

# *Mycobacterium smegmatis* is a suitable cell factory for the production of steroidal synthons

Beatriz Galán,<sup>1,#</sup> Iria Uhía,<sup>1,2,#</sup> Esther García-Fernández,<sup>1</sup> Igor Martínez,<sup>1</sup> Esther Bahillo,<sup>3</sup> Juan L. de la Fuente,<sup>3</sup> José L. Barredo,<sup>3</sup> Lorena Fernández-Cabezón<sup>1</sup> and José L. García<sup>1,\*</sup>

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

<sup>2</sup>MRC Centre for Molecular Bacteriology and Infection, Department of Medicine, Imperial College London, London SW7 2AZ, UK.

<sup>3</sup>Department of Biotechnology, Gadea Biopharma, Parque Tecnológico de León, Nicostrato Vela s/n, 24009 León, Spain.

## Summary

A number of pharmaceutical steroid synthons are currently produced through the microbial side-chain cleavage of natural sterols as an alternative to multi-step chemical synthesis. Industrially, these synthons have been usually produced through fermentative processes using environmental isolated microorganisms or their conventional mutants. *Mycobacterium smegmatis* mc<sup>2</sup>155 is a model organism for tuberculosis studies which uses cholesterol as the sole carbon and energy source for growth, as other mycobacterial strains. Nevertheless, this property has not been exploited for the industrial production of steroidal synthons. Taking advantage of our knowledge on the cholesterol degradation pathway of *M. smegmatis* mc<sup>2</sup>155 we have demonstrated that the *MSMEG\_6039* (*kshB1*) and *MSMEG\_5941* (*kstD1*) genes encoding a reductase component of the 3-ketosteroid 9 $\alpha$ -hydroxylase (KshAB) and a ketosteroid  $\Delta^1$ -dehydrogenase (KstD), respectively, are indispensable enzymes for

the central metabolism of cholesterol. Therefore, we have constructed a *MSMEG\_6039* (*kshB1*) gene deletion mutant of *M. smegmatis* MS6039 that transforms efficiently natural sterols (e.g. cholesterol and phytosterols) into 1,4-androstadiene-3,17-dione. In addition, we have demonstrated that a double deletion mutant *M. smegmatis* MS6039-5941 [ $\Delta$ *MSMEG\_6039* ( $\Delta$ *kshB1*) and  $\Delta$ *MSMEG\_5941* ( $\Delta$ *kstD1*)] transforms natural sterols into 4-androstene-3,17-dione with high yields. These findings suggest that the catabolism of cholesterol in *M. smegmatis* mc<sup>2</sup>155 is easy to handle and equally efficient for sterol transformation than other industrial strains, paving the way for valuating this strain as a suitable industrial cell factory to develop *à la carte* metabolic engineering strategies for the industrial production of pharmaceutical steroids.

## Introduction

Androstenedione (4-androstene-3,17-dione; AD) and androstadienedione (1,4-androstadiene-3,17-dione; ADD) are key intermediates of microbial steroid metabolism (García *et al.*, 2012). These compounds belong to the 17-keto steroid family and are used as the starting materials for the preparation of different pharmaceutical steroids (Malaviya and Gomes, 2008; Donova and Egorova, 2012; García *et al.*, 2012). These synthons can be produced by microbial side-chain cleavage of cholesterol or phytosterols as an alternative to multi-step chemical synthesis based on digoxigenin, a steroid found exclusively in the flowers and leaves of *Digitalis* plants, as a starting material. Industrially, AD and ADD have been produced through fermentative processes using wild microorganisms that have been subsequently modified and optimized by conventional mutagenic procedures (Donova *et al.*, 2005; Andor *et al.*, 2006; Donova and Egorova, 2012; García *et al.*, 2012). *Mycobacterium* spp. NRRL 3805B and 3683B capable of forming AD and ADD from sterols, respectively, are examples of these mutants used at industrial scale (Donova *et al.*, 2005; Donova and Egorova, 2012). A drawback of the AD and ADD industrial production based on these wild strains is the usual and concomitant accumulation of unwanted by-products which hinder downstream processes. The use of molecularly defined mutants has been envisioned to avoid such drawbacks, but the lack of genetic data on the microbial catabolism of steroids has hampered the

Received 21 June, 2016; revised 19 September, 2016; accepted 21 September, 2016. \*For correspondence. E-mail: jlgarcia@cib.csic.es; Tel. +34 918373112; Fax +34 915360432.

*Microbial Biotechnology* (2017) 10(1), 138–150  
doi:10.1111/1751-7915.12429

<sup>#</sup>These authors contributed equally to this work.

## Funding Information

The technical work of A. Valencia is greatly appreciated. We thank I. Calvillo for assistance in the analytical GC/MS experiments. This study 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).

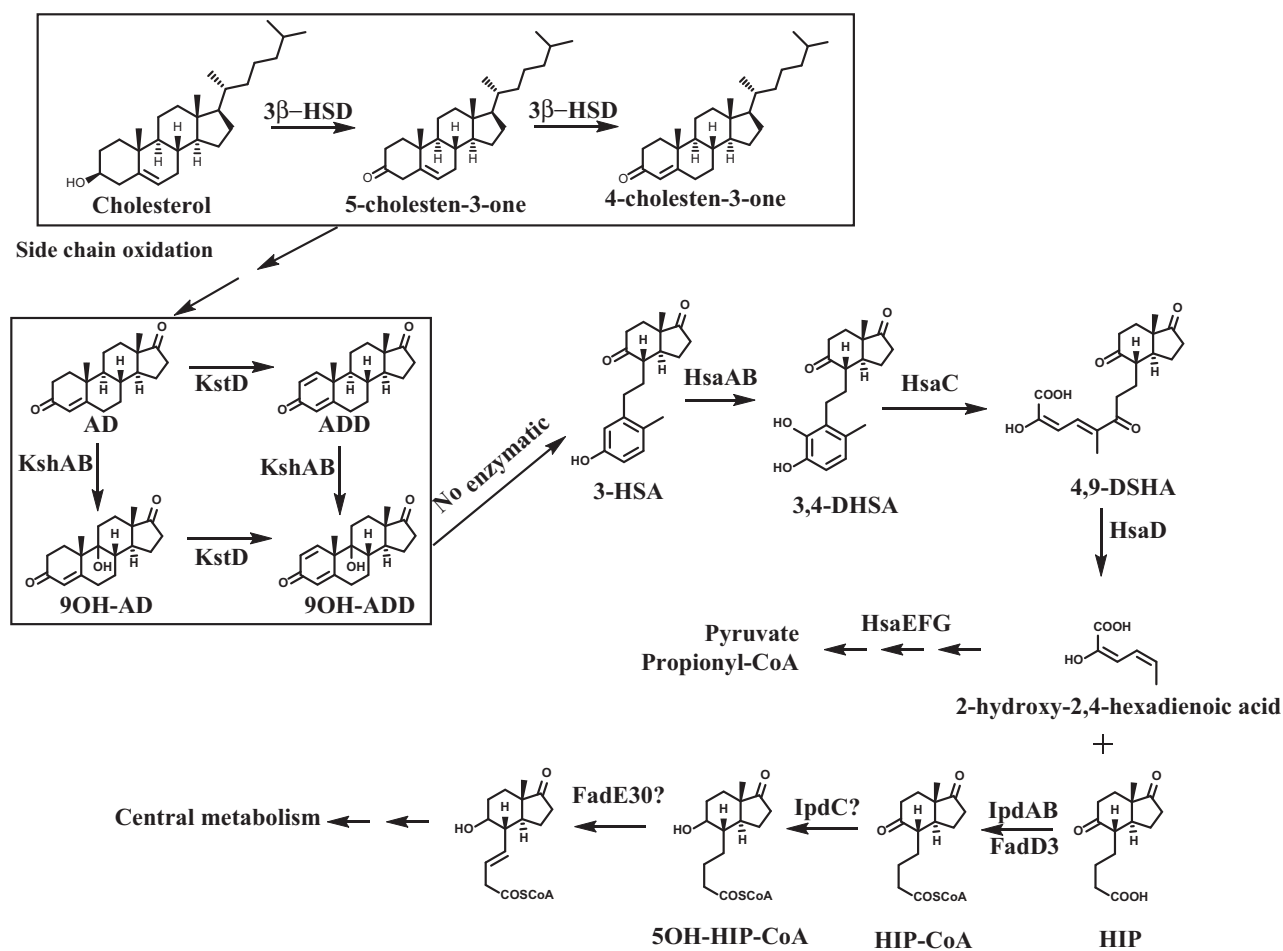
© 2016 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

construction of genetically engineered mutants so far. Nevertheless, some attempts have been conducted to construct site-directed mutant strains of *Rhodococcus* (i.e. the best characterized cholesterol-degrading organism) to produce AD, ADD and 9 $\alpha$ -hydroxy-4-androstene-3,17-dione (9OH-AD) from natural sterols (e.g. cholesterol or phytosterols), but these mutants have not been used at industrial scale yet (van der Geize *et al.*, 2000, 2001a,b, 2002a,b, 2008; Wilbrink *et al.*, 2011; Yeh *et al.*, 2014).

To create *à la carte* mutants able to produce AD and ADD from cholesterol or phytosterols we have tested *Mycobacterium smegmatis* mc<sup>2</sup>155 as a model strain

based on our current knowledge on sterol catabolism in this microorganism (Fig. 1) (García *et al.*, 2012). The 3-ketosteroid 9 $\alpha$ -hydroxylase (Ksh) has been proposed as the key enzyme for ring-B opening, and therefore, the removal of this activity should render ADD as end-product. The Ksh enzymes of *Rhodococcus* and *Mycobacterium* have been characterized as two-component monooxygenases, composed of an oxygenase (KshA) and a ferredoxin reductase (KshB) (van der Geize *et al.*, 2002a; Capyk *et al.*, 2009, 2011; Petrusma *et al.*, 2009, 2012, 2014; Hu *et al.*, 2010; Bragin *et al.*, 2013; Penfield *et al.*, 2014). Therefore, theoretically, a deletion mutant of one of the Ksh encoding genes, i.e. *kshA* or *kshB*,



**Fig. 1.** Proposed pathway for cholesterol degradation in *Mycobacterium smegmatis*. Cholest-4-en-3-one or any of the subsequent metabolites from degradation of the side-chain up to (and including) AD may undergo a dehydrogenation reaction to introduce a double bond in the position 1, leading to compound cholest-1,4-diene-3-one in the case of cholest-4-en-3-one, or to the corresponding 1,2-dehydro derivatives for other molecules. The side-chain degradation of these compounds will be identical to that of the cholest-4-en-3-one to the common intermediate 9 $\alpha$ -hydroxyandrost-1,4-diene-3,17-dione (9OHAD). 3 $\beta$ -hydroxysteroid dehydrogenase (3 $\beta$ -HSD), 3-ketosteroid- $\Delta^1$ -dehydrogenase (KstD), 3-ketosteroid-9 $\alpha$ -hydroxylase (KshAB), 3-hydroxy-9,10-secoandrost-1,3,5(10)-trien-9,17-dione (3-HSA), 3,4-dihydroxy-9,10-secoandrost-1,3,5(10)-trien-9,17-dione (3,4-HSA), 4,5,9,10-diseco-3-hydroxy-5,9,7-trioxoandrost-1(10),2-diene-4-oic acid (4,9-DSHA), 3 $\alpha$ -H-4 $\alpha$ (3'-propionic acid)-7 $\beta$ -methylhexahydro-1,5-indanone (HIP), 3 $\alpha$ -H-4 $\alpha$ (3'-propionic acid)-5 $\alpha$ -hydroxy-7 $\beta$ -methylhexahydro-1-indanone (5OH-HIP), 3-hydroxy-9,10-secoandrost-1,3,5(10)-trien-9,17-dione dioxygenase (HsaC), 4,5,9,10-diseco-3-hydroxy-5,9,7-trioxoandrost-1(10),2-diene-4-oic acid hydroxylase (HsaD), 2-hydroxy-2,4-hexadienoic acid hydratase (HsaE), 4-hydroxy-2-hydroxy-2-ketovalerate aldolase (HsaF), acetaldehyde dehydrogenase (HsaG), HIP-CoA transferase (LpdAB and FadD3), FadE30 (Acyl-CoA dehydrogenase).

should accumulate ADD. On the other hand, 3-ketosteroid  $\Delta^1$ -dehydrogenase (KstD) is the key enzyme that transforms AD into ADD and thus, strains lacking KstD should theoretically accumulate 9OH-AD. Finally, double mutants in Ksh and KstD should render AD as the main product for sterol degradation (Fig. 1) (García *et al.*, 2012).

Nevertheless, these previous assumptions cannot be always easily confirmed by gene inactivation since cholesterol catabolic pathway usually presents functional redundancy for some steps of the pathway; this means that one catalytic step can be carried out by several homologous enzymes. Two KstD enzymes were reported in *Rhodococcus erythropolis* SQ1 (van der Geize *et al.*, 2000) and accordingly, targeted disruption of only one of the *kstD* genes did not result on the accumulation of intermediates, but the strain lacking both dehydrogenase activities was able to convert AD into 9OH-AD stoichiometrically (van der Geize *et al.*, 2002b). Similarly, Fernández de las Heras *et al.* (2012) have demonstrated the existence of three KstD activities in *Rhodococcus ruber* strain Chol-4 and showed that the triple KstD mutant was able to convert AD into 9OH-AD. However, these mutants were unable to accumulate 9OH-AD from cholesterol or phytosterols (van der Geize *et al.*, 2001a; Fernández de las Heras *et al.*, 2012). Brzostek *et al.* (2009) have identified six putative KstD enzymes within the *M. smegmatis* genome and a targeted disruption of one of them (KstD1) resulted in partial inactivation of the cholesterol degradation pathway and the consequent accumulation of AD. More recently, Wei *et al.* (2010) have identified and deleted a *kstD* gene (named *ksdD<sub>M</sub>*) in *Mycobacterium neoaurum* NwIB-01, a strain isolated from a sterol contaminated soil that naturally accumulates ADD from phytosterols. Interestingly, the resulting deleted mutant (named NwIB-2) was able to accumulate AD from phytosterols, but ADD is still present in the culture medium, suggesting that this strain contains other KstD enzymes (Wei *et al.*, 2010). Three homologues of KstD have been characterized in *M. neoaurum* ATCC 25795 and thus, single deletion of these genes failed to result in a stable and maximum accumulation of 9OH-AD due to residual KstD activities (Yao *et al.*, 2014).

On the other hand, van der Geize *et al.* (2002a, 2008) have demonstrated that Ksh mutants of *R. erythropolis* SQ1 deleted in *kshA* or *kshB* genes can produce ADD from AD, but surprisingly they do not accumulate AD, ADD or other metabolites during sterol conversion. Interestingly, the degradation of phytosterols was not impaired in a *kshA*<sup>-</sup> mutant and rates of degradation were comparable to those of the parent strain, suggesting that *Rhodococcus* has alternative enzymes to catabolize phytosterols. Moreover, the *kshB*<sup>-</sup> mutant failed to

cleave the side-chain of sterols and, although phytosterols were oxidized to their stenone derivatives, they were not metabolized further, suggesting that *kshB* could be also involved in the degradation of sterol side-chain in *Rhodococcus* (van der Geize *et al.*, 2002a). The situation can be more complex taking into account that some strains of *Rhodococcus* may contain up to five KshA homologous proteins, each displaying unique steroid induction patterns and substrate ranges, confirmed that the 9 $\alpha$ -hydroxylation can take place at different steps of steroid oxidation (Petrusma *et al.*, 2011).

On the basis of our knowledge on cholesterol metabolism in *M. smegmatis* (Uhía *et al.*, 2012) we detected that this organism has several important differences with the sterol catabolism in *Rhodococcus* and we anticipated that these differences might render *M. smegmatis* as a suitable cell factory for the metabolic manipulation of this pathway. The results presented below demonstrate this hypothesis and show that *M. smegmatis* is an ideal cell factory to develop metabolic engineering strategies for the industrial production of AD, ADD or other steroid intermediates using natural sterols as feedstock.

## Results

### Identification of Ksh and KstD enzymes in *M. smegmatis*

In contrast to the frequent observation that the components of bacterial multicomponent oxygenases are encoded in a single operon, the *kshA* and *kshB* genes encoding Ksh activity in *Rhodococcus* (van der Geize *et al.*, 2002a) or in *M. tuberculosis* (Cole *et al.*, 1998) are located far from each other in the genome. Nevertheless, using the sequence of the annotated *kshA* and *kshB* genes from *Rhodococcus* (van der Geize *et al.*, 2002a, 2008) and *M. tuberculosis* (Capyk *et al.*, 2009) as probes, we have localized the corresponding orthologues in *M. smegmatis* (Table 1). This analysis revealed that there are at least two genes encoding putative oxygenase components, named *kshA1* and *kshA2*, and two genes encoding putative reductase components, named *kshB1* and *kshB2*. This means that *M. smegmatis* could produce theoretically four different Ksh hydroxylases, i.e. Ksh1 (KshA1B1), Ksh2 (KshA2B2), Ksh3 (KshA1B2), Ksh4 (KshA2B1). However, a detailed analysis of the upstream sequences of these genes showed that only the *MSMEG\_5925* (*kshA1*) and *MSMEG\_6039* (*kshB1*) genes are preceded by promoters inducible by cholesterol, having the consensus operator sequence for the binding of the KstR repressor, one of the regulators of cholesterol catabolism (named KstR regulon) (Kendall *et al.*, 2007). Moreover, microarray expression experiments carried out on *M. smegmatis* indicate that *kshA1* and *kshB1* are induced 1.9-fold and 7.0-fold, respectively, in cholesterol

**Table 1.** *In silico* analysis of *M. smegmatis* mc<sup>2</sup> 155 genome.

Protein	Gene <sup>a</sup> <i>M. smegmatis</i> mc <sup>2</sup> 155	Protein (aa length)	Identity <sup>b</sup> (%)	KstR-regulon operator	Induction fold <sup>c</sup>
KshA	<i>MSMEG_5925</i>	383	67	Yes	7.00
	<i>MSMEG_2870</i>	386	53	No	1.00
KshB	<i>MSMEG_6039</i>	353	63	Yes	1.88
	<i>MSMEG_2893</i>	351	56	No	1.21
KstD	<i>MSMEG_5941</i>	566	65	Yes	13.00
	<i>MSMEG_4864</i>	587	39	No	-1.46
	<i>MSMEG_2869</i>	558	40	No	-1.17
	<i>MSMEG_2867</i>	522	31	No	-1.14
	<i>MSMEG_4870</i>	546	30	No	-1.35
	<i>MSMEG_5835</i>	547	17	No	1.14

a. Gene identifications correspond to the last annotation of *M. smegmatis* genome and are different to the annotations described by Brzostek *et al.* (2005).

b. Identities were established using the proteins of *R. jostii* RHA1 as reference.

c. Induction fold was calculated comparing gene expression in *M. smegmatis* cultured in cholesterol versus glycerol containing media (Uhía *et al.*, 2012).

with respect to glycerol (Table 1) (Uhía *et al.*, 2012). On the contrary, *kshA2* and *kshB2* are not induced by cholesterol under the tested conditions (Table 1) (Uhía *et al.*, 2012). These results strongly suggest that most probably only Ksh1 encoded by *kshA1B1* is involved in the metabolism of cholesterol in *M. smegmatis*.

Among the six putative *kstD* genes identified in the *M. smegmatis* genome by Brzostek *et al.* (2009) that might encode the KstD-like activity only the *MSMEG\_5941* (*kstD1*) gene is controlled by a promoter containing the consensus KstR operator sequence and is differentially expressed in the presence of cholesterol (Table 1) (Uhía *et al.*, 2012). This observation suggests that in *M. smegmatis* most probably only *kstD1* was involved in the catabolism of cholesterol.

Therefore, based on these analyses we anticipated the hypothesis that in *M. smegmatis* a single deletion of *kshB1* (or *kshA1*) and a double deletion of *kshB1* (or *kshA1*) and *kstD1* might be sufficient to generate mutants able to accumulate ADD or AD, respectively, when cultured in the presence of sterols.

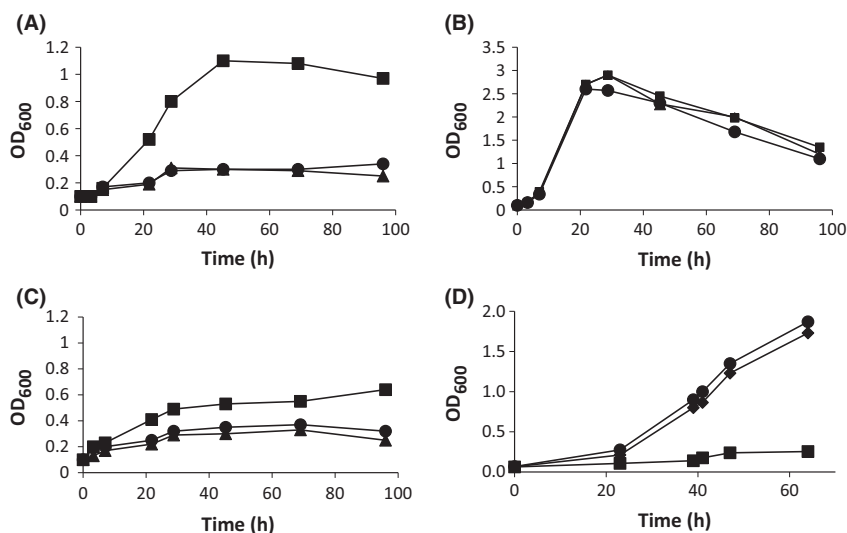
#### Construction of the $\Delta$ kshB1 *M. smegmatis* mutant

As mentioned above, to eliminate the two-component Ksh1 hydroxylase activity (KshA1B1) in *M. smegmatis*, we assumed that it was enough to delete the *MSMEG\_6039* (*kshB1*) gene encoding the reductase component, thus we constructed the MS6039 mutant ( $\Delta$ kshB1) (Fig. S1). Ksh1 activity can be eliminated either by deleting *kshB1* or *kshA1* (or both), but we decided to test the single deletion of *kshB1* instead of *kshA1*, because in the absence of the corresponding oxygenase component, the reductase component of multicomponent oxygenases usually acts as a gratuitous scavenger of reduction power, generating futile cycles that might reduce the efficiency of the biotransformation

process (Galán *et al.*, 2000; Blank *et al.*, 2010) and therefore, in this sense, it can be more convenient to suppress the reductase instead the oxygenase component. However, it is also true that other reductase enzymes present in the cell could replace their function. Nevertheless, we also decided to delete *kshB1* because the effect of the deletion of the *kshB* homologous gene has already been analysed in *Rhodococcus* (van der Geize *et al.*, 2002a,b) and we would like to compare the performance of *M. smegmatis* having the same gene deletion.

The engineered MS6039 mutant ( $\Delta$ kshB1) was unable to efficiently grow in cholesterol or phytosterols as a sole carbon and energy source when compared with the wild-type strain (Fig. 2A). However, the mutant perfectly grows using glycerol as a carbon source (Fig. 2B). To confirm that the deletion of *kshB1* was responsible for the observed phenotype, we transformed the MS6039 mutant with the plasmid pMV6039 harbouring the *kshB1* gene. Figure 2D shows that the complemented MS6039 (pMV6039) strain recovered the capacity to grow in sterols as a sole carbon and energy source. These results confirmed that *kshB1* plays an essential role in the catabolism of sterols in *M. smegmatis*. In addition, considering that the MS6039 mutant is unable to efficiently grow in sterols even at long incubation periods, we can conclude that the KshB1 reductase activity of Ksh1 cannot be replaced by other mycobacterial reductases, i.e. KshB2 or other KshB-like reductases, in the tested conditions, either because they are not expressed in these conditions or because they cannot interact with the KshA1 oxygenase component.

These experiments also revealed that wild-type *M. smegmatis* is able to use efficiently different phytosterols as a carbon and energy source (Fig. 2C), and this is remarkable because the capacity of this strain to grow



**Fig 2.** Growth curves of *M. smegmatis* mc<sup>2</sup>155, mutants and complemented strains cultured in shake flasks with different carbon sources. (A) Strains cultured with 0.4 g l<sup>-1</sup> of cholesterol (mc<sup>2</sup>155 (squares), MS6039 (circles) and MS6039-5941 (triangles)); (B) Strains cultured with 18 mM glycerol (mc<sup>2</sup>155 (squares), MS6039 (circles) and MS6039-5941 (triangles)); (C) Strains cultured with 0.4 g l<sup>-1</sup> of phytosterols (mc<sup>2</sup>155 (squares), MS6039 (circles) and MS6039-5941 (triangles)); (D) Gene complementation of MS6039 mutant strain cultured with 1.8 mM of cholesterol (mc<sup>2</sup>155 (pMV261) (diamonds), MS6039 (pMV261) (squares) and MS6039 (pMV6039) (circles)). The data reported are the averages of three different assays.

on a mixture of phytosterols has not been well documented in the literature so far.

On the other hand, the slight growth of MS6039 in the presence of sterols suggested that the mutant was able to degrade their side-chain and predicted a possible accumulation of metabolic intermediates in the culture medium. As expected, Figure 4 shows the accumulation of ADD in the culture medium of MS6039 mutant when cultured in shake flasks, whereas *M. smegmatis* wild-type strain completely mineralizes cholesterol (or phytosterols) and does not accumulate any intermediate in these culture conditions (data not shown). The ADD molar yield for the transformation of cholesterol or phytosterols was 91% for both feedstocks. Nevertheless, we have also detected the presence of small amounts of 22-hydroxy-23,24-bisnorcholesterol-1,4-dien-3-one (1,4-HBC, 20-hydroxymethylpregna-1,4-dien-3-one) as a by-product (Fig. 4C and D).

In addition, we have tested the production of ADD from phytosterols in 5- and 20-l stirred jar bioreactors using higher concentrations of phytosterols (20 g l<sup>-1</sup>) as substrate to study the industrial potential of MS6039 mutant strain. By carrying several replicates of these experiments in different operational conditions, we concluded that between 55% and 70% of added phytosterols are consumed and the molar yield to ADD varies from 67% to 80%, depending on the bioreactor scale. AD was also obtained in a yield from consumed phytosterols between 22% and 30%; other by-products, such as 1,4-HBC were detected as traces (Fig. S2). These

experiments carried out at high sterol concentrations showed a significant contamination of unconverted AD suggesting that KstD activity is a bottleneck for the complete transformation of phytosterols into ADD at industrial scale. This problem can be overcome by changing the operational conditions (e.g. preinoculation growth, inoculation conditions, bioreactor agitation configuration, aeration flow, composition of culture medium, pH control, substrate additions, etc.) in combination with the overexpression of *kstD* genes in MS6039 that lead to a significant increase in ADD/AD ratio (Gadea personal communication).

The results contrast with the data obtained with the equivalent *kshB* mutant of *R. erythropolis* strain SQ1 (RG4 mutant), which only accumulates sitostenone from  $\beta$ -sitosterol, because apparently, the RG4 mutant does not degrade the side-chain of  $\beta$ -sitosterol. To explain this performance, it was proposed that the KshB reductase from *R. erythropolis* was not only involved in the production of 9OH-AD, but also in sterol side-chain removal (van der Geize *et al.*, 2002b). In *M. smegmatis*, the absence of KshB1 does not hinder the complete degradation of the sterol side-chain (see Discussion).

#### Construction of the $\Delta$ kshB1 and $\Delta$ kstD1 *M. smegmatis* double mutant

According to our previous genomic analysis, it should be possible to produce AD from sterols in *M. smegmatis* by eliminating the KstD activity in the MS6039 mutant.

Therefore, to test this hypothesis, we constructed the *M. smegmatis* MS6039-5941 double mutant ( $\Delta kstD1$ ,  $\Delta kshB1$ ) by deleting the *MSMEG\_5941* (*kstD1*) gene in the MS6039 ( $\Delta kshB1$ ) mutant (Fig. S1). As expected, the MS6039-5941 ( $\Delta kshB1$ - $\Delta kstD1$ ) double mutant was unable to efficiently grow in cholesterol or phytosterols as the sole carbon and energy source, but the double mutant grows in glycerol at the same rate than the wild-type strain (Fig. 2B).

Accordingly to our predictions, when the MS6039-5941 double mutant was cultured in minimal media containing glycerol as carbon and energy source and cholesterol (or phytosterols) as feedstock, the culture yields a large accumulation of AD in shake flasks (Fig. 4A and B). The AD molar yields for sterol transformations were 90% and 84% when cholesterol or phytosterols were used as substrates respectively. Interestingly, in these culture conditions, small amounts of 22-hydroxy-23,24-bisnorcholesterol-4-en-3-one (4-HBC, 20-hydroxymethylpregna-4-en-3-one) were accumulated during the biotransformation (Fig. 4C and D).

These results support the hypothesis that although *M. smegmatis* has six putative KstD dehydrogenases, the KstD1 enzyme encoded by the *MSMEG\_5941* gene is the main KstD used for sterol catabolism in this bacterium. This observation contrasts with other cholesterol-degrading bacteria, where several KstD dehydrogenases are involved in the metabolism of sterols (van der Geize *et al.*, 2000, 2002b; Fernández de las Heras *et al.*, 2012; Yao *et al.*, 2014).

The MS6039-5941 double mutant was tested in 2-l jar bioreactor in the presence of 10 g l<sup>-1</sup> of phytosterols. An almost complete transformation of phytosterols into AD (88–90%), 4-HBC (10–11%) and very small amounts of ADD and 1,4-HBC were detected in several replicates of this experiment (Fig. S3).

Interestingly, the accumulation of small amounts of ADD observed in the bioreactor experiments suggests that one of the other identified *kstD2*, *kstD3*, *kstD4*, *kstD5* or *kstD6* homologous genes (Table 1) is somewhat active in the mutant. Thus, to further improve the AD production yield, this residual KstD activity should be identified and eliminated. Moreover, in this sense, it will be also important to understand why the three last carbons at C-17 are not efficiently converted into propionyl-CoA, rendering 4-HBC and 1,4 HBC as by-products of the pathway.

## Discussion

The first remarkable finding presented above is the experimental demonstration that *M. smegmatis* mc<sup>2</sup>155 is able to efficiently metabolize mixtures of phytosterols as carbon and energy sources through the same

catabolic pathway utilized to mineralize cholesterol (Fig. 1) (Uhía *et al.*, 2012). This finding is important because it supports the proposal of considering *M. smegmatis* as a useful cell factory for the industrial production of steroidal synthons from raw sterols. Surprisingly, in spite of the large accumulated knowledge on *M. smegmatis* mc<sup>2</sup>155 biology, this organism has not been used as a cell factory for this industrial purpose so far. The only case reported in the literature is an antibiotic resistant mutant deposited as *M. smegmatis* VKPM Ac-1552 in the Russian National Collection of Industrial Microorganisms (VKPM) that appears to transform sterols into AD (Russian Federation Patent no. 2 126 837 (1999), reviewed in Donova (2007)).

The observation that the growth of *M. smegmatis* mc<sup>2</sup>155 in sterols is impaired by a deletion of the *kshB1* gene (*MSMEG\_6039*) suggests that KshB1 is the main reductase component of the two-component KshAB 9 $\alpha$ -hydroxylase in the cholesterol degradative pathway of this organism. Apparently, this reductase activity is not redundant in *M. smegmatis* and thus, it cannot be replaced by similar enzymes, as it has been demonstrated for other key enzymes of the pathway (Uhía *et al.*, 2011). In addition, the *kshB1* gene does not appear to be critical for the degradation of the side-chain of sterols in *M. smegmatis*, as suggested in *Rhodococcus* (van der Geize *et al.*, 2002a), and therefore, its absence does not impair an efficient transformation of sterols into AD or ADD in the mutant strain.

On the other hand, the significant accumulation of AD in the MS6039-5941 ( $\Delta kshB1$ - $\Delta kstD1$ ) double mutant ascribes a fundamental role to the *kstD1* gene (*MSMEG\_5941*) in the catabolism of cholesterol. Nevertheless, in this mutant, we have observed some enzyme redundancy because we were able to detect small amounts of ADD when high concentrations of phytosterols are transformed. Therefore, we can conclude that at least one of the other five putative KstD enzymes identified in *M. smegmatis* mc<sup>2</sup>155 (Table 1) can also replace this function, but this alternative KstD activity is very low and does not appear to fulfil the relevant function observed in other cholesterol-degrading strains (van der Geize *et al.*, 2000, 2002b; Fernández de las Heras *et al.*, 2012; Yao *et al.*, 2014).

The results presented in this work suggest that the metabolism of sterols in *M. smegmatis* mc<sup>2</sup>155 concerning to the two central/key enzymes investigated, i.e. KstD and Ksh, is less redundant than in *Rhodococcus*. Perhaps, the low redundancy of these central enzymes can explain why many strains currently used at industrial scale to transform phytosterols into AD, ADD or 9OH-AD are mycobacterial mutants obtained by conventional mutagenesis, i.e. because single mutations like those produced in this work might render a producer strain.

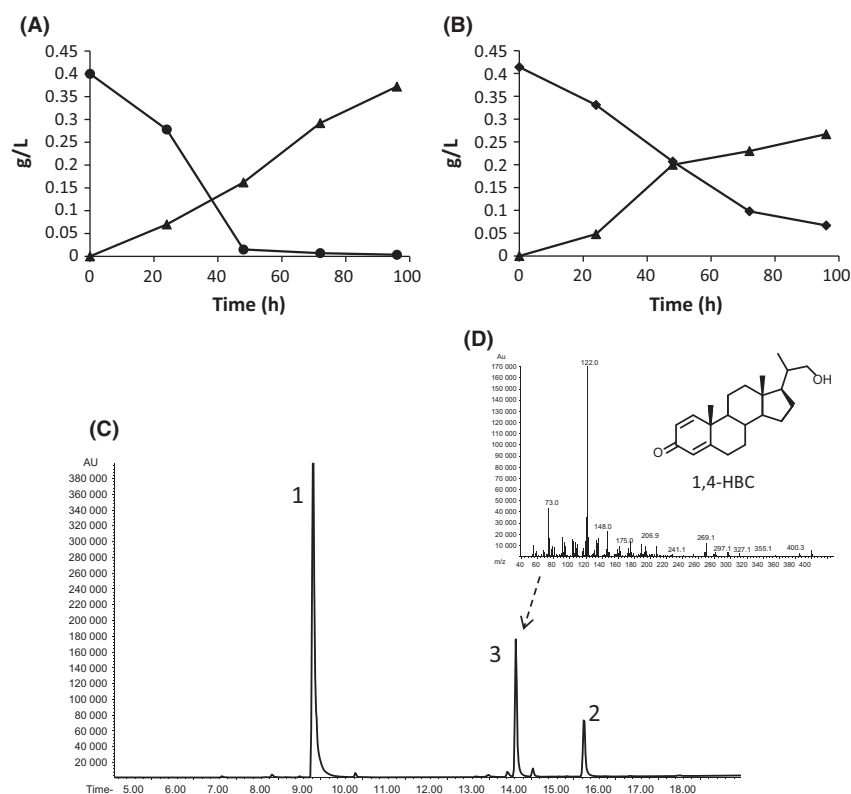
Genomic analyses of some of these industrial mutants have confirmed this hypothesis (data not shown).

Another interesting difference observed between the sterol catabolism of *Mycobacterium* and *Rhodococcus* is the presence of 4-HBC and 1,4-HBC alcohols as metabolic by-products in the cultures of our mutants (Figs 3C, D and 4C, D). It has been described that the cultures of equivalent *Rhodococcus* mutants only accumulate the corresponding acids (Wilbrink *et al.*, 2011; Yeh *et al.*, 2014). The RG32 mutant of *Rhodococcus rhodochrous* DSM43269, a mutant completely devoid of Ksh by inactivation of five *kshA* genes, was able to produce very small amounts of ADD from  $\beta$ -sitosterol (7% molar) but large amounts of 3-oxo-23,24-bisnorchola-1,4-dien-22-oic acid (1,4-BNC) (67% molar) (Wilbrink *et al.*, 2011). A *kshB* mutant of *Rhodococcus equi* USA-18 was also able to produce ADD and 1,4-BNC from sterols in similar molar ratios (Yeh *et al.*, 2014).

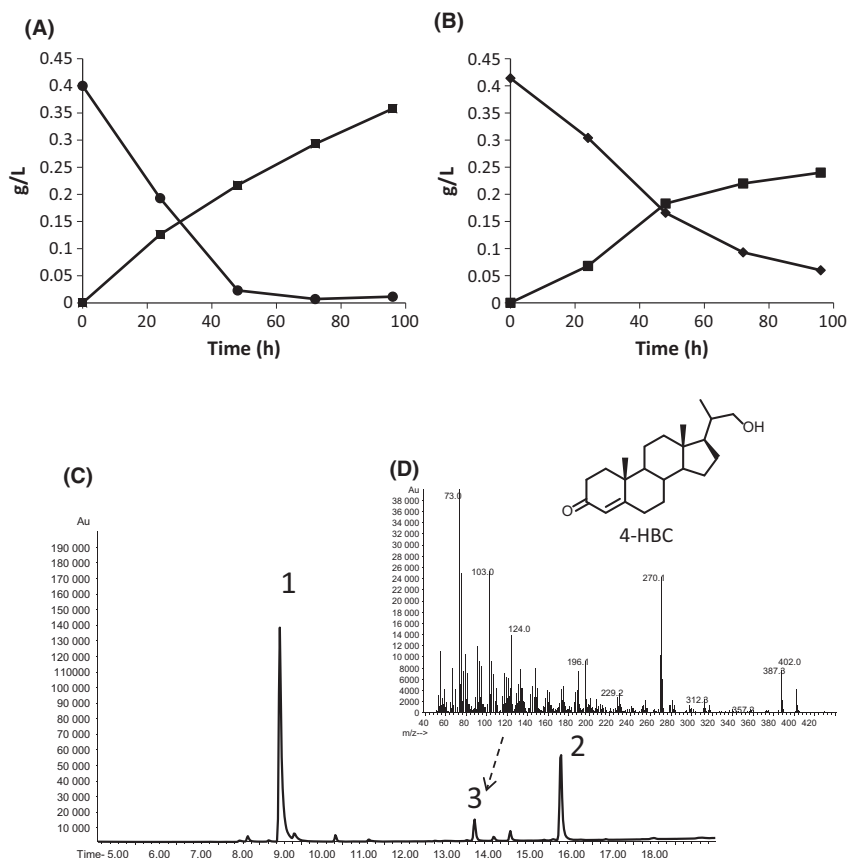
In this regard, 4-HBC alcohol was detected many years ago together with AD during phytosterol transformations performed by the industrial strain of *Mycobacterium* sp. NRRL B-3805 (Marsheck *et al.*, 1972). This alcohol was postulated as a side reaction product but

not as a physiological intermediate of the major pathway leading to production of C-17-ketonic products (Marsheck *et al.*, 1972). It has been proposed that 4-HBC and 1,4-HBC derive from 4-BNC and 1,4-BNC, respectively, by the action of a carboxyl-reductase by one or two consecutive enzymatic steps (Szentirmai, 1990; Xu *et al.*, 2016). Nevertheless, these carboxyl-reductases have not been identified in mycobacteria yet (Xu *et al.*, 2016). Assuming this hypothesis, these enzymes should not be active, or very low active, in *Rhodococcus*, since only the 4-BNC/1,4-BNC acids are detected in this organism (Yeh *et al.*, 2014).

The accumulation of these C-22 intermediates suggests that the elimination of the last isopropyl group of the sterol side-chain seems to be highly dependent of the  $9\alpha$ -hydroxylation by Ksh. We assume that in the absence of  $9\alpha$ -hydroxylation, the enzymes responsible for removing the last propionic acid of the side-chain at C-17, i.e. the postulated acyl-CoA-dehydrogenase, enoyl-CoA-hydratase and aldol-lyase enzymes (García *et al.*, 2012), do not work efficiently. This inefficiency can be caused because they are feedback inhibited by the accumulation of AD and/or ADD, or because the real



**Fig 3.** Production of ADD by the MS6039 mutant in shake flasks. ADD is represented by triangles. (A) 9 mM glycerol + 0.4 g l<sup>-1</sup> of cholesterol (circles) used as substrates; (B) 9 mM glycerol + 0.4 g l<sup>-1</sup> of phytosterols (diamonds) used as substrates; (C) Analysis by CG-MS of the products after 96 h of incubation on phytosterols. (1) ADD, (2) cholestenone (internal standard), and (3) 1,4-HBC. (D) Chemical structure and fragmentation pattern of 1,4-HBC. The data reported are the averages of three different assays.



**Fig 4.** Production of AD by the MS6039-5941 mutant in shake flasks. AD is represented by squares. (A) 9 mM glycerol + 0.4 g l<sup>-1</sup> of cholesterol (circles) used as substrates; (B) 9 mM glycerol + 0.4 g l<sup>-1</sup> of phytosterols (diamonds) used as substrates; (C) Analysis by GC/MS of the products after 96 h of culture on phytosterols. (1) AD, (2) cholestenone (internal standard) and (3) 4-HBC. (D) Chemical structure and fragmentation pattern of 4-HBC. The data reported are the averages of three different assays.

substrates of these enzymes are the 9 $\alpha$ -hydroxy derivatives of 1,4-BNC-CoA or 4-BNC-CoA. In this sense, it has been demonstrated that the best substrate of Ksh from *M. tuberculosis* is 4-BNC-CoA and 1,4-BNC-CoA (Capyk *et al.*, 2011), suggesting that 9 $\alpha$ -hydroxylation occurs before the release of the last propionic acid at C-17. If the previous 9 $\alpha$ -hydroxylation of the steroid is critical to be recognized as a substrate for some of the enzymes involved in the release of the side-chain, this might explain why in *Rhodococcus* the absence of KshB impairs the complete degradation of the sterol side-chain (van der Geize *et al.*, 2002a). Then, most probably the homologous enzymes of *Mycobacterium* are efficient enough on non-hydroxylated substrates and thus, the 9 $\alpha$ -hydroxylation of sterols is not so critical to allow the complete side-chain degradation. In *M. neoaurum* ATCC 25795, the accumulation of 4-HBC and 1,4-HBC was only detected after the deletion of the *hsd4A* gene coding for a dual-functional enzyme with both 17 $\beta$ -hydroxysteroid dehydrogenase and  $\beta$ -hydroxyacyl-CoA dehydrogenase activities. However, these compounds are not the substrates of the Hsd4A *in vitro* (Xu *et al.*, 2016).

This means that some partial or complete blockage of the side-chain degradation might cause the accumulation of 4-HBC or 1,4-HBC as by-products. This blockage can be caused not only by specific mutations, but also by retro-inhibitions of the enzymes due to the accumulation of certain intermediates when the cells are cultured with high concentrations of sterols.

During the course of this work, Xu *et al.* (2015) have described a double ( $\Delta kshA1$ ,  $\Delta kstD1$ ) mutant of *Mycobacterium* sp. (apparently a derivative of *M. neoaurum* NwIB-02 ( $\Delta kstD1$ ) (Wei *et al.*, 2010) that accumulates AD, ADD, 1,4-HBC and 4-HBC at different concentrations depending of fermentation temperature when cultured on phytosterols. This result confirmed that mycobacteria accumulate C-22 alcohols instead of the corresponding C-22 acids. Interestingly, they have also demonstrated that the residual KstD activity is not active on 9OH-AD (Xu *et al.*, 2015) suggesting that a  $\Delta^1$ -dehydrogenation by KstD precedes Ksh hydroxylation. Remarkably, in contrast with our MS6039-5941 ( $\Delta kshB1$ - $\Delta kstD1$ ) double mutant, this *M. neoaurum* mutant still retains other Ksh and KstD active isoenzymes, because this mutant completely metabolized high amounts



of the sterol nucleus (39.8% by moles), which considerably reduces the final conversion yield of the process (Xu *et al.*, 2015). This result strongly reinforces our proposal that *M. smegmatis* can be considered as a good cell factory to produce steroid synthons by metabolic engineering.

The finding that MS6039 and MS6039-5941 mutants can produce large amounts of ADD and AD from natural sterols, respectively, constitutes an experimental demonstration that metabolic engineering can be implemented in a model mycobacterial system like *M. smegmatis* mc<sup>2</sup>155 to generate pharmaceutical steroid synthons at industrial scale using a rational metabolic engineering approach. In fact, these engineered strains are already competing at industrial scale, i.e. under industrial operational conditions, with the existing strains isolated from environmental sources that were modified for their industrial use long time ago by conventional mutagenic procedures. Thus, our results lay the foundations for the valuation of *M. smegmatis* as a useful tool to develop *à la carte* engineered strains as cell factories to transform with high efficiency natural sterols into valuable pharmaceutical steroids.

## Experimental procedures

### Chemicals

Cholest-4-en-3-one was purchased from Fluka (Steinheim, Germany). ADD and AD were purchased from TCI

America. Chloroform and glycerol were purchased from Merck (Darmstadt, Germany). Cholesterol, N,O-bis(trimethylsilyl) trifluoroacetamide (BSTFA), gentamicin, pyridine, Tween 80 and tyloxapol were from Sigma (Steinheim, Germany). Oligonucleotides were from Sigma-Genosys.

### Bacterial strains and culture conditions

The strains as well as the plasmids used in this work are listed in Table 2. *M. smegmatis* mc<sup>2</sup>155 and its mutant strains were grown at 37°C in an orbital shaker at 200 r.p.m. Middlebrook 7H9 broth medium (Difco) containing 10% albumin-dextrose-catalase supplement (Becton Dickinson, New Jersey, USA) and 0.05% Tween 80 was used as rich medium. 7H10 agar (Difco) plates supplemented with albumin-dextrose-catalase were also used for solid media. Middlebrook 7H9 broth medium (Difco) without albumin-dextrose-catalase supplements containing 18 mM glycerol was used as a minimal medium. The minimal culture media used for the production of AD or ADD contained a mixture of 9 mM glycerol and 0.4 g l<sup>-1</sup> cholesterol, or a mixture of 9 mM glycerol and 0.4 g l<sup>-1</sup> of phytosterols. Commercial phytosterols (provided by Gadea Biopharma, Spain) contained a mixture of different sterols (w/w percentage): brassicasterol (2.16%), stigmasterol (8.7%), campesterol (36.8%) and  $\beta$ -sitosterol (54.4%). Cholesterol and phytosterols were

**Table 2.** Bacterial strains and plasmids used in this study.

Strains or plasmids	Genotype and/or description	Source or reference
<b>Strains</b>		
<i>Mycobacterium smegmatis</i>		
mc <sup>2</sup> 155	<i>ept-1</i> , mc <sup>2</sup> 6 mutant efficient for electroporation	Snapper <i>et al.</i> (1990)
MS6039	<i>M. smegmatis</i> mc <sup>2</sup> 155 $\Delta$ MSMEG_6039	This study
MS6039-5941	<i>M. smegmatis</i> mc <sup>2</sup> 155 $\Delta$ MSMEG_6039 $\Delta$ MSMEG_5941	This study
mc <sup>2</sup> 155 (pMV261)	mc <sup>2</sup> 155 strain harbouring plasmid pMV261	This study
MS6039 (pMV261)	MS6039 strain harbouring plasmid pMV261	This study
MS6039 (pMV6039)	MS6039 strain harbouring plasmid pMV6039	This study
<i>Escherichia coli</i>		
DH10B	F <sup>-</sup> , <i>mcrA</i> , $\Delta$ ( <i>mrrhsdRMS-mcrBC</i> ), $\Phi$ 80d <i>lacZ</i> $\Delta$ M15, $\Delta$ <i>lacX74</i> , <i>deoR</i> , <i>recA1</i> , <i>araD139</i> , $\Delta$ ( <i>ara-leu</i> )7697, <i>galU</i> , <i>galK</i> , $\lambda$ <sup>-</sup> , <i>rpsL</i> , <i>endA1</i> , <i>nupG</i>	Invitrogen
<b>Plasmids</b>		
pJQ200x	Suicide vector used to perform allelic exchange mutagenesis in <i>Mycobacterium</i> , P15A <i>ori</i> , <i>sacB</i> , Gm <sup>r</sup> , <i>xyIE</i>	Jackson <i>et al.</i> (2001)
pJQ6039	pJQ200x containing one fragment upstream and another downstream of MSMEG_6039 gene	This study
pJQ5941	pJQ200x containing one fragment upstream and another downstream of MSMEG_5941 gene	This study
pGEM <sup>®</sup> -T Easy	<i>E. coli</i> cloning vector; Amp <sup>r</sup>	Promega
pGEMT6039	pGEMT-Easy harbouring the MSMEG_6039 gene encoding KshB1 from <i>M. smegmatis</i> mc <sup>2</sup> 155	This study
pMV261	<i>Mycobacterium/E. coli</i> shuttle vector with the kanamycin resistance <i>aph</i> and the promoter from the <i>hsp60</i> gene from <i>M. tuberculosis</i>	Stover <i>et al.</i> (1991)
pMV6039	pMV261 harbouring the MSMEG_6039 gene encoding KshB1 from <i>M. smegmatis</i> mc <sup>2</sup> 155	This study

dissolved in 10% tyloxapol prior to its addition to the minimal medium when assayed in flasks. Due to the low solubility of cholesterol and phytosterols, stock solutions were warmed at 80°C in agitation, sonicated in a bath for 1 h and then autoclaved. Gentamicin (5 µg/ml) was used for selection of *M. smegmatis* mutant strains when appropriate.

Bioreactor experiments were performed using the culture media previously described (Herrington and Spasov, 2003) containing a vegetal oil to dissolve phytosterols and corn steep liquor as supplementary carbon and nitrogen sources. The bioconversion experiments were performed in stainless steel (20-l) or glass (2- and 5-l) jar bioreactors with efficient stirring. The entire process was performed at 37°C.

*Escherichia coli* DH10B strain was used as a host for cloning. It was grown in rich LB medium at 37°C in an orbital shaker at 200 r.p.m. LB agar plates were used for solid media. Gentamicin (10 µg ml<sup>-1</sup>), ampicillin (100 µg ml<sup>-1</sup>) or kanamycin (50 µg ml<sup>-1</sup>), were used for plasmid selection and maintenance in this strain.

#### Gene deletions

The knock-out strains of *M. smegmatis* named MS6039 and MS6039-5941 were constructed by homologous recombination using the pJQ200x plasmid, a derivative of the suicide pJQ200 vector that does not replicate in *Mycobacterium* (Jackson *et al.*, 2001). The strategy consist, for each gene, in generating two fragments of ~700 bp, the first one containing the upstream region and few nucleotides of the 5' end of the gene and the second one containing the downstream region and few nucleotides of the 3' end of the gene, that are amplified by PCR using the oligonucleotides described in Table 3 and *M. smegmatis* genomic DNA as template (Fig. S1). *M. smegmatis* genomic DNA extraction was performed as described (Uhía *et al.*, 2011). The two fragments generated were digested with the corresponding enzymes and cloned into the plasmid pJQ200x using *E. coli* DH10B competent cells as described (Uhía *et al.*, 2011).

Plasmid DNA from *E. coli* DH10B recombinant strains was extracted using the High Pure Plasmid Purification Kit (Roche, Basel, Switzerland), according to the manufacturer's instructions. This procedure was performed for genes *MSMEG\_6039* and *MSMEG\_5941*, generating the pJQ6039 and pJQ5941 plasmids respectively. Plasmid pJQ6039 was electroporated into competent *M. smegmatis* mc<sup>2</sup>155 to obtain strain MS6039. Plasmid pJQ5941 was electroporated into competent MS6039 cells to obtain MS6039-5941 strain. Single cross-overs were selected on 7H10 agar plates containing gentamicin and the presence of the *xylE* gene encoded in pJQ200x was confirmed by spreading catechol over the single colonies of electroporated *M. smegmatis*. The appearance of a yellow coloration indicates the presence of the *xylE* gene. Colonies were also contra-selected in 10% sucrose. A single colony was grown in 10 ml of 7H9 medium with 5 µg ml<sup>-1</sup> gentamicin up to an optical density of 0.8–0.9 and 20 µl of a 1:100 dilution was plated onto 7H10 agar plates with 10% sucrose to select for double cross-overs. Potential double cross-overs (sucrose-resistant colonies) were screened for gentamicin sensitivity and the absence of the *xylE* gene. The mutant strains MS6039 and MS6039-5941 were analysed by PCR and DNA sequencing to confirm the deletions of *MSMEG\_6039* and *MSMEG\_5941* genes.

#### Construction of pMV6039 plasmid

To isolate the *kshB1* gene from *M. smegmatis* mc<sup>2</sup> 155 genomic DNA was extracted and amplified by PCR using the primers MSMEG\_6039F (CGGAATTCTGACCTAAGGAGGTGAATGTGACTGATGAGCCCCTGGG) and MSMEG\_6039R (CGAAGCTTCTATTCGTCGTAGGTGAC TTCG). The amplified fragment of 1092 bp was cloned into the commercial plasmid pGEM<sup>®</sup>-T Easy to generate pGEMT6039 plasmid that was transformed in *E. coli* DH10B competent cells. The cloned fragment was further digested with *EcoRI* and *HindIII* to clone into pMV261, a shuttle plasmid that replicates in *E. coli* and *Mycobacterium* generating the plasmid pMV6039 that was

**Table 3.** Primers used in this study.

Primer	Sequence (5'–3') <sup>a</sup>	Use
MSMEG_6039 up F	ctagctcgcagccagttgtgcacaccgatg	<i>MSMEG_6039</i> deletion (amplification of upstream region)
MSMEG_6039 up R	ctagactagtcagcgcaggaactggc	<i>MSMEG_6039</i> deletion (amplification of upstream region)
MSMEG_6039 down F	ctagactagtggtgcacatggagatcaacg	<i>MSMEG_6039</i> deletion (amplification of downstream region)
MSMEG_6039 down R	ctagctagagtaccagtcgatcgggtgc	<i>MSMEG_6039</i> deletion (amplification of downstream region)
MSMEG_5941 up F	ctagactagtgatgtgcgaatgtcgtatg	<i>MSMEG_5941</i> deletion (amplification of upstream region)
MSMEG_5941 up R	ctagctagaccaccacaacgtctgtactcc	<i>MSMEG_5941</i> deletion (amplification of upstream region)
MSMEG_5941 down F	ctagctagaccatgacattcggttacctgg	<i>MSMEG_5941</i> deletion (amplification of downstream region)
MSMEG_5941 down R	ctaggagctcgcagggagatctcgaatcgc	<i>MSMEG_5941</i> deletion (amplification of downstream region)

a. Restriction sites are underlined.

transformed into *E. coli* DH10B to generate the recombinant strain *E. coli* DH10B (pMV6039). Once the sequence of the pMV6039 was checked it was used to transform electrocompetent cells of *M. smegmatis* MS6039 generating the *M. smegmatis* MS6039 (pMV6039) recombinant strain.

#### GC/MS analyses

To perform GC/MS analysis, culture aliquots (0.2 ml) were extracted twice at various extents of incubation with an equal volume of chloroform. Previously to its extraction, 10  $\mu$ l of a solution of 10 mM cholesterol (when phytosterols were used as substrate) or 10 mM cholestenone (when cholesterol was used as substrate) dissolved in chloroform were added to the aliquots as internal standards. The chloroform fraction was concentrated by evaporation and the trimethylsilyl ether derivatives were formed by reaction with 50  $\mu$ l of BSTFA and 50  $\mu$ l of pyridine and heating at 60°C for 45 min. Calibration standards were derivatized in the same way. The GC/MS analysis was carried out using an Agilent 7890A gas chromatograph coupled to an Agilent 5975C mass detector (Agilent Technologies, Palo Alto, CA, USA). Mass spectra were recorded in electron impact (EI) mode at 70 eV within the *m/z* range 50–550. The chromatograph was equipped with a 30 m  $\times$  0.25 mm i.d. capillary column (0.25\_μ film thickness) HP-5MS (5% diphenyl 95% dimethylpolysiloxane from Agilent Technologies). Working conditions in the sample were as follows: split ratio (20:1), injector temperature, 320°C; column temperature 240°C for 3 min, then heated to 320°C at 5°C min<sup>-1</sup>. For quantification of the peak area, the quantitative masses were 329 + 458 *m/z* for cholesterol; 343 + 384 *m/z* for cholestenone; 382 + 472 *m/z* for campesterol; 394 + 484 *m/z* for stigmastanol; 357 + 486 *m/z* for  $\beta$ -sitosterol; 244 + 286 *m/z* for AD and 122 + 284 *m/z* for ADD in selected ions of monitoring. EI mass spectra and retention data were used to assess the identity of compounds by comparing them with those of standards in the NIST Mass Spectral Database and commercial standards (NIST 2011).

#### High-performance liquid chromatography analyses

To perform high-performance liquid chromatography (HPLC) analysis, samples (1 g) were withdrawn and extracted with 10 mL of ethyl acetate during two hours in a magnetic stirrer. One aliquot was centrifugated 10 min at 10 000 r.p.m. and supernatant was diluted 1:10 in acetonitrile. Samples were filtered (0.2  $\mu$ m pore size) prior to the chromatographic analysis. The HPLC analysis was carried out using a Waters liquid chromatograph with a PDA detector system. A Phenomenex

C18 column (Nucleosil C18, 100 Å, 250  $\times$  4.6 mm, 5  $\mu$ m particles) was employed and the temperature of the column was fixed in 50°C. Working conditions were as follows: the mobile phase was a mixture water: acetonitrile: acetic acid (48: 52: 0.1 v/v); flow rate was 1.1 mL min<sup>-1</sup>; the injection volumen was 10  $\mu$ l. Peaks were monitorized at 240 nm and calibrations were performed using highly purified standards of each compound.

#### Conflict of interest

None declared.

#### References

- Andor, A., Jekkel, A., Hopwood, D.A., Jeanplong, F., Ilkoy, E., Kónya, A., *et al.* (2006) Generation of useful insertionally blocked sterol degradation pathway mutants of fast-growing mycobacteria and cloning, characterization, and expression of the terminal oxygenase of the 3-ketosteroid 9 $\alpha$ -hydroxylase in *Mycobacterium smegmatis* mc<sup>2</sup>155. *Appl Environ Microbiol* **72**: 6554–6559.
- Blank, L.M., Ebert, B.E., Buehler, K., and Bühler, B. (2010) Redox biocatalysis and metabolism: molecular mechanism and metabolic network analysis. *Antioxid Redox Signal* **13**: 349–394.
- Bragin, E.Y., Shtratnikov, V.Y., Dovbnya, D.V., Schelkunov, M.I., Pekov, Y.A., Malakho, S.G., *et al.* (2013) Comparative analysis of genes encoding key steroid core oxidation enzymes in fast-growing *Mycobacterium* spp. strains. *J Steroid Biochem Mol Biol* **138**: 41–53.
- Brzostek, A., Sliwiński, T., Rumijowska-Galewicz, A., Korycka-Machała, M., and Dziadek, J. (2005) Identification and targeted disruption of the gene encoding the main 3-ketosteroid dehydrogenase in *Mycobacterium smegmatis*. *Microbiology* **151**: 2393–2402.
- Brzostek, A., Pawelczyk, J., Rumijowska-Galewicz, A., Dziadek, B., and Dziadek, J. (2009) *Mycobacterium tuberculosis* is able to accumulate and utilize cholesterol. *J Bacteriol* **191**: 6584–6591.
- Capyk, J.K., D'Angelo, I., Strynadka, N.C., and Eltis, L.D. (2009) Characterization of 3-ketosteroid 9 $\alpha$ -hydroxylase, a Rieske oxygenase in the cholesterol degradation pathway of *Mycobacterium tuberculosis*. *J Biol Chem* **284**: 9937–9946.
- Capyk, J.K., Casabon, I., Gruninger, R., Strynadka, N.C., and Eltis, L.D. (2011) Activity of 3-ketosteroid 9 $\alpha$ -hydroxylase (KshAB) indicates cholesterol side chain and ring degradation occur simultaneously in *Mycobacterium tuberculosis*. *J Biol Chem* **286**: 40717–40724.
- Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., *et al.* (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**: 537–544. Erratum in: *Nature* **396**: 190.
- Donova, M.V. (2007) Transformation of steroids by Actinobacteria: a review. *Appl Biochem Microbiol* **43**: 1–14.

- Donova, M.V., and Egorova, O.V. (2012) Microbial steroid transformations: current state and prospects. *Appl Microbiol Biotechnol* **94**: 1423–1447.
- Donova, M.V., Gulevskaya, S.A., Dovbnaya, D.V., and Puntus, I.F. (2005) *Mycobacterium* sp. mutant strain producing 9 $\alpha$ -hydroxyandrostenedione from sitosterol. *Appl Microbiol Biotechnol* **67**: 671–678.
- Fernández de las Heras, L., van der Geize, R., Drzyzga, O., Perera, J., and Navarro-Llorens, J.M. (2012) Molecular characterization of three 3-ketosteroid- $\Delta^1$ -dehydrogenase isoenzymes of *Rhodococcus ruber* strain Chol-4. *J Steroid Biochem Mol Biol* **132**: 271–281.
- Galán, B., Díaz, E., Prieto, M.A., and García, J.L. (2000) Functional analysis of the small component of the 4-hydroxyphenylacetate 3-monooxygenase of *Escherichia coli* W: a prototype of a new Flavin:NAD(P)H reductase subfamily. *J Bacteriol* **182**: 627–636.
- García, J.L., Uhía, I., and Galán, B. (2012) Catabolism and biotechnological applications of cholesterol degrading bacteria. *Microbial Biotechnol* **5**: 679–699.
- van der Geize, R., Hessels, G.I., and Dijkhuizen, L. (2000) 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.
- van der Geize, R., Hessels, G. and Dijkhuizen, L. (2001a) Preparing genetically modified organism unable to degrade the steroid nucleus, useful for producing corticosteroid intermediates, comprises inactivating the dehydrogenase genes. Patent WO 200131050.
- van der Geize, R., Hessels, G.I., van Gerwen, R., van der Meijden, P., and Dijkhuizen, L. (2001b) 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.
- van der Geize, R., Hessels, G.I., van Gerwen, R., van der Meijden, P., and Dijkhuizen, L. (2002a) 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.
- van der Geize, R., Hessels, G.I., and Dijkhuizen, L. (2002b) Molecular and functional characterization of the *kstD2* gene of *Rhodococcus erythropolis* SQ1 encoding a second 3-ketosteroid  $\Delta^1$ -dehydrogenase isoenzyme. *Microbiology* **148**: 3285–3292.
- van der Geize, R., Hessels, G.I., Nienhuis-Kuiper, M., and Dijkhuizen, L. (2008) 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.
- Herrington, E.J. and Spassov, G. (2003) Process for fermentation of phytosterols to androstadienedione. Patent WO 2003/064674.
- Hu, Y., van der Geize, R., Besra, G.S., Gurucha, S.S., Liu, A., Rohde, M., et al. (2010) 3-ketosteroid 9 $\alpha$ -hydroxylase is an essential factor in the pathogenesis of *Mycobacterium tuberculosis*. *Mol Microbiol* **75**: 107–121.
- Jackson, M., Camacho, L.R., Gicquel, B., and Guilhot, C. (2001) Gene replacement and transposon delivery using the negative selection marker *sacB*. In *Mycobacterium Tuberculosis Protocols*, Vol. **54**, Parish, T., and Stocker, N.G. (eds). Totowa, NJ: Humana Press, pp. 59–75.
- Kendall, S.L., Withers, M., Soffair, C.N., Moreland, N.J., Gurucha, S., Sidders, B., et al. (2007) A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *Mol Microbiol* **65**: 684–699.
- Malaviya, A., and Gomes, J. (2008) Androstenedione production by biotransformation of phytosterols. *Bioresour Technol* **99**: 6725–6737.
- Marsheck, W., Kraychy, S., and Muir, R. (1972) Microbial degradation of sterols. *Appl Microbiol* **23**: 72–77.
- Penfield, J.S., Worrall, L.J., Strynadka, N.C., and Eltis, L.D. (2014) Substrate specificities and conformational flexibility of 3-ketosteroid 9 $\alpha$ -hydroxylases. *J Biol Chem* **289**: 25523–25536.
- Petrusma, M., Dijkhuizen, L., and van der Geize, R. (2009) *Rhodococcus rhodochrous* DSM 43269 3-ketosteroid 9 $\alpha$ -hydroxylase, a two-component iron-sulfur-containing monooxygenase with subtle steroid substrate specificity. *Appl Environ Microbiol* **75**: 5300–5307.
- Petrusma, M., Hessels, G., Dijkhuizen, L., and van der Geize, R. (2011) Multiplicity of 3-ketosteroid-9 $\alpha$ -hydroxylase enzymes in *Rhodococcus rhodochrous* DSM 43269 for specific degradation of different classes of steroids. *J Bacteriol* **193**: 3931–3940.
- Petrusma, M., Dijkhuizen, L., and van der Geize, R. (2012) Structural features in the KshA terminal oxygenase protein that determine substrate preference of 3-ketosteroid 9 $\alpha$ -hydroxylase enzymes. *J Bacteriol* **194**: 115–121.
- Petrusma, M., van der Geize, R., and Dijkhuizen, L. (2014) 3-Ketosteroid 9 $\alpha$ -hydroxylase enzymes: Rieske non-heme monooxygenases essential for bacterial steroid degradation. *Antonie Van Leeuwenhoek* **106**: 157–172.
- Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T., and Jacobs, W.R. Jr (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**: 1911–1919.
- Stover C.K., de la Cruz V.F., Fuerst T.R., Burlein J.E., Benson L.A., Bennett L.T., et al. (1991) New use of BCG for recombinant vaccines. *Nature* **351**: 456–460.
- Szentirmai, A. (1990) Microbial physiology of sidechain degradation of sterols. *J Ind Microbiol* **6**: 101–115.
- Uhía, I., Galán, B., Morales, V., and García, J.L. (2011) Initial step in the catabolism of cholesterol by *Mycobacterium smegmatis* mc<sup>2</sup>155. *Environ Microbiol* **13**: 943–959.
- Uhía, I., Galán, B., Kendall, S.L., Stoker, N.G., and García, J.L. (2012) Cholesterol metabolism in *Mycobacterium smegmatis*. *Environ Microbiol Rep* **4**: 168–182.
- Wei, W., Wang, F.Q., Fan, S.Y., and Wei, D.Z. (2010) 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.
- Wilbrink, M.H., Petrusma, M., Dijkhuizen, L., and van der Geize, R. (2011) FadD19 of *Rhodococcus rhodochrous*

- DSM 43269, a steroid-Coenzyme A ligase essential for degradation of C-24 branched sterol side chains. *Microbiology* **158**: 3054–3062.
- Xu, X.W., Gao, X.Q., Feng, J.X., Wang, X.D., and Wei, D.Z. (2015) Influence of temperature on nucleus degradation of 4-androstene-3, 17-dione in phytosterol biotransformation by *Mycobacterium* sp. *Lett Appl Microbiol* **61**: 63–68.
- Xu, L.Q., Liu, Y.J., Yao, K., Liu, H.H., Tao, X.Y., Wang, F.Q., et al. (2016) Unraveling and engineering the production of 23,24-bisnorcholenic steroids in sterol metabolism. *Sci Rep* **6**: 21928.
- Yao, K., Xu, L.Q., Wang, F.Q., and Wei, D.Z. (2014) 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.
- Yeh, C.H., Kuo, Y.S., Chang, C.M., Liu, W.H., Sheu, M.L., and Meng, M. (2014) Deletion of the gene encoding the reductase component of 3-ketosteroid 9 $\alpha$ -hydroxylase in *Rhodococcus equi* USA-18 disrupts sterol catabolism, leading to the accumulation of 3-oxo-23,24-bisnorchola-1,4-dien-22-oic acid and 1,4-androstadiene-3,17-dione. *Microb Cell Fact* **13**: 130.

### Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

**Fig. S1.** Construction of the mutant strains MS6039 and MS6039-5941. The polylinker restriction sites of the suicide plasmid pJQ200x are indicated (B, *Bam*HI; Sp, *Spe*I; X, *Xba*I; N, *Not*I; Bs, *Bst*XI; S, *Sac*I).

**Fig. S2.** Production of ADD from phytosterols by the MS6039 mutant in 5-L jar bioreactor. Analysis by HPLC of the transformation products at 120 h of culture. (1) solvent front; (2) ADD; (3) AD; (4) 1,4-HBC.

**Fig. S3.** Production of AD from phytosterols by the MS6039-5941 mutant in 2-L jar bioreactor. Analysis by GC/MS of the transformation products at 96 h of culture. (1) AD; (2) ADD; (3) 4- HBC; (4) 1,4-HBC; (5) campesterol; (6) stigmaterol; (7)  $\beta$ -sitosterol.