation, an Example of Reverse Evolution among Triterpene Cyclases

Cheng-Hsiang Chang, Yuan-Ting Liu, Wen-Hsuan Li, Hao-Yu Wen and Tung-Kung Wu

Department of Biological Science and Technology, National Chiao Tung University, 300, Hsin-Chu, Taiwan

Abstract. A *Saccharomyces cerevisia*e oxidosqualene-lanosterol cyclase double mutant, simultaneously mutated His234 to Trp and Tyr510 to Trp, converts its biosynthetic characteristic from lanosterol formation into that of achilleol A synthase activity and makes achilleol A as the sole product. Product profile and homology modeling analyses suggested the functional roles of Tyr510 and His234 both in cation-stabilization and hydrogen-bonding dyad interaction, resulting in synergistic product specificity of achilleol A formation. Moreover, the evolved achilleol A synthase provides an example of reverse evolutionary order among triterpene cyclases and also opens a new avenue of engineering ERG7 for unnatural enzyme designs.

Keywords: Oxidosqualene-lanosterol cyclase, Reverse evolution, Achilleol A, Site-saturated mutagenesis

1. Introduction

Oxidosqualene cyclases (OSCs) cyclize an acyclic substrate, (3S)-2,3-oxidosqualene (OS, **1**), into diverse tetracyclic or pentacyclic sterols and triterpenoids.[1-5] The intricate and yet highly stereo- and regio-specific cationic mechanism comprises prefolded substrate conformation, oxirane ring protonation and opening, four consecutive ring annulations, 1,2-shifted hydride and methyl groups migration, and final specific deprotonation. The generation of premature or non-specifically water nucleophilic addition products are normally rigorously prevented in the OSCs-mediated reaction. We and others previously characterized several plasticity residues within either first- or second-tier active site of OSCs that influence the cyclization/rearrangement cascade and generate diverse truncated and rearranged products.[6-21]

Achilleol A (**2**), the first monocyclic triterpenoid found in the nature, was originally isolated from *Achillea odorata* and later found to be a relatively widespread metabolite in the plant kidgdom.[22] The biosynthesis of **2** is of considerable interest both in considerations on the OSC-catalyzed concerted or non-concerted cyclization mechanism and biological activities.[22,23] The first but multiple step chemical synthesis of **2** was developed using titanium(III)-mediated free-radical cyclization of epoxypolyprene. However, an overall yield of only about 5% from starting material was obtained.[24] Alternatively, a zeolite NaY-promoted monocyclization of **1** yielded **2**, camelliol C (**3**), together with the production of one bicyclic ether and one allylic alcohol in the ratio of 44 : 11: 40 : 15.[25] In parallel, site-specific mutations of functional residues within either OSCs or squalene cyclase resulted in altering their respective product specificity to generate trace or moderate amout of **2** and **3**, along with other truncated polycyclic products.[6-15] Additionally, a camelliol C synthase (CASM1) that makes 98% of 3 was isolated from *Arabidopsis thaliana*.[26] However, no achilleol A synthase with accurate achilleol A production has been reported.

Recently, the concept of iteratively additional controls from critical amino acid motifs that assist in the formation of polycyclic rings has been challenged.[26] For example, phylogenetic analysis indicates that CASM1 is a descendent of LUP clade enzymes that form pentacyclic amyrin or lupeol.[26] Sequence comparison of a marneral synthase (MRN1) from *Arabidopsis thaliana*, which catalyzes annulation of **1** to a

rearranged bicyclic intermediate with a rare evolutionary transition from a B-ring chair to B-ring boat conformation, with related cyclases also indicates that MRN1 arose within the eudicot PEN clade from enzymes that generate the all-chair dammarenyl cation.[27] Furthermore, a unique OSC yielding C-ring seco-triterpenes was characterized in *A. thaliana* OSC homologues, supporting the direct production of other natural seco-triterpenes by OSCs.[28] Finally, the cloning and functional characterization of an achilleol B synthase from *Oryza sativa* also showed that achilleol B synthase shares the same clade as β-amyrin synthase from *Avn strigosa*.[29] The identification of achilleol B synthase, camelliol C synthase, marneral synthase, and C-ring seco-triterpene synthase suggest the potential of functional evolution of polycyclic triterpene cyclases into monocyclic or few-ring triterpene cyclases.

We have been engaged in structure-function-reaction mechanism studies of *Saccharomyces cerevisiae* oxidosqualene-lanosterol cyclase (ERG7) and protein engineering of the putative active site residues to direct toward desired structure with improved product specificity. Previous studies indicate potential involvement of the ERG7 Tyr510 residue both in influencing monocyclic to tetracyclic formation and stabilizing C8/C9 cation for deprotonation, and that the His234 residue is involved in influencing deprotonation via a hydrogen-bonding network to the hydroxyl group of the Tyr510. Specifically, site-directed mutation of Tyr510 to Lys or Trp produced achilleol A and camelliol C in ratios of 86 : 14 and 96 : 4, respectively. In parallel, the His234 to Trp mutation produced minor amount of parkeol as the sole product. These observations suggested that steric or electrostatic interaction between these two residues might play a functional role in determining the deprotonation step for monocyclic, few-ring formation, or altered deprotonation. We have previously performed site-saturated mutagenesis of His234 to investigate its functional role in ERG7 activity and to determine its effect on the oxidosqualene cyclization/rearrangement reaction. In continue to elucidate the functional role of Tyr510, we performed the site-saturated mutagenesis of Tyr510 and characterized the complete product profile of each mutation. Specifically, we report the construction of an ERG7 double mutant, simultaneously mutated His234 to Trp and Tyr510 to Trp, results in generation of achilleol A as its sole product, thus altering the ERG7 function into the achilleol A synthase activity. These findings further supports the concept that few-ring triterpene synthase could be derived from the cyclases that polycyclize **OS** by reverse evolution.

The ERG7^{Y510X} site-saturated and ERG7$^{Y510W/H234W}$ double mutations were constructed, using the QuikChange site-directed mutagenesis, and transformed into a yeast TKW14 strain for genetic selection and product characterization, as previously described.[11,12] Following genetic selection with ergosterol complementation experiments, the nonsaponifable lipid (NSL) from each mutant were extracted and applied to AgNO$_3$-impregnated silica gel column for product profiles characterization, using GC-MS and NMR (^{1}H, ^{13}C NMR, DEPT, 1H-1H COSY, HMQC, HMBC, and NOE) spectroscopic techniques.

The genetic selection results and product profiles of the ERG7^{Y510X} site-saturated and ERG7$^{Y510W/H234W}$ double mutants with molecular mass of 426 Da were listed in Table 1. Based on the retention time and mass spectrum of products from GC-MS as well as NMR data compared with the authentic samples, truncated cyclized and altered deprotonated compounds including achilleol A (**2**), camelliol C (**3**), (13αH)-isomalabarica-14(26),17,21-trien-3β-ol (**4**), lanosterol (**5**), and parkeol (**6**) were identified from various mutants. The genetic selection results showed that Tyr510 to Lys, Arg, Thr, Pro, and Trp mutations failed to complement the ERG7 activity in the absence of exogenous ergosterol. Neither lanosterol nor truncated intermediates could be detected from the non-viable ERG7$^{Y510P/R/T}$ mutants, which is consistent with the genetic selection results. Nevertheless, the ERG7^{Y510K} and the ERG7^{Y510W} nonviable mutants produced monocyclic products, achilleol A and camelliol C in the ratios of 87 : 13 and 94 : 6, respectively, without any polycyclic products formation. Most of the viable mutants, except for the ERG7$^{Y510A/L/S/C}$, produced identical product profile but with different ratios. Neither **3** nor **4** were produced by the ERG7^{Y510A} mutation. The ERG7^{Y510C} mutation produced tricyclic compound **4** as major product, whereas the ERG7^{Y510H} mutation mainly terminated reaction at either monocyclic or tricyclic to yield **2** and **4** as major products. Notably, the ERG7$^{Y510W/H234W}$ double mutant produced achilleol A as the sole product.

A general mechanism for the ERG7^{Y510X} site-saturated mutants that lead to specific product profiles is shown in Scheme 1. For the ERG7^{Y510X} nonviable mutants, no **OS** could be cyclized for the ERG7$^{Y510P/R/T}$ mutants to generate any product, whereas the ERG7$^{Y510K/W}$ mutants underwent an epoxide ring opening and

subsequent A-ring annulation to generate the C-10 cation (**I**). Subsequent abstraction of the proton from the C-25 and C-1 positions generated achilleol A and camelliol C as the end product, respectively. In most of the ERG7^{Y510X} viable mutants, **OS** can be cyclized to either the C-10 or the tricyclic Markonikov C-14 cation (**II**) as the first stopping point. Subsequent abstraction of the proton from the C-26 position produced **4** as the end product. Alternatively, the cation **II** proceeds with a C-ring expansion and is followed by D-ring annulations to generate the protosteryl C-20 cation (**III**), which is followed by a series of skeletal rearrangements of a hydride shift (H-13α→H-17α), two methyl-group shifts (Me-14β→Me-13β and Me-8α→Me-14α), and another hydride shift (H-9β→H-8β) to generate the lanosteryl C-8/C-9 cation (**IV**). Finally, deprotonation at C-8 and C-11 yielded **5** and **6**, respectively.

Characterization of the product profiles of ERG7^{Y510X} site-saturated mutants revealed that the ERG7^{Y510X} mutants exhibit certain deprotonation profiles, mainly monocyclic achilleol A, tricyclic (13αH)-isomalabarica-14(26),17,21-trien-3 -ol, altered deprotonated parkeol, in conjunction with lanosterol. Neither truncated rearranged products derived from hydride or methyl group shift nor the altered deprotonation product from C-7 cation was identified from the ERG7^{Y510X} mutants. The increased production of monocyclic products was observed when Tyr510 was substituted with bulky or basic amino acids, whereas polycyclic products were observed with smaller or acidic amino acids substitution. Alternatively, previous studies showed that substitution of His234 with small nonpolar hydrophobic residues facilitated the production of polycyclic products such as parkeol, protosta-12,24-dien-3β-ol, and protosta-20,24-dien-3β-ol, but interfered the monocyclic achilleol A formation; whereas substitution of His234 with Tyr or Phe introduced steric hindrance or electrostatic repulsion to the Tyr510, causing the shift of Tyr510 slightly out of active site pocket, and resulted in the production of achilleol A as well as other altered products. In this study, the ERG7$^{Y510W/H234W}$ double mutant resulted in the production of achilleol A as the sole product.

A plausible explanation for the product profile is the hypothetical hydrogen bonding basic dyad between His234 and Tyr510 in affecting various cations stabilization, thus directing electronic density for specific deprotonation. The homology modeling of the *S. cerevisiae* ERG7, based on the human OSC X-ray structure, showed a distance of approximately 4.0 Å from the phenolic oxygen of Tyr510 to the C-10 cation of the monocyclic intermediate.[30] In addition, either His234 or Tyr510 can act as the base and influence other active site interactions for the deprotonation reaction. The abundant generation of the monocyclic products in the Tyr510 to basic or bulky amino acids substitution might be derived from partial disruption of the transient dipole interaction between C-10 cationic intermediate and the hydroxyl group of the Tyr510 as well as steric or electronic repulsion to His234 residue. Alternatively, transient disturbance of His234:Tyr510 hydrogen-bonding dyad by the Tyr510 mutation might cause shift of Nε2 orientation of the His234 imidazole group away from the C-13 and C-20 protosteryl cations, which resulted in the production of tricyclic and altered deprotonation products. Finally, the His234 to Trp and Tyr510 to Trp double mutation might tune the proper orientation of cationic intermediate for deprotonation reaction to produce achilleol A as sole product.

In summary, we demonstrate that the site-saturated mutagenesis not only provide the complete analysis of the role of specific amino acid within OSCs active site but also example how nature evolves enzyme by mutation and selection. Our results also exemplify the power of saturated mutagenesis to accelerate the rate of evolution by exchanging specific position for all proteinoic amino acids.
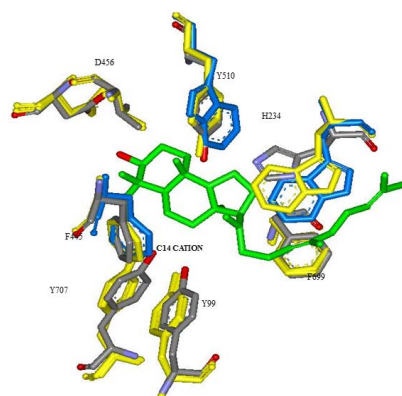
2. Acknowledgements

Fig. 1: The superposition of ERG7[H234W] (yellow) and ERG7[H234W/Y510W] (blue) with the wildtype ERG7 cyclase, in complex with C-14 cation intermediate. The substitutions of Tyr510 with Trp pushed the Trp234 residue away from the original position. The distance change from respective residue to the C14 cation was also observed.

AA substitution	2	3	4	5	6
Y510G	15	—	17	27	41
Y510A	27	—	—	39	34
Y510V	4	—	17	38	41
Y510L	—	—	26	74	—
Y510I	4	—	11	78	7
Y510D	9	—	11	40	40
Y510N	5	—	9	9	77
Y510E	6	—	27	18	49
Y510Q	5	—	26	14	55
Y510H	45	—	24	5	26
Y510S	—	—	4	40	56
Y510C	—	—	67	33	—
Y510M	1	—	15	43	41
Y510F	4	—	10	50	36
Y510W	94	6	—	—	—
Y510K	87	13	—	—	—
Y510P			No product		
Y510R			No product		
Y510T			No product		
H234W[1]	—	—	—	—	100(Minor)
H234Y/Y510A[1]	—	—	—	100	—
H234W/Y510V	2	—	8	90	—
H234W/Y510W	100	—	—	—	—

Table 1: The product profiles of *S. cerevisiae* ERG7[Y510X] site-saturated and ERG7[H234XX/Y510X] double mutants
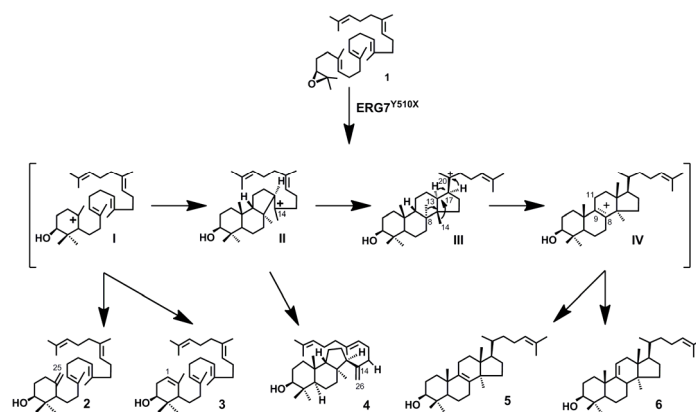
Scheme 1: The proposed cyclization/rearrangement pathways of oxidosqualene within *S. cerevisiae* ERG7[Y510X] mutants.

3. References

[1] I. Abe. Enzymatic synthesis of cyclic triterpenes. *Nat. Prod. Rep.* 2007, **24**, 1311-1331.

[2] I. Abe, M. Rohmer, and G.D. Prestwich. Enzymatic cycilzation of squalene and oxidosqualene to sterols and triterpenes. *Chem. Rev.* 1993, **93**, 2189-2206.

[3] K. U. Wendt, G. E. Schulz, E. J. Corey, and D. R. Liu. Enzyme mechanisms for polycyclic triterpene formation. *Angew. Chem. Int. Ed.* 2000, **39**, 2812-2833.

[4] T. K. Wu, C. H. Chang, Y. T. Liu, and T. T. Wang. Saccharomyces cerevisiae Oxidosqualene-Lanosterol Cyclase: A Chemistry-Biology Interdisciplinary Study of the Protein's Structure-Function-Mechanism Relationships. *Chem. Rec.* 2008, **8**, 302-325.

[5] R. Xu, G. C. Fazio, and S. P. T. Matsuda. On the origins of triterpenoid skeletal diversity. *Phytochemistry* 2004, **65**, 261-291.

[6] T. K. Wu, and C. H. Chang. Enzymatic formation of multiple triterpenes by mutation of tyrosine 510 of the oxidosqualene-lanosterol cyclase from Saccharomyces cerevisiae. *ChemBioChem* 2004, **5**, 1712-1715.

[7] T. K. Wu, C. H. Chang, H. Y. Wen, Y. T. Liu, W. H. Li, T. T. Wang, and W. S. Shie. Alteration of the substrate's prefolded conformation and cyclization stereochemistry of oxidosqualene-lanosterol cyclase of Saccharomyces cerevisiae by substitution at phenylalanine 699. *Org. Lett.* 2010, **12**, 500-503.

[8] T. K. Wu, Y. C. Chang, Y. T. Liu, C. H. Chang, H. Y. Wen, W. H. Li, and W. S. Shie. Mutation of isoleucine 705 of the oxidosqualene-lanosterol cyclase from Saccharomyces cerevisiae affects lanosterol's C/D-ring cyclization and 17α/β-exocyclic side chain stereochemistry. *Org. Biomol. Chem.* 2011, **9**, 1092-1097.

[9] T. K. Wu, W. H. Li, C. H. Chang, H. Y. Wen, and Y. T. Liu. Tyrosine 99 of Saccharomyces cerevisiae Oxidosqualene-Lanosterol Cyclase Influences Tricyclic Terpenoid Moiety Formation with Differential Stereochemical Control. *Eur. J. Org. Chem.* 2009, 5731-5737.

[10] T. K. Wu, Y. T. Liu, and C. H. Chang. Histidine Residue at Position 234 of Oxidosqualene-Lanosterol Cyclase from Saccharomyces cerevisiae Simultaneously Influences Cyclization, Rearrangement, and Deprotonation Reactions. *ChemBioChem* 2005, **6**, 1177-1181.

[11] T. K. Wu, Y. T. Liu, C. H. Chang, M. T. Yu, and H. J. Wang. Site-Saturated Mutagenesis of Histidine 234 of Saccharomyces cerevisiae Oxidosqualene-Lanosterol Cyclase Demonstrates Dual Functions in Cyclization and Rearrangement Reactions. *J. Am. Chem. Soc.* 2006, **128**, 6414-6419.

[12] T. K. Wu, Y. T. Liu, F. H. Chiu, and C. H. Chang. Phenylalanine 445 within oxidosqualene-lanosterol cyclase from Saccharomyces cerevisiae influences C-Ring cyclization and deprotonation reactions. *Org. Lett.* 2006, **8**, 4691-4694.

[13] T. K. Wu, T. T. Wang, C. H. Chang, and Y. T. Liu. Importance of Saccharomyces cerevisiae oxidosqualene-lanosterol cyclase tyrosine 707 residue for chair-boat bicyclic ring formation and deprotonation reactions. *Org. Lett.* 2008, **10**, 4959-4962.

[14] T. K. Wu, H. Y. Wen, C. H. Chang, and Y. T. Liu. Protein Plasticity: A Single Amino Acid Substitution in the Saccharomyces cerevisiae Oxidosqualene-Lanosterol Cyclase Generates Protosta-13(17),24-dien-3β-ol, a Rearrangement Product. *Org. Lett.* 2008, **10**, 2529-2532.

[15] T. K. Wu, M. T. Yu, Y. T. Liu, C. H. Chang, H. J. Wang, and E. W. G. Diau. Tryptophan 232 within oxidosqualene-lanosterol cyclase from Saccharomyces cerevisiae influences rearrangement and deprotonation but not cyclization reactions. *Org. Lett.* 2006, **8**, 1319-1322.

[16] E. J. Corey, H. Cheng, C. H. Baker, S. P. T. Matsuda, D. Li, and X. Song. Studies on the substrate binding segments and catalytic action of lanosterol synthase. Affinity labeling with carbocations derived from mechanism-based analogs of 2,3-oxidosqualene and site-directed mutagenesis probes. *J. Am. Chem. Soc.* 1997, **119**, 1289-1296.

[17] S. Lodeiro, M. J. Segura, M. Stahl, T. Schulz-Gasch, and S. P. T. Matsuda. Oxidosqualene cyclase second-sphere residues profoundly influence the product profile. *ChemBioChem* 2004, **5**, 1581-1585.

[18] S. Lodeiro, Q. Xiong, W. K. Wilson, Y. Ivanova, M. L. Smith, G. S. May, and S. P. T. Matsuda. Protostadienol Biosynthesis and Metabolism in the Pathogenic Fungus Aspergillus fumigatus. *Org. Lett.* 2009, **11**, 1241-1244.

[19] S. Lodeiro, Q. Xiong, W. K. Wilson, M. D. Kolesnikova, C. S. Onak, and S. P. T. Matsuda. An oxidosqualene cyclase makes numerous products by diverse mechanisms: a challenge to prevailing concepts of triterpene biosynthesis. *J. Am. Chem. Soc.* 2007, **129**, 11213-11222.

[20] M. M. Meyer, M. J. R. Segura, W. K. Wilson, and S. P. T. Matsuda. Oxidosqualene cyclase residues that promote formation of cycloartenol, lanosterol, and parkeol. *Angew. Chem. Int. Ed.* 2000, **39**, 4090-4092.

[21] M. J. R. Segura, S. Lodeiro, M. M. Meyer, A. J. Patel, and S. P. T. Matsuda. Directed evolution experiments reveal mutations at cycloartenol synthase residue His477 that dramatically alter catalysis. Org. Lett. 2002, 4, 4459-4462.

[22] A. F. Barrero, R. E. J. Alvarez-Manzaneda, and R. Alvarez-Manzaneda. Achilleol A: A new monocyclic triterpene skeleton from Achillea Odorata L. *Tetrahedron Lett.* 1989, **30**, 3351-3352.

[23] J. S. Yadav, K. Satyanarayana, P. Sreedhar, P. Srihari, T. B. Shaik, and S. V. Kalivendi. Total synthesis of (+/-)-elegansidiol, (+/-)-farnesiferol B, and (+/-)-farnesiferol D. *Bioorg. Med. Chem. Lett.* 2010, **20**, 3814-3817.

[24] A. F. Barrero, J. M. Cuerva, E. J. Alvarez-Manzaneda, J. E. Oltra, and R. Chahboun. A first synthesis of achilleol A using titanium(III) chemistry. *Tetrahedron Lett*. 2002, **43**, 2793-2796.

[25] C. Tsangarakis, C. Raptis, E. Arkoudis, and M. Stratakis. Zeolite NaY-promoted monocyclization of epoxy polyene terpenes: A unique route for the direct synthesis of incompletely cyclized naturally occurring terpenols. *Adv. Synth. Catal.* 2008, **350**, 1587-1600.

[26] M. D. Kolesnikova, W. K. Wilson, D. A. Lynch, A. C. Obermeyer, and S. P. T. Matsuda. Arabidopsis camelliol C synthase evolved from enzymes that make pentacycles. *Org. Lett.* 2007, **9**, 5223-5226.

[27] Q. Xiong, W. K. Wilson, and S. P. T. Matsuda. An Arabidopsis oxidosqualene cyclase catalyzes iridal skeleton formation by Grob fragmentation. *Angew. Chem. Int. Ed. Engl.* 2006, **45**, 1285-1288.

[28] M. Shibuya, T. Xiang, Y. Katsube, M. Otsuka, H. Zhang, and Y. Ebizuka. Origin of structural diversity in natural triterpenes: direct synthesis of seco-triterpene skeletons by oxidosqualene cyclase. *J. Am. Chem. Soc.* 2007, **129**, 1450-1455.

[29] R. Ito, K. Mori, I. Hashimoto, C. Nakano, T. Sato, and T. Hoshino. Triterpene Cyclases from Oryza sativa L.: Cycloartenol, Parkeol and Achilleol B Synthases. *Org. Lett.* 2011, **13**, 2678-2781.

[30] R. Thoma, T. Schulz-Gasch, B. D'Arcy, J. Benz, J. Aebi, H. Dehmlow, M. Hennig, M. Stihle, and A. Ruf. Insight into steroid scaffold formation from the structure of human oxidosqualene cyclase. *Nature* 2004, **432**, 118-122.