Histidine Residue at Position 234 of Oxidosqualene-Lanosterol Cyclase from *Saccharomyces cerevisiae* Simultaneously Influences Cyclization, Rearrangement, and Deprotonation Reactions

Tung-Kung Wu,* Yuan-Ting Liu, and Cheng-Hsian Chang[a]

Oxidosqualene cyclases catalyze the biotransformation of acyclic (3*S*)-2,3-oxidosqualene (OS) to a variety of polycyclic sterols and triterpenoids, generating over 100 distinct triterpenoid skeletons with the formula $C_{30}H_{50}O$.[1–3, 4 and references therein] Product specificity is species-dependent and precisely controlled by the prefolded substrate conformation as well as by interactions between the carbocationic intermediate for deprotonation and the functional groups of catalytic amino acid residues of the enzyme. The transformation mechanisms of this single class of enzymes can vary widely. For example, the triterpenes lanosterol, cycloartenol, and parkeol are formed from a preorganized chair–boat–chair substrate conformation of OS, and cationic cyclization to the protosteryl cation is followed by skeletal rearrangements until the final deprotonation step. Formation of the pentacyclic β-amyrin and lupeol proceed similarly except that OS is in the chair–chair–chair conformation (this results in stereochemical differences in the products relative to the chair–boat–chair substrate conformation), and the cationic cyclization to the dammarenyl cation is followed by annulation of a fifth ring.

Various strategies have been used to probe the complex cyclization/rearrangement reaction mechanism, both for the purpose of understanding these complex enzymes and also to engineer cyclases to generate new product profiles. For example, site-directed mutagenesis was used to identify the residues responsible for the product specificity of β-amyrin synthase (PNY) and lupeol synthase (OEW). Two residues of PNY from *Panax ginseng*, Trp259 and Tyr261, were found to play important roles in the reaction mechanism to direct β-amyrin and/or lupeol formation.[5] We and others independently identified several critical residues from oxidosqualene-lanosterol cyclase (ERG7) from *Saccharomyces cerevisiae* and oxidosqualene-cycloartenol synthase (CAS) from *Arabidopsis thaliana*, and demonstrated their roles in facilitating tetracyclic formation and/or stabilizing the lanosteryl cation for deprotonation, as

[a] Prof. Dr. T.-K. Wu, Y.-T. Liu, C.-H. Chang
   Department of Biological Science and Technology
   Center for Interdisciplinary Molecular Science
   National Chiao Tung University
   300, Hsin-Chu, Taiwan (Republic of China)
   Fax: (+ 886) 35-725-700
   E-mail: tkwmll@mail.nctu.edu.tw

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

well as in changing product specificity from cycloartenol to lanosterol production.[4,6–11] However, no residues or cationic intermediates that act catalytically in substrate prefolding or the rearrangement process have been identified.

We have initiated studies designed to yield a detailed understanding of the mechanism of product differentiation exhibited by ERG7 and CAS, both of which utilize chair–boat–chair substrate conformations and proceed from cationic cyclizations to a protosteryl cation. We are engineering mutant enzymes with conservative amino acid substitutions at residues putatively involved in differentiating chair–boat–chair from chair–chair–chair conformations, or in defining the rearrangement mechanism, and characterizing their catalytic reaction mechanism and product profiles. A residue of particular interest is His234 of ERG7, which some studies suggest is the catalytic base, and others suggest acts to stabilize reaction intermediates.[2,12–15] CAS also has a histidine at the corresponding position, while PNY and OEW both have Tyr (Tyr261 in PNY). On the one hand, mutagenesis studies of His234 in ERG7, demonstrated that His234Ala was nonfunctional while His234Phe was active, and therefore suggested a model in which the residue at that position functioned as the catalytic base.[12,13] However, no product isolation and characterization was performed with this mutant enzyme to investigate its effects on catalysis. The corresponding residue in the related enzyme squalene-hopene cyclase (SHC) of *Alicyclobacillus acidocaldarius* is Trp169. Various mutations at Trp169 of SHC led to the production of several abortive cyclization products; this implicated that residue and His234 of ERG7 in the stabilization of the C20 protosteryl cation, rather than as a catalytic base.[16] On the other hand, the crystal structure of human oxidosqualene-lanosterol cyclase (OSC) suggests a functional role of the His234 residue in stabilizing the C-ring formation, whereas another residue, Tyr510, would be positioned to function as the catalytic base when hydrogen-bonded to the His234 residue.[15] We have previously reported that mutation of Tyr510 in ERG7 leads to the formation of multiple triterpene products including achilleol A, camelliol C, lanosterol, and parkeol.[4]

To further examine the role of His234 in catalysis and product selection, and its interaction with Tyr510, we examined two mutant ERG7 proteins, His234Tyr and the double mutant His234Tyr Tyr510Ala. We report the isolation of multiple triterpenes, including incompletely cyclized, truncated rearranged, and alternatively deprotonated products, from a strain expressing ERG7[H234Y] as its only oxidosqualene cyclase. Moreover, a truncated rearrangement product, protosta-12,24-dien-3β-ol, was isolated and characterized for the first time from an ERG7/CAS-catalyzed reaction; this suggests an important role for H234 in stabilizing the cationic intermediate generated in the rearrangement process. The ERG7[H234Y/Y510A] mutant, in contrast, yielded only lanosterol; this suggests that there is rigorous control of product distribution upon alteration of the active site structure. Our data are consistent with a role for His234 in intermediate stabilization rather than as a catalytic base.

Plasmids carrying the ERG7[H234Y] or ERG7[H234Y/Y510A] mutants were constructed by PCR and transformed into a cyclase-deficient *S. cerevisiae* strain, TKW14. TKW14 is a CBY57-derived HEM1 ERG7 double-knockout strain that is only viable when supplied with exogenous ergosterol.[4,17,18] TKW14 expressing wild-type ERG7 on a plasmid does not need exogenous ergosterol due to the cyclase activity of ERG7.[4] TKW14 carrying the ERG7[H234Y] or ERG7[H234Y/Y510A] mutant was also viable in the absence of ergosterol, this indicated that both mutants retained the essential activity of ERG7. To analyze the product profile resulting from expression of the mutant ERG7 proteins, we grew the strains in liquid media, harvested the cells, prepared nonsaponifiable lipid (NSL) extracts, and used thin-layer chromatography and gas chromatography–mass spectrometry (GC-MS) to detect products.

In the case of the strain expressing ERG7[H234Y], several triterpenoids with a molecular mass of $m/z = 426$ were detected. Chromatographic purification isolated a NSL product that migrated between oxidosqualene and lanosterol, and was later identified as the monocyclic triterpene achilleol A by comparison of its GC-MS and NMR data with those of standard compounds.[11] In addition, a major-product mixture was observed that was indistinguishable from lanosterol (LA) when examined by TLC and high-performance liquid chromatography (HPLC). The compounds in this major-product mixture were acetylated and loaded onto a AgNO3-impregnated silica gel column for separation. Preparative argentic column chromatography yielded three pure compounds, two of which were indistinguishable from authentic lanosteryl acetate and parkeyl acetate standards by ^{1}H and ^{13}C NMR as well as by GC-MS.[4]

The third product from the mixture was deacetylated and characterized de novo by using GC-MS, ^{1}H and ^{13}C NMR, distortionless enhancement by polarization transfer (DEPT), HMBC, HSQC, and NOE spectra. The 150 MHz ^{13}C NMR spectrum of this compound revealed the presence of two tertiary–quaternary substituted double bonds ($\delta = 119.78$, 148.76, and 125.12, 130.99 ppm). The 600 MHz ^{1}H NMR spectra also showed the presence of seven methyl singlets ($\delta = 1.662$, 1.580, 1.013, 0.954, 0.943, 0.878, 0.833, s, 3H) and one methyl doublet ($\delta = 0.905$, d, $J = 6.76$ Hz, 3H), as well as two vinylic protons ($\delta = 5.277$ and 5.076, t, 1H). The HSQC spectrum showed that the methyl protons at $\delta = 0.954$ are attached to the carbon at 22.65 ppm and that the methyl protons at $\delta = 0.878$ are attached to the carbon at 19.92 ppm. For the olefinic protons, the HSQC spectrum established that the vinylic proton at $\delta = 5.277$ is attached to the carbon at 119.78 ppm. Further, NOEs observed between Me-19/Me-28, Me-30/H-5, and Me-18/H-9 as well as the absence of NOEs between Me-18/H-17 and Me-19/Me-30 confirmed the structure as protosta-12,24-dien-3β-ol, a product with $\Delta^{12,24}$ double bonds.[19]

Once these products had been identified, GC-MS analysis of the original NSL extract was used to determine the ratios of products present in the *S. cerevisiae* ERG7[H234Y] mutant extract: achilleol A (monocyclic), protosta-12,24-dien-3β-ol (C12 deprotonation), lanosterol (C8 deprotonation), and parkeol (C11 deprotonation) in relative ratio of 14:26:51:9 (Scheme 1). This is the first report of a truncated rearrangement product (protosta-12,24-dien-3β-ol) generated from substrates in the chair–boat–chair prefolded conformation by ERG7/CAS cyclases.
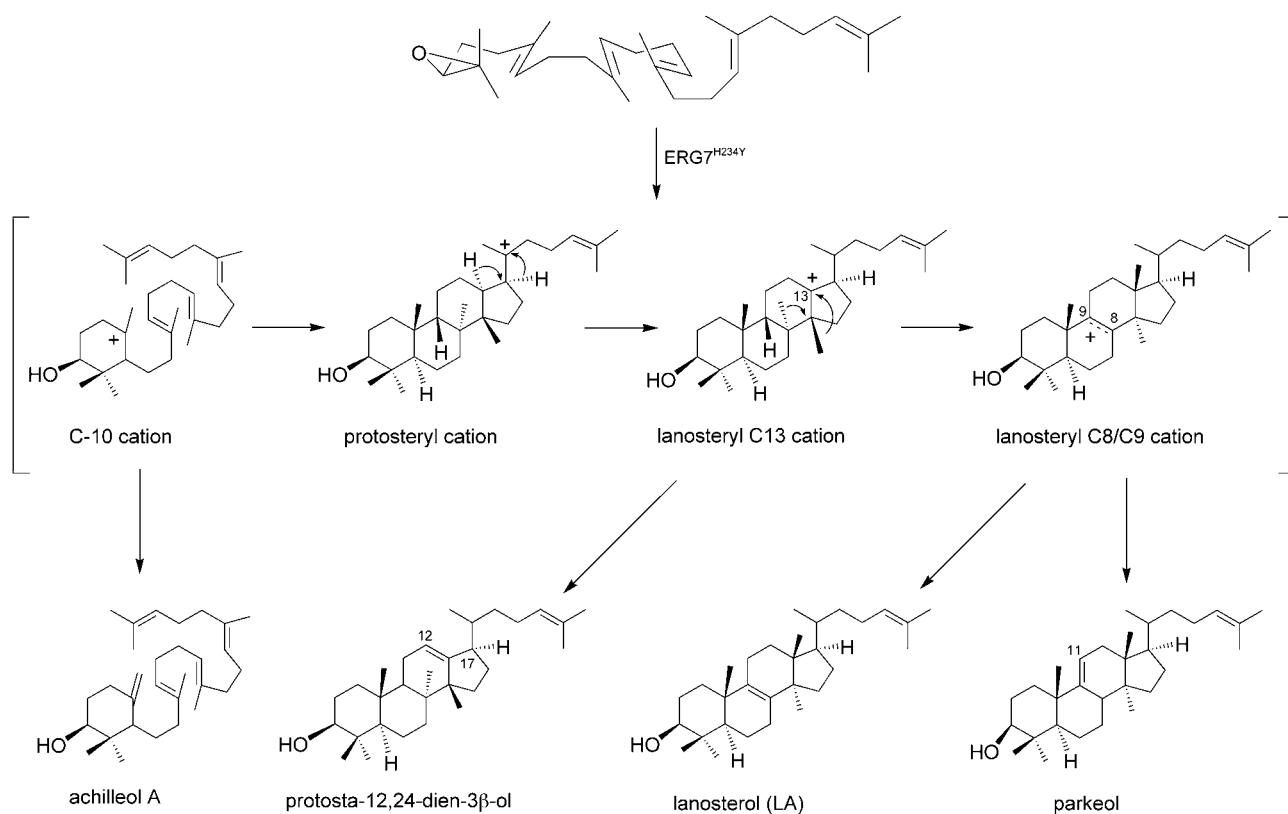
Scheme 1. Cyclization of oxidosqualene in *S. cerevisiae* TKW14 expressing ERG7[H234Y].

In contrast to this complex product yield from the ERG7[H234Y] extract, lanosterol was the sole product observed in the AgNO$_3$-impregnated TLC and GC-MS results from the NSL extract of the strain expressing ERG7[H234Y/Y510A]; this is consistent with the notion that the absence of basic residues in the proximity of the charge of the protosterol cation avoids premature truncation by deprotonation.[15]

We were intrigued by the question of how the enzymatic activity of ERG7 was altered by the H234Y mutation to yield the different cyclization products shown in Scheme 1. It appears that the enzyme first initiates cyclization of OS from a chair–boat–chair conformation to a monocyclic C10 cation (lanosterol numbering). Specific deprotonation at C10 results in the production of achilleol A. Further cyclizations generate the tetracyclic protosteryl C20 cation, which then rotates about 60° to generate the natural C-20*R* configuration, with the 17β-side chain. Then, a backbone rearrangement of H-17α→20α, H13α→17α 1,2-hydride shift to generate the C13 cation with elimination of H-12β yields protosta-12,24-dien-3β-ol with a Δ^{12} double bond. The formation of lanosterol from the C13 cation involves two additional methyl-group shifts (Me-14β→Me-13β, Me-8α→Me-14α) to form the lanosteryl C8 cation. Lanosterol formation is completed by a final deprotonation, either abstracting the proton originally at C9 or after a hydride shift from C9 to C8. Parkeol is formed after a hydride shift from C9 to C8, followed by C11 proton abstraction.

His234 of *S. cerevisiae* ERG7 corresponds to His232 in human oxidosqualene-lanosterol cyclase (OSC) and to Tyr261 in *P. gin-seng* β-amyrin synthase. The *S. cerevisiae* ERG7 was homology modeled by using the human OSC X-ray structure as the template to correlate the active-site structure and catalytic mechanism of the enzymes.[15] Good accordance in the distribution of secondary structure and 3D profile was observed. The model revealed that His234 is positioned in the bottom half of the active site cavity and may influence the rearrangement process, since the observed Nε2 of the His234 imidazole group was found at a distance of about 4.24 and 3.96 Å to the C13 and C20 protosteryl cations, respectively, and the π-electron-rich pocket of His is optimal for the stabilization of the electron-deficient cationic intermediate (Figure 1A). The substitution of His234 with Tyr introduced more steric hindrance and altered the electronic contribution to the active-site cavity. Moreover, the hydrogen bond between Tyr510 and His234 is replaced by electrostatic repulsion between Tyr510 and Tyr234. This shifts Tyr510 away from the potential proton-acceptor position or alters the Tyr234 side-chain orientation (Figure 1B) and, we predict, results in the formation of altered products. The finding that lanosterol is the sole product produced by the strain expressing ERG7[H234Y/Y510A] is consistent with a role for Tyr510–His234 interactions in product formation. The substitution of Tyr510 with alanine, in the presence of the ERG7[H234Y] mutation, releases the electrostatic repulsion and inverts the active-site cavity environment (Figure 1C), thus restabilizing the rearrangement cascade. Perhaps a new hydrogen bond is formed between the hydroxyl group of Tyr234 and an adjacent hydrogen-acceptor residue in the proper position, thus ena-
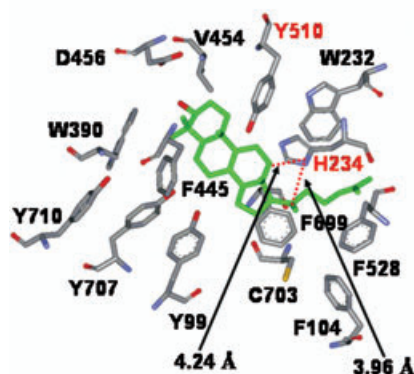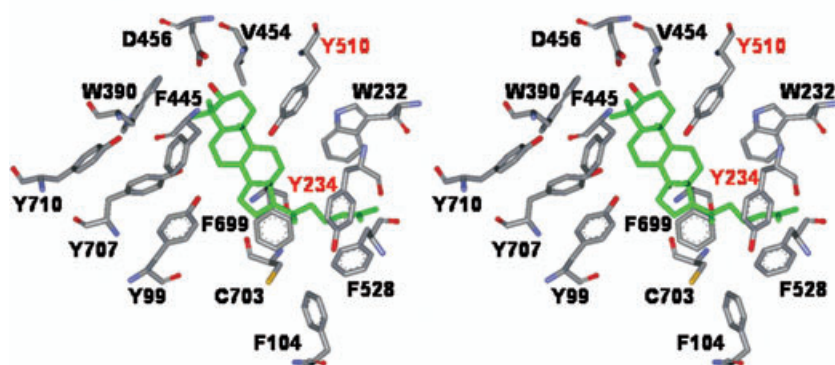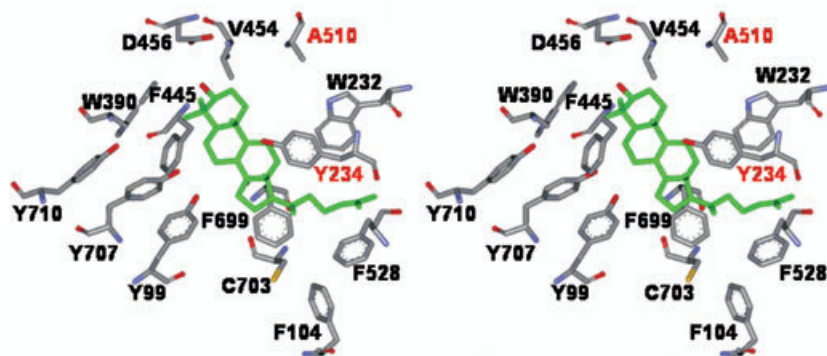
Figure 1. Stereo representations of *S. cerevisiae* ERG7 homology model structures. a) ERG7 wild-type, b) ERG7[H234Y] mutant, and c) ERG7[H234Y/Y510A] double mutant based on the X-ray structure of lanosterol-complexed human OSC and determined by using the Insight II Homology program. Putative active-site residues (stick representation) participating in the active-site formation of modeled ERG7 structures are included. Lanosterol is shown in green; while red dotted lines show interactions of C13 and C20 of lanosterol with Nε2 of the His234 imidazole in the wild-type.

suggest a role for His234 in stabilizing a cationic intermediate, probably at or after protosteryl cation formation, and/or determining deprotonation position, rather than a role acting as a catalytic base. In addition, the enzymatic formation of various regioisomers with the Δ^{10}, $\Delta^{9(11)}$, Δ^{8}, and Δ^{12} double bond suggests that the replacement of the imidazole ring with a phenolic ring at position 234 affects the interactions between spatially adjacent amino acid residues and the stabilization of cationic intermediates as well as the formation of diverse deprotonation products. Perhaps the substitution of His234 by other amino acid residues will affect the enzyme structure differently and result in the production of different deprotonation products. Further ERG7 mutagenesis studies at the 234 position are in progress to examine if diverse product profiles can be generated from other substitutions and to elucidate if one base or multiple proton acceptors are responsible for rearrangement and deprotonation steps.

Experimental Section

The ERG7[H234Y] and ERG7[H234Y/Y510A] mutant plasmids were constructed by PCR by using the QuickChange site-directed mutagenesis kit and following manufacturer's protocols (Stratagene Inc., La Jolla, CA). Recombinant plasmids were electroporated into the yeast strain TKW14, selecting for growth on SD+Ade+Lys+His+Met+Ura+hemin+G418+Erg plates, and then reselecting on SD+Ade+Lys+His+Met+Ura+hemin+G418+5-FOA plates to detect whether the ERG7 mutant protein could complement the cyclase-deficiency of TKW14, as described previously.[4] Transformants were grown in SD+Ade+Lys+His+Ura+Met+Hemin+Erg medium for nonsaponifiable lipid extraction and column chromatography.[4] The extract was fractionated by silica gel column chromatography with hexane/ethyl acetate (19:1) to obtain products that migrated between oxidosqualene and lanosterol and lanosterol-positioned compounds. The acetylation modification of LA-positioned compounds was performed as described previously.[20] The reactions were monitored by TLC and isolated by AgNO₃-impregnated silica gel chromatography with 15% diethyl ether in hexane.[21,22] Products were identified by 600 MHz NMR and by GC-MS analyses. GC-MS was performed on an Agilent 6890N chromatograph equipped with a DB-5HT column (30 m×0.25 mm i.d., 0.1 μm film; oven gradient: 50 °C for 1 min, and then +10 °C per min until 300 °C, held at 300 °C for 8 min, 250 °C inlet; splitless, flow rate 1 mL min⁻¹). Molecular-modeling studies were performed by using the Insight II Homology program with the X-ray structure of lanosterol-complexed human OSC as the template.[15] The MODELER program extracts spatial constraints

bling the phenolic oxygen of Tyr234 to act as the catalytic base for the final deprotonation step.

In summary, the expression of the ERG7[H234Y] mutant protein in cyclase-deficient yeast allowed the isolation and characterization of multiple triterpenes presumed to be formed by ERG7[H234Y], including incomplete cyclization, truncated rearrangement, and altered deprotonation products. The results

such as stereochemistry, main- and side-chain conformations, distance, and dihedral angle from the template structure. The resulting structure was optimized by using an objective function that included spatial constraints and a CHARMM energy function. The objective function combines free-energy perturbation, correlation analysis, and combined quantum and molecular mechanics (QM/MM) to obtain a better description of molecular-level structure, interactions, and energetics. The homologous model structure with the lowest objective function was further evaluated by using the Align2D algorithm for sequence-structure alignment.[23,24]

[1] I. Abe, M. Rohmer, G. D. Prestwich, *Chem. Rev.* **1993**, *93*, 2189–2206.

[2] K. U. Wendt, G. E. Schulz, E. J. Corey, D. R. Liu, *Angew. Chem.* **2000**, *112*, 2930–2952; *Angew. Chem. Int. Ed.* **2000**, *39*, 2812–2833.

[3] R. Xu, G. C. Fazio, S. P. T. Matsuda, *Phytochemistry* **2004**, *65*, 261–291.

[4] T. K. Wu, C. H. Chang, *ChemBioChem* **2004**, *5*, 1712–1715.

[5] T. Kushiro, M. Shibuya, K. Masuda, Y. Ebizuka, *J. Am. Chem. Soc.* **2000**, *122*, 6816–6824.

[6] E. A. Hart, L. Hua, L. B. Darr, W. K. Wilson, J. Pang, S. P. T. Matsuda, *J. Am. Chem. Soc.* **1999**, *121*, 9887–9888.

[7] B. M. Joubert, L. Hua, S. P. T. Matsuda, *Org. Lett.* **2000**, *2*, 339–341.

[8] S. P. T. Matsuda, L. B. Darr, E. A. Hart, J. B. R. Herrera, K. E. McCann, M. M. Meyer, J. Pang, H. G. Schepmann, W. K. Wilson, *Org. Lett.* **2000**, *2*, 2261–2263.

[9] J. B. R. Herrera, W. K. Wilson, S. P. T. Matsuda, *J. Am. Chem. Soc.* **2000**, *122*, 6765–6766.

[10] M. M. Meyer, R. Xu, S. P. T. Matsuda, *Org. Lett.* **2002**, *4*, 1395–1398.

[11] T. K. Wu, J. H. Griffin, *Biochemistry* **2002**, *41*, 8238–8244.

[12] E. J. Corey, H. Cheng, H. Baker, S. P. T. Matsuda, D. Li, X. Song, *J. Am. Chem. Soc.* **1997**, *119*, 1289–1296.

[13] S. Lodeiro, M. J. R. Segura, M. Stahl, T. Schulz-Gasch, S. P. T. Matsuda, *ChemBioChem* **2004**, *5*, 1581–1585.

[14] T. Schulz-Gasch, M. Stahl, *J. Comput. Chem.* **2003**, *24*, 741–753.

[15] R. Thoma, T. Schulz-Gasch, B. D'Arcy, J. Benz, J. Aebi, H. Dehmlow, M. Hennig, M. Stihle, A. Ruf, *Nature* **2004**, *432*, 118–122.

[16] T. Sato, T. Abe, T. Hoshino, *Chem. Commun.* **1998**, 2617–2618.

[17] Z. Shi, C. J. Buntel, J. H. Griffin, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7370–7374.

[18] R. S. Sikorski, J. D. Boeke, *Methods Enzymol.* **1991**, *194*, 302–318.

[19] Chemical shift data of protosta-12,24-dien-3β-ol, referenced to Si(CH$_3$)$_4$ and generally accurate to +0.001 ppm: ^{1}H NMR (600 MHz, CDCl$_3$): $\delta =$ 5.276 (t, $J = 1.8$ Hz, 1H; H-12), 5.075 (t, $J = 7$ Hz, 1H; H-24), 3.196 (dd, $J = 11.5, 4.7$ Hz, 1H; H-3α), 2.170 (br q, $J = 7.8, 7.8$ Hz, 1H; H-17α), 2.029 (m, 1H; H-23S), 1.938–1.909 (m, 1H), 1.850–1.829 (m, 4H), 1.790–1.706 (m, 3H), 1.662 (s, 3H; Me-26), 1.644 (m, 2H; H-2), 1.588 (s, 3H; Me-27), 1.620–1.554 (m, 3H), 1.527–1.489 (m, 3H), 1.479–1.398 (m, 5H), 1.232–1.168 (m, 2H), 1.013 (s, 3H; Me-19), 0.954 (s, 3H; Me-18), 0.943 (s, 3H; Me-29), 0.905 (d, $J = 6.76$ Hz, 3H; Me-21), 0.878 (s, 3H; Me-30), 0.833 (s, 3H; Me-28); ^{13}C NMR (150.77 MHz, 30 mM solution in CDCl$_3$, 25 °C) $\delta = 148.76$ (C-13), 130.98 (C-25)„ 125.15 (C-24), 119.78 (C-12), 79.33 (C-3), 51.95 (C-8), 49.43 (C-17), 45.42 (C-9), 45.19 (C-5), 39.61 (C-4), 37.59 (C-14), 37.56 (C-20), 36.08 (C-10), 35.04 (C-1), 34.11 (C-22), 31.59 (C-15), 30.21 (C-7), 28.48 (C-29), 28.44 (C-2), 26.99 (C-16), 26.29 (C-19), 25.89 (C-23), 25.72 (C-26), 25.67 (C-11), 22.65 (C-18), 19.92 (C-30), 18.27 (C-21), 18.17 (C-6), 17.68 (C-27) 15.82 (C-28); MS: $m/z = 426$ ([M]$^+$), 410, 393, 357, 341, 313, 295, 274, 257, 229, 204, 175, 147, 129, 107, 86, 69 (100), 52. The GC relative retention time (measured relative to lanosterol) was −0.21.

[20] S. Freimund, S. Kopper, *Carbohydr. Res.* **2004**, *339*, 217–220.

[21] H. E. Vroman, A. F. Cohen, *J. Lipid Res.* **1967**, *8*, 150–152.

[22] R. A. J. Pascal, C. L. Farris, G. J. J. Schroepfer, *Anal. Biochem.* **1980**, *101*, 15–22.

[23] A. Sali, T. L. Blundell, *J. Mol. Biol.* **1990**, *212*, 403–428.

[24] A. Sali, T. L. Blundell, *J. Mol. Biol.* **1993**, *234*, 779–815.