A Journal of the Gesellschaft Deutscher Chemiker

Angewandte International Edition GDCh Chemie

www.angewandte.org

Accepted Article

Title: A Dual Role Reductase from Phytosterols Catabolism Enables the Efficient Production of Valuable Steroid Precursors

Authors: Haidong Peng, Yaya Wang, Kai Jiang, Xinru Chen, Wenlu Zhang, Yanan Zhang, Zixin Deng, and Xudong Qu

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.202015462

Link to VoR: https://doi.org/10.1002/anie.202015462

WILEY-VCH

10.1002/anie.202015462

WILEY-VCH

RESEARCH ARTICLE

A Dual Role Reductase from Phytosterols Catabolism Enables the Efficient Production of Valuable Steroid Precursors

Haidong Peng,[a][b] Yaya Wang,[a] Kai Jiang,[a] Xinru Chen,[a] Wenlu Zhang,[a] Yanan Zhang,[a] Zixin Deng,[a][b] and Xudong Qu*[a][b]

[a] Dr. H. Peng, Dr. Y. Wang, K. Jiang, X. Chen, W. Zhang, Dr. Y. Zhang, Prof. Dr. Z. Deng, Prof. Dr. X. Qu
Key Laboratory of Combinatorial Biosynthesis and Drug Discovery Ministry of Education, School of Pharmaceutical Sciences
Wuhan University
1 Luojiashan Rd., Wuhan 430071, China
E-mail: quxd19@sjtu.edu.cn

[b] Dr. H. Peng, Prof. Dr. Z. Deng, Prof. Dr. X. Qu
State Key Laboratory of Microbial Metabolism and School of Life Sciences and Biotechnology
Shanghai Jiao Tong University,
800 Dongchuan Rd. Shanghai 200240, China

Supporting information for this article is given via a link at the end of the document.

Abstract: 4-androstenedione (4-AD) and progesterone (PG) are two of the most important precursors for synthesis of steroid drugs, however their current manufacturing processes suffer from low efficiency and severe environmental issues. In this study, we decipher a dual-role reductase (mnOpccR) in the phytosterols catabolism, which engages in two different metabolic branches to produce the key intermediate 20-hydroxymethyl pregn-4-ene-3-one (4-HBC) through a 4-e reduction of 3-oxo-4-pregnene-20-carboxyl-CoA (3-OPC-CoA) and 2-e reduction of 3-oxo-4-pregnene-20-carboxyl aldehyde (3-OPA), respectively. Inactivation or overexpression of mnOpccR in the *Mycobacterium neoaurum* can achieve exclusive production of either 4-AD or 4-HBC from phytosterols. By utilizing a two-step synthesis, 4-HBC can be efficiently converted to PG in a scalable manner (100 gram scale). This study deciphers a pivotal biosynthetic mechanism of phytosterol catabolism and provides very efficient production routes of 4-AD and PG.

Introduction

Steroid drugs are the second largest catalogue of pharmaceuticals, containing more than 400 drug entities (~17% of FDA approved drugs), used in the clinics for a wide range of diseases.[1,2] To meet the large demand, over 1,000,000 tons of steroid drugs with a sale value of ~100 billion USD are produced annually.[3,4] Due to their structural complexity production of steroid drugs primarily relies on semi-synthesis, with progesterone (PG, **1**) and 4-androstenedione (4-AD, **2**) (Figure 1) being two of the most important precursors for manufacturing.[4] In industrial settings, 4-AD and PG are mainly produced by biodegradation of phytosterols and chemical degradation of diosgenin respectively.[2,4-6] Despite being used for many years these processes still have problems that require solutions, including: i) the co-production of side product 20-hydroxymethyl pregn-4-ene-3-one (4-HBC, also named HMP and 4-BNA, **3** in Figure 1) during 4-AD fermentation, [7] which reduces the conversion of phytosterols and complicates 4-AD purification due to their identical chemical properties; and ii) poor overall production yield of PG (50%) and serious environmental issues of the diosgenin chemical degradation process, which requires 8 synthetic-steps and the use of strong acids and oxidants (Figure S1).[5-6]

To replace the polluting diosgenin degradation process, several approaches using semi-synthesis and synthetic biology techniques were developed.[8-11] One particularly relevant method involves the semi-synthesis of PG from 4-HBC. This process uses only 2 synthetic steps, with less hazardous chemical reagents (Figure S1), and is more efficient than the diosgenin degradation process. Additionally, 4-HBC can be produced from the biodegradation of phytosterols; an abundant side-product from the industrial refinement of vegetable oil, whereas diosgenin needs to be isolated from commercial plants. For this reason, production of PG using the 4-HBC route is both more economical and environmentally friendly. The major obstacle for the application of this strategy is the low production yield of 4-HBC from phytosterols biodegradation. Interestingly, the solution to increase 4-HBC production yield is essentially opposite to the other aim to abolish 4-HBC production in 4-AD production process from phytosterol biodegradation. Thus, the key for addressing these two long-term issues within the steroid industry requires the investigation into an improved method to regulate the production of 4-HBC.

Recently, the 17-hydroxysteroid/ 22-OH-BNC-CoA dehydrogenase; Hsd4A, was found to be relevant to 4-HBC formation.[12] Inactivation of *hsd4A* and 3-ketosteroid Δ1-dehydrogenases (*kstDs*) in the *Mycobacteria neoaurum* ATCC 25795 enabled the production of 4-HBC from phytosterols and resulted in a conversion efficiency of 47-49% (Molar yield).[12] Although manipulation of nutrient and physiological factors can further improve the conversion of 4-HBC,[1,13,14] the lack of a detailed biosynthetic mechanism is a major bottleneck for metabolic engineering that aims to either eliminate (for 4-AD production) or enhance the production of 4-HBC (for PG synthesis). To address this, we performed an investigation on the biosynthesis of 4-HBC. Our study identified an unusual reductase (mnOpccR) engaging in two independent branches of the phytosterols catabolism pathway to produce 4-HBC. Manipulation of this gene in the *Mycobacterium neoaurum* CCTCC AB2019054 can shift the metabolic flux between 4-AD and 4-HBC, leading to the exclusive production of 4-AD or 4-HBC from phytosterols. 4-HBC is then transformed into PG through an optimized two-step synthetic procedure. Therefore, we provide a practical solution to both PG and 4-AD manufacturing by addressing the aforementioned long-standing issues.

10.1002/anie.202015462
WILEY-VCH

RESEARCH ARTICLE

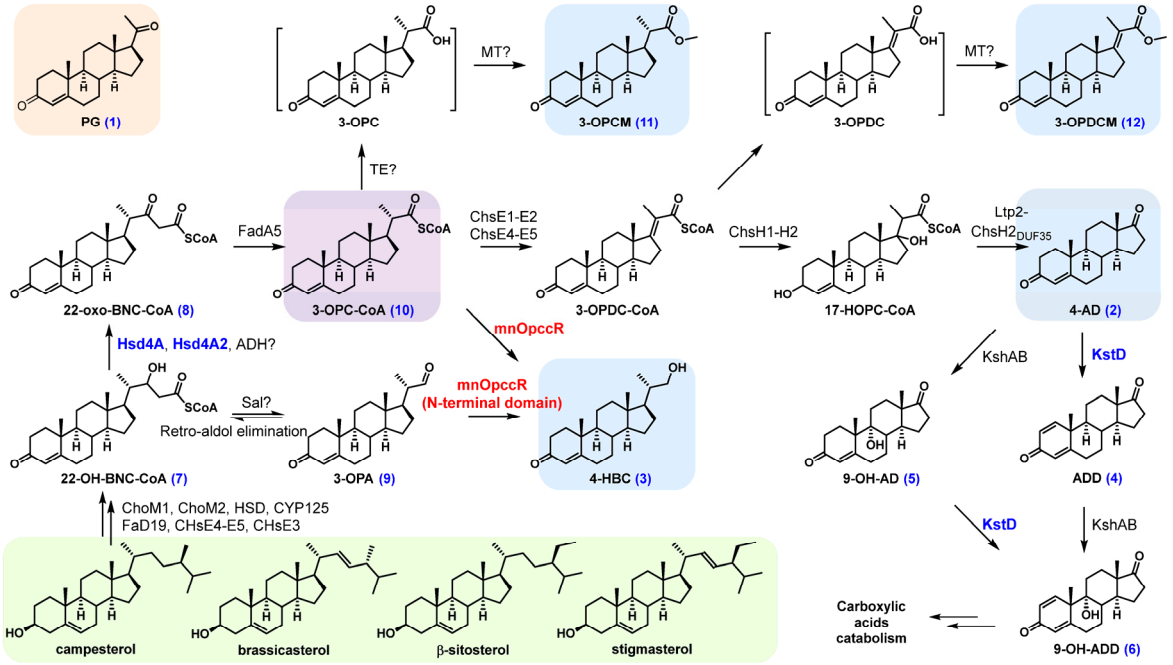


Figure 1. The Physterol Biodegradation Pathway in *Mycobacterium neoaurum*. Physterols (green) which are a mixture of β-sitosterol, stigmasterol, campersterol and brassicasterol (contents see supplymentary inforamtion), 3-OPC-CoA (purple), steroid compounds produced by physterols degradation (light blue) and semi-synthesis (mageta) are shaded in colors. Enzymes with question marks and intermediates with brackets are putative and haven't been verified in *Mycobacterium*. ADH, alchohol dehydrogenase; Sal, steroid aldolase; MT, methyltransferase; TE, thioesterase.

Results and Discussions

Analysis of the Genome of *M. neoaurum* CCTCC AB2019054 and Elimination of ADD by Deleting *kstD*.

Industrial phytosterol biodegradation is mainly achieved using the fast growing *Mycobacterium* strain *M. neoaurum*. This process requires a set of enzymes (Figure 1)[15] in which cholesterol oxidase (ChoM1, ChoM2) and 3*β*-hydroxysteroid dehydrogenase/ isomerase (Hsds) first introduce a 3-ketone group and a Δ5→4 isomerization in the A-ring of phytosterols.[16-20] The intermediates are then oxygenated at C27 by P450 enzymes (CYP125)[21-23] and converted into their CoA form by a ligase (FaD19).[24] These activated intermediates enter into a *β*-oxidation-like route including dehydrogenases (ChsE4-E5, ChsE3, ChsE1-E2, Hsd4A),[12,25-26] hydratase (ChsH1-ChsH2),[27] thiolase (FadA5),[28] and an aldolase complex (Ltp2-ChsH2 $_{DUF35}$)[29-30] that sequentially remove the side chain to generate 4-AD. Further processing by KstD[31-32] and 9*α*-hydroxylase KshAB[33-34], in an individual or collaborative manner, results in 4-AD conversion into 1, 4-androstadienedione (ADD, **4**), 9α-hydroxyandrostenedione (9-OH-AD, **5**) and 9α-hydroxyandrostadienedione (9-OH-ADD, **6**). Finally, 9-OH-ADD is unstable and readily undergoes ring-opening, further degrading into small carboxylic acids.[4]

M. neoaurum CCTCC AB2019054 is a laboratory strain able to transform phytosterols into 4-AD (68%), ADD (15%) and 4-HBC (17%) (Figure 2, trace I). To access the genetic information of phytosterol catabolism, its genome was sequenced. CTCC AB 2019054 shows an identical sequence pattern to the another *M. neoaurum* strain VKM Ac-1815D.[7] Unlike the *M. neoaurum* ATCC 25795, which has three KstD genes (*kstD1-3*),[35] both CCTCC AB2019054 and VKM Ac-1815D only bear a single KstD gene (97% identity to KstD1 of ATCC 25795). Since ADD is not a desired component, we deleted the *kstD* gene. As expected, the resulting strain *M. neoaurum mJTU1* only produces 4-AD and 4-HBC as the two major degradation products (Figure 2, trace II), thus it was used as the starting strain for further study.

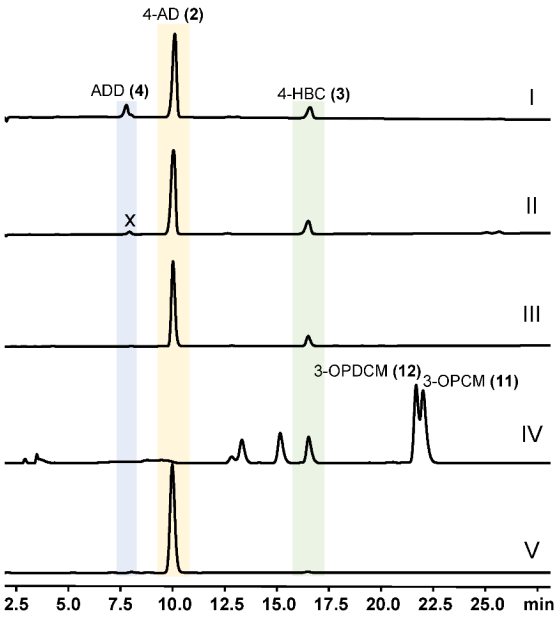


Figure 2. Production of Steroid Precursors in the *M. neoaurum* Wild Type and Mutant Strains. (I) CCTCC AB2019054; (II) *mJTU1* (*ΔkstD*), x is a marginal steroid product different from ADD; (III) *mJTU2* (*ΔkstD-Δhsd4A*); (IV) *mJTU3* (*ΔkstD-ΔchsE1-E2-ΔchsH1-H2*); (V) *mJTU4* (*ΔkstD-ΔmnOpccR*).

10.1002/anie.202015462

Unveiling 3-OPC-CoA Reduction Involves the Formation of 4-HBC and Inactivation of ChsE1-E2 and ChsH1-H2 to Produce 3-OPC and 3-OPDC.

With the objective of improving the production of 4-HBC, we first replicated the reported strategy in ATCC 25795 to delete *hsd4A* in *mJTU1*, generating the strain *mJTU2*.[12] Blocking the conversion of 22-OH-BNC-CoA (**7**) to 22-oxo-BNC-CoA (**8**) was assumed to direct **7** into entering the retro-aldol elimination route. Subsequently, this would yield 3-oxo-4-pregnene-20-aldehyde (3-OPA, **9**), leading to the conversion of 4-HBC by aldehyde reductase (Figure 1). However, our experimental results of *hsd4A* deletion (97.4% identity to the Hsd4A in ATCC 25795), show no effect on the profile of products (Figure 2, trace II-III). These results suggest that CCTCC AB2019054 may contain other enzymes which can compensate for the function of Hsd4A.

The failure to boost 4-HBC production through inactivation of *hsd4A* motivated us to focus on the 4-HBC biosynthesis itself. Except for the retro-aldol elimination route,[12] we hypothesize it is also possible to synthesize 4-HBC via reduction of 3-oxo-4-pregnene-20-carboxyl-CoA (3-OPC-CoA, **10**), although the presently known CoA reductases only accept alkyl acyl-CoAs. To verify this assumption, we performed an in vitro reduction assay of 3-OPC-CoA (**10**) using the crude lysate of *mJTU1* cells. HPLC analysis reveals that both 4-AD and 4-HBC can be produced when synthetic 3-OPC-CoA (**10**) and reducing equivalents of NADPH were both added into the cell lysate (Figure 3, trace V). The ratio of 4-HBC to 4-AD (12%: 88%) is lower than that observed in the phytosterols fermentation broth (17%: 83%) indicating that 4-HBC is indeed partially produced via the reduction of 3-OPC-CoA.

Next, considering that the metabolic branch of 4-HBC resides in front of the dehydrogenation of 3-OPC-CoA (**10**), we envision that abolishing the late stage steps in the *β*-oxidation route will divert the production from 4-AD to 4-HBC. Therefore, both the 3-OPC-CoA dehydrogenase (ChsE1-E2) and 3-OPDC-CoA hydratase (ChsH1-H2) genes were deleted from the strain *mJTU1*.

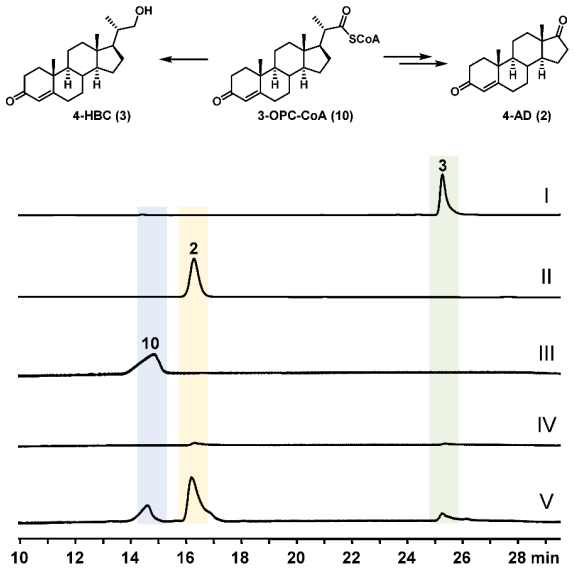


Figure 3. Biochemical Assay Investigating the Conversion of 3-OPC-CoA Using Crude Cell Lysate of *mJTU1*. (I) standard 4-HBC; (II) standard 4-AD; (III) standard 3-OPC-CoA; (IV) control reaction, omitting the 3-OPC-CoA; (V) standard reaction system, converting the 3-OPC-CoA into 4-HBC and 4-AD in 2 hours.

Surprisingly, the resulting mutant *mJTU3* not only increased 4-HBC by 12%, but also produced two new major steroidal products (Figure 2, trace IV). Structural characterization revealed that these products are 3-OPC methyl ester (3-OPCM, **11**)[36] and 3-oxo-4,17-pregnadiene-20-carboxylic methyl ester (3-OPDCM, **12**). These are possibly generated through successive thioester hydrolysis and carboxyl methylation from their CoA intermediates (Figure 1). Although (**11**) and (**12**) are not expected products, they are very useful for the synthesis of many steroidal 17-carboxylates.[37] These products, or their free acids, are not able to be detected in the *mJTU1* suggesting the branched pathways to their biosynthesis are silenced in the ordinary phytosterol degradation condition. The accumulation of 3-OPC-CoA (**10**) in high concentrations could activate these pathways to produce (**11)** and (**12**), which appears to be heavily favored over the 4-HBC pathway. The low conversion of 3-OPC-CoA (**10**) to 4-HBC suggests that identifying the putative 3-OPC-CoA reductase is necessary to improve the conversion and subsequent yield of 4-HBC.

Identification of mnOpccR and Elimination of 4-HBC.

Synthesis of alcohol from acyl-CoA normally requires two consecutive two-electron (2-e) reduction steps: reduction of acyl-CoA to aldehyde and reduction of aldehyde to alcohol.[38] In a few cases, this conversion can be fulfilled through a single four-electron (4-e) reduction.[38,39] As no aldehyde intermediate was detected in the crude lysate assay system, we assume that 4-HBC is formed through the 4-e reduction route from a acyl-CoA precursor by a 4e reductase. So far, only two types of four-electron fatty acyl-CoA reductase have been identified: I) the first, existing in prokaryotes and eukaryotes, can catalyze the conversion of both fatty acyl-CoA and fatty acyl-ACP (acyl carrier protein) to fatty acyl alcohols;[40-44] II) the second type contains an N- and C- terminal didomain, observed in a few bacterial species,[39,45] e.g. maFACoAR (Maqu_2507) in *Marinobacter aquaeolei* VT8 which only accepts fatty acyl-CoAs.[39]

Through blast analysis, we found a maFACoAR homologous gene (*mnOpccR*, accession no. MT747422, 43% identity to maFACoAR, Figure S2) in CCTCC AB2019054. To test its function, mnOpccR was overexpressed in the model strain *Mycobaterium smegmatis* MC^2 155. The purified protein was able to efficiently convert 3-OPC-CoA (**10**) into 4-HBC using NADPH as an exclusive reductant (Figure 4, trace IV) (NADH is not active, Figure S3). Considering that maFACoAR-type enzymes are also able to reduce aldehydes, we further assayed its activity on 3-OPA (**9**). To our delight, mnOpccR is even more efficient in the reduction of 3-OPA (**9**) (K_m = 0.2736 ± 0.06 mM, k_{cat} = 18. 13 ± 1.36 min^{-1}, k_{cat} / K_m = 66.25 ± 13.84 mM^{-1} min^{-1}) than 3-OPC-CoA (**10**) (K_m=0.08 ± 0.006 mM, k_{cat} = 2. 74 ± 0.08 min^{-1}, k_{cat}/ Km = 34.33 ± 3.6 mM^{-1} min^{-1}) (Figure 4, trace V and Figure S4). This activity suggests that the mnOpccR also engages in the retro-aldol elimination route to produce 4-HBC by reduction of 3-OPA (**9**). As expected, deletion of this dual role gene in *mJTU1* can completely abolish the production of 4-HBC (Figure 2, trace V). Additionally, the blockage of the metabolic flux to 4-HBC increases the yield of 4-AD by 16% compared to wild-type. Furthermore, the elimination of 4-HBC in the degradation products also drastically facilitates the purification of 4-AD. Approximately 95% chemical purity of 4-AD can be achieved by a simple ethyl acetate extraction, creating an applicable solution for industrial 4-AD manufacturing.

10.1002/anie.202015462

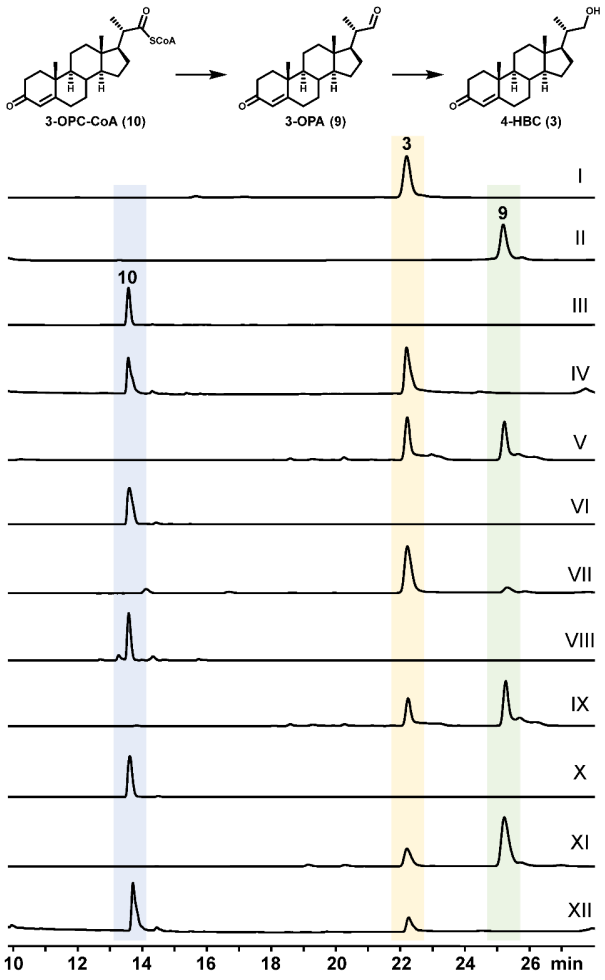


Figure 4. Biochemical Assay Investigating the Conversion of 3-OPC-CoA by mnMnOpccR. (I) standard 4-HBC; (II) standard 3-OPA; (III) standard 3-OPC-CoA; (IV) mnOpccR reaction, converting 3-OPC-CoA into 4-HBC; (V) mnOpccR reaction, converting 3-OPA into 4-HBC; (VI) mnOpccR-NtD reaction, showing no activity on 3-OPA-CoA; (VII) mnOpccR-NtD reaction, converting 3-OPA into 4-HBC; (VIII) mnOpccR-S375A-G377A-R398A reaction, showing no activity on 3-OPC-CoA; (IX) mnOpccR-S375A-G377A-R398A reaction, converting 3-OPA into 4-HBC; (X) msOpccR reaction, showing no activity on 3-OPC-CoA; (XI) msOpccR reaction, converting 3-OPA into 4-HBC; (XII) msOpccR-NtD-mnOpccR-CtD chimeric protein's reaction, converting 3-OPC-CoA to 4-HBC.

Deciphering the Catalytic Mechanism of mnOpccR.

Currently, the reaction mechanisms of the maFACoAR-type enzyme remain poorly understood, particularly the catalytic roles of the N- and C- terminal domains (NtD and CtD).[39,45] Sequence analysis revealed that both the NtD and CtD of mnOpccR have a conserved pyridine nucleotide binding region (Figure S5) and resemble the Re domain of the non-ribosomal peptide synthetase (NRPS) MxaA[46] (30% identity) and the acyl-CoA reductase Acr1[47] (42% identity) respectively (Figure S5). As both MxaA-Re and Acr1 are able to catalyze the reduction of thioesters, the function of the NtD and CtD of mnOpccR cannot be readily distinguished. To verify their exact roles, we dissected the mnOpccR by individually expressing the NtD (1-361aa) and the CtD (362-667aa) (Figure S5). Efforts to solubilize the CtD or introduce inactive mutations (G13A) within the N-terminal NADPH-binding motif were unsuccessful; both for the full length mnOpccR and the NtD (Table S1). Fortunately, expression of NtD alone and inactivation of the NADPH binding motif in the CtD (S375A-G377A-R398A) of the full-length enzyme was successful. Activity assays reveal that these two mutants lose activity on 3-OPC-CoA (**10**), but retain activity to efficiently reduce the aldehyde 3-OPA (**9**) (Figure 4, trace VI to IX). This confirms that the NtD is responsible for the conversion of aldehyde to alcohol and the CtD appears to catalyze the reduction of 3-OPC-CoA (**10**) to aldehyde. In addition, the 3-OPA (**9**) was not able to be detected during the reduction of 3-OPC-CoA (**10**), suggesting that this aldehyde intermediate was shuttled within an internal tunnel between the N- and C- terminal domains.

The previously known maFACoAR-type enzymes are all specific to the medium-to-long chain fatty acyl-CoAs for wax biosynthesis.[39,45] To verify whether mnOpccR is also active towards fatty acyl-CoA or other type of CoAs, including lauryl-CoA (C12 fatty acyl-CoA), butyryl-CoA, crotonyl-CoA and cinnamic-CoA were employed for the activity assay. LC-MS and GC-MS analysis shows that mnOpccR is not able to convert these substrates at all (Figure S6), suggesting it is exclusively involved in steroid catabolism. Exploration of its distribution in other organisms revealed that the mnOpccR homologues widely exist in many other *Mycolicibacteria* (Figure S7), including both the slow growth pathogenic and fast growth non-pathogenic strains such as *M. smegmatis*. To verify this discovery, the homologous gene msOpccR (identity 78%, accession number: AIU06854) of *M. smegmatis* was overexpressed and purified. Interestingly, this enzyme shows no reductive activity of 3-OPC-CoA (**10**) (Figure 4, trace X) as well as above the aforementioned fatty acyl-CoA substrates (data not shown). However, it does appear capable of reducing 3-OPA (**9**) to 4-HBC (Figure 4, trace XI). By replacing the CtD (365-681aa) of msOpccR with the mnOpccR CtD (362-667aa), the reducing activity of 3-OPC-CoA (**10**) into 4-HBC can be restored (Figure 4, trace XII). This confirms that the msOpccR's CtD is naturally inactive. It suggests that *M. smegmatis* may only have the retro-aldol elimination route to produce 4-HBC and msOpccR catalyzes the conversion of 3-OPA to 4-HBC. In addition, the interchangeable CtDs between mnOpccR and msOpccR suggest that the two half-reactions are performed independently in the two domains. Based on these results, two possible modes for maFACoAR-type enzyme reactions exist, these are: (I) the 4-e reduction mode, where 3-OPC-CoA (**10**) first enters into the cavity of the CtD. After the first 2-e reduction, the yielding aldehyde intermediates 3-OPA (**9**) are shuttled from the C-terminal cavity to the N-terminal cavity through an internal tunnel, where they are further reduced into alcohol through the second 2-e reduction; (II) The 2-e reduction mode, 3-OPA (**9**) may directly enter into the N-terminal cavity for reduction or transport through the C-terminal cavity and the internal tunnel. In the phytosterols catabolism of CCTCC AB2019054 both modes co-exists and mnOpccR plays dual roles for the generation of 4-HBC.

Efficient Production of 4-HBC and the Dual Role of mnOpccR in the Phytosterol Catabolism.

After elucidating mnOpccR, efforts were directed towards enhancing its activity to overproduce 4-HBC. The *Mycobacterium* system has a unique gene organization, with around half of its genes containing no 5'-UTR, e.g. the Shine-Dalgarno (SD) sequence.[48,49] From sequence analysis, mnOpccR seems to have no 5'-UTR (Figure S8). To ensure the success of overexpression of mnOpccR, we designed three types of

10.1002/anie.202015462

constructs: (I) with the native 100 bp sequence upstream of the start codon; (II) with a standard mycobacterial SD sequence $_{(-16)}$GAGAAAGGG$_{(-8)}$[48,49] and (III) without 5'-UTR. Those three cassettes were placed under the strong constitutive promoter pG13 and transferred into the *mJTU1* for integration into its genome. Surprisingly, none of these strains show significant improvement in 4-HBC production. The highest increase in 4-HBC conversion was only by about 10%, using the strain with native 100 bp upstream sequence (*mJTU5*) compared with *mJTU1* (Figure 5, trace II). This suggests the 4-AD pathway is still stronger than the accumulation of the enhanced 4-HBC and the retro-aldol elimination branch.

Considering the inactivation of *hsd4A* can weaken the mainstream pathway to 4-AD by diverting part of metabolic flux into the retro-aldol elimination route, the constructs with native 100 bp sequence pG13-mnOpccR were further introduced into the strain *mJTU2* (*ΔkstD1-hsd4A)* to generate mJTU8. The fermentation of this recombinant strain revealed a dramatic increase in the yield of 4-HBC, almost completely diverting the production of 4-AD to 4-HBC (93% conversion from phytosterols in molar ratio, Figure 5, trace III). These results demonstrate that the 2-e reduction by mnOpccR is indeed more crucial for the production of 4-HBC than the 4-e reduction route, and the cumulative effect of both routes is able to surpass the 4-AD pathway when *hsd4A* was inactivated.

The in vitro and in vivo results provide a comprehensive explanation of the catalytic role mnOpccR plays in phytosterols catabolism. Evidently, mnOpccR plays a dual role in the conversion of 3-OPC-CoA and 3-OPA by performing both 4-e and 2-e reduction respectively. Within the wild type strain, a preference for the *β*-oxidation pathway is favored, with only 17% of metabolic products being converted into 4-HBC through the 3-OPC-CoA reduction and retro-aldol elimination branches. From the ratio of 4-AD and 4-HBC in the crude lysate reaction, it is estimated about 12% of 4-HBC was generated through the 3-OPC-CoA reduction and the remaining 5% of 4-HBC being produced through retro-aldol elimination. Due to the preferential direction towards the 4-AD pathway, it is necessary to weaken it by inactivation of *hsd4A*, however, the sole inactivation of *hsd4A* is not able to direct the metabolic flux towards the formation of 4-HBC. Analysis of the CCTCC AB2019054 genome revealed that there are 6 proteins with >38% identity and 10 proteins with 31%-38% identity to the Hsd4A (accession no. MW228416-MW228431) which may compensate its function. To confirm this assumption, we deleted the most identical gene *hsd4A2* (45% identity to Hsd4A) on the basis of strain mJTU2 (ΔkstD-Δhsd4A). Fermentation analysis of the resulting mutant strain mJTU9 (ΔkstD-Δhsd4A-Δhsd4A2) revealed that it indeed can produce 4-HBC with a significant increased ratio (85% in the products profile) comparing to the stain of mJTU2 (with 17% 4-HBC in the products profile) (Figure 5, trace IV). This result indicated that Hsd4A2 plays a more important role than Hsd4A in conversion of **7** to **8** (Figure 1). Furthermore, as mJTU9 still produces 4-AD (15% in the products profile), other isoenzymes should still exist in the strain, waiting for further study to verify their function.

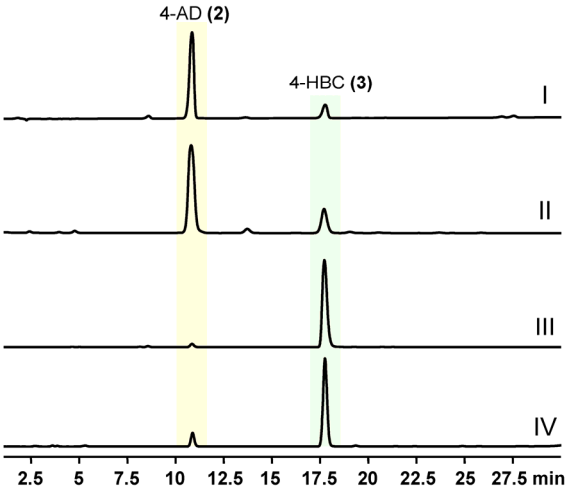
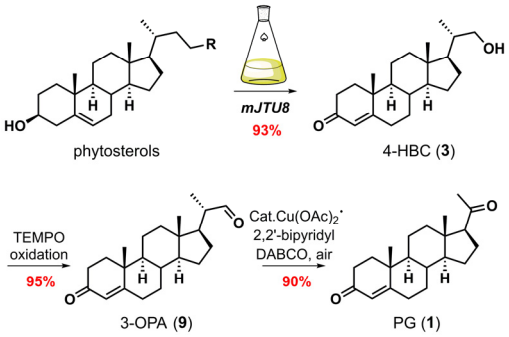


Figure 5. Production of 4-HBC in the *M. neoaurum* mutant strains. (I) *mJTU1* (*ΔkstD*); (II) *mJTU5* (*ΔkstD* and *pG13-mnOpccR*); (III) *mJTU8* (*ΔkstD-Δhsd4A and pG13-mnOpccR*); (IV) *mJTU9* (*ΔkstD-Δhsd4A-Δhsd4A2*)



Scheme 1. Production of PG from phytosterols.

Although the strain mJTU9 can significantly improve the ratio of 4-HBC in the products profile, its conversion (68% from phytosterols in molar ratio) is still much lower than the strain mJTU8 (ΔkstD-Δhsd4A and pG13-mnOpccR, 93% from phytosterols in molar ratio). This is likely due to the toxicity of the aldehyde 3-OPA (**9**) and the reversibility of retro-aldol elimination reaction, restricting the flow of metabolic flux. The overexpression of mnOpccR rapidly removes the accumulation of 3-OPA (**9**), and the metabolic flux can be driven towards retro-aldol elimination, resulting in a dramatic increase in 4-HBC production. These results confirm that the mnOpccR plays a determinant role in the retro-aldol elimination branch.

Preparation of PG from 4-HBC and Phytosterols.

To demonstrate the practical utility of the 4-HBC route, we performed a preparative scale biotransformation of phytosterols using flask fermentation. 10 g L^{-1} of phytosterols were fed to the culture uniformly at an interval of 24 h between the 1rd - 3th day of the fermentation period, resulting in a total of 7.41 g L^{-1} of pure 4-HBC (93% molar ratio) obtained by day 6 (Scheme 1). Using a simple ethyl acetate method of extraction from the fermentation broth 4-HBC was obtained > 95% purity, substantially improving the existing purification process. To the best of our knowledge, this is the highest yield of 4-HBC ever reported. It almost doubles the observed conversion (47-49%) previously achieved using ATCC 25795-*Δhsd4A-ΔkstD1-3*,[12] addressing a major bottleneck for the 4-HBC route in the production of PG.

Synthesis of PG from 4-HBC has been previously documented.[9-11] The synthetic procedure contains two separate steps; the conversion of 4-HBC to 4-HBC aldehyde, and the conversion of the aldehyde to PG. Although pyridinium dichromate (PDC) has proven to be efficient in oxidizing the

alcohol of 4-HBC (73% yield),[11] this reagent is highly toxic and highly restricted by EHS in China. By exploring a few reagents, we found that TEMPO oxidation can provide excellent conversion to produce 4-HBC aldehyde with 95% yield (Scheme 1). Furthermore, through minor optimization of the second reaction condition,[10] PG can be efficiently produced (90% yield) through the copper-mediated catalytic radical oxygenation of 4-HBC aldehyde (Scheme 1). Finally, to demonstrate its practical utility, we scaled up the reaction to a preparative level. Based on these conditions, 80.5 g PG (84.6% yield) was obtained from 100 g 4-HBC. By accounting for the yield of phytosterols degradation, the total yield of phytosterols-based PG can be up to 78.7%, which is much more efficient than the diosgenin degradation route (50% yield over 8 steps,[5,6] Figure S1). Therefore, this process is both more economical and sustainable than the diosgenin degradation route for industrial production of PG.

Conclusion

In summary, we have deciphered the biosynthetic mechanisms of 4-HBC in the phytosterols biodegradation pathway of *M. neoaurum* CCTCC AB2019054. 4-HBC is formed by two independent pathways, in which mnOpccR catalyzes the pivot reactions between 4-e reduction of 3-OPC-CoA (**10**) and 2-e reduction of 3-OPA for its generation. The 4-e reduction uses the C-terminal domain of MnOpccR to convert 3-OPC-CoA into 3-OPA and the N-terminal domain further reduces 3-OPA into 3-HBC. 3-OPA is translocated between the two domains and is not released during catalysis. For 2-e reduction, 3-OPA is solely reduced by the N-terminal domain. Inactivation of *mnOpccR* can eliminate 4-HBC while, overexpression of *mnOpccR* together with *hsd4A* inactivation can lead to exclusively producing 4-HBC from phytosterols. By combining this with a two-step chemical process, 4-HBC can be efficiently converted into PG in a scalable way. Therefore, our results provide two highly efficient processes in which to manufacture PG and 4-AD, providing a much greener and economical method than the current routes.

Acknowledgements

The authors are very grateful to Prof. Xiaoyong Fan and Prof. Jiaoyu Deng for kindly providing the plasmid pMF406 and *M. smegmatis* mc^2 155, respectively. This work was financially supported by the National Key R&D Program of China (Grant Nos. 2018YFC1706200) and the National Natural Science Foundation of China (Grant Nos. 31970054 and 31570057).

Keywords: biosynthesis • biosynthetic pathway • steroid • acyl-CoA reductase • semi-synthesis

[1] L.-B. Xiong, H.-H. Liu, M. Zhao, Y.-J. Liu, L. Song, Z.-Y. Xie, Y.-X. Xu, F.-Q. Wang, D.-Z Wei, *Microb. Cell Fact.* **2020,** *19*, 80.

[2] W.-Y. Tong, X. Dong, *Recent Pat. Biotechnol.* **2009,** *3*, 141-153.

[3] M. Bureik, R. Bernhardt, *Steroid hydroxylation: microbial steroid biotransformations using cytochrome P450 enzymes*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, **2007**

[4] M. V. Donova, O. V. Egorova, *Appl. Microbiol. Biotechnol.* **2012,** *94*, 1423-1447.

[5] R. E. Marker, J. Krueger, *J. Am. Chem. Soc.* **1940,** *62*, 3349-3350.

[6] R. K. Mackie, D. M. Smith, R. A. Aitken, *Guidebook to organic synthesis,* Longman, Harlow, **1982**.

[7] E. Y. Bragin, V. Y. Shtratnikova, D. V. Dovbnya, M. I. Schelkunov, Y. A. Pekov, S. G. Malakho, O. V. Egorova, T. V. Ivashina, S. L. Sokolov, V. V. Ashapkin, M. V. Donova, *J. Steroid Biochem. Mol. Biol.* **2013,** *138*, 41-53.

[8] C. Duport, R. Spagnoli, E. Degryse, D. Pompon, *Nat. Biotechnol.* **1998,** *16*, 186-189.

[9] F. W. Heyl, M. E. Herr, *J. Am. Chem. Soc.* **1950,** *72*, 2617-2619.

[10] V. Van Rheenen, *Tetrahedron Lett.* **1969,** *10*, 985-988.

[11] J. J. Poza, R. Fernández, F. Reyes, J. Rodríguez, C. Jiménez, *J. Org. Chem.* **2008,** *73*, 7978-7984.

[12] L.-Q. Xu, Y.-J. Liu, K. Yao, H.-H. Liu, X.-Y. Tao, F.-Q. Wang, D.-Z. Wei, *Sci. Rep.* **2016,** *6*, 21928.

[13] W.-J. Sun, L. Wang, H.-H. Liu, Y.-J. Liu, Y.-H. Ren, F.-Q. Wang, D.-Z. Wei, *Metab. Eng.* **2019,** *56*, 97-110.

[14] L.-B. Xiong, H.-H. Liu, L.-Q. Xu, W.-J. Sun, F.-Q. Wang, D.-Z. Wei, *Microb. Cell Fact.* **2017,** *16*, 89.

[15] R. Van der Geize, K. Yam, T. Heuser, M. H. Wilbrink, H. Hara, M. C. Anderton, E. Sim, L. Dijkhuizen, J. E. Davies, W. W. Mohn, L. D. Eltis, *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104*, 1947-1952.

[16] K. Yao, F.-Q. Wang, H.-C. Zhang, D.-Z. Wei, *Metab. Eng.* **2013,** *15*, 75-87.

[17] X. Yang, E. Dubnau, I. Smith, N. S. Sampson, *Biochemistry* **2007,** *46*, 9058-9067.

[18] V. M. Nikolayeva, O. V. Egorova, D. V. Dovbnya, M. V. Donova, *J. Steroid Biochem. Mol. Biol.* **2004,** *91*, 79-85.

[19] S. Horinouchi, H. Ishizuka, T. Beppu, *Appl. Environ. Microbiol.* **1991,** *57*, 1386-1393.

[20] J. Li, A. Vrielink, P. Brick, D. M. Blow, *Biochemistry* **1993,** *32*, 11507-11515.

[21] J. K. Capyk, R. Kalscheuer, G. R. Stewart, J. Liu, H. Kwon, R. Zhao, S. Okamoto, W. R. Jacobs, L. D. Eltis, W. W. Mohn, *J. Biol. Chem.* **2009,** *284*, 35534-35542.

[22] K. Z. Rosłoniec, M. H. Wilbrink, J. K. Capyk, W. W. Mohn, M. Ostendorf, R. Van Der Geize, L. Dijkhuizen, L. D. Eltis, *Mol. Microbiol.* **2009,** *74*, 1031-1043.

[23] M. D. Driscoll, K. J. McLean, C. Levy, N. Mast, I. A. Pikuleva, P. Lafite, S. E. J. Rigby, D. Leys, A. W. Munro, *J. Biol. Chem.* **2010,***285*, 38270-38282.

[24] M. H. Wilbrink, M. Petrusma, L. Dijkhuizen, R. van der Geize, *Appl. Environ. Microbiol.* **2011,** *77*, 4455-4464.

[25] M. Yang, R. Lu, K. E. Guja, M. F. Wipperman, J. R. St. Clair, A. C. Bonds, M. Garcia-Diaz, N. S. Sampson, *ACS Infect. Dis.* **2015,** *1*, 110-125.

[26] S. T. Thomas, B. C. VanderVen, D. R. Sherman, D. G. Russell, N. S. Sampson, *J. Biol. Chem.* **2011,** *286*, 43668-43678.

[27] M. Yang, K. E. Guja, S. T. Thomas, M. Garcia-Diaz, N. S. Sampson, *ACS Chem. Biol.* **2014,** *9*, 2632-2645.

[28] N. M. Nesbitt, X. Yang, P. Fontán, I. Kolesnikova, I. Smith, N. S. Sampson, E. Dubnau, *Infect. Immun.* **2010,** *78*, 275-282.

[29] S. Gilbert, L. Hood, S. Y. K. Seah, *J. Bacteriol.* **2018**, 200, e00512-17.

[30] R. Aggett, E. Mallette, S. E. Gilbert, M. A. Vachon, K. L. Schroeter, M. S. Kimber, S. Y. K. Seah, *J. Biol. Chem.* **2019**, 294, 11934-11943.

[31] C. J. Shi, *Biochim Biophys Acta* **1962,** *62*, 541-547.

[32] W. Wei, F.-q. Wang, S.-y. Fan, D.-z. Wei, *Appl. Environ. Microbiol.* **2010,** *76*, 4578-4582.

[33] M. Petrusma, L. Dijkhuizen, R. van der Geize, *Appl. Environ. Microbiol.* **2009,** *75*, 5300-5307.

[34] J. K. Capyk, I. Casabon, R. Gruninger, N. C. Strynadka, L. D. Eltis, *J. Biol. Chem.* **2011,** *286*, 40717-40724.

[35] K. Yao, L.-Q. Xu, F.-Q. Wang, D.-Z. Wei, *Metab. Eng.* **2014,** *24*, 181-191.

[36] X. Li, X. Chen, Y. Wang, P. Yao, R. Zhang, J. Feng, Q. Wu, D. Zhu, Y. Ma, *Steroids* **2018**, *132*, 40-45.

[37] B. B. Shingate, B. G. Hazra, *Chem. Rev.* **2014,** *114*, 6349-6382.

[38] S. Fillet, J. L. Adrio, *World J. Microbiol. Biotechnol.* **2016,** *32*, 152.

[39] R. M. Willis, B. D. Wahlen, L. C. Seefeldt, B. M. Barney, *Biochemistry* **2011,** *50*, 10550-10558.

[40] J. G. Metz, M. R. Pollard, L. Anderson, T. R. Hayes, M. W. Lassner, *Plant Physiol.* **2000,** *122*, 635-644.

10.1002/anie.202015462

[41] O. Rowland, H. Zheng, S. R. Hepworth, P. Lam, R. Jetter, L. Kunst, *Plant Physiol.* **2006,** *142*, 866-877.

[42] J. B. Cheng, D. W. Russell, *J. Biol. Chem.* **2004,** *279*, 37789-37797.

[43] P. Hofvander, T. T. P. Doan, M. Hamberg, *FEBS Lett.* **2011,** *585*, 3538-3543.

[44] B. D. Wahlen, W. S. Oswald, L. C. Seefeldt, B. M. Barney, *Appl. Environ. Microbiol.* **2009,** *75*, 2758-2764.

[45] J. Round, R. Roccor, S.-N. Li, L. D. Eltis, *Appl. Environ. Microbiol.* 2017, *83*, e00902-17.

[46] Jesus F. Barajas, Ryan M. Phelan, Andrew J. Schaub, Jaclyn T. Kliewer, Peter J. Kelly, David R. Jackson, R. Luo, Jay D. Keasling, S.-C. Tsai, *Chem. Biol.* **2015,** *22*, 1018-1029.

[47] S. Reiser, C. Somerville, *J. Bacteriol.* **1997,** *179*, 2969-2975.

[48] S. S. Shell, J. Wang, P. Lapierre, M. Mir, M. R. Chase, M. M. Pyle, R. Gawande, R. Ahmad, D. A. Sarracino, T. R. Ioerger, S. M. Fortune, K. M. Derbyshire, J. T. Wade, T. A. Gray, *PLOS Genet.* **2015,** *11*, e1005641.

[49] T. Cortes, O. T. Schubert, G. Rose, K. B. Arnvig, I. Comas, R. Aebersold, D. B. Young, *Cell Rep.* **2013,** *5*, 1121-1131.

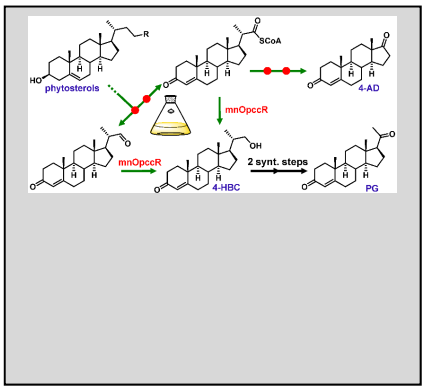
10.1002/anie.202015462

WILEY-VCH

RESEARCH ARTICLE

Entry for the Table of Contents

Insert graphic for Table of Contents here. ((Please ensure your graphic is in **one** of following formats))



An unusual reductase (mnOpccR) in phytosterol catabolism is found to catalyze the formation of 4-HBC through two different routes. Inactivation or overexpression of mnOpccR leads to exclusive production of 4-androstenedione (4-AD) and 4-HBC from phytosterols. By a two-step synthesis, 4-HBC can be further efficiently converted into progesterone (PG). This work provides two green and economical methods for 4-AD and PG manufacturing.

Table of Contents