

# SCIENTIFIC REPORTS



OPEN

## Unraveling and engineering the production of 23,24-bisnorcholesterol in sterol metabolism

Li-Qin Xu, Yong-Jun Liu, Kang Yao, Hao-Hao Liu, Xin-Yi Tao, Feng-Qing Wang & Dong-Zhi Wei

Received: 23 November 2015

Accepted: 02 February 2016

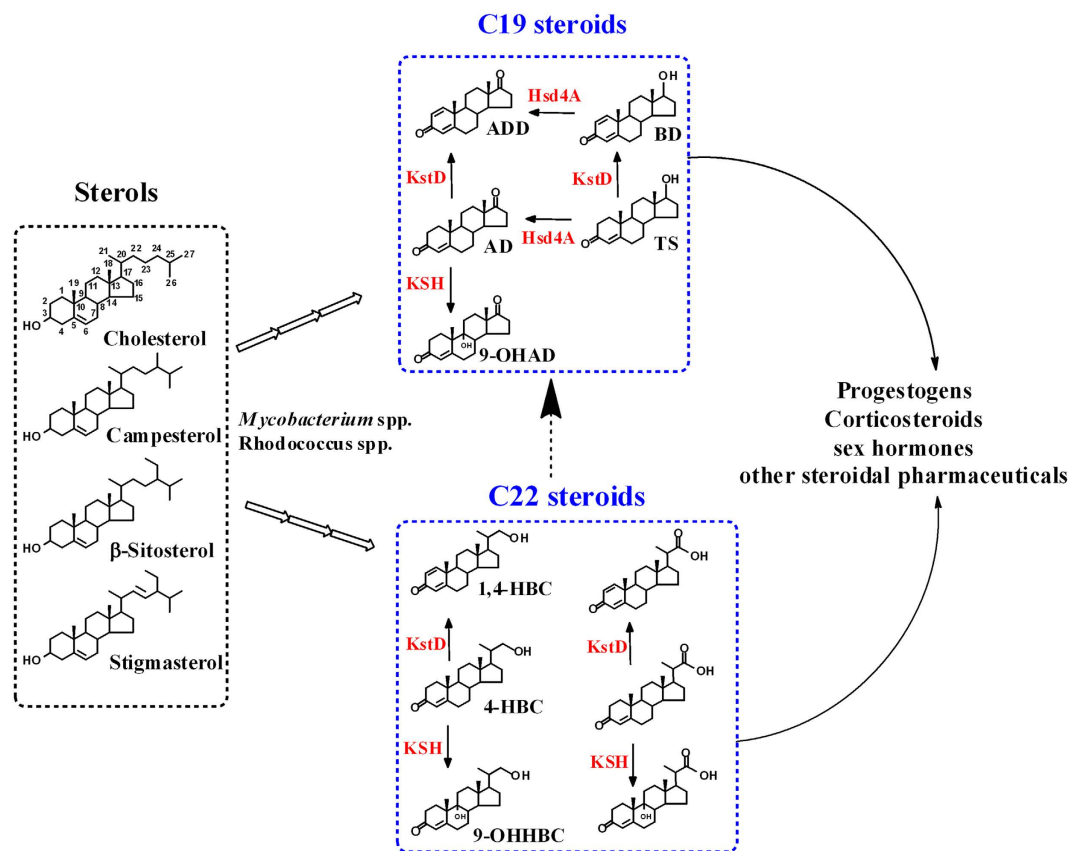
Published: 22 February 2016

The catabolism of sterols in mycobacteria is highly important due to its close relevance in the pathogenesis of pathogenic strains and the biotechnological applications of nonpathogenic strains for steroid synthesis. However, some key metabolic steps remain unknown. In this study, the *hsd4A* gene from *Mycobacterium neoaurum* ATCC 25795 was investigated. The encoded protein, Hsd4A, was characterized as a dual-function enzyme, with both 17 $\beta$ -hydroxysteroid dehydrogenase and  $\beta$ -hydroxyacyl-CoA dehydrogenase activities *in vitro*. Using a *kshAs*-null strain of *M. neoaurum* ATCC 25795 (NwIB-XII) as a model, Hsd4A was further confirmed to exert dual-function in sterol catabolism *in vivo*. The deletion of *hsd4A* in NwIB-XII resulted in the production of 23,24-bisnorcholesterol (HBCs), indicating that *hsd4A* plays a key role in sterol side-chain degradation. Therefore, two competing pathways, the AD and HBC pathways, were proposed for the side-chain degradation. The proposed HBC pathway has great value in illustrating the production mechanism of HBCs in sterol catabolism and in developing HBCs producing strains for industrial application via metabolic engineering. Through the combined modification of *hsd4A* and other genes, three HBCs producing strains were constructed that resulted in promising productivities of 0.127, 0.109 and 0.074 g/l/h, respectively.

Utilizing natural sterols, such as cholesterol and phytosterols, as carbon and energy sources is a common physiological attribute of certain actinomycetes, such as mycobacteria and rhodococci<sup>1–7</sup>. The catabolism of sterols in pathogenic strains, such as *M. tuberculosis*, is highly concerned due to its close relevance to pathogenesis and persistence<sup>1,2</sup>. Interestingly, sterol metabolism in nonpathogenic microorganisms generates metabolites that can be used as ideal precursors to synthesize steroidal pharmaceuticals (Fig. 1)<sup>8,9</sup>. Generally, two major valuable intermediates can be achieved from the metabolism of sterols, including C19 steroids and C22 steroids (23,24-bisnorcholesterol, HBCs), of which the main difference is the side chain located at carbon 17 (Fig. 1)<sup>8,10,11</sup>. In industrial manufacturing, C19 steroids, including androst-4-ene-3,17-dione (AD), androst-1,4-dien-3,17-dione (ADD) and 9 $\alpha$ -hydroxy-androst-4-ene-3,17-dione (9-OHAD), can be used to synthesize sex and adrenocortical hormones, which have been massively produced due to the successful development of industrial strains by metabolic engineering or mutation breeding (Fig. 1)<sup>4,10</sup>. However, the production of C22 steroids is less optimal, as ideal industrial strains have yet to be developed, although some C22 steroids, including 22-hydroxy-23,24-bisnorchol-4-ene-3-one (4-HBC), 22-hydroxy-23,24-bisnorchol-1,4-dien-3-one (1,4-HBC) and 9,22-dihydroxy-23,24-bisnorchol-4-ene-3-one (9-OHHBC), are highly valuable precursors in synthesizing progestational and adrenocortical hormones (Fig. 1)<sup>8,9</sup>.

The metabolism of sterols in actinomycetes follows a similar pathway, which is initiated by the conversion of sterols to 4-ene-3-sterone, catalyzed by cholesterol oxidases or 3 $\beta$ -hydroxy steroid dehydrogenases<sup>3,12,13</sup>. The catabolic pathway then can be divided into two parts: the cleavage of the steroid skeleton and the degradation of the side chain<sup>1,14,15</sup>. The opening of the steroid skeleton has been well studied and is performed by the combined 1,2-desaturation and 9-hydroxylation catalyzed by 3-ketosteroid- $\Delta^1$ -dehydrogenases (KstDs) and 3-ketosteroid-9 $\alpha$ -hydroxylases (KSHs), respectively<sup>16,17</sup>. The blockage of 1,2-desaturation or/and 9-hydroxylation by genetic manipulation or mutagenesis blocks complete sterol catabolism and results in the production of C19 steroids<sup>4,18</sup>. However, the side-chain degradation of sterols is quite complicated and has yet to be completely understood<sup>1,19</sup>. The sterol side-chain degradation is proposed to proceed in a manner similar to the  $\beta$ -oxidation of

State Key Laboratory of Bioreactor Engineering, Newworld Institute of Biotechnology, East China University of Science and Technology, Shanghai 200237, China. Correspondence and requests for materials should be addressed to F.-Q.W. (email: fqwang@ecust.edu.cn) or D.-Z.W. (email: dzhwei@ecust.edu.cn)

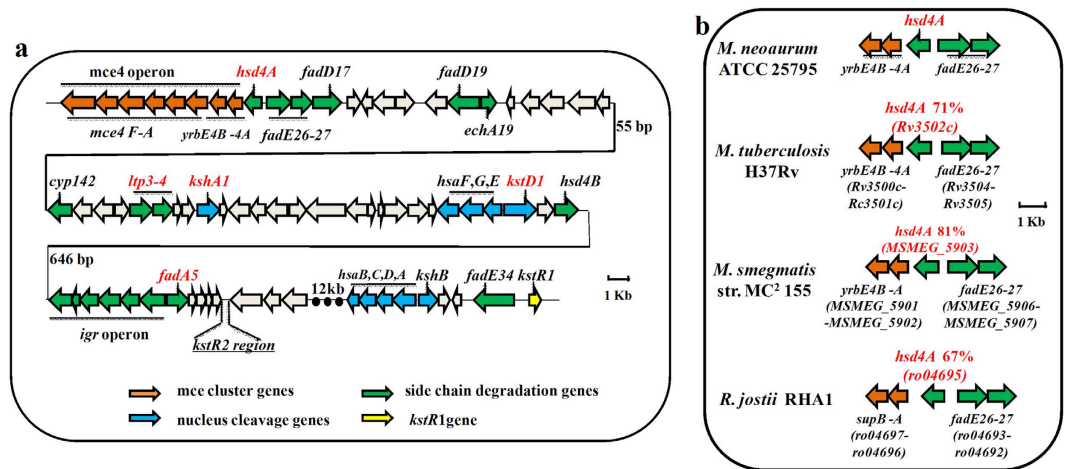


**Figure 1. An overview of the biotransformation of sterols to steroidal pharmaceuticals.** Sterols, including cholesterol and phytosterols, can be degraded to two major valuable intermediates, such as C19 and C22 steroids, through multiple biotransformation steps by some actinomycetes, such as *Mycobacterium* spp. and *Rhodococcus* spp. Well-studied C19 steroids are AD (androst-4-ene-3,17-dione), ADD (androst-1,4-dien-3,17-dione), 9-OHAD (9 $\alpha$ -hydroxy-androst-4-ene-3,17-dione), BD (boldenone) and TS (testosterone). Common C22 steroids are 4-HBC (22-hydroxy-23,24-bisnorchol-4-ene-3-one), 1,4-HBC (22-hydroxy-23,24-bisnorchol-1,4-dien-3-one) and 9-OHHBC (9,22-dihydroxy-23,24-bisnorchol-4-ene-3-one). C22 steroids can be degraded to C19 steroids through multiple undefined biotransformation steps. These two intermediates can be used as ideal precursors to synthesize steroidal pharmaceuticals, such as progesterones and corticosteroids. The abbreviations in red represent the key enzymes that determine product selectivity. KstD, 3-ketosteroid- $\Delta^1$ -dehydrogenases; KSH, 3-ketosteroid-9 $\alpha$ -hydroxylases; Hsd4A, a short chain dehydrogenase. Arrows in arrays signify multiple biotransformation steps. Curved arrows indicate the application value of C22 steroids and C19 steroids to produce steroidal pharmaceuticals. The dashed arrow indicates undefined biotransformation steps.

fatty acids, which is initiated by a terminal carboxylation by Cyp125<sup>20,21</sup>. The side-chain degradation then is proposed to comprise three  $\beta$ -oxidation-like cycles, which each cycle containing successive enzymatic steps catalyzed by acyl-CoA ligases, acyl-CoA dehydrogenases, enoyl-CoA hydratases,  $\beta$ -hydroxyacyl-CoA-dehydrogenases and acyl-CoA thiolases<sup>1,18</sup>. To date, several steps in the  $\beta$ -oxidation-like cycles have been characterized, such as the thiolysis of 3,22-dioxo-chol-4-ene-24-oyl-CoA to 3-OPC-CoA catalyzed by FadA5 and the dehydrogenation of 3-oxo-4-pregnene-20-carboxyl-CoA catalyzed by FadE28-FadE29<sup>2,22</sup>. However, some key steps involved in the sterol side-chain degradation remain unknown, restricting a complete understanding of this process.

*Hsd4A* is a putative gene located in the sterol catabolic gene cluster, which was deduced to encode a short chain dehydrogenase (Hsd4A). As Hsd4A shares intriguing similarity with the N-terminal domain of eukaryotic 17 $\beta$ -hydroxysteroid dehydrogenase IV (17 $\beta$ HSD4), Hsd4A has been speculated to be a 17 $\beta$ -hydroxysteroid dehydrogenase<sup>15,23</sup>. However, the *hsd4A* gene in *R. jostii* RHA1 was highly up-regulated along with other genes that are involved in the side-chain degradation of sterols, and thus was speculated to be essential for sterol side-chain degradation<sup>15</sup>. Therefore, Hsd4A was further proposed to be a  $\beta$ -hydroxyacyl-CoA dehydrogenase in the degradation of the sterol side chain because the N-terminal domain of 17 $\beta$ HSD4 also could act as a D-3-hydroxyacyl-CoA dehydrogenase in degrading branched fatty acids and bile acids<sup>15,23,24</sup>. Although these two possible biochemical functions have been assigned to Hsd4A, neither have been validated<sup>1</sup>. As these two putative functions may have underlying associations in the production of C19 or C22 steroids, the exact function of Hsd4A requires elucidation.

With the characterization of *kstDs* and *kshAs* homologous genes in some actinomycetes<sup>17,25–28</sup>, it has become feasible to develop C19 steroids producing strains with ideal productive performance by metabolic



**Figure 2.** (a) Partial gene cluster encoding the catabolism of sterols in *M. neoaurum* ATCC 25795. Genes in the map are color-coded according to assigned or proposed function: orange, mce cluster genes for steroids transportation; green, side-chain degradation genes; blue, nucleus cleavage genes; gray, unassigned function; yellow, gene encoding the transcriptional repressor KstR1. The numbers, e.g. 55 bp and 646 bp, beside the meander lines indicate the spaces between adjacent genes. (b) Schematic of the genomic organization of *hsd4A* homologues in *M. neoaurum* ATCC 25795 and other mycobacteria. Percentages, such as 71%, 81% and 67%, indicate the amino acid sequence identity of Hsd4A<sub>MN</sub> in *M. neoaurum* ATCC 25795 with homologs from other mycobacteria, including *M. tuberculosis* H37Rv, *M. smegmatis* str. MC<sup>2</sup> 155 and *R. jostii* RHA1. The location of *hsd4A* is between the *yrbE4B-4A* and *fadE26-27*. The size and direction of genes are indicated by arrows with corresponding length to scale.

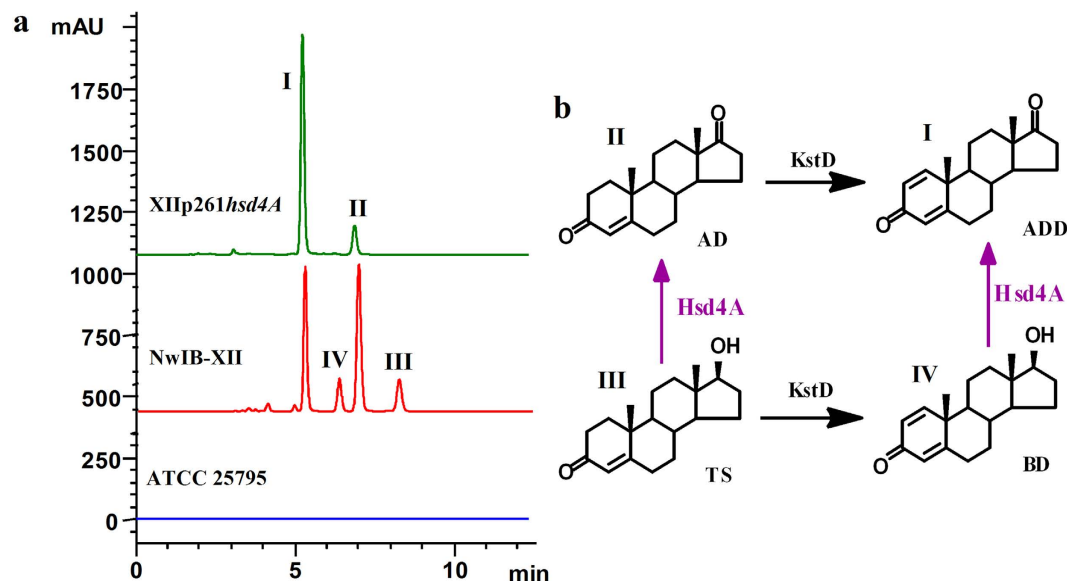
engineering<sup>4,29</sup>. However, C22 steroids producing strains cannot be developed by rational genetic engineering due to the unknown formation mechanism of C22 steroids. Early studies identified C22 steroids as the byproducts of the transformation of sterols to C19 steroids in several *Mycobacterium* strains acquired by mutagenesis, such as *Mycobacterium* sp. NRRL B-3683 and *M. fortuitum* NRRL B-8119<sup>30-32</sup>. For example, only about 55 μmol bisnorcholelic acid derivatives and 9-OHHBC can be produced from transforming β-sitosterols to 700 μmol of 9-OHAD by *M. fortuitum* NRRL B-8119<sup>32,33</sup>. Obviously, these C19 steroids producing mycobacteria are not the desired strains to produce C22 steroids. To develop attractive C22 steroids producing strains, mutation breeding method has been utilized. For example, *M. parafortuitum* complex MCI 0617 was developed using ultraviolet mutagenesis, and this strain was able to produce 5.3 g/l of 1,4-HBC, along with two byproducts 4-HBC (about 0.44 g/l) and ADD (0.06 g/l) from 10 g/l of cholesterol<sup>11</sup>. However, the development of C22 steroids producing strains has been largely unsuccessful because of the low productivity and the presence of multiple byproducts, which greatly limits the commercial application of these strains.

While investigating the role of *hsd4A* in sterol metabolism, we discovered that *hsd4A* might be a key gene relating to the formation of C22 steroids. In this study, therefore, we attempted to investigate the physiological role of Hsd4A and its possible mechanism in the conversion of sterols to C22 steroids, based on which we also tried to establish a rational strategy to modify *hsd4A* and other key genes to develop attractive HBCs producing strains.

## Results

**Hsd4A<sub>MN</sub> gene from *M. neoaurum* ATCC 25795.** The sterol catabolic gene cluster in *M. neoaurum* ATCC 25795 is described schematically in Fig. 2a. In the gene cluster, a putative gene *hsd4A* (*hsd4A<sub>MN</sub>*) was annotated as an ortholog of the *hsd4A* gene (Rv3502c) from *M. tuberculosis* H37Rv. Its encoded protein, Hsd4A<sub>MN</sub>, was estimated to be 30.9 kDa and shares high similarities with its counterparts, such as 97%, 71% and 67% amino acid identities with the Hsd4As from *Mycobacterium* sp. VKM Ac-1815D, *M. tuberculosis* H37Rv and *R. jostii* RHA1, respectively (Fig. 2b). As shown in Fig. 2, *hsd4A<sub>MN</sub>* is located between the *mce4* operon and *fadE26-27*. The *mce4* operon encodes a multicomponent ABC-like ATP-dependent transport system, which plays an important role in the cellular uptake of sterols<sup>1,15,34</sup>. The *fadE26-27* locus encodes acyl-CoA dehydrogenase enzymes that catalyze the dehydrogenation of 3-oxo-cholest-4-en-26-oyl CoA in the first cycle of cholesterol side-chain β-oxidation<sup>35</sup>. Obviously, *hsd4A<sub>MN</sub>* and its genetic organization are highly conserved in some mycobacteria and rhodococci (Fig. 2b), indicating the conserved function of Hsd4A in these strains for sterol catabolism. Hsd4As from *M. tuberculosis* H37Rv and *R. jostii* RHA1 have been proposed to be a β-hydroxyacyl-CoA dehydrogenase or a 17β-hydroxysteroid dehydrogenase<sup>1,15,23,36</sup>, indicating a possibly similar function for Hsd4A<sub>MN</sub>. However, the exact role of Hsd4A in these microorganisms has not been characterized<sup>1</sup>.

**Enzymatic characteristics of Hsd4A<sub>MN</sub> in vitro.** To explore the possible function, *hsd4A<sub>MN</sub>* was expressed heterologously in *Escherichia coli* BL21(DE3) (*E-hsd4A*, Table S1). The proposed activities of Hsd4A<sub>MN</sub> were then investigated using purified protein (Fig. S1). As indicated in Table 1, Hsd4A<sub>MN</sub> catalyzed

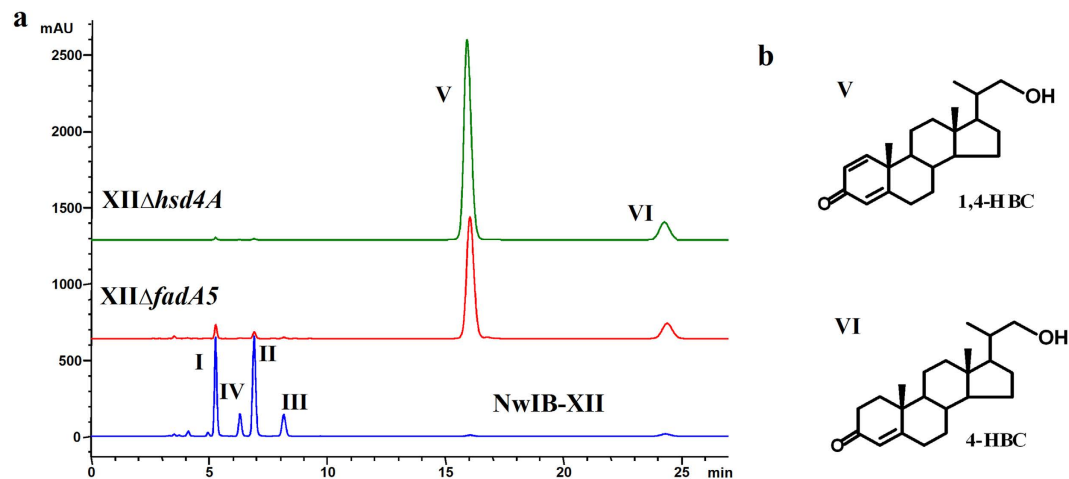


**Figure 3. Phenotypic analyses of the metabolism of sterol by *M. neoaurum* ATCC25795 and its derivative strains.** (a) HPLC chromatogram comparison of the products from the transformation of 2 g/l of cholesterol in MYC/02 media at 30 °C by strains *M. neoaurum* ATCC 25795 (blue), NwIB-XII (red) and XIIp261*hsd4A* (green). Cholesterol can be completely degraded by *M. neoaurum* ATCC 25795 without obvious accumulation of intermediates. The catabolism of cholesterol in NwIB-XII was blocked to accumulate multiple metabolites, including ADD (I), AD (II), TS (III) and BD (IV). The plasmid pMV261-*hsd4A* was electro-introduced into NwIB-XII, resulting in the XIIp261*hsd4A* strain. Strain XIIp261*hsd4A* transformed cholesterol to ADD (I) and AD (II). (b) Conversion relationship between the metabolites (I to IV). Arrows in purple signify the function of Hsd4A<sub>MN</sub> deduced from the product phenotypes of strains NwIB-XII and XIIp261*hsd4A*. Hsd4A<sub>MN</sub> can irreversibly catalyze the oxidation of TS to AD and BD to ADD, respectively. AD, androst-4-ene-3,17-dione; ADD, androst-1,4-dien-3,17-dione; BD, boldenone; TS, testosterone.

the dehydrogenation of testosterone (TS) and boldenone (BD) in the presence of NAD<sup>+</sup> to produce AD and ADD, respectively. Interestingly, it did not catalyze the 17 $\beta$ -reduction of AD and ADD in the presence of NADH, demonstrating that Hsd4A<sub>MN</sub> does not have 17 $\beta$ -hydroxysteroid dehydrogenase function without reversible activity. As for the proposed  $\beta$ -hydroxyacyl-CoA dehydrogenase activity of Hsd4A<sub>MN</sub>, two commercially available fatty acyl-CoA esters, acetoacetyl-CoA and DL- $\beta$ -hydroxybutyryl-CoA, were tested as substrates for Hsd4A<sub>MN</sub>. The results showed that Hsd4A<sub>MN</sub> catalyzed the conversion of  $\beta$ -hydroxybutyryl-CoA to acetoacetyl-CoA using NAD<sup>+</sup> as coenzyme and also carried out the reverse reaction in the presence of NADH, which unambiguously demonstrated that Hsd4A<sub>MN</sub> also had  $\beta$ -hydroxyacyl-CoA dehydrogenase function with reversible catalytic activity (Table 1). According to the above results, therefore, Hsd4A<sub>MN</sub> was characterized as a dual-function enzyme *in vitro*, possessing both the activities of  $\beta$ -hydroxyacyl-CoA dehydrogenase and 17 $\beta$ -hydroxysteroid dehydrogenase.

To further characterize Hsd4A<sub>MN</sub>, its enzymatic properties were determined (Table 1). Compared to TS and BD, 4-androstene-17 $\alpha$ -ol-3-one (epi-TS) and 4-androstene-17 $\alpha$ -methyl-17 $\beta$ -ol-3-one (17 $\alpha$ -MT) could not be catalyzed by Hsd4A<sub>MN</sub>, confirming that Hsd4A<sub>MN</sub> has the typical activity of 17 $\beta$ -hydroxysteroid dehydrogenase with high regio and stereoselectivity. Compared to acetoacetyl-CoA and DL- $\beta$ -hydroxybutyryl-CoA, acetoacetate and DL- $\beta$ -hydroxybutyric acid could not be converted by Hsd4A<sub>MN</sub>, indicating that the enzyme is a typical  $\beta$ -hydroxyacyl-CoA dehydrogenase similar to the counterparts involved in fatty acid  $\beta$ -oxidation. The kinetic parameters  $K_m$  and  $V_{max}$  of Hsd4A<sub>MN</sub> for TS and BD were 19–23  $\mu$ M and 0.169–0.217  $\mu$ mol/min/mg, respectively; and for acetoacetyl-CoA and DL- $\beta$ -hydroxybutyryl-CoA were 101–124  $\mu$ M and 0.091–0.113  $\mu$ mol/min/mg, respectively.

**Physiological role of Hsd4A<sub>MN</sub> *in vivo*.** Hsd4A<sub>MN</sub> acts as a dual-function enzyme *in vitro*, but its physiological role *in vivo* is unknown. As *hsd4A<sub>MN</sub>* is highly conserved in the sterol catabolic gene cluster, the investigation mainly focused on its role in sterol metabolism. Corresponding to the enzymatic activities *in vitro*, Hsd4A<sub>MN</sub> were deduced to have two roles in sterol metabolism *in vivo*: the oxidation of 17 $\beta$ -hydroxysteroids by 17 $\beta$ -hydroxysteroid dehydrogenase and the  $\beta$ -hydroxyacyl-CoA dehydrogenation in the  $\beta$ -oxidation of the sterol side chain by  $\beta$ -hydroxyacyl-CoA dehydrogenase. Both of these reactions can occur before the cleavage of the sterol nucleus in sterol metabolism. Given that the natural sterols can be completely degraded by *M. neoaurum* ATCC 25795 to CO<sub>2</sub> and H<sub>2</sub>O without obvious accumulation of intermediates (Fig. 3a), the function of Hsd4A<sub>MN</sub> is difficult to determine via identifying the metabolites of sterols after genetic manipulation of *hsd4A<sub>MN</sub>* in *M. neoaurum* ATCC 25795. To facilitate the functional analysis of Hsd4A<sub>MN</sub>, therefore, it was necessary to build a model strain which could accumulate suitable metabolites with intact steroid skeletons before the genetic manipulation of *hsd4A<sub>MN</sub>*.



**Figure 4. Phenotypic analyses of the sterol metabolism of *M. neoaurum* NwIB-XII and its derivative strains.** (a) HPLC chromatogram comparison of the products from the transformation of 2 g/l of cholesterol in MYC/02 media at 30 °C by strains NwIB-XII (blue), XII $\Delta$ *fadA5* (red) and XII $\Delta$ *hsd4A* (green). The deletion of *hsd4A<sub>MN</sub>* or *fadA5* in NwIB-XII greatly changed the metabolic products of cholesterol. Both XII $\Delta$ *hsd4A* and XII $\Delta$ *fadA5* converted cholesterol to 1,4-HBC (V) and 4-HBC (VI). (b) 23,24-Bisnorcholesterols characterized from the metabolites accumulated in *M. neoaurum hsd4A<sub>MN</sub>* or *fadA5* deleted strains. 1,4-HBC, 22-hydroxy-23,24-bisnorchol-1,4-dien-3-one; 4-HBC, 22-hydroxy-23,24-bisnorchol-4-ene-3-one.

3-Ketosteroid-9 $\alpha$ -hydroxylase (encoded by two genes, *kshA* and *kshB*) is one of the key enzymes in dissociation of the steroid skeleton, and inactivating this enzyme would block the cleavage of the steroid skeleton<sup>4,28</sup>. Based on the genome sequence, two putative *kshA* (*kshA1* and *kshA2*) genes were identified in *M. neoaurum* ATCC 25795 (Fig. 2a). Therefore, a model to investigate the function of Hsd4A<sub>MN</sub> *in vivo* was constructed by the successive deletion of *kshA1* and *kshA2* in *M. neoaurum* ATCC 25795, which is a *kshA*-null mutant (NwIB-XII, Table S1). As expected, sterol catabolism in strain NwIB-XII was blocked, accumulating multiple metabolites with intact steroid skeletons, including ADD (I), AD (II), and two 17 $\beta$ -hydroxysteroids TS (III) and BD (IV) (Figs 3a,b and S2).

To determine whether Hsd4A<sub>MN</sub> possesses 17 $\beta$ -hydroxysteroid dehydrogenase activity *in vivo*, *hsd4A<sub>MN</sub>* was augmented in NwIB-XII using plasmid pMV261-*hsd4A*, resulting in the XIIp261*hsd4A* strain. The vacant plasmid pMV261 was introduced into NwIB-XII as a control (designated as XIIp261). In contrast to NwIB-XII and XIIp261, strain XIIp261*hsd4A* transformed sterols to AD and ADD without any detected TS and BD (Figs 3a and S3), which could be ascribed to the conversion of TS and BD to AD and ADD, respectively, by the enhanced Hsd4A<sub>MN</sub> activity (Fig. 3b). Therefore, this suggests that Hsd4A<sub>MN</sub> can also act as a 17 $\beta$ -hydroxysteroid dehydrogenase to irreversibly catalyze the oxidation of TS to AD and BD to ADD *in vivo*.

To further characterize the function of Hsd4A<sub>MN</sub>, *hsd4A<sub>MN</sub>* was deleted in NwIB-XII, leading to the XII $\Delta$ *hsd4A* strain (Table S1). Surprisingly, the metabolic products of sterols in XII $\Delta$ *hsd4A* were significantly different from those in NwIB-XII. XII $\Delta$ *hsd4A* converted sterols to two novel products, V and VI (Fig. 4a). The major product V was identified as 1,4-HBC and the minor product VI was identified as 4-HBC (Figs 4b and S4). Both 1,4-HBC and 4-HBC are the incompletely degraded products of the sterol side chain, which still contain three carbon atoms at carbon 17. To confirm whether the production of these two compounds could be attributed to the deletion of *hsd4A<sub>MN</sub>*, *hsd4A<sub>MN</sub>* was then complemented in XII $\Delta$ *hsd4A* (strain Cp-4A, Table S1), which restored the same metabolic phenotype observed in NwIB-XII (Fig. S5). These results confirmed that deletion of *hsd4A<sub>MN</sub>* resulted in incomplete degradation of the sterol side chain, demonstrating that Hsd4A<sub>MN</sub> plays a significant role in sterol side-chain degradation. Therefore, Hsd4A<sub>MN</sub> exhibited two enzymatic activities *in vivo*, similar to results observed *in vitro*. According to the characterized enzymatic properties of Hsd4A<sub>MN</sub> *in vitro*, the involvement of Hsd4A<sub>MN</sub> in side-chain degradation should be attributed to its  $\beta$ -hydroxyacyl-CoA dehydrogenase activity, which has been proposed in previous studies<sup>1,15,36</sup>.

Although 1,4-HBC and 4-HBC were the products derived from the deletion of *hsd4A<sub>MN</sub>*, they were verified to be not the substrates of Hsd4A<sub>MN</sub> *in vitro* (Table 1), indicating that the two compounds are further derivatives of the physiological substrate of Hsd4A<sub>MN</sub>. According to the postulated  $\beta$ -oxidation of the sterol side chain, the most probable precursor of 1,4-HBC and 4-HBC is 22-hydroxy-3-oxo-25,26-bisnorchol-4-en-24-oyl CoA (22HOBNC-CoA), which contains a  $\beta$ -hydroxybutyryl-CoA moiety that has been determined to be the active substrate of Hsd4A<sub>MN</sub> *in vitro* (Table 1).

**Engineering of the metabolism of sterols to produce HBCs.** In the industry, HBCs (including 1,4-HBC, 4-HBC and 9-OHHBC, Fig. 1) can be used as precursors to produce progestational and adrenocortical hormones. HBCs are mainly produced from the modified metabolism of sterols in microorganisms. However, the formation mechanism of HBCs in sterol metabolism has not been clarified, and industrial strains used to produce HBCs are usually achieved by mutation breeding<sup>11</sup>. Therefore, it is a great challenge to rationally develop

Substrate <sup>b</sup>	Hsd4A <sub>MN</sub>		
	V <sub>max</sub> (μmol/min/mg)	K <sub>m</sub> (μM)	Co-factor
AD	n.d. <sup>c</sup>	n.d.	NADH
TS	0.169 ± 0.052	23 ± 7	NAD <sup>+</sup>
ADD	n.d.	n.d.	NADH
BD	0.217 ± 0.068	19 ± 4	NAD <sup>+</sup>
epi-TS	n.d.	n.d.	NAD <sup>+</sup>
17α-MT	n.d.	n.d.	NAD <sup>+</sup>
4-HBC	n.d.	n.d.	NAD <sup>+</sup>
1,4-HBC	n.d.	n.d.	NAD <sup>+</sup>
acetoacetate	n.d.	n.d.	NADH
DL-β-hydroxybutyric acid	n.d.	n.d.	NAD <sup>+</sup>
acetoacetyl-CoA	0.091 ± 0.012	101 ± 27	NADH
DL-β-hydroxybutyryl-CoA	0.113 ± 0.025	124 ± 39	NAD <sup>+</sup>

**Table 1. Kinetic characterization of Hsd4A<sub>MN</sub> from *M. neoaurum* ATCC 25795<sup>a</sup>.** <sup>a</sup>The maximal specific activity V<sub>max</sub> and substrate affinity constant K<sub>m</sub> are indicated as average ± standard error (n = 3). Hsd4A<sub>MN</sub> displayed distinct activities in distinct co-factor environments (NADH/NAD<sup>+</sup>), which is specified in the rightmost column. The activity of Hsd4A<sub>MN</sub> was determined by the rate of NAD<sup>+</sup>/NADH formation at 340 nm (ε = 6.22 mM<sup>-1</sup> cm<sup>-1</sup>) and 30 °C. <sup>b</sup>Abbreviations: AD, androst-4-ene-3,17-dione; TS, testosterone, ADD, androst-1,4-dien-3,17-dione; BD, boldenone; epi-TS, 4-androstene-17α-ol-3-one; 17α-MT, 4-androstene-17α-methyl-17β-ol-3-one; 4-HBC, 22-hydroxy-23,24-bisnorchol-4-ene-3-one; 1,4-HBC, 22-hydroxy-23,24-bisnorchol-1,4-dien-3-one. <sup>c</sup>n.d., not detected.

HBCs producing strains. In this study, the production of HBCs in the *hsd4A<sub>MN</sub>*-deleted strain opened the door to unravel the production mechanism of HBCs and paved the way to rationally construct HBCs producing strains.

As indicated in Fig. 4a, strain XIIΔ*hsd4A* could transform sterols to HBCs. To evaluate the production performance, a resting cell biotransformation of phytosterols was employed. When 40 g/l of phytosterols was fermented by XIIΔ*hsd4A* resting cells for 144 h, a 39–40% molar yield (12.48–13.04 g/l) of 1,4-HBC was achieved, along with an 18–20% yield (6.02–6.47 g/l) of 4-HBC and a 1–2% yield (0.41–0.53 g/l) of ADD (Table 2). This performance made the strain XIIΔ*hsd4A* not ideal for industrial application because the low yield ratio of 4-HBC or 1,4-HBC among the products would lead to a low recovery rate at the industrial scale. Therefore, a combined strategy to further modify the metabolic pathway of sterols was necessary to develop promising strains for industrial application.

1,4-HBC and 4-HBC differ at the C1, 2 double bond, which is catalyzed by KstDs<sup>4,6,25</sup>. There are three KstD homologues in *M. neoaurum* ATCC 25795, and KstD1 was characterized as the major enzyme in our previous work<sup>4</sup>. Therefore, strain XIIΔ*hsd4A* could be further optimized to better produce 1,4-HBC or 4-HBC by strengthening or weakening the activities of KstDs (Fig. 5a,b).

To develop the 1,4-HBC producing strain, *kstD1* was overexpressed in XIIΔ*hsd4A* (pMV261-*kstD1*), resulting in the XIIΔ*hsd4A*-p261*kstD1* strain (Fig. 5a). The transformation of 40 g/l of phytosterols by resting cells for 144 h resulted in a 54–57% molar yield of 1,4-HBC (17.56–18.23 g/l) with a 4–7% yield of 4-HBC (1.57–2.15 g/l) and a 1–2% yield (0.35–0.52 g/l) of ADD (Table 2). The overexpression of *kstD1* in XIIΔ*hsd4A* significantly increased the yield ratio of 1,4-HBC among the products from 65–66% to about 90% and the productivity of 1,4-HBC from 0.087–0.091 g/l/h to 0.122–0.127 g/l/h, greatly improving the industrial applicability for 1,4-HBC production.

To construct the 4-HBC producing strain, *kstD1* was deleted in XIIΔ*hsd4A*, resulting in a strain XIIΔ*hsd4A*Δ*kstD1* (Fig. 5b). The deletion of *kstD1* increased the production of 4-HBC from an 18–20% molar yield (6.02–6.47 g/l) to a 33–35% molar yield (10.68–11.22 g/l) after the transformation of 40 g/l of phytosterols. However, there was still a 15–18% molar yield of 1,4-HBC (5.02–5.77 g/l) (Table 2). In our previous study, *kstD3* was demonstrated to be another important gene encoding KstD<sup>4</sup>. Therefore, *kstD3* was further deleted in XIIΔ*hsd4A*Δ*kstD1* to weaken the yield of 1,4-HBC. Unexpectedly, in the transformation of 40 g/l of phytosterols for 144 h, strain XIIΔ*hsd4A*Δ*kstD13* still accumulated significant amounts of 1,4-HBC (4.35–4.86 g/l), indicating that there was significant residual KstD activity. Therefore, *kstD2*, which was characterized as a weak KstD gene<sup>4</sup>, was also deleted, resulting in a *kstDs*-null mutant XIIΔ*hsd4A*Δ*kstD123*. The transformation of 40 g/l of phytosterols by XIIΔ*hsd4A*Δ*kstD123* achieved a 47–49% molar yield (15.24–15.75 g/l) of 4-HBC, along with a 1–2% yield (0.52–0.61 g/l) of 1,4-HBC, a 2–3% yield (0.93–1.04 g/l) of compound VII and a small amount of AD (0.31–0.42 g/l) (Table 2). Among these products, the yield ratio of 4-HBC reached about 90%, and its productivity was increased to 0.106–0.109 g/l/h, indicating that XIIΔ*hsd4A*Δ*kstD123* was a good 4-HBC producing strain (Table 2).

The compound VII was identified as 9-OHHBC (Fig. S4), which is a useful precursor to synthesize adrenocorticoids. Therefore, the feasibility in developing a 9-OHHBC producing strain was investigated. According to the steroid nucleus metabolic mechanism (Fig. 5c), *kstD1*, *kstD2*, *kstD3* and *hsd4A<sub>MN</sub>* were progressively deleted in *M. neoaurum* ATCC 25795, resulting in the final strain MNΔ*hsd4A*Δ*kstD123*. In the transformation of 40 g/l of phytosterols, strain MNΔ*hsd4A*Δ*kstD123* produced a 30–32% molar yield of 9-OHHBC (10.25–10.72 g/l) (Table 2), demonstrating that a 9-OHHBC producing strain could be developed by combined modification of the

Strains	Substrates	Substrate molar conversion rate (%)	Metabolites <sup>b</sup>	Molar yield (%)
XIIΔ <i>hsd4A</i>	phytosterols, 40 g/l	70–73 in 144 h	<b>1,4-HBC</b>	<b>39–40</b>
			4-HBC	18–20
			ADD	1–2
XIIΔ <i>hsd4A</i> -p261 <i>kstD1</i>	phytosterols, 40 g/l	70–73 in 144 h	<b>1,4-HBC</b>	<b>54–57</b>
			4-HBC	4–7
			ADD	1–2
XIIΔ <i>hsd4A</i> Δ <i>kstD1</i>	phytosterols, 40 g/l	63–66 in 144 h	<b>4-HBC</b>	<b>33–35</b>
			1,4-HBC	15–18
			9-OHHBC	3–4
			AD	1–2
XIIΔ <i>hsd4A</i> Δ <i>kstD13</i>	phytosterols, 40 g/l	63–66 in 144 h	<b>4-HBC</b>	<b>36–37</b>
			1,4-HBC	13–15
			9-OHHBC	3–5
			AD	1–2
XIIΔ <i>hsd4A</i> Δ <i>kstD123</i>	phytosterols, 40 g/l	60–62 in 144 h	<b>4-HBC</b>	<b>47–49</b>
			9-OHHBC	2–3
			1,4-HBC	1–2
			AD	1–2
MNΔ <i>hsd4A</i> Δ <i>kstD123</i>	phytosterols, 40 g/l	50–54 in 144 h	<b>9-OHHBC</b>	<b>30–32</b>
			9-OHAD	14–15
<i>Mycobacterium</i> sp. NRRL B-3683 <sup>30</sup>	Tall oil phytosterols 1 g/l	n.m. in 192h	<b>1,4-HBC</b>	3
			<b>4-HBC</b>	Trace
			ADD	48
			AD	1
<i>Mycobacterium</i> sp. 2-4M <sup>10</sup>	sitosterol, 5 g/l	95–97 in 120 h	<b>9-OHHBC</b>	1.5–1.6
			<b>4-HBC</b>	n.m. <sup>c</sup>
			9-OHAD	48–50
			AD	21–22
			9-HCBC	0.5–0.7
			9-OHT	0.10–0.14
<i>M. parafortuitum</i> complex MCI 0617 <sup>11</sup>	cholesterol 10 g/l	99 in 160 h	<b>1,4-HBC</b>	62
			<b>4-HBC</b>	5
			ADD	0.8
<i>M. neoaurum</i> Mut <sub>MN-kstD1e-2e-3</sub> <sup>4</sup>	phytosterols, 15 g/l	70–75 in 144 h	<b>4-HBC</b>	1–2
			9-OHAD	50–55
			AD	10–15

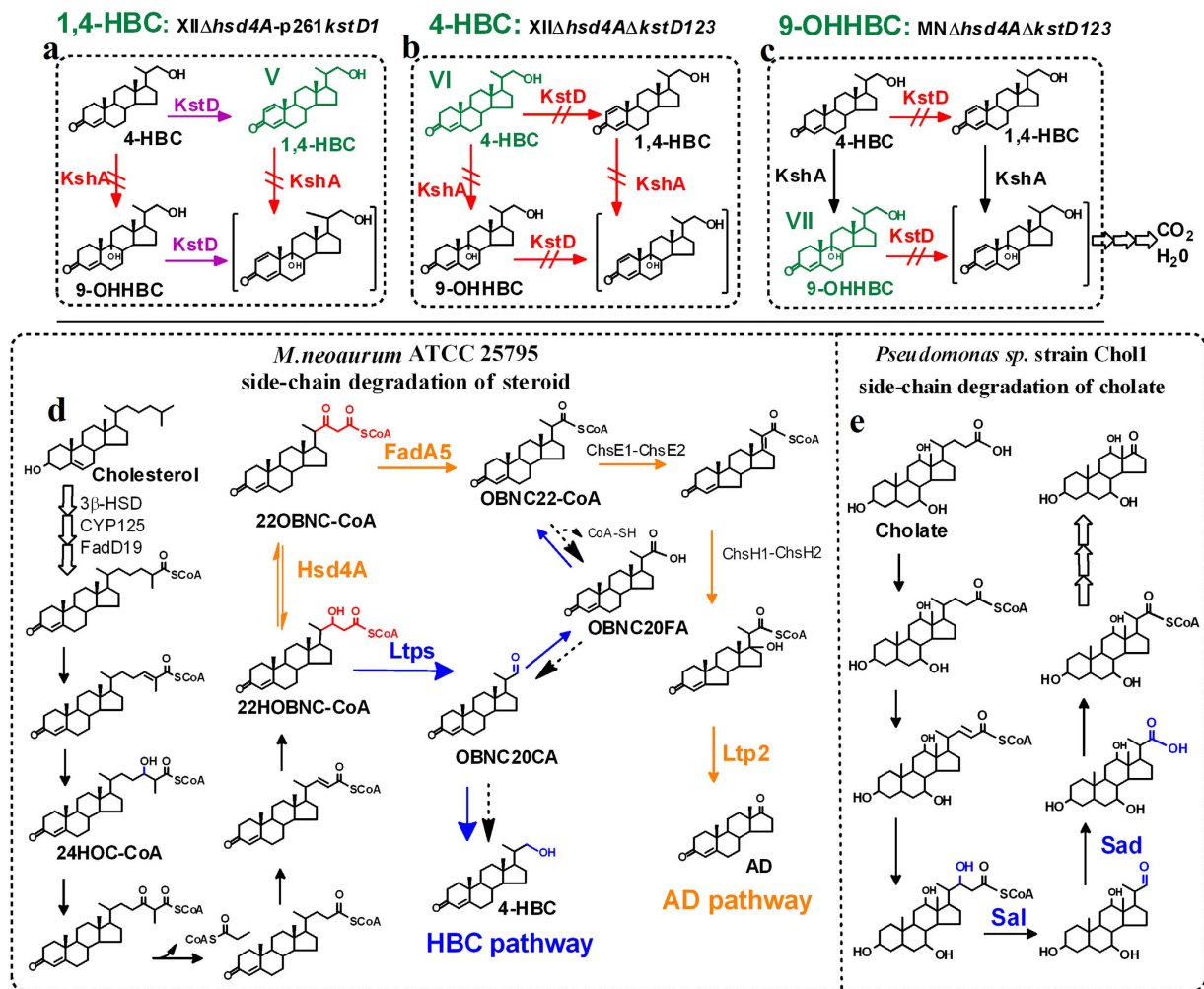
**Table 2. Production performance of the constructed HBCs producing strains in this study and some previously reported HBCs producing strains<sup>a</sup>.** <sup>a</sup>The biotransformation was performed in a resting cell transformation system supplemented with 40 g/l of phytosterols for 144 h at 30 °C. Data are obtained from three independent experiments performed upon triplicate samples. <sup>b</sup>Abbreviations: AD, androst-4-ene-3,17-dione; ADD, androst-1,4-dien-3,17-dione; 9-OHAD, 9 $\alpha$ -hydroxy-androst-4-ene-3,17-dione; 4-HBC, 22-hydroxy-23,24-bisnorchol-4-ene-3-one; 1,4-HBC, 22-hydroxy-23,24-bisnorchol-1,4-dien-3-one; 9-OHHBC, 9,22-dihydroxy-23,24-bisnorchol-4-ene-3-one; 9-OHT, 9 $\alpha$ -hydroxy-4-androsten-17 $\beta$ -ol-3-one; 9-HCBC, 9 $\alpha$ -hydroxy-22-carboxy-23,24-bisnorchol-4-en-3-one. The structural information of 9-OHT and 9-HCBC is provided in Fig. S7. <sup>c</sup>n.m., not mentioned.

key enzyme genes *hsd4A*<sub>MN</sub> and *kstDs*. In terms of industrial applications, strain MNΔ*hsd4A*Δ*kstD123* should be further optimized because of the massive accumulation of the byproduct 9-OHAD.

In conclusion, *hsd4A*<sub>MN</sub> plays a key role in the production of HBCs during sterol metabolism, and the combined modification of *hsd4A*<sub>MN</sub> and other key genes is a useful strategy to develop HBCs producing strains. Compared to C19 steroids producing strains in the industry (typically produce 12–13 g/l of C19 steroids from 30 g/l of phytosterols after 120 h), these HBCs producing strains displayed to be promising for industrial applications.

## Discussion

The catabolism of sterols is a long and complex process. Although components of the metabolic process have been identified, several key steps remain unknown, seriously limiting our understanding of sterol metabolism. *Hsd4A*<sub>MN</sub> is a conserved gene in the gene cluster of sterol catabolism, the function of which has been speculated



**Figure 5. Schematic diagram of sterol catabolism and the combinatorial construction of HBCs producing strains.** (a) The construction of strain XII $\Delta$ hsd4A-p261kstD1. To construct XII $\Delta$ hsd4A-p261kstD1, *hsd4A* was deleted in a *kshA*s-null mutant and then *kstD1* was overexpressed. The resulting strain XII $\Delta$ hsd4A-p261kstD1 transformed phytosterols to 1,4-HBC as the major product. (b) The construction of strain XII $\Delta$ hsd4A $\Delta$ kstD123. To construct XII $\Delta$ hsd4A $\Delta$ kstD123, *hsd4A* was deleted in a *kshA*s-null mutant and then *kstD1*, *kstD2* and *kstD3* were knocked out sequentially. The resulting strain XII $\Delta$ hsd4A $\Delta$ kstD123 transformed phytosterols to 4-HBC as the major product. (c) The construction of strain MN $\Delta$ hsd4A $\Delta$ kstD123. To construct MN $\Delta$ hsd4A $\Delta$ kstD123, *hsd4A* was deleted in *M. neoaurum* ATCC 25795 and then *kstD1*, *kstD2* and *kstD3* were knocked out sequentially. The resulting strain MN $\Delta$ hsd4A $\Delta$ kstD123 transformed phytosterols to 9-OHHBC as the major product. 9-OHHBC, 9,22-dihydroxy-23,24-bisnorchol-4-ene-3-one. (d) Proposed metabolism of the sterol side chain for the production of C19 and C22 steroids in *M. neoaurum* ATCC 25795. The conversion from 22HOBNC-CoA to AD was designated as the AD pathway (orange arrows) and has been proposed as the sole pathway of sterol side-chain degradation. Here, a new pathway is proposed: the conversion from 22HOBNC-CoA to 4-HBC (designated as the HBC pathway, blue arrows). Between the two pathways, 22HOBNC-CoA is the branching-node, which leads to AD via the catalysis of Hsd4A<sub>MN</sub> and leads to 4-HBC via an aldolytic reaction. The  $\beta$ -hydroxybutyryl-CoA moiety of 22HOBNC-CoA and acetoacetyl-CoA moiety of 22OBNC-CoA are labelled in red, which were the active substrates of Hsd4A<sub>MN</sub> *in vitro*. 24HOC-CoA, 24-hydroxy-3-oxo-chol-4-en-26-oyl CoA; 22HOBNC-CoA, 22-hydroxy-3-oxo-25,26-bisnorchol-4-en-24-oyl CoA; 22OBNC-CoA, 3,22-dioxo-25,26-bisnorchol-4-ene-24-oyl CoA; OBNC22-CoA, 3-oxo-22,23-bisnorchol-4-ene-22-oyl-CoA; OBNC20FA, 3-oxo-23,24-bisnorchol-4-ene-20-formic acid; OBNC20CA, 3-oxo-23,24-bisnorchol-4-ene-20-carbaldehyde. (e) The proposed pathway of cholate side-chain degradation in *Pseudomonas* sp. strain Chol1<sup>37</sup>. *sal*, an aldolase; *sad*, an aldehyde dehydrogenase. Intercepted arrows (red) indicate the blockage of the target reaction by gene deletion. Arrows in purple signify the enhanced KstD1 activity. Compounds in green are the major products accumulated in constructed strains. The compound in brackets is an unstable compound that will lead to the decomposition of steroid nucleus and further degradation. Dashed arrows indicate the proposed way for HBCs formation previously suggested by Szentirmai<sup>32</sup>. Arrows in arrays signify multiple reaction steps.



but not verified. This gene has been speculated as a  $17\beta$ -hydroxysteroid dehydrogenase or a  $\beta$ -hydroxyacyl-CoA dehydrogenase in sterol metabolism<sup>15,23,36</sup>. In this study, Hsd4A<sub>MN</sub> was first characterized as a dual-function enzyme that possesses both  $17\beta$ -hydroxysteroid dehydrogenase and  $\beta$ -hydroxyacyl-CoA dehydrogenase activities *in vitro*. Subsequently, the dual-role of Hsd4A<sub>MN</sub> in the metabolism of sterols *in vivo* was also confirmed using a *kshA*-null strain NwIB-XII as a model. The  $17\beta$ -hydroxysteroid dehydrogenase activity of Hsd4A<sub>MN</sub> *in vivo* was identified as the augmentation of Hsd4A<sub>MN</sub> in NwIB-XII significantly decreased the production of TS and BD (Fig. 3a). However, the identification of the  $\beta$ -hydroxyacyl-CoA dehydrogenase activity of Hsd4A<sub>MN</sub> *in vivo* was a difficult process. The deletion of *hsd4A<sub>MN</sub>* in NwIB-XII resulted in the production of 4-HBC and 1,4-HBC, rather than TS and BD, suggesting that Hsd4A<sub>MN</sub> could play another enzymatic function other than as a  $17\beta$ -hydroxysteroid dehydrogenase *in vivo*. As 4-HBC and 1,4-HBC do not act as substrates for Hsd4A<sub>MN</sub>, these must be further derivatives of the physiological substrate of Hsd4A<sub>MN</sub> during sterol side-chain degradation. Given that sterol side-chain degradation is a consensus  $\beta$ -oxidation, according to the characterized enzymatic properties of Hsd4A<sub>MN</sub> *in vitro*, the observed  $\beta$ -hydroxyacyl-CoA dehydrogenase activity of Hsd4A<sub>MN</sub> is likely responsible for *in vivo* sterol side-chain degradation. In the side-chain  $\beta$ -oxidation of sterols (Fig. 5d), there are two steryl fatty acyl-CoA esters that are considered as the substrates of  $\beta$ -hydroxyacyl-CoA dehydrogenase: 24-hydroxy-3-oxo-cholesterol-4-en-26-oyl CoA (24HOC-CoA) and 22HOBNC-CoA. Compared to 24HOC-CoA, 22HOBNC-CoA seemed to be the most likely substrate of Hsd4A<sub>MN</sub>, as 22HOBNC-CoA could theoretically be converted to HBCs via a common retro-aldol cleavage, while the conversion of 24HOC-CoA to HBCs would be a complex and difficult biochemical process (Fig. 5d). Furthermore, the fact that Hsd4A<sub>MN</sub> catalyzes the dehydrogenation of  $\beta$ -hydroxybutyryl-CoA *in vitro* further suggested that 22HOBNC-CoA would be a suitable substrate for Hsd4A<sub>MN</sub>, as  $\beta$ -hydroxybutyryl-CoA is an intrinsic moiety of 22HOBNC-CoA. If this was the case, 3,22-dioxo-25,26-bisnorcholesterol-4-ene-24-oyl CoA (22OBNC-CoA) should be the direct derivative of 22HOBNC-CoA after catalysis by Hsd4A<sub>MN</sub>, which is a reaction known to be catalyzed by a  $\beta$ -hydroxyacyl-CoA dehydrogenase in the side-chain  $\beta$ -oxidation of sterols (Fig. 5d). Acetoacetyl-CoA and  $\beta$ -hydroxybutyryl-CoA are the side chain moieties of 22OBNC-CoA and 22HOBNC-CoA, respectively. As  $\beta$ -hydroxybutyryl-CoA and acetoacetyl-CoA could be reversibly interconverted by Hsd4A<sub>MN</sub>, it is conceivable that 22OBNC-CoA might also be a possible substrate of Hsd4A<sub>MN</sub>. 22OBNC-CoA has been characterized as a substrate of FadA5, a  $\beta$ -ketoacyl-CoA thiolase that catalyzes the thiolytic cleavage of 22OBNC-CoA to 3-oxo-22,23-bisnorcholesterol-4-ene-22-oyl-CoA (OBNC22-CoA)<sup>2</sup>. To help decipher the above, *fadA5* was deleted in NwIB-XII (XII $\Delta$ *fadA5*, Table S1), which resulted in a similar phenotype to XII $\Delta$ *hsd4A*, producing 1,4-HBC and 4-HBC from the metabolism of sterols (Fig. 4a). These results demonstrated that 22OBNC-CoA similar to 22HOBNC-CoA is also a precursor of HBCs, indicating a close relationship among 22HOBNC-CoA, 22OBNC-CoA, Hsd4A<sub>MN</sub> and HBCs. Structurally, the transformation of 22HOBNC-CoA or 22OBNC-CoA to HBCs requires a cleavage between the carbon<sub>22</sub>-carbon<sub>23</sub> bond of the side chain. Biochemically, the carbon<sub>22</sub>-carbon<sub>23</sub> bond cleavage of 22HOBNC-CoA could be performed by an aldolytic reaction, and the cleavage of 22OBNC-CoA could be achieved by a  $\beta$ -ketoacyl-CoA thiolysis. As the  $\beta$ -ketoacyl-CoA thiolysis of 22OBNC-CoA is carried out by FadA5<sup>2</sup>, the direct carbon<sub>22</sub>-carbon<sub>23</sub> cleavage of 22OBNC-CoA was impossible in XII $\Delta$ *fadA5*. However, the direct carbon<sub>22</sub>-carbon<sub>23</sub> cleavage of 22HOBNC-CoA was feasible in XII $\Delta$ *fadA5* as a very similar reaction has been identified in cholate side-chain degradation in *Pseudomonas* sp strain Chol1 (Fig. 5e)<sup>37</sup>. In contrast to 22OBNC-CoA, therefore, the conversion of 22HOBNC-CoA to HBCs was more plausible by carbon<sub>22</sub>-carbon<sub>23</sub> cleavage. In this way, the conversion of 22OBNC-CoA to HBCs in XII $\Delta$ *fadA5* could be reasonably attributed to the reversible conversion of 22OBNC-CoA to 22HOBNC-CoA by Hsd4A<sub>MN</sub>.

The conversion from 22HOBNC-CoA to AD (designated as the AD pathway, Fig. 5d) has long been proposed as the sole metabolic pathway of sterol side-chain degradation. Based on the results in this study, an additional pathway was proposed: the conversion from 22HOBNC-CoA to HBCs (designated as the HBC pathway, Fig. 5d). Between the two pathways, 22HOBNC-CoA is the branching-node, which leads to AD via the catalysis of Hsd4A<sub>MN</sub> and leads to HBCs via an aldolytic reaction (Fig. 5d). Chemically, the aldolytic reaction of 22HOBNC-CoA to 3-oxo-23,24-bisnorcholesterol-4-ene-20-carbaldehyde (OBNC20CA) could be plausibly speculated according to the similar reaction catalyzed by a *sal*-encoded aldol-lyase within the degradation of cholate side chain in *Pseudomonas* sp. strain Chol1 (Fig. 5e)<sup>37,38</sup>. Further, two aldolytic reactions have been proposed and genetically characterized in sterol side-chain degradation, including carbon17-carbon20 cleavage by aldol-lyase Ltp2<sup>39</sup> and C24-branched chain cleavage by aldol-lyases Ltp3 and Ltp4<sup>40</sup>. For many organisms, aldehydes are a group of active compounds that are readily converted to alcohols by reductive enzymatic reactions or to carboxylic acids by oxidative enzymatic reactions. Therefore, the terminal aldehyde group of OBNC20CA has two possible further conversions. One is the oxidative transformation to 3-oxo-23,24-bisnorcholesterol-4-ene-20-formic acid (OBNC20FA), and the other is the reductive reaction to HBCs. Obviously, the reductive conversion correlates with the results of this study, suggesting the presence of this reaction in *M. neoaurum* ATCC 25795. It must be noted that the presence of an oxidative conversion could not be determined from the results of this study. However, it is the established conversion in cholate side-chain degradation in *Pseudomonas* sp. strain Chol1 as a subsequent reaction catalyzed by an aldehyde dehydrogenase encoded by *sad* after the aldolytic reaction (Fig. 5e)<sup>37</sup>. Under the conditions of this study, HBCs were shown to be the dead-end products of sterols in XII $\Delta$ *hsd4A* and XII $\Delta$ *fadA5*, suggesting that the reductive conversion is a terminal reaction, negating further degradation of the sterol side chain. In contrast, the oxidative conversion may not affect the complete side-chain degradation, as the terminal formic acid group can be further converted to its CoA ester (i.e. OBNC22-CoA) by acyl-CoA synthetases, which will return the metabolite to the AD pathway. As the reduction and oxidation of aldehydes are common reactions in cell metabolism, the existence of the reductive conversion does not exclude the existence of the oxidative conversion in *M. neoaurum* ATCC 25795. The minor yield of C19 sterols in HBCs producing strains, such as the 1–2% molar yield of ADD in XII $\Delta$ *hsd4A* and a small amount of AD in XII $\Delta$ *hsd4A* $\Delta$ *kstD123*, implied that the oxidative conversion may also be active in these strains (Table 2). The

molar yields of HBCs were much higher than the C19 steroids in these engineered strains; therefore, the reductive conversion was likely the dominant way to yield HBCs in *M. neoaurum* ATCC 25795, which may be closely related to the oxidation-reduction potential in mycobacterial cells.

As HBCs are valuable precursors in steroid synthesis, mycobacteria have been developed to produce HBCs<sup>8,11</sup>. Previously, 23,24-bisnorcholelic acids were considered as derivatives from the blocked AD pathway, and HBCs were speculated to be the reduced derivatives of 23,24-bisnorcholelic acids via 23,24-bisnorcholelic aldehydes (Fig. 5d)<sup>32</sup>. However, this study disclosed another possibility to produce HBCs in mycobacterial mutants, which has practical significance in clarifying the metabolic mechanism of C22 steroids and in the development of engineered strains for producing specific C22 steroids.

In this study, three HBCs producing strains were constructed and the productivity of each was assessed using a resting cell system that was supplemented with 40 g/l of phytosterols (Table 2). To the best of our knowledge, the production performances of these HBCs producing strains were significantly superior to other documented strains (Table 2), and the productivity of 4-HBC (0.106–0.109 g/l/h) and 1,4-HBC (0.122–0.127 g/l/h) was comparable to industrial productivity of C19 steroids (about 0.100–0.108 g/l/h), such as AD and 9-OHAD. For industrial application, the genetic and performance stability of engineered strains is highly important. In this study, gene inactivation was performed by in-frame deletion in the core region of gene, resulting in permanent inactivation of the target gene. Gene overexpression was performed using a universal mycobacterial plasmid pMV261 with kanamycin resistance. The plasmid stability of pMV261-*kstD1* in strain XIIΔ*hsd4A*-p261*kstD1* was tested under conditions lacking kanamycin addition. The results showed that the loss rate of pMV261-*kstD1* averaged less than 7% during 144 h (Fig. S6). In addition, the engineered strains developed in this study, as well as our previous study<sup>3,4</sup>, displayed stable production performances after irregular subculture for years. These results clearly demonstrated that the metabolic engineering strategies used in this study were effective in developing HBCs producing strains that were promising for industrial application, although their performances require further improvement.

## Methods

**Bacterial strains, plasmids, and chemicals.** Phytosterols were purchased from Davi Biochemistry (Zhejiang, China). Hydroxypropyl-β-cyclodextrin (HP-β-CD) was purchased from Cyclochem Chemicals Co., Ltd (Kunshan, China). Cholesterol, AD, TS, ADD, BD, epi-TS, 17α-MT, acetoacetyl-CoA sodium salt, DL-β-hydroxybutyryl-CoA lithium salt, lithium acetoacetate and DL-β-hydroxybutyric acid sodium salt were obtained from Sigma-Aldrich (Shanghai, China). Other molecular biology reagents and chemicals were of the highest grade available from Sigma-Aldrich (Shanghai, China) and Thermo Scientific Fermentas (Shanghai, China).

All strains used in this study are listed in Supplementary Table S1. Oligonucleotides and plasmids are listed in Table S2. Mycobacteria were cultivated aerobically in MYC/01 medium, as described previously<sup>3</sup>. The stock solutions of sterols were prepared as reported<sup>3</sup>. *E. coli* DH5α used for molecular cloning and *E. coli* BL21(DE3) used for heterologous expression were grown in Luria-Bertani medium with appropriate amounts of antibiotics at 37 °C with shaking at 200 rpm.

**Gene deletion and complementation in *M. neoaurum*.** The genome was isolated from *M. neoaurum* ATCC 25795 and sequenced as indicated previously<sup>4</sup>. Basic bioinformatics tools for sequence alignments and searching for protein homologues were performed as previously described<sup>3,4</sup>.

Unmarked in-frame gene deletion of *M. neoaurum* ATCC 25795 was based on a non-replicative suicide plasmid, p2NIL, and combined with the selectable marker cassette from pGOAL19, following a procedure described previously<sup>3,41</sup>. To delete *kshA1* and *kshA2* from *M. neoaurum*, pDEL*kshA1* and pDEL*kshA2* were constructed. *M. neoaurum* NwIB-XII was an engineered strain with the unmarked deletion of *kshA1* and *kshA2*, which served as the host to generate further gene knock-outs. The deletion of *hsd4A<sub>MN</sub>* in *M. neoaurum* ATCC 25795 and NwIB-XII (designated as MNΔ*hsd4A* and XIIΔ*hsd4A*, respectively) was carried out using plasmid pDEL*hsd4A*. A deletion mutant of *fadA5* was constructed in NwIB-XII (XIIΔ*fadA5*) using plasmid pDEL*fadA5*. Three *kstD* genes (*kstD1*, *kstD2* and *kstD3*) were knocked out sequentially in XIIΔ*hsd4A* or MNΔ*hsd4A* using p2N-*k1*, p2N-*k2* and p2N-*k3*, respectively. Functional complementation through recombinant pMV306 (pMV306-*hsd4A*) was established as indicated<sup>3</sup>. A strain harboring a vacant pMV306 was used as the blank control (XIIΔ*hsd4A*-p306). All desirable mutants and complements were validated by PCR and gene sequencing. The gene sequences of *kshA1*, *kshA2*, *hsd4A<sub>MN</sub>* and *fadA5* have been deposited in the GenBank database with the accession numbers KF573736, KF573737, KP642512 and KP642515, respectively.

**In vivo expression of Hsd4A<sub>MN</sub> and KstD1.** The coding sequences of *hsd4A<sub>MN</sub>* and *kstD1* (GenBank ID: GQ411074.1) were amplified using PrimeSTAR<sup>®</sup> High-fidelity DNA polymerase (Takara, Dalian, China) and the primers shown in Table S2, and then were digested by *Pst*I–*Hind*III (for *hsd4A<sub>MN</sub>*) or *Bam*HI–*Hind*III (for *kstD1*) before being inserted into the corresponding sites in pMV261<sup>42</sup>. The resulting plasmids pMV261-*hsd4A* and pMV261-*kstD1* were electro-introduced into NwIB-XII and XIIΔ*hsd4A*, respectively. The resulting strains were then selected under 30 μg/ml kanamycin. A strain with a vacant pMV261 was used as the blank control. The plasmid stability of pMV261-*kstD1* was assessed as indicated in Supplementary Method.

**Expression and purification of Hsd4A<sub>MN</sub>.** The coding region of *hsd4A<sub>MN</sub>* was ligated into vector pET-28a(+) via PCR amplification and restriction digestion (*Nco*I–*Xho*I). The resulting plasmid pET28-*hsd4A* was first replicated in *E. coli* DH5α and then transformed into *E. coli* BL21(DE3) to generate E-*hsd4A* for expression. E-*hsd4A* was cultured at 37 °C and shaken at 220 rpm in LB medium to OD<sub>600</sub> 0.4–0.8 and IPTG at a concentration of 0.5 mM was added to induce Hsd4A<sub>MN</sub> expression. The induced cells were further cultured for 24 h at

25 °C and shaken at 220 rpm until harvested by centrifuge at  $8,000 \times g$  for 20 min at 4 °C. The pellets were washed in buffer A (50 mM Tris-HCl pH7.5, 500 mM NaCl, 0.1 mM EDTA, 20 mM imidazole) plus 1 mM PMSF, centrifuged at  $8,000 \times g$  for 20 min at 4 °C and resuspended in 80 mL of buffer A. The cell-free extract of E-*hsd4A* was prepared by sonication for 15 min and centrifugation at  $13,000 \times g$  for 30 min at 4 °C. Purification of Hsd4A<sub>MN</sub> was conducted using an AKTA Prime system (GE Healthcare, Shanghai, China) as indicated in Supplementary Method. The standard Bradford method, with bovine serum albumin as the standard, was applied to determine the final protein concentration<sup>43</sup>.

**Characterization of Hsd4A<sub>MN</sub>.** The 17 $\beta$ -hydroxysteroid dehydrogenase activity of Hsd4A<sub>MN</sub> was determined in triplicate as the rate of NAD<sup>+</sup>/NADH formation at 340 nm ( $\epsilon = 6.22 \text{ mM}^{-1} \text{ cm}^{-1}$ ) and 30 °C, and recorded by a SpectraMax 190 (Molecular Devices; CA, USA) with the Soft-max PRO program. The assay mixture (200  $\mu$ l) was comprised of 100 mM Tris-HCl (pH 7.8 for oxidation or pH 6.6 for reduction), 100  $\mu$ M NAD<sup>+</sup>/NADH, 0.15 mM substrate and purified Hsd4A<sub>MN</sub><sup>44,45</sup>. The  $\beta$ -hydroxyacyl-CoA dehydrogenase activity of Hsd4A<sub>MN</sub> was measured by monitoring the reduction of NAD<sup>+</sup> as described<sup>46</sup>. The assay mixture (200  $\mu$ l) was comprised of 125 mM glycine-NaOH buffer (pH 9.0), 125  $\mu$ M NAD<sup>+</sup>, 0.125 mM substrate (DL- $\beta$ -hydroxybutyryl-CoA and DL- $\beta$ -hydroxybutyric acid) and purified Hsd4A<sub>MN</sub><sup>46</sup>. The acetoacetyl-CoA reductase activity of Hsd4A<sub>MN</sub> was assayed by measuring the oxidation of NADH<sup>46</sup>. The assay mixture (200  $\mu$ l) was comprised of 125 mM Tris-HCl (pH 8.0), 125  $\mu$ M NADH, 0.02 mM substrate (acetoacetyl-CoA and acetoacetate) and purified Hsd4A<sub>MN</sub><sup>46</sup>. The  $\beta$ -hydroxyacyl-CoA dehydrogenase and acetoacetyl-CoA reductase activities were also recorded by the SpectraMax 190 (Molecular Devices). In these enzyme reactions, one unit of specific enzyme activity was indicated as the reduction or oxidation of 1  $\mu$ M NADH/NAD<sup>+</sup> within 1 min per 1 mg protein at 30 °C. SigmaPlot graphical software (version 10.0, Systat Software Inc., Chicago, IL, USA) was employed to formulate the relation between [V] (reaction velocity) and [S] (substrate concentration) in terms of the Michaelis-Menten formula.

**Steroid bioconversion and analysis.** In this study, two methods, including the vegetative cell biotransformation and the resting cell biotransformation, were used to perform the conversion of steroids. Vegetative cell biotransformation was carried out using MYC/02 medium with 2 g/l of cholesterol<sup>3</sup>. The hydroxypropyl- $\beta$ -cyclodextrin resting cells were fermented with 40 g/l of phytosterols, according to the model biotransformation process<sup>47</sup>. The conversion was monitored for several days and sampled every 24 h. After acidification by 10% H<sub>2</sub>SO<sub>4</sub>, the samples were extracted twice using the same volume of ethyl acetate. The extracts were collected, evaporated to dryness and redissolved in methanol for analysis.

GC system 6890N (Agilent Technologies, Palo Alto, CA, USA) was used to determine cholesterol or the mixture of phytosterols<sup>3</sup>. High performance liquid chromatography (HPLC, Agilent Technologies) was performed on an Agilent XDB-C18 column (4.6  $\times$  250 mm; 40 °C), as previously described<sup>4</sup>. Liquid chromatography-mass spectrometry (LC/MS, an Agilent Technologies Series 1100 LC/MSD SL system with an Agilent XDB-C18, 4.6  $\times$  250 mm column, Agilent Technologies) was used to identify the metabolites of AD, ADD, BD and TS. Electrospray ionization (ESI<sup>+</sup>) was carried out as follows: temperature 100 °C, desolvation temperature 300 °C, sourcecone 30 V, and desolvation 500 L nitrogen/h.

Purification of the products was performed on a preparative thin layer chromatography (HF254–366, 20  $\times$  20 cm, Qingdao Marine Chemical Factory, China) using petroleum ether/ethyl acetate (6:4; v/v) as a mobile phase. The purified products were subjected to high resolution mass spectroscopy (MS) and nuclear magnetic resonance (NMR) analyses. High resolution mass spectroscopy was performed on a Micromass GCT instrument (Micromass UK, UK) via electron ionization (EI<sup>+</sup>) measurements: source temperature, 250 °C; electron energy, 70 eV. The <sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded at 300 K on a Bruker Avance 500 spectrometer (Bruker, Germany) at 125 and 500 MHz, respectively, with tetramethylsilane (TMS) as the internal standard in CDCl<sub>3</sub>. Chemical shifts ( $\delta$ ) are shown as parts per million (ppm) relative to TMS.

22-hydroxy-23,24-bisnorchole-4-en-3-one (4-HBC): <sup>1</sup>H NMR:  $\delta = 5.73$  (1H, s, H-4); 3.64 (1H, dd,  $J_1 = 10.3$  Hz,  $J_2 = 2.2$  Hz, H-22); 3.37 (1H, dd,  $J_1 = 10.2$  Hz,  $J_2 = 7.0$  Hz, H-22); 1.19 (3H, s, CH<sub>3</sub>-19); 1.05 (3H, d,  $J = 6.7$  Hz, CH<sub>3</sub>-21); 0.74 (3H, s, CH<sub>3</sub>-18). <sup>13</sup>C NMR:  $\delta = 35.7$  (C1); 34.0 (C2); 199.8 (C3); 123.8 (C4); 171.7 (C5); 32.9 (C6); 32.0 (C7); 35.6 (C8); 53.8 (C9); 38.6 (C10); 21.0 (C11); 39.5 (C12); 42.5 (C13); 55.6 (C14); 24.3 (C15); 27.7 (C16); 52.4 (C17); 12.1 (C18); 17.4 (C19); 38.7 (C20); 16.7 (C21); 67.9 (C22). MS: EI: 330.2558 [M]<sup>+</sup>.

22-hydroxy-23,24-bisnorchole-1,4-dien-3-one (1,4-HBC): <sup>1</sup>H NMR:  $\delta = 7.06$  (1H, d,  $J = 10.1$  Hz, H-1); 6.23 (1H, d,  $J = 10.1$  Hz, H-2); 6.07 (1H, s, H-4); 3.65 (1H, dd,  $J_1 = 10.3$  Hz,  $J_2 = 2.6$  Hz, H-22); 3.37 (1H, dd,  $J_1 = 10.3$  Hz,  $J_2 = 7.0$  Hz, H-22); 1.23 (3H, s, CH<sub>3</sub>-19); 1.04 (3H, d,  $J = 5.1$  Hz, CH<sub>3</sub>-21); 0.76 (3H, s, CH<sub>3</sub>-18). <sup>13</sup>C NMR:  $\delta = 156.0$  (C1); 127.5 (C2); 186.5 (C3); 123.8 (C4); 169.4 (C5); 33.7 (C6); 32.9 (C7); 35.6 (C8); 52.3 (C9); 43.6 (C10); 22.9 (C11); 39.3 (C12); 42.8 (C13); 55.2 (C14); 24.5 (C15); 27.7 (C16); 52.3 (C17); 12.2 (C18); 18.7 (C19); 38.7 (C20); 16.7 (C21); 67.9 (C22). MS: EI: 328.2406 [M]<sup>+</sup>.

9,22-dihydroxy-23,24-bisnorchole-4-en-3-one (9-OHHBC): <sup>1</sup>H NMR:  $\delta = 5.87$  (1H, d,  $J = 1.6$  Hz, H-4); 3.65 (1H, dd,  $J_1 = 10.5$  Hz,  $J_2 = 3.2$  Hz, H-22); 3.37 (1H, dd,  $J_1 = 10.5$  Hz,  $J_2 = 7.0$  Hz, H-22); 1.33 (3H, s, CH<sub>3</sub>-19); 1.06 (3H, d,  $J = 6.6$  Hz, CH<sub>3</sub>-21); 0.75 (3H, s, CH<sub>3</sub>-18). <sup>13</sup>C NMR:  $\delta = 28.5$  (C1); 34.1 (C2); 199.2 (C3); 126.8 (C4); 168.9 (C5); 31.8 (C6); 25.4 (C7); 37.4 (C8); 76.4 (C9); 44.3 (C10); 26.7 (C11); 34.9 (C12); 42.3 (C13); 49.2 (C14); 24.2 (C15); 27.6 (C16); 52.2 (C17); 11.2 (C18); 19.9 (C19); 38.7 (C20); 16.7 (C21); 67.9 (C22). MS: EI: 346.3 [M]<sup>+</sup>.

## References

1. Wipperman, M. F., Sampson, N. S. & Thomas, S. T. Pathogen roid rage: cholesterol utilization by *Mycobacterium tuberculosis*. *Crit. Rev. Biochem. Mol. Biol.* **49**, 269–293 (2014).
2. Schaefer, C. M. et al. FadA5 a thiolase from *Mycobacterium tuberculosis*: a steroid-binding pocket reveals the potential for drug development against tuberculosis. *Structure*. **23**, 21–33 (2015).

3. Yao, K., Wang, F. Q., Zhang, H. C. & Wei, D. Z. Identification and engineering of cholesterol oxidases involved in the initial step of sterols catabolism in *Mycobacterium neoaurum*. *Metab. Eng.* **15**, 75–87 (2013).
4. Yao, K., Xu, L. Q., Wang, F. Q. & Wei, D. Z. Characterization and engineering of 3-ketosteroid- $\Delta^1$ -dehydrogenase and 3-ketosteroid-9 $\alpha$ -hydroxylase in *Mycobacterium neoaurum* ATCC 25795 to produce 9 $\alpha$ -hydroxy-4-androstene-3,17-dione through the catabolism of sterols. *Metab. Eng.* **24**, 181–191 (2014).
5. Andriushina, V. A. *et al.* Conversion of soybean sterols into 3,17-diketosteroids using actinobacteria *Mycobacterium neoaurum*, *Pimelobacter simplex*, and *Rhodococcus erythropolis*. *Prikl Biokhim Mikrobiol.* **47**, 297–301 (2011).
6. Fernandez de las Heras, L., van der Geize, R., Drzyzga, O., Perera, J. & Navarro Llorens, J. M. Molecular characterization of three 3-ketosteroid- $\Delta^1$ -dehydrogenase isoenzymes of *Rhodococcus ruber* strain Chol-4. *J. Steroid Biochem. Mol. Biol.* **132**, 271–281 (2012).
7. McLeod, M. P. *et al.* The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proc. Natl. Acad. Sci.* **103**, 15582–15587 (2006).
8. Donova, M. V. & Egorova, O. V. Microbial steroid transformations: current state and prospects. *Appl. Microbiol. Biotechnol.* **94**, 1423–1447 (2012).
9. Toró, A. & Ambrus, G. Oxidative decarboxylation of 17(20)-dehydro-23,24-dinorcholanoic acids. *Tetrahedron Lett.* **31**, 3475–3476 (1990).
10. Donova, M. V., Gulevskaya, S. A., Dovbnaya, D. V. & Puntus, I. F. *Mycobacterium* sp. mutant strain producing 9 $\alpha$ -hydroxyandrostenedione from sitosterol. *Appl. Microbiol. Biotechnol.* **67**, 671–678 (2005).
11. Imada, Y. & Takahashi, K. US Patent 4,223,091 (1980).
12. Brzostek, A., Dziadek, B., Rumijowska-Galewicz, A., Pawelczyk, J. & Dziadek, J. Cholesterol oxidase is required for virulence of *Mycobacterium tuberculosis*. *FEMS Microbiol. Lett.* **275**, 106–112 (2007).
13. Yang, X. X., Dubnau, E., Smith, I. & Sampson, N. S. *Rv1106c* from *Mycobacterium tuberculosis* is a 3 $\beta$ -hydroxysteroid dehydrogenase. *Biochem.* **46**, 9058–9067 (2007).
14. García, J. L., Uhía, I. & Galán, B. Catabolism and biotechnological applications of cholesterol degrading bacteria. *Microb. Biotechnol.* **5**, 679–699 (2012).
15. van der Geize, R. *et al.* A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages. *Proc. Natl. Acad. Sci.* **104**, 1947–1952 (2007).
16. Petrusma, M., Hessels, G., Dijkhuizen, L. & van der Geize, R. Multiplicity of 3-ketosteroid-9 $\alpha$ -hydroxylase enzymes in *Rhodococcus rhodochrous* DSM43269 for specific degradation of different classes of sterols. *J. Bacteriol.* **193**, 3931–3940 (2011).
17. van der Geize, R. *et al.* Targeted disruption of the *kstD* gene encoding a 3-ketosteroid  $\Delta^1$ -dehydrogenase isoenzyme of *Rhodococcus erythropolis* strain SQ1. *Appl. Environ. Microbiol.* **66**, 2029–2036 (2000).
18. Wilbrink, M. H., Petrusma, M., Dijkhuizen, L. & van der Geize, R. FadD19 of *Rhodococcus rhodochrous* DSM43269, a steroid-coenzyme A ligase essential for degradation of C-24 branched sterol side chains. *Appl. Environ. Microbiol.* **77**, 4455–4464 (2011).
19. Ouellet, H., Johnston, J. B. & de Montellano, P. R. O. Cholesterol catabolism as a therapeutic target in *Mycobacterium tuberculosis*. *Trends Microbiol.* **19**, 530–539 (2011).
20. Capyk, J. K. *et al.* Mycobacterial cytochrome P450 125 (Cyp125) catalyzes the terminal hydroxylation of C27 steroids. *J. Biol. Chem.* **284**, 35534–35542 (2009).
21. Rosloniec, K. Z. *et al.* Cytochrome P450 125 (CYP125) catalyses C26-hydroxylation to initiate sterol side-chain degradation in *Rhodococcus jostii* RHA1. *Mol. Microbiol.* **74**, 1031–1043 (2009).
22. Thomas, S. T. & Sampson, N. S. *Mycobacterium tuberculosis* utilizes a unique heterotetrameric structure for dehydrogenation of the cholesterol side chain. *Biochem.* **52**, 2895–2904 (2013).
23. Mindnich, R., Möller, G. & Adamski, J. The role of 17 beta-hydroxysteroid dehydrogenases. *Mol. Cell. Endocrinol.* **218**, 7–20 (2004).
24. Uhía, I., Galán, B., Kendall, S. L., Stoker, N. G. & García, J. L. Cholesterol metabolism in *Mycobacterium smegmatis*. *Environ. Microbiol.* **4**, 168–182 (2012).
25. van der Geize, R., Hessels, G. I. & Dijkhuizen, L. Molecular and functional characterization of the *kstD2* gene of *Rhodococcus erythropolis* SQ1 encoding a second 3-ketosteroid  $\Delta^1$ -dehydrogenase isoenzyme. *Microbia.* **148**, 3285–3292 (2002).
26. van der Geize, R., Hessels, G. I., Nienhuis-Kuiper, M. & Dijkhuizen, L. Characterization of a second *Rhodococcus erythropolis* SQ1 3-ketosteroid 9 $\alpha$ -hydroxylase activity comprising a terminal oxygenase homologue, KshA2, active with oxygenase-reductase component KshB. *Appl. Environ. Microbiol.* **74**, 7197–7203 (2008).
27. van der Geize, R., Hessels, G. I., van Gerwen, R., van der Meijden, R. & Dijkhuizen, L. Unmarked gene deletion mutagenesis of *kstD*, encoding 3-ketosteroid  $\Delta^1$ -dehydrogenase, in *Rhodococcus erythropolis* SQ1 using *sacB* as counter-selectable marker. *FEMS Microbiol. Lett.* **205**, 197–202 (2001).
28. van der Geize, R., Hessels, G. I., van Gerwen, R., van der Meijden, R. & Dijkhuizen, L. Molecular and functional characterization of *kshA* and *kshB*, encoding two components of 3-ketosteroid 9 $\alpha$ -hydroxylase, a class IA monooxygenase, in *Rhodococcus erythropolis* strain SQ1. *Mol. Microbiol.* **45**, 1007–1018 (2002).
29. Wei, W., Wang, F. Q., Fan, S. Y. & Wei, D. Z. Inactivation and augmentation of the primary 3-ketosteroid- $\Delta^1$ -dehydrogenase in *Mycobacterium neoaurum* NwIB-01: biotransformation of soybean phytosterols to 4-androstene-3,17-dione or 1,4-androstadiene-3,17-dione. *Appl. Environ. Microbiol.* **76**, 4578–4582 (2010).
30. Conner, A. H., Nagaoka, M., Rowe, J. W. & Perlman, D. Microbial conversion of tall oil sterols to C<sub>19</sub> steroids. *Appl. Environ. Microbiol.* **32**, 310–311 (1976).
31. Marsheck, W. J., Kraychy, S. & Muir, R. D. Microbial degradation of sterols. *Appl. Microbiol.* **23**, 72–77 (1972).
32. Szentirmai, A. Microbial physiology of sidechain degradation of sterols. *J. Ind. Microbiol. Biotechnol.* **6**, 101–115 (1990).
33. Wovcha, M. G., Antosz, F. J., Beaton, J. M., Garcia, A. B. & Kominak, L. A. US Patent 4,214,052 (1980).
34. Mohn, W. W. *et al.* The actinobacterial *mce4* locus encodes a steroid transporter. *J. Biol. Chem.* **283**, 35368–35374 (2008).
35. Yang, M. *et al.* Unraveling cholesterol catabolism in *Mycobacterium tuberculosis*: ChsE4-ChsE5  $\alpha$ 2 $\beta$ 2 acyl-CoA dehydrogenase initiates  $\beta$ -oxidation of 3-oxo-cholest-4-en-26-oyl CoA. *ACS Infect. Dis.* **1**, 110–125 (2015).
36. Griffin, J. E. *et al.* High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog.* **7**, e1002251 (2011).
37. Holert, J., Jagmann, N. & Philipp, B. The essential function of genes for a hydratase and an aldehyde dehydrogenase for growth of *Pseudomonas* sp strain Chol1 with the steroid compound cholate indicates an aldolytic reaction step for deacetylation of the side chain. *J. Bacteriol.* **195**, 3371–3380 (2013).
38. Holert, J. *et al.* Degradation of the acyl side chain of the steroid compound cholate in *Pseudomonas* sp. strain Chol1 proceeds via an aldehyde intermediate. *J. Bacteriol.* **195**, 585–595 (2013).
39. Thomas, S. T., VanderVen, B. C., Sherman, D. R., Russell, D. G. & Sampson, N. S. Pathway profiling in *Mycobacterium tuberculosis*: elucidation of cholesterol-derived catabolite and enzymes that catalyze its metabolism. *J. Biol. Chem.* **286**, 43668–43678 (2011).
40. Wilbrink, M. H., van der Geize, R. & Dijkhuizen, L. Molecular characterization of *ltp3* and *ltp4*, essential for C24-branched chain sterol-side-chain degradation in *Rhodococcus rhodochrous* DSM 43269. *Microbiology.* **158**, 3054–3062 (2012).
41. Gordhan, B. G. & Parish, T. in *Mycobacterium tuberculosis protocols from methods in molecular medicine*, Vol. 54. (eds. T. Parish & N. G. Stoker) 77–92 (Humana Press Inc., Totowa, New Jersey; 2001).
42. Stover, C. *et al.* New use of BCG for recombinant vaccines. *Nature.* **351**, 456–460 (1991).
43. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).

44. Egorova, O. V., Nikolayeva, V. M. & Donova, M. V. 17-Hydroxysteroid dehydrogenases of *Mycobacterium* sp. VKM Ac-1815D mutant strain. *J. Steroid Biochem. Mol. Biol.* **81**, 273–279 (2002).
45. Leenders, F. *et al.* Porcine 80-kDa protein reveals intrinsic 17-hydroxysteroid dehydrogenase, fatty acyl-CoA-hydratase/dehydrogenase, and sterol transfer activities. *J. Biol. Chem.* **271**, 5438–5442 (1996).
46. Saito, T., Fukui, T., Ikeda, F., Tanaka, Y. & Tomita, K. An NADP-linked acetoacetyl CoA reductase from *Zoogloea ramigera*. *Arch. Microbiol.* **114**, 211–217 (1977).
47. Gao, X., Feng, J., Hua, Q., Wei, D. & Wang, X. Investigation of factors affecting biotransformation of phytosterols to 9-hydroxyandrost-4-ene-3,-17-dione based on the HP- $\beta$ -CD-resting cells reaction system. *Biocatal. Biotransform.* **32**, 343–347 (2014).

## Acknowledgements

We thank T. Parish (Department of Infectious and Tropical Diseases, United Kingdom) for providing plasmids, p2NIL and pGOAL19, and W. R. Jacobs, Jr. (Howard Hughes Medical Institute) for providing plasmid pMV261. This research was financially supported by the National Natural Science Foundation of China (No. 31370080), the National Special Fund for State Key Laboratory of Bioreactor Engineering, the National Basic Research Program of China (973, Program No. 2012CB721003) and the National Major Science and Technology Projects of China (No. 2012ZX09304009).

## Author Contributions

L.Q.X. and F.Q.W. designed experiments. L.Q.X., Y.J.L., K.Y. and H.H.L. performed the experiments. L.Q.X. and F.Q.W. analyzed the data. X.Y.T., F.Q.W. and D.Z.W. contributed reagents and materials. L.Q.X. and F.Q.W. wrote the manuscript. All authors have read and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Xu, L.-Q. *et al.* Unraveling and engineering the production of 23,24-bisnorcholesterol steroids in sterol metabolism. *Sci. Rep.* **6**, 21928; doi: 10.1038/srep21928 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>