Unravelling a new catabolic pathway of C-19 steroids in Mycobacterium smegmatis

Fernández-Cabezón, Lorena; Galán, Beatriz; García, José L.

Published in:
Environmental Microbiology

Link to article, DOI:
10.1111/1462-2920.14114

Publication date:
2018

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Unravelling a new catabolic pathway of C-19 steroids in
*Mycobacterium smegmatis*

Lorena Fernández-Cabezón†, Beatriz Galán¹ and José L. García†

¹Department of Environmental Biology. Centro de Investigaciones Biológicas. Consejo
Superior de Investigaciones Científicas. Ramiro de Maeztu 9, 28040 Madrid, Spain

*For correspondence: E-mail jlgarcia@cib.csic.es; Tel. +34 918373112; Fax +34 915360432

†Present address: The Novo Nordisk Foundation Center for Biosustainability, Technical
University of Denmark, Kemitorvet 220, 2800 Lyngby, Denmark
Summary

In this work, we have characterized the C-19+ gene cluster (MSMEG_2851 to MSMEG_2901) of Mycobacterium smegmatis. By in silico analysis we have identified the genes encoding enzymes involved in the modification of the A/B steroid rings during the catabolism of C-19 steroids in certain M. smegmatis mutants mapped in the PadR-like regulator (MSMEG_2868), that constitutively express the C-19+ gene cluster. By using gene complementation assays, resting-cell biotransformations and deletion mutants, we have characterized the most critical genes of the cluster, i.e, kstD2, kstD3, kshA2, kshB2, hsaA2, hsaC2 and hsaD2. These results have allowed us to propose a new catabolic route named C-19+ pathway for the mineralization of C-19 steroids in M. smegmatis. Our data suggest that the deletion of the C-19+ gene cluster may be useful to engineer more robust and efficient M. smegmatis strains to produce C-19 steroids from sterols. Moreover, the new KshA2, KshB2, KstD2 and KstD3 isoenzymes may be useful to design new microbial cell factories for the 9α-hydroxylation and/or Δ1-dehydrogenation of 3-ketosteroids.
Introduction

Steroids are naturally occurring hydrophobic molecules that present a structure core formed by four fused alicyclic rings named gonane. Steroid compounds are frequently found in the biosphere and can be used as carbon and energy sources by different bacteria (e.g., Mycobacterium, Pseudomonas, Sterolibacterium, Sphingomonas, Novosphingobium, Comamonas) (Fujii et al., 2002, 2003; Tarlera and Denner, 2003; Horinouchi et al., 2003a, 2004a; Philipp et al., 2006; van der Geize et al., 2007; Roh and Chu, 2010; Leu et al., 2011; Merino et al., 2013). Most bacteria capable of degrading aerobically sterols (e.g., cholesterol and phytosterols), which are one of the most abundant steroid compounds in nature, belong to the phylum Actinobacteria (e.g., Mycobacterium, Rhodococcus, Gordonia) (van der Geize et al., 2007; Kendall et al., 2007, 2010; Drzyzga et al., 2009; Uhía et al., 2012; Fernández de las Heras et al., 2013; Li et al., 2014; Bergstrand et al., 2016). The bacterial catabolism of cholesterol has been investigated in detail because of its relevance in the pathogenicity of certain bacteria such as Mycobacterium tuberculosis or Rhodococcus equi. These bacteria are able to metabolize cholesterol located in the membranes of the host cells during the course of infection, being the degradation of this sterol crucial for bacterial persistence (van der Geize et al., 2007, 2011; Pandey and Sassetti, 2008). The investigation of cholesterol catabolism has also attracted increasing interest because low-cost natural sterols (e.g., phytosterols) are currently biotransformed at an industrial scale into valuable pharmaceutical steroidal intermediates such as the C-19 steroids 4-androstene-3,17-dione (AD), 1,4-androstadiene-3,17-dione (ADD) or 9α-hydroxy-4-androstene-3,17-dione (9OH-AD) (García et al., 2012; Donova, 2017).

The aerobic 9,10-seco degradation for cholesterol pathway has been postulated based on the compilation of multiple genetic and biochemical investigations carried out on different Gram-positive bacteria, but this pathway is also partially shared by some Gram-negative
bacteria that degrade bile acids and testosterone (Fig. S1) (Barrientos et al., 2015; García et al., 2012; Galán et al., 2017a). Briefly, in Actinobacteria, the cholesterol is transported by a specific ATP-dependent transport system named Mce4 (Casali and Riley 2007; Pandey and Sassetti, 2008; Mohn et al., 2008; Klepp et al., 2012; García-Fernández et al., 2017a) and is subsequently transformed into cholest-4-en-3-one (cholestenone) by 3β-hydroxysteroid dehydrogenase/isomerases enzymes (3β-HSD) and/or cholesterol oxidases (ChOx) (Kreit, 2017). Concomitantly with the formation of cholestenone, at least two P450 cytochromes, named CYP125 and CYP142, initiate the side-chain degradation by performing sequential oxidations of the methyl group at C-26 to generate a carboxylic acid (Capyk et al., 2009a; McLean et al., 2009; Rosłoniec et al., 2009; Driscoll et al., 2010; Johnston et al., 2010; Ouellet et al., 2010; García-Fernández et al., 2013). This modification enables the β-oxidation of the side-chain that yields two molecules of propionyl-CoA, an acetyl-CoA and the 17-ketosteroid AD (Fig. S1). Thereafter, the degradation of the A/B steroid rings of the intermediate AD occurs via a central catabolic route called 9,10-seco pathway that results in the generation of 2-hydroxyhexa-2,4-dienoic acid (HHD) and 9,17-dioxo-1,2,3,4,10,19-hexanorandrostan-5-oic acid (HIP) (Fig. S1) (García et al., 2012; Galán et al., 2017a). The degradation of HDD probably occurs through the action of enzymes similar to the TesE, TesF and TesG proteins described in C. testosteroni (Horinouchi et al., 2005), while the degradation of HIP (i.e., the mineralization of C/D steroid rings) takes place in a lower catabolic pathway, recently studied in detail in M. tuberculosis and other Actinobacteria, which leads to the formation of metabolites that enter the central metabolic pathways (Crowe et al., 2017). The expression of the genes responsible for the catabolism of C/D rings is controlled by the KstR2 repressor, a TetR-like transcriptional regulator (Kendall et al., 2010; Casabon et al., 2013; Crowe et al., 2015; García-Fernández et al., 2015), while the expression of the genes encoding the enzymes involved in the side-chain and the A/B rings degradation
as well as those encoding the sterol uptake system, is regulated by another TetR-like transcriptional regulator, the KstR repressor (Kendall et al., 2007; 2010; Uhía et al., 2011).

Conventionally, it has been proposed that the C-19 steroids AD, ADD and 9OH-AD are key intermediates of the (chole)sterol catabolic pathway due to multiple lines of evidence (Fig. S1). Firstly, the C-19 compounds have been found in the culture medium of several naturally sterols-degrading bacteria (Marsheck et al., 1972; Wei et al., 2010). Recently, various mutant strains producing C-19 steroids from sterols have also been constructed through metabolic engineering approaches by redirecting the metabolic flux of sterol catabolism to the accumulation of C-19 compounds by gene deletions (Wilbrink et al., 2011; Yeh et al., 2014; Yao et al., 2014; Galán et al., 2017b). On the other hand, several examples of bacteria capable of using both cholesterol and C-19 steroids as the only carbon and energy source have been described (e.g., Rhodococcus erythropolis strain SQ1, Rhodococcus ruber strain Chol-4, Rhodococcus rhodochrous DSM 43269, Gordonia neofellicaeis NRRL B-59395 (van der Geize et al., 2000; Petrusma et al., 2011; Fernández de las Heras et al., 2012; Li et al., 2014). However, these bacterial strains have other sets of steroid-degrading enzymes, different from those participating in the cholesterol catabolism, that are specifically induced by different C-19 steroids. Therefore, in these strains, the different steroid catabolic pathways could be partially overlapped creating a complex network of metabolic intermediates that are degraded by a battery of redundant enzymes that have possibly led to some misinterpretations of the pathways for sterol catabolism.

In a previous manuscript, we proposed that the C-19 steroids AD, ADD and 9OH-AD are not the natural intermediates of the (chole)sterol catabolic pathway based on several lines of evidence (Fernández-Cabezón et al., 2017). Firstly, different authors showed that certain modifications introduced into the A/B rings during the catabolism of cholesterol (e.g., 9α-hydroxylation and Δ¹-dehydrogenation) might occur simultaneously with the side-chain
degradation. The enzymes involved in these catabolic steps appear to use as preferred substrates the acyl-CoA derivatives resulting from the β-oxidation of the side-chain (i.e., most probably compounds that still retain 3 or more carbons of the sterol side-chain) (Capyk et al., 2011; Penfield et al., 2014). Secondly, by inhibiting the B-ring opening during the degradation of cholesterol, it is possible to detect in the culture medium, in addition to the C-19 steroids, other steroid intermediates that still retain part of the side-chain like 3-oxo-23,24-bisnorchol-1,4-dien-22-oic acid (1,4-BNC), 3-oxo-23,24-bisnorchol-4-en-22-oic acid (4-BNC), 22-hydroxy-23,24-bisnorchol-1,4-dien-3-one (1,4-HBC), or 22-hydroxy-23,24-bisnorchol-4-en-3-one (4-HBC) (Szentirmai, 1990; Donova et al., 2005; Wilbrink et al., 2011; Yeh et al., 2014; Galán et al., 2017b). Thirdly, we demonstrated that the C-19 steroids AD, ADD and 9OH-AD cannot be efficiently mineralized through the cholesterol pathway in M. smegmatis mc2155, since neither the wild-type strain nor its unregulated mutant ΔkstR are able to degrade efficiently these compounds (Fernández-Cabezón et al., 2017). All these results reinforce the hypothesis that certain modifications introduced into the A/B rings of cholesterol occur simultaneously with the side-chain degradation and therefore, the C-19 steroids AD, ADD or 9OH-AD are actually side products of the cholesterol catabolic route (Fig. S1).

On the other hand, we also reported the existence of a new silent and tightly-regulated catabolic pathway in M. smegmatis, that is not induced by sterols (e.g., cholesterol, phystosterols) nor by C-19 steroids (e.g., AD, ADD, 9OH-AD), but which, after its desilencing/activation, confers to this bacterium the capacity to mineralize these and other C-19 steroids (Fernández-Cabezón et al., 2017). The activation of this catabolic pathway is found in some M. smegmatis spontaneous mutants mapped in the PadR-like regulator (MSMEG_2868) that acquire the ability to efficiently metabolize C-19 steroids. The PadR inactivation allows the expression of a set of genes named C-19+ cluster (MSMEG_2851 to
that encode putative steroid-degrading enzymes. The C-19+ cluster would have evolved independently from the upper cholesterol kstR-regulon, since C-19+ cluster has an independent regulation, but both C-19+ cluster and kstR-regulon would converge on the lower cholesterol kstR2-regulon responsible for the metabolism of C/D steroid rings (Fig. S1). However, this is only a working hypothesis that requires further confirmation. Other steroid-degrading Actinobacteria have homologous C-19+ clusters (e.g., Mycobacterium neoaurum ATCC 25795, Rhodococcus jostii RHA1, Gordonia neofelicaeis NRRL-B59395, Nocardiooides simplex VKM Ac-2033D), but it is interesting to note that it is absent in M. tuberculosis and in certain industrial mycobacterial strains used to produce C-19 steroids from sterols (e.g., Mycobacterium sp. NRRL B-3805) (Fernández-Cabezón et al., 2017).

In this work, we have characterized in more detail the C-19+ gene cluster in M. smegmatis. Using different methodologies, we have investigated the biological function and/or the essentiality of those genes of the C-19+ cluster that encode the enzymes involved in the modification of A/B rings during the catabolism of C-19 steroids in this bacterium.

Results

Identification of the key genes involved in the degradation of C-19 steroids in M. smegmatis

Based on the knowledge of the 9,10-seco pathway for the bacterial catabolism of steroids (García et al., 2012; Galán et al., 2017a), we have proposed a catabolic route named C-19+ pathway for the catabolism of C-19 steroids in M. smegmatis (Fig. 1A) (Fernández-Cabezón et al., 2017). By in silico analysis, we identified in the C-19+ gene cluster homologous genes for almost all the postulated enzymatic steps involved in the A/B steroid ring modification (Fig. 1B). The members of the kstR regulon involved in the cholesterol catabolism in M. smegmatis have been used as reference genes (Kendall et al., 2007; Uhía et al., 2012). As it is shown in the Fig. 1B, despite the existence of homologous genes in the two steroid catabolic gene clusters found in M. smegmatis they do not share any synteny.
The genes MSMEG_2867 (kstD2) and MSMEG_2869 (kstD3) encode putative 3-ketosteroid-Δ1-dehydrogenases sharing a 36% and 44% protein identity to KstD1 (MSMEG_5941), respectively. The MSMEG_2893 (kshB2) gene encodes a putative reductase component of a 3-ketosteroid-9α-hydroxylase showing 55% protein identity to KshB1 (MSMEG_6039), whereas the MSMEG_2870 (kshA2) gene would encode the oxygenase component of the same enzyme since it has a 59% protein identity to KshA1 (MSMEG_5925). The MSMEG_2892 (hsaA2) and MSMEG_2891 (hsaC2) genes may encode a 3-HSA monooxygenase (oxygenase component) and a 3,4-HSA dioxygenase, that share between 40-50% protein identity to HsaA1 (MSMEG_6038) and HsaC1 (MSMEG_6036), respectively. Finally, the MSMEG_2900 (hsaD2) gene encodes a putative 4-9-DSHA hydrolase with 41% protein identity to HsaD1 (MSMEG_6037). There are not homologous genes to the hsaB1 (MSMEG_6035), hsaE (MSMEG_5940), hsaF (MSMEG_5939), and hsaG (MSMEG_5937) genes within the C19+ cluster.

**Determination of the biological function of the kstD2, kstD3, kshA2 and kshB2 genes**

To investigate the role of the MSMEG_2893 (kshB2) gene, the mutant *M. smegmatis* MS6039 (ΔMSMEG_6039, ΔkshB1) was complemented in trans with an original copy of the *kshB2* gene. Galán et al. (2017b) previously demonstrated that the MS6039 mutant is unable to efficiently grow in sterols as the sole carbon and energy source, unlike the wild-type strain of *M. smegmatis*, since the mutant lacks the KshB1 reductase activity necessary to mineralize cholesterol. Analogously, to study the biological function of the MSMEG_2867 (kstD3) and MSMEG_2869 (kstD2) genes, the mutant *M. smegmatis* MS5941 (ΔMSMEG_5941, ΔkstD1) was constructed and complemented in trans with an original copy of the *kstD2* or *kstD3* genes (Table S2). The KstD1 activity required to mineralize cholesterol is eliminated in the MS5941 mutant, and thus, this mutant is unable to efficiently grow in sterols as the sole carbon and energy (Fig. 2A). However, when the MS5941 mutant is cultured in minimal media
containing glycerol as carbon and energy source and cholesterol (or phytosterols) as feedstock, a large accumulation of 9OH-AD is observed (data not shown). According to our hypothesis, the expression of a kshB gene (e.g., kshB1 or kshB2) or a kstD gene (e.g., kstD1, kstD2 or kstD3) in the MS6039 or MS5941 mutants, respectively, will restore its capability to grow in cholesterol. For this purpose, the MSMEG_2867 (kstD3), MSMEG_2869 (kstD2) and MSMEG_2893 (kshB2) genes were cloned into the pMV261 plasmid generating the corresponding recombinant plasmids pMV2867 (harbouring kstD3), pMV2869 (harbouring kstD2) and pMV2893 (harbouring kshB2). The recombinant plasmids pMV5941 (harbouring kstD1) and pMV6039 (harbouring kshB1) were also constructed and used as controls for the complementation experiments. The plasmids harboring the kshB genes were transformed in the MS6039 mutant (ΔkshB1), whereas the plasmids harbouring the kstD genes were transformed into the MS5941 mutant (ΔkstD1).

The overexpression of the kstD2 gene in the MS5941 strain restored its capability of growing in cholesterol as the sole carbon and energy source, exhibiting a similar growth pattern to the wild-type strain mc^2155 and to the MS5941 (pMV5941) that is complemented with the kstD1 gene (Fig. 2A). Otherwise, the overexpression of the kstD3 gene in the MS5941 (pMV2867) recombinant did not restore the ability of MS5941 of growing in cholesterol, suggesting that KstD3 present another enzymatic activity or a different substrate specificity. To investigate this hypothesis, we monitored the bioconversion of several steroid substrates by resting-cell biotransformation using the recombinant MS5941 (pMV2867). The Figure 3A shows that this recombinant was able to transform AD into ADD, whereas the control strain MS5941 (pMV261) harboring the empty plasmid was not able to, confirming that KstD3 present 3-ketosteroid-Δ1-dehydrogenase activity. On the other hand, the overexpression of the kshB2 gene in the MS6039 strain restored its capability of growing in 1.8 mM cholesterol as the sole carbon and energy source (Fig. 2B; Table S2)).
this recombinant strain MS6039 (pMV2893) was similar to the wild-type strain mc²155 and to the MS6039 (pMV6039) that is complemented with the kshB1 gene. Finally, to investigate the activity of MSMEG_2870 (kshA2) gene, we constructed the plasmid pMV2870 (harbouring kshA2) and followed the same resting cell approach. As expected, the recombinant strain MS5941 (pMV2870) that produces constitutively the KshA2 protein, was able to transform AD into 9OH-AD (Fig. 3B; Table S2).

These results confirmed that the kshA2 and kshB genes encode the oxygenase subunit (KshA) and the ferredoxin reductase subunit (KshB) of a 3-ketosteroid-9α-hydroxylase enzyme, while the kstD2 and kstD3 genes encode 3-ketosteroid-Δ1-dehydrogenase proteins.

**Determination of the biological function of the hsaA2, hsaC2 and hsaD2 genes**

To investigate the biological role of the hsaA2 (MSMEG_2892), hsaC2 (MSMEG_2891) and hsaD2 (MSMEG_2900) genes in the catabolism of C-19 steroids in *M. smegmatis*, we constructed by double homologous recombination three mutant strains in *M. smegmatis* ADD+ named MS2891 (ΔhsaC2), MS2892 (ΔhsaA2) and MS2900 (ΔhsaD2) (Table 1). The three engineered mutants were unable to efficiently grow in C-19 steroids (e.g., 1.8 mM AD) as the sole carbon and energy source when compared with the parental strain ADD+ (Fig. 4A), although they perfectly grew using glycerol as a carbon source (data not shown). Interestingly, a color alteration of the C-19 steroid-containing medium was observed during the culture of the MS2891 and MS2900 mutants (Fig. 4B). A pink coloration of the medium was developed in the culture of the MS2891 mutant (ΔhsaC2), possibly indicating the accumulation of a catecholic derivate (e.g., 3,4-HSA) and its subsequent non-enzymatic oxidation to quinone derivatives. This phenotype was previously observed in ΔhsaC mutants of the cholesterol catabolism in *R. jostii* RHA1 and *M. tuberculosis* H37Rv (van der Geize *et al.*, 2007; Yam *et al.*, 2009), as well as in the ΔtesB mutant of *C. testosteroni* TA441 in the
presence of ADD or cholic acid (Horinouchi et al., 2004b). On the other hand, a yellowish-brown coloration in the medium was observed in the culture of the MS2900 mutant (ΔhsaD2) likely due to the accumulation of 4,9-HSA (Gibson et al., 1966). This characteristic color was also described in the mutant ΔtesD of *C. testosteroni* TA441 in the presence of testosterone (Horinouchi et al., 2003b).

To confirm the absence of polar effects, the three mutant strains were complemented *in trans* with an original copy of the corresponding deleted gene (*MSMEG_2891*, *MSMEG_2892* or *MSMEG_2900*). For that, each gene was cloned into the pMV261 plasmid generating the recombinant plasmids pMV2891, pMV2892 and pMV2900 and then, they were transformed into the mutant strains MS2891, MS2892 and MS2900, respectively (Table 1). Fig. 5B shows that the complemented MS2891 and MS2900 mutants (MS2891c and MS2900c, respectively) recovered the capacity to grow in C-19 steroids as the sole carbon and energy source (e.g., 1.8 mM AD) (Table S2). In contrast, the complemented MS2892 strain (MS2892c) was not able to grow in 1.8 mM AD as the sole carbon and energy source, although it showed a normal growth in other substrates (e.g., 18 mM glycerol) (Fig. 5A). However, the same pink coloration in the AD-containing medium described above with the MS2891 mutant was also observed in the culture broths of MS2892c strain (Fig. 5C). This result supports the hypothesis that the *MSMEG_2892* gene encodes a 3-HSA monooxygenase, which under overexpression conditions, could generate high amounts of 3,4-HSA to inhibit bacterial growth. The high toxicity of these compounds to the cells has been previously described (Yam et al., 2009).

Our results confirmed that the *hsaA2*, *hsaC2* and *hsaD2* genes play an essential role in the catabolism of C-19 steroids in *M. smegmatis*. In addition, considering that the engineered mutants were unable to efficiently grow in C-19 steroids, we can conclude that the HsaA2, HsaC2 and HsaD2 activities cannot be replaced in these conditions by other enzymatic
activities. Notice that their corresponding homologous proteins involved in cholesterol catabolism are not expressed in the tested conditions since the cholesterol pathway is not induced by C-19 compounds.

**Discussion**

In this work, we have characterized the C-19+ gene cluster recently identified in *M. smegmatis* (Fernández-Cabezón *et al.*, 2017). First, we have identified by *in silico* analysis several genes encoding enzymes putatively involved in the modification of the A/B steroid rings for the catabolism of C-19 steroids in this bacterium. Then, we have used different genetic and biochemical approaches to determine the biological function of these genes that has allowed us to postulate the C-19+ pathway in *M. smegmatis* (Fig. 1A).

According to the classical 9,10-seco pathway, the transformation of AD into ADD is the first enzymatic step necessary for the opening of A/B rings for the aerobic catabolism of steroids. This reaction is catalyzed by 3-ketosteroid-Δ1-dehydrogenases named KstD in *M. smegmatis* (Brzostek *et al.*, 2005), *R. erythropolis* (van der Geize *et al.*, 2000, 2001, 2002) and *M. tuberculosis* (Brzostek *et al.*, 2009), but named TesH in *C. testosteroni* (Horinouchi *et al.*, 2003a). The 3-ketosteroid-Δ1-dehydrogenase are FAD-dependent proteins that catalyze the trans-axial removal of the hydrogen atoms C-1(α) and C-2(β) of 3-ketosteroids (EC 1.33.99.4) (Knol *et al.*, 2008). We have demonstrated that the KstD2 protein is able to replace the biological role of KstD1 in the catabolism of cholesterol in *M. smegmatis* MS5941 (ΔMSMEG_5941, ΔkstD1) (at least, under the tested conditions of kstD2 gene overexpression), even though KstD2 is not involved in the catabolism of sterols in this bacterium (Fig. 2A). This result agrees with the previous findings of Brzostek *et al.* (2005). In contrast, the overexpression of the kstD3 gene (MSMEG_2867) in the MS5941 mutant is not able to restore this capability (Fig. 2A), although the KstD3 protein has 3-ketosteroid-Δ1-dehydrogenase activity in resting-cell biotransformations (Fig. 3A). This result suggests that
KstD3 is not able to efficiently recognize the intermediates with side-chain that are possibly the true substrates of KstD1 during the cholesterol catabolism (Fernández-Cabezón et al., 2017).

Then, the 9α-hydroxylation step takes place by the action of the 3-ketosteroid-9α-hydroxylases (Andor et al., 2006; van der Geize et al., 2007; Capyk et al., 2009b). These enzymes are two-component monooxygenases consisting of a terminal oxygenase subunit (KshA) and a ferredoxin reductase subunit (KshB) (E.C. 1.14.13.142) (Petrusma et al., 2014). The KshA subunit specifically recognizes the steroid substrates, whereas the KshB component transfers the reducing power required for the progression of the enzymatic reaction. Petrusma et al. (2011) described the presence of 5 kshA genes involved in the catabolism of different classes of steroids (e.g., cholesterol, cholate, AD and progesterone) in R. rhodochrous DSM43269, each displaying unique steroid induction patterns and substrate ranges. They demonstrated that the overexpression of only certain kshA genes in a null-kshA mutant restores its growth on any of the different classes of steroids tested. The KshA/KshB interaction does not seem to be very specific, since it is possible to find more kshA genes than kshB genes in the genomes of actinobacterial species (e.g. R. jostii RHA1, N. simplex VKM Ac-2033D or G. neofelifaecis NRRL B-59395) (Petrusma et al., 2011). This observation suggests that a particular KshB protein can interact with several KshAs or even that, other non-specific reductases present in the bacterial cell can act as donors of reducing power. In agreement with this observation, we have demonstrated that the KshB2 of M. smegmatis is able to replace KshB1 in the catabolism of cholesterol (Fig. 2B). In the same line, Yuan et al. (2015) were able to reconstitute a functional KSH in Mycobacterium sp. NRRL B-3805 by overexpressing the kshA gene of M. smegmatis and the kshB gene of G. neofelifaecis NRRL B-59395. Moreover, we have observed that the co-expression of a kshB gene does not seem to
be always essential for the 9α-hydroxylations of 3-ketosteroids (Fig. 3B), as previously
demonstrated by other authors (Andor et al., 2005; Arnell et al., 2007).

The 3-HSA intermediate obtained after the action of the 3-ketosteroid-Δ¹-

dehydrogenase and the 3-ketosteroid-9α-hydroxylase enzymes, is hydroxylated by a two-

component oxygenase (HsaA2B1), similar to TesA1A2 from C. testosteroni (Horinouchi et al.,

2004b) or to HsaAB from M. tuberculosis and R. jostii RHA1 (Dresen et al., 2010)

leading to the production of 3,4-DHSA. This catechol derivative is subsequently cleaved by a

meta-extradial dioxygenase (HsaC2), similar to TesB in C. testosteroni (Horinouchi et al.,

2001) or to HsaC in R. jostii RHA1 (van der Geize et al., 2007) and M. tuberculosis H37Rv

(Yam et al., 2009), that yields 4,9-DSHA. This compound is subsequently hydrolyzed by a

meta-cleavage product (MCP) hydrolase (HsaD2) similar to TesD in C. testosteroni

(Horinouchi et al., 2003b) or to HsaD in M. tuberculosis and R. jostii RHA1 (van der Geize et

al., 2007; Lack et al., 2008, 2010) yielding HIP and HHD. To investigate the biological role

of the HsaA2, HsaC2 and HsaD2 proteins in the catabolism of C-19 steroids, in this work we

have constructed by double homologous recombination the mutants ADD+ MS2891

(ΔMSMEG_2891; ΔhsaC2), ADD+ MS2892 (ΔMSMEG_2892; ΔhsaA2) and ADD+ MS2900

(ΔMSMEG_2900; ΔhsaD2). The three mutants are unable to efficiently grow in C-19 steroids

as the sole carbon and energy source, confirming that the hsaA2, hsaC2 and hsaD2 genes are

essential for the catabolism of C-19 steroids in M. smegmatis (Fig. 4A). Moreover, the color

changes of culture media observed when the mutant strains are cultured in the presence of C-

19 compounds strongly support the C-19+ pathway postulated above.

Finally, the degradation of HDD should occur through the action of the enzymes

HsaE, HsaF and HsaG encoded by the unique hsaEFG operon of M. smegmatis (Fernández-

Cabezón et al., 2017). These genes are not always present in all steroid-degrading gene

clusters in Actinobacteria, because they encode enzymes that are common in many catabolic
pathways of aromatic compounds and thus, these enzymes can replace those (Bergstrand et al., 2016).

Summarizing, the results presented above confirm that the C-19 steroids AD, ADD and 9OH-AD are mineralized through the C-19+ catabolic pathway that is constitutively expressed in the padR mutants of M. smegmatis. Although we have identified some of the key enzymatic steps of the C-19+ pathway, additional studies are required to determine the roles of other genes found in the C-19+ gene cluster of M. smegmatis (MSMEG_2851 to MSMEG_2901), as well as to identify the chemical compound that induces this cluster in vivo. A better understanding of the complexity of steroid catabolism in M. smegmatis and other actinobacterial species can also be fundamental to optimize the current industrial bioproduction of steroidal intermediates from sterols. The deletion of the C-19+ gene cluster or at least of the most relevant genes located therein (e.g., kstD, kshA, kshB) will be useful to increase the bioconversion yields in producer strains derived from M. smegmatis and other cholesterol degrading bacteria. Yao et al. (2014) demonstrated that two residual enzymatic activities KstD (KstD2 and KstD3) present in the mutant Mut(MN-KSTD1) (∆kstD1) of M. neoaurum ATCC 25795, markedly decreases the conversion yield of 9OH-AD from sterols. Interestingly, KstD2 and KstD3 are encoded by genes located in a homologous C-19+ cluster that is not found in the close phylogenetic strains M. neoaurum VKM Ac-1815D and Mycobacterium sp. NRRL B-3805, both used in industrial bioprocesses. On the other hand, given the industrial relevance of the 9α-hydroxylation and Δ1-dehydrogenation reactions of 3-ketosteroids, the new KshA2, KshB2, KstD2 and KstD3 enzymes identified could be used to design new microbial cell factories for the production of steroids à la carte.
Experimental procedures

Chemicals

4-Androstene-3,17-dione (AD) and 1,4-androstadien-3,17-dione (ADD) were purchased from TCI America. Cholesterol, Tween 80, tyloxapol, gentamicin and kanamycin were from Sigma (Steinheim, Germany). Glycerol and glucose were purchased from Merck (Darmstadt, Germany). 9α-Hydroxy-4-androstene-3,17-dione (9OH-AD) was kindly provided by Gadea Biopharma (León, Spain).

Bacterial strains, plasmids and culture conditions

The plasmids and bacterial strains used in this work are listed in Table 1. M. smegmatis strains were grown at 37°C in an orbital shaker at 200 r.p.m. Middlebrook 7H9 broth medium (Difco) supplemented with 0.4% glycerol and 0.05% Tween 80 was used as rich medium. 7H9 broth without any supplement was used as minimal medium. 7H10 agar (Difco) plates supplemented with 10% albumin-dextrose-catalase (Becton Dickinson) were used for solid media. Kanamycin (20 µg ml\(^{-1}\)) or gentamicin (5 µg ml\(^{-1}\)), were used for strain selection when appropriate. All the steroids were dissolved in 10% tyloxapol prior to its addition to the minimal medium. Because of the low solubility of the steroid compounds, the stock solutions were warmed at 80 ºC in agitation, sonicated in a bath for 1 h and then autoclaved.

Escherichia coli DH10B strain was used as a host for cloning. It was grown in LB medium at 37°C in an orbital shaker at 200 r.p.m. LB agar plates were used for solid media. Kanamycin (50 µg ml\(^{-1}\)) or Gentamicin (10 µg ml\(^{-1}\)), were used for plasmid selection and maintenance.
DNA manipulations and sequencing

DNA manipulations and other molecular biology techniques were essentially as described by Sambrook and Russell (2001). Oligonucleotides used at this work were purchased from Sigma-Aldrich and are listed at Table S1. Genomic DNA purification from mycobacterial strains was performed as previously described (Uhía et al., 2011). DNA amplification was performed on a Mastercycler Gradient (Eppendorf) using Phusion High-Fidelity DNA Polymerase from ThermoFisher Scientific. DNA fragments were purified with High Pure PCR System Product Purification Kit (Roche). Restriction enzymes were obtained from various suppliers and were used according to their specifications. Plasmid DNA was prepared with a High Pure Plasmid Isolation Kit (Roche Applied Science). *E. coli* was transformed by the rubidium chloride method (Wirth et al., 1989). *M. smegmatis* cells were transformed by electroporation (Gene Pulser; Bio-Rad) (Parish and Stoker, 1998). All cloned inserts and DNA fragments were confirmed by DNA sequencing through an ABI Prism 377 automated DNA sequencer (Applied Biosystems Inc.) at Secugen S.L. (Madrid, Spain).

Gene deletion

The knock-out strains were constructed by double homologous recombination using the plasmid pJQ200x, a derivative of the suicide vector pJQ200 that does not replicate in *Mycobacterium* (Jackson et al., 2001). The followed approach consisted in amplifying by PCR two fragments of ~700 bp containing the upstream region (UP fragment) and the downstream region (DOWN fragment) of the target gene to be deleted (Table S1). The two PCR fragments were purified, digested with the corresponding enzymes and ligated. Using the ligation product as template, the UP+DOWN fragment was amplified by PCR using the most external primers. The resulting amplicon was cloned into the plasmid pJQ200x using *E. coli* DH10B competent cells. Then, the resulting pJQ200x-derivate plasmid was
electroporated into competent *M. smegmatis* ADD+ (Table 1). The single cross-overs were selected on 7H10 agar plates containing gentamicin. The obtained single colonies were contra-selected in 10% sucrose and in addition, the presence of the *xylE* gene marker was confirmed by the appearance of yellow coloration after spreading catechol over each colony replica. To select for double cross-overs, a single colony was grown in rich medium up to an optical density of 0.8–0.9 and several dilutions of this culture were plated onto 7H10 agar plates with 10% sucrose. Potential double cross-overs (i.e., sucrose-resistant colonies) were subsequently screened for gentamicin sensitivity and the gene deletion was confirmed by PCR.

**Complementation of the knock-out mutants**

To isolate the genes (e.g., *MSMEG_2891*) from *M. smegmatis* mc²155 for the complementation experiments, its genomic DNA was extracted and amplified by PCR using the corresponding primers forward and reverse (e.g., *MSMEG_2891F* and *MSMEG_2891R*) (Table S1). The resulting amplicons were purified and digested with the corresponding restriction enzymes to be cloned into the expression vector pMV261 (Stover et al., 1991), a shuttle plasmid that replicates in *E. coli* and *Mycobacterium*. The resulting plasmids (e.g., pMV2891) were transformed into *E. coli* DH10B for cloning purposes and once their sequences were verified, they were used to transform electrocompetent cells of the corresponding *M. smegmatis* mutant strain (e.g., MS2891). The mutant strains harboring the plasmid pMV261 were used as controls.

**Resting-cell biotransformations**

The recombinant strains were grown in rich medium at 37 °C during 24 h. The cells were harvested by centrifugation at 5000 x g for 20 min at 4 °C and washed once with 0.85% NaCl.
The biotransformation was carried out with an optical density (OD$_{600}$) of 5 in a 100 mL shake flask containing 40 mL of reaction mixture: 0.1 M phosphate buffer (pH 8.0), 2 mM AD (substrate) and 0.05% Tween 80. An additional carbon source (1% glucose) was added in the 9alpha-hydroxylation biotransformations. The substrate AD was incorporated into the medium as a solution with randomly methylated β-cyclodextrin (1:10.3, molar ratio) (Klein et al., 1995) or alternatively, was dissolved in 10% tyloxapol as mentioned above. The steroid bioconversion was monitored by thin layer chromatography (TLC) (García-Fernández et al., 2017b).

**Bioinformatic analysis**

Sequence alignments were carried out using Clustal W (Thompson et al., 1994) and different BLAST algorithms from the National Centre of Biotechnology Information Server (NCBI) were also used.

**Acknowledgements**

The technical work of A. Valencia is greatly appreciated. This work was supported by grants from the Ministry of Science and Innovation (BFU2006-15214-C03-01, BFU2009-11545-C03-03) and Ministry of Economy and Competitiveness (BIO2012-39695-C02-01). LFC was supported by an FPU fellowship from the Spanish Ministry of Education, Culture and Sports.

**Conflict of interest**

None declare
References


### Table 1. Bacterial strains and plasmids used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype and/or description</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Mycobacterium smegmatis</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mc²155</td>
<td>ept-1, mc²6 mutant efficient for electroporation</td>
<td>Snapper <em>et al.</em> (1990)</td>
</tr>
<tr>
<td>MS6039</td>
<td>mc²155 ΔMSMEG_6039 (ΔkshB1); high-ADD producer</td>
<td>Galán <em>et al.</em> (2017)</td>
</tr>
<tr>
<td>ADD+</td>
<td>mc2155 mutant mapped in the PadR-like regulator (MSMEG_2868) that can efficiently metabolize all C-19 steroids (i.e., AD, ADD, 9OH-AD)</td>
<td>Fernández-Cabezón <em>et al.</em> (2017)</td>
</tr>
<tr>
<td>MS5941</td>
<td>mc²155 ΔMSMEG_5941 (ΔkstD1); high-9OH-AD producer</td>
<td>This study</td>
</tr>
<tr>
<td>MS2891</td>
<td>ADD+ ΔMSMEG_2891 (AhsaC2)</td>
<td>This study</td>
</tr>
<tr>
<td>MS2892</td>
<td>ADD+ ΔMSMEG_2892 (AhsaA2)</td>
<td>This study</td>
</tr>
<tr>
<td>MS2900</td>
<td>ADD+ ΔMSMEG_2900 (AhsaD2)</td>
<td>This study</td>
</tr>
<tr>
<td>mc²155 (pMV261)</td>
<td>mc²155 strain harbouring plasmid pMV261</td>
<td>This study</td>
</tr>
<tr>
<td>MS2891 (pMV261)</td>
<td>MS2891 strain harbouring plasmid pMV261</td>
<td>This study</td>
</tr>
<tr>
<td>MS2891 (pMV2891)</td>
<td>MS2891 strain harbouring plasmid pMV2891</td>
<td>This study</td>
</tr>
<tr>
<td>MS2892 (pMV261)</td>
<td>MS2892 strain harbouring plasmid pMV261</td>
<td>This study</td>
</tr>
<tr>
<td>MS2892 (pMV2892)</td>
<td>MS2892 strain harbouring plasmid pMV2892</td>
<td>This study</td>
</tr>
<tr>
<td>Strain/Plasmid</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>MS2900 (pMV261)</td>
<td>MS2900 strain harbouring plasmid pMV261</td>
<td>This study</td>
</tr>
<tr>
<td>MS2900 (pMV2900)</td>
<td>MS2900 strain harbouring plasmid pMV2900</td>
<td>This study</td>
</tr>
<tr>
<td>MS5941 (pMV261)</td>
<td>MS5941 strain harbouring plasmid pMV261</td>
<td>This study</td>
</tr>
<tr>
<td>MS5941 (pMV2867)</td>
<td>MS5941 strain harbouring plasmid pMV2867</td>
<td>This study</td>
</tr>
<tr>
<td>MS5941 (pMV2869)</td>
<td>MS5941 strain harbouring plasmid pMV2869</td>
<td>This study</td>
</tr>
<tr>
<td>MS5941 (pMV2870)</td>
<td>MS5941 strain harbouring plasmid pMV2870</td>
<td>This study</td>
</tr>
<tr>
<td>MS5941 (pMV5941)</td>
<td>MS5941 strain harbouring plasmid pMV5941</td>
<td>This study</td>
</tr>
<tr>
<td>MS6039 (pMV261)</td>
<td>MS6039 strain harbouring plasmid pMV261</td>
<td>Galán et al. (2017)</td>
</tr>
<tr>
<td>MS6039 (pMV2893)</td>
<td>MS6039 strain harbouring plasmid pMV2893</td>
<td>This study</td>
</tr>
<tr>
<td>MS6039 (pMV6039)</td>
<td>MS6039 strain harbouring plasmid pMV6039</td>
<td>Galán et al. (2017)</td>
</tr>
</tbody>
</table>

**Escherichia coli**

DH10B

- F, mcrA, Δ (mrrhsdRMS-mcrBC), Φ80dlacZΔM15, ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)7697, galU, galK, λ, rpsL, endA1, nupG

**Plasmids**

- pMV261: *Mycobacterium/E. coli* shuttle vector with the kanamycin resistance *aph* gene from transposon Tn903 and the promoter from the *hsp60* gene from *M. tuberculosis*

- pMV2867: pMV261 harbouring the *MSMEG_2867* gene encoding KstD3

- pMV2869: pMV261 harbouring the *MSMEG_2869* gene encoding KstD2

Invitrogen

Stover et al. (1991)

This study
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMV2870</td>
<td>pMV261 harbouring the <em>MSMEG_2870</em> gene encoding KshA2</td>
<td>This study</td>
</tr>
<tr>
<td>pMV2891</td>
<td>pMV261 harbouring the <em>MSMEG_2891</em> gene encoding HsaC2</td>
<td>This study</td>
</tr>
<tr>
<td>pMV2892</td>
<td>pMV261 harbouring the <em>MSMEG_2892</em> gene encoding HsaA2</td>
<td>This study</td>
</tr>
<tr>
<td>pMV2893</td>
<td>pMV261 harbouring the <em>MSMEG_2893</em> gene encoding KshB2</td>
<td>This study</td>
</tr>
<tr>
<td>pMV2900</td>
<td>pMV261 harbouring the <em>MSMEG_2900</em> gene encoding HsaD2</td>
<td>This study</td>
</tr>
<tr>
<td>pMV5941</td>
<td>pMV261 harbouring the <em>MSMEG_5941</em> gene encoding KstD1</td>
<td>This study</td>
</tr>
<tr>
<td>pMV6039</td>
<td>pMV261 harbouring the <em>MSMEG_6039</em> gene encoding KshB1</td>
<td>Galán <em>et al.</em> (2017)</td>
</tr>
<tr>
<td>pJQ200x</td>
<td>Suicide vector used to perform allelic exchange mutagenesis in mycobacteria, <em>P15A ori, sacB, Gm&lt;sup&gt;R&lt;/sup&gt;, xylE</em></td>
<td>Jackson <em>et al.</em> (2001)</td>
</tr>
<tr>
<td>pJQ2891</td>
<td>pJQ200x containing one fragment upstream and another fragment downstream of <em>MSMEG_2891</em> gene</td>
<td>This study</td>
</tr>
<tr>
<td>pJQ2892</td>
<td>pJQ200x containing one fragment upstream and another fragment downstream of <em>MSMEG_2892</em> gene</td>
<td>This study</td>
</tr>
<tr>
<td>pJQ2900</td>
<td>pJQ200x containing one fragment upstream and another fragment downstream of <em>MSMEG_2900</em> gene</td>
<td>This study</td>
</tr>
<tr>
<td>pJQ5941</td>
<td>pJQ200x containing one fragment upstream and another fragment downstream of <em>MSMEG_5941</em> gene</td>
<td>Galán <em>et al.</em> (2017)</td>
</tr>
</tbody>
</table>
FIGURE LEGENDS

Fig. 1. Catabolic pathway of C-19 steroids in *M. smegmatis* mc2155. A) Scheme of the C-19+ route proposed for the degradation of C-19 steroids (i.e., AD, ADD and 9OH-AD) in this bacterium. The upper catabolic pathway consisting of the degradation of A/B rings of C-19 steroids is carried out by enzymes encoded by the C-19+ gene cluster. The resulting metabolic intermediates (HHD and HIP) converge with the cholesterol catabolic pathway where the degradation of C/D steroid rings occurs. The upper catabolic pathway is transcriptionally regulated by the PadR repressor (*padR* regulon), while the lower catabolic pathway is controlled by the KstR2 regulator (*kstR2* regulon). B) Comparative genomic analysis of genes encoding A/B steroid ring modification enzymes in the catabolism of cholesterol or C-19 steroids in *M. smegmatis*. Black numbers indicate the relative position of these gene clusters in the genome of *M. smegmatis* (NC_008596.1). Homologous genes are shown with the same color. Gene annotation: *MSMEG_* (number). Protein identities are indicated in parentheses. Abbreviations of enzymes: 3-ketosteroid-Δ1-dehydrogenase (KstD); 3-ketosteroid-9α-hydroxylase (KshAB); 3-HSA monooxygenase (HsaAB); 3,4-HSA dioxygenase (HsaC); (4,9-DSHA) hydroxylase (HsaD); 2-hydroxy-2,4-hexadienoic acid hydratase (HsaE); 4-hydroxy-2-hydroxy-2-ketovalerate aldolase (HsaF); acetaldehyde dehydrogenase (HsaG). Abbreviations of metabolites: 4-androstene-3,17-dione (AD); 1,4-androstadiene-3,17-dione (ADD); 9α-hydroxy-4-androstene-3,17-dione (9OH-ADD); 3-hydroxy-9,10-secoandrosta-1,3,5(10)-trien-9,17-dione (3-HSA); 3,4-dihydroxy-9,10-secoandrosta-1,3,5(10)-trien-9,17-dione (3,4-HSA); 4,5,9,10-diseco-3-hydroxy-5,9,7-trioxandrosta-1(10),2-diene-4-oic acid (4,9-DSHA); 2-hydroxy-2,4-hexadienoic acid (HHD); 3αα-H-4α(3′-propionic acid)-7β-methylhexahydro-1,5-indanedione (HIP).
Fig. 2. Gene complementation assays of MS5941 and MS6039 mutant strains in shake flasks with minimal medium containing 1.8 mM cholesterol as the only carbon and energy source. (A) Gene complementation of MS5941 (ΔMSMEG_5941, ΔkstD1) mutant: mc^2155 (pMV261) [wt, black]; MS5941 (pMV261) [MS5941, red]; MS5941 (pMV5941) [KstD1, blue]; MS5941 (pMV2869) [KstD2, green]; and MS5941 (pMV2867) [KstD3, grey]. (B) Gene complementation of the MS6039 (ΔMSMEG_6039, ΔkshB1) mutant: mc^2155 (pMV261) [wt, black]; MS6039 (pMV261) [MS6039, red]; MS6039 (pMV6039) [KshB1, blue]; and MS6039 (pMV2893) [KshB2, green].

Fig. 3. Determination of the biological function of enzymes KstD and KshA by resting-cell biotransformations. The bioconversion of 1.8 mM AD into ADD (A) or 9-OHAD (B) by recombinant strains of M. smegmatis was monitored by thin layer chromatography (TLC). Nomenclature of M. smegmatis strains used: MS5941 (pMV261) [Control]; MS5941 (pMV2867) [KstD3]; and MS5941 (pMV2870) [KshA2].

Fig. 4. Growth curves of mutant strains of M. smegmatis ADD+ on C-19 steroids. (A) The strains ADD+ [black], ADD+ MS2891 (ΔMSMEG_2891, ΔhsaC2) [red], ADD+ MS2892 (ΔMSMEG_2892, ΔhsaA2) [blue] and ADD+ MS2900 (ΔMSMEG_2900, ΔhsaD2) [green] were cultivated in shake flasks with minimal medium containing 1.8 mM AD as the sole carbon and energy source. (B) Photograph taken of a final sample of the bacterial cultures in the presence of AD.

Fig. 5. Gene complementation of MS2891, MS2892 and MS2900 mutant strains. The knockout mutants were complemented in trans with an original copy of the corresponding deleted
gene [MS2891c, MS2892c and MS2900c strains; dotted lines] or transformed with the empty plasmid as control [MS2891v, MS2892v and MS2900v strains; continuous lines]. The recombinant strains were cultured in shake flasks with minimal medium containing 18 mM glycerol (A) or 1.8 mM AD (B) as the sole carbon and energy sources. (C) Photograph taken of a final sample of the bacterial cultures in the presence of AD. (D) Scheme of several key enzymatic steps postulated in the catabolism of C-19 steroids in *M. smegmatis*. 

Accepted Article
Supplementary information

Table S1. Oligonucleotides used in this study.

Fig. S1. Scheme of the cholesterol catabolic pathway conventionally proposed in Actinobateria. The cholesterol pathway has been traditionally considered to take place through three consecutive stages. Firstly, the side-chain degradation and $3\beta$-dehydrogenation of the cholesterol molecules occurs in the upper catabolic pathway apparently yielding two molecules of propionyl-CoA, an acetyl-CoA and the 17-ketosteroid AD. Then, the modification of A/B rings of AD in the central catabolic pathway leads to the generation of the HHD and HIP acids. This last compound is mineralized through the lower catabolic pathway. The expression of the genes encoding the enzymes involved in the side-chain and the A/B rings degradation as well as those encoding the sterol uptake system, is regulated by the KstR repressor ($kstR$ regulon), while the expression of the genes responsible for the catabolism of C/D rings is controlled by the KstR2 repressor ($kstR2$ regulon). Abbreviations of enzymes: 3-ketosteroid-$\Delta 1$-dehydrogenase (KstD); 3-ketosteroid-$9\alpha$-hydroxylase (KshAB); 3-HSA monooxygenase (HsaAB); 3,4-HSA dioxygenase (HsaC); (4,9-DSHA) hydroxylase (HsaD). Abbreviations of metabolites: 4-androstene-3,17-dione (AD); 1,4-androstadiene-3,17-dione (ADD); $9\alpha$-hydroxy-4-androstene-3,17-dione (9OH-AD); $9\alpha$-hydroxy-1,4-androstadiene-3,17-dione (9OH-ADD); 3-hydroxy-9,10-secoandrosta-1,3,5(10)-trien-9,17-dione (3-HSA); 3,4-dihydroxy-9,10-secoandrosta-1,3,5(10)-trien-9,17-dione (3,4-HSA); 4,5,9,10-diseco-3-hydroxy-5,9,7-trioxaandrosta-1(10),2-diene-4-oic acid (4,9-DSHA); 2-hydroxyhexa-2,4.dienoic acid (HHD); 3aα-H-4a(3’-propionic acid)-7αβ-methylhexahydro-1,5-indanedione (HIP).

Fig. S2. Chemical structures of relevant steroidal intermediates detected in the culture medium of several naturally sterols-degrading bacteria. Abbreviations: AD (4-androstene-3,17-dione); ADD (1,4-androstadiene-3,17-dione); ADP (4-androstene-3,17-dionopropionic acid); ADDP (1,4-androstadiene-3,17-dionopropionic acid); LOS (4-androstene-3,17-dione-6aHydroxypropionic acid); ADDOS (1,4-androstadiene-3,17-dione-6aHydroxypropionic acid); 4,5,9,10-diseco-3-hydroxy-5,9,7-trioxaandrosta-1(10),2-diene-4-oic acid (4,9-DSHA); 2-hydroxyhexa-2,4.dienoic acid (HHD); 3αα-H-4α(3’-propionic acid)-7αβ-methylhexahydro-1,5-indanedione (HIP).
3,17-dione); ADD (1,4-androstadiene-3,17-dione); 9OH-AD (9α-hydroxy-4-androstene-3,17-dione); 4-HBC (22-hydroxy-23,24-bisnorchol-4-en-3-one); 1,4-HBC (22-hydroxy-23,24-bisnorchol-1,4-dien-3-one); 4-BNC (3-oxo-23,24-bisnorchol-4-en-3-oic acid); 1,4-BNC (3-oxo-23,24-bisnorchol-1,4-dien-3-oic acid).
Fig. 1 Catabolic pathway of C-19 steroids in *M. smegmatis* mc²155.

(A)  
KstD2/ KstD3  
ADD  
KshA2B2  
9OH-ADD  

9OH-ADD  
KstD2/ KstD3  

9OH-ADD  
KshA2B2  

3-HSA  
Non-enzymatic  
HsaA2B1  

3,4-HSA  

4,9-DSHA  

HsaD2  
HsaE, HsaF, HsaG  
Propionyl-CoA  
Pyruvate  

(B)  
Catabolism of cholesterol  

Catabolism of C-19 steroids  

Fig. 2. Gene complementation assays of MS5941 and MS6039 mutants.
Fig. 3. Determination of the biological function of enzymes KstD and KshA by resting-cell biotransformations.

<table>
<thead>
<tr>
<th></th>
<th>(A) AD → ADD</th>
<th>(B) AD → 9OH-AD</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS5941</td>
<td>Control KstD3</td>
<td>Control KshA2</td>
</tr>
<tr>
<td>5 h</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This article is protected by copyright. All rights reserved.
Fig. 4 Growth curves of mutant strains of *M. smegmatis* ADD+ on C-19 steroids.
Fig. 5. Gene complementation of MS2891, MS2892 and MS2900 mutant strains.