Systems Microbiology and Biomanufacturing (2024) 4:127–137
https://doi.org/10.1007/s43393-022-00103-w

ORIGINAL ARTICLE

Whole-genome analyses and metabolic modification of *Mycobacterium* sp. LY-1 to enhance yield of 9α-OH-AD

W. Liu[1] · H. Li[2] · J. X. Zhang[2] · Y. N. Xu[2] · X. M. Zhang[2] · J. S. Shi[2] · M. A. G. Koffas[3] · Z. H. Xu[1,4,5]

Received: 19 February 2022 / Revised: 13 April 2022 / Accepted: 17 April 2022 / Published online: 11 May 2022

Abstract

With the increasing application of steroid drugs as therapeutics, the demand for steroid drugs is increasing. In recent years, biological synthesis has become the standard approach to produce steroid intermediates, while this method still faces some problems such as unclear metabolic pathway and low yield. *Mycobacterium* sp. LY-1 can convert phytosterols into 9α-hydroxyandrost-4-ene-3,17-dione (9α-OH-AD) which is a key intermediate for the synthesis of steroid drugs with long effective time and significant pharmacological activity. In this work, the whole-genome sequence of the *Mycobacterium* sp. LY-1 was analyzed, and the side-chain degradation pathway of phytosterols in *Mycobacterium* sp. LY-1 was proposed. Meanwhile, the related key enzymes of phytosterol metabolism were identified through qRT-PCR. Through overexpressing the key enzymes including KshA2, KshB, and HsdB, the yield of 9α-OH-AD increased by 12.7% compared to that of the control. Furthermore, by optimizing the medium and culture conditions, the yield of 9α-OH-AD reached 50.4%. The maximum yield was 30.7% higher than that of the original strain. The results are of significance for the industrial production of 9α-OH-AD using metabolic engineering methods.

Keywords *Mycobacterium* sp. LY-1 · Phytosterols · 9α-Hydroxyandrost-4-ene-3,17-dione · Biotransformation

Introduction

Steroid drugs are the second largest class of drugs in the world after antibiotics [1, 2]. In recent years, the clinical application of steroid drugs has shown an upward trend [3]. They are widely used to treat tumors [4], arthritis, asthma, and regulate hormone levels in the body [5, 6]. The oxidation state of the steroid nucleus and the attached functional groups determine the specific biological characteristics of each steroid compound [7]. 9α-Hydroxyandrost-4-ene-3,17-dione (9α-OH-AD) is an important steroidal drug intermediate which is difficult to obtain by the conventional chemical synthesis [8], with a special hydroxyl group at the C9 position of its structure. The C9 hydroxyl group can be used for the synthesis of corticosteroids after being chemically substituted by halogen atoms [9]. The corticosteroids produced by 9α-OH-AD have a long action time and significant pharmacological functions [10, 11]. In the mid-1970s, due to the gradual increase in the price of diosgenin, some pharmaceutical companies and research institutes turned to research new raw materials such as phytosterols. Phytosterols are natural sterols extracted from by-products of oil crops [12]. Using them as precursors, steroidal drug intermediates can be prepared by biotransformation of *Mycobacterium* [13]. *Mycobacterium* have been used to produce steroid drug intermediates on an industrial scale. For example, *Mycobacterium smegmatis* was engineered to produce AD and ADD [14], while *Mycobacterium neoaurum* ATCC 25795 was engineered to produce 9α-OH-AD [15]. The microbial transformation method has the advantages of less

✉ H. Li
lihui@jiangnan.edu.cn

✉ M. A. G. Koffas
koffam@rpi.edu

1 School of Biotechnology, Jiangnan University, Wuxi 214122, China

2 School of Life Sciences and Health Engineering, Jiangnan University, Wuxi 214122, China

3 Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

4 National Engineering Laboratory for Cereal Fermentation Technology, Jiangnan University, Wuxi 214122, People's Republic of China

5 Jiangsu Provincial Research Center for Bioactive Product Processing Technology, Jiangnan University, Wuxi 214122, People's Republic of China

environmental pollution and fewer reaction steps, and thus, the current production of 9α-OH-AD mainly depends on the microbial transformation method [16, 17].

Previously, Shtratnikova et al. [18] reported the genome of a *Mycobacterium* sp. that could transform phytosterols to produce 9α-OH-AD for the first time. In 2016, Luthra et al. [19] improved the transformation efficiency of 9α-OH-AD by optimizing the medium resulting in a maximum titer of 9α-OH-AD of 9.10 mg/g. *Mycobacterium* sp. LY-1 that can transform phytosterols to produce 9α-OH-AD was isolated in our laboratory. Through phylogenetic tree analysis, *Mycobacterium* sp. LY-1 is the most closely related to *M. neoaurum* VKM Ac-1817D, which was reported to be a non-pathogenic mycobacteria. Zhou et al. [20] improved the solubility of the substrate phytosterols using the nonionic surfactant TX-40 and the yield of 9α-OH-AD reached 42.5% in 2019. Although methods such as media optimization and substrate solubilization can increase the yield of 9α-OH-AD, there are still problems such as low conversion efficiency and low substrate utilization.

Metabolic engineering [21, 22] or using resting cell catalysis [23, 24] are methods that have been used to resolve some of these problems; however, the microbial degradation pathways of phytosterols in some strains such as *Mycobacterium* sp. LY-1 are not clear, and the function of key enzymes has yet to be elucidated [25]. Therefore, it is particularly important to analyze the sterol metabolism in this strain, identify the key enzymes that affect the synthesis of the target product, and further improve the production efficiency of the target product through metabolic engineering methods. In the present study, to improve the ability of the strain to produce 9α-OH-AD, the whole genome of *Mycobacterium* sp. LY-1 was analyzed and the genes related to phytosterols metabolic pathway were identified. Thus, the putative sterol metabolic pathway of *Mycobacterium* sp. LY-1 was predicted. Differential transcription-level analysis was used to study the changes of metabolic pathway-related enzymes under different conditions with or without soya bean oil, and the key enzymes affecting the synthesis of 9α-OH-AD were determined. Furthermore, metabolic engineering methods were used to enhance the expression of single or multiple enzyme genes to construct genetically engineered recombinant strain with high yield of 9α-OH-AD. Combined with the optimization of the transformation process, the yield of 9α-OH-AD was further increased.

Materials and methods

Strains and plasmids

All strains and plasmids used in this study are shown in Table 1. The microorganism used is *Mycobacterium* sp. LY-1. This strain was stored in the China General Microbiological Culture Collection Center with the depositing No. 13031. *Escherichia coli* JM109 was used for plasmid construction. Plasmid pMV261 was used to overexpress in *Mycobacterium* sp. LY-1.

Medium and growth conditions

Phytosterols (β-sitosterol 47.0%, campesterol 24.6%, stigmasterol 15.5%, and brassicasterol 3.4%) was purchased from Hubei Jusheng Technology Co., Ltd., China; 9α-OH-AD was purchased from Shanghai Hanxiang Biotechnology (98.0% of purity), China; Soybean oil was purchased from Shanghai Kerry Food Industries Co., Ltd, China; Corn steep liquor was purchased from Anhui Huaheng Biotechnology Co., Ltd., China [20].

The $(NH_4)_2HPO_4$ (0.6 g/L), $NaNO_3$ (5.4 g/L), yeast extract (15 g/L), and glycerin (2%) medium was the seed medium. *Mycobacterium* sp. LY-1 was cultivated at 30 °C and 120 rpm for 3–5 days. Then, 1% seed culture was transferred into 50 mL of fermentation medium ($NaNO_3$ 5.4 g/L, $(NH_4)_2HPO_4$ 0.6 g/L, corn steep liquor 20.0 g/L, phytosterols 15 g/L) in 250-mL shake flasks and cultivated at 120 rpm and 30 °C for 168 h.

High-throughput sequencing and accession numbers

Genomic DNA was sheared, and then, 20 Kb double-stranded DNA fragments were selected. DNA fragments were end repaired and ligated with universal hairpin adapters. Subsequent steps were followed as per the manufacturer's instruction to prepare SMRTbell library. The library was sequenced in PacBio RSII SMRT instrument, and then, the PacBio reads were assembled using PBcR of WGS-Assembler 8.2. The Glimmer gene-finding software has been used for finding coding genes in bacteria. Transfer RNAs (tRNAs) were detected in the genome using the program tRNAscan-SE with default parameter settings. rRNA were identified using RNAmmer. The coding genes were annotated with National Center for Biotechnology Information (NCBI) nr database by BLAST. Then, the functions of genes were annotated by GO (Gene Ontology) database, and the pathways were annotated using KEGG (Kyoto Encyclopedia of Genes and Genomes) database. The proteins encoded by genes were classificated on a phylogenetic classification by the database of COG (Clusters of Orthologous Groups).

The whole-genome sequence data reported in this paper have been deposited in the *Genome Warehouse in National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation*, under accession number GWHBHSJ00000000 that is publicly accessible at https://ngdc.cncb.ac.cn/gwh.

Table 1 Bacterial strains and plasmids used in this study

Strain/plasmid	Description	Source
Strains		
Escherichia coli		
E.coli JM109	*recA1, endA1, gyrA96, thi-1, hsdR17(rk− mk+)supE44*	Invitrogen
Mycobacterium sp.		
Mycobacterium sp. LY-1	Strain for production of 9α-OH-AD	Preliminary research
M. LY-1-KshA2	*M*. LY-1 harboring the plasmid pMV261-KshA2	This study
M. LY-1-KshB	*M*. LY-1 harboring the plasmid pMV261-KshB	This study
M. LY-1-Cho	*M*. LY-1 harboring the plasmid pMV261-Cho	This study
M. LY-1-HsdB	*M*. LY-1 harboring the plasmid pMV261-HsdB	This study
M. LY-1-FadE30	*M*. LY-1 harboring the plasmid pMV261-FadE30	This study
M. LY-1-Hsd1	*M*. LY-1 harboring the plasmid pMV261-Hsd1	This study
M. LY-1-FadA5	*M*. LY-1 harboring the plasmid pMV261-FadA5	This study
M. LY-1-KshA2/KshB	*M*. LY-1 harboring the plasmid pMV261-KshA2/KshB	This study
M. LY-1-KshA2/HsdB	*M*. LY-1 harboring the plasmid pMV261-KshA2/HsdB	This study
M. LY-1-KshA2/Hsd1	*M*. LY-1 harboring the plasmid pMV261-KshA2/Hsd1	This study
M. LY-1-KshA2/KshB/HsdB	*M*. LY-1 harboring the plasmid pMV261-KshA2/KshB/HsdB	This study
Plasmids		
pMD19-T	*E. coli* cloning vector, Amp^R	TaKaRa
pMV261	*E. coli-Mycobacterium* sp. *shuttle* vector with heat shock promoter hsp60, Kan^R	YouBio
pMV261-Cho	Cloning *cho* gene into pMV261	This study
pMV261-KshB	Cloning *kshB* gene into pMV261	This study
pMV261-HsdB	Cloning *hsdB* gene into pMV261	This study
pMV261-KshA2	Cloning *kshA2* gene into pMV261	This study
pMV261-FadE30	Cloning *fadE30* gene into pMV261	This study
pMV261-Hsd1	Cloning *hsd1* gene into pMV261	This study
pMV261-FadA5	Cloning *fadA5* gene into pMV261	This study
pMV261-KshA2/KshB	Cloning *kshA2* and *kshB* gene into pMV261	This study
pMV261-KshA2/HsdB	Cloning *kshA2* and *hsdB* gene into pMV261	This study
pMV261-KshA2/Hsd1	Cloning *kshA2* and *hsd1* gene into pMV261	This study
pMV261-KshA2/KshB/HsdB	Cloning *kshA2/kshB/hsdB* gene into pMV261	This study

RNA extraction and quantitative real-time PCR analysis

For RNA extraction, we centrifuged the fermentation broth samples at different fermentation periods at 8000 rpm for 10 min, removed the supernatant, and washed the sample 2–3 times with PBS (pH = 7.0), and stored it at −80 °C. The frozen bacteria were then poured into a pre-cooled mortar and grind, where they were grinded with the addition of liquid nitrogen. Afterward, 1 mL of Trizol was added and the liquid was transferred to a 2 mL sterile EP tube. Total RNA was purified using a MasterPure™ RNA purification kit (Shanghai sangon, China). The first-strand cDNAs were synthesized by HisScript R II Q RT SuperMix (Takara, China) using total RNA (200 ng). For RT-qPCR, the 16S rRNA was chosen as the endogenous control. The reaction mixture was prepared in a qPCR tube and consisted of 10 μL of 2 × Power qPCR PreMix, 4 μL of forward and reverse primers (0.2 mM) respectively, 2 μL of cDNA, 0.4 μL 50 × Rox Reference Dye, and 3.6 μL of ddH_2O. The RT-qPCR was performed on a Bio-Rad CFX96 Manager PCR system (Bio-Rad, United States) using the following parameters: 95 °C for 10 min followed by 40 cycles of denaturation at 95 °C for 10 s and 55 °C for 20 s and extension at 72 °C for 15 s. The $2^{-\Delta\Delta Ct}$ method was applied to analyze the data, which were normalized to the transcription level of 16S rRNA.

Plasmid construction and electroporation

Genes were cloned from the genome of *Mycobacterium* sp. LY-1 using f1&r1, PrimeSTAR enzyme for PCR amplification. The PCR products were run on an agarose gel and purified according to a kit instructions (Generay, China). Purified fragments were digested with enzymes *Hind* III–*BamH*

I and ligated into plasmid pMV261at the *Hind* III–*BamH* I site. The ligation product was transformed into *E. coli* JM109 competent cells. As for polygenic overexpression, the one Step Cloning Kit (Vazyme, China) was used to construct the plasmids.

For electroporation, 2 μg of the expression plasmid was added to *Mycobacterium* sp. LY-1 competent cells and left at 4 °C for 20–30 min. We used 2.5 kV as electroporation voltage and electric shock frequency of 5 ms; two electric shocks were applied followed by the addition of fresh seed medium and incubation at 30 °C for 3–4 h. We then collected the bacteria and spread it on resistant solid medium and cultured at 30 °C in a constant temperature incubator for 5–7 days until a clearly visible yellow colony was formed.

Analytical methods

During fermentation, we pipetted 1 mL of fermentation broth into a sterile 2 mL EP tube every 12 h, followed by centrifugation at 12,000 rpm for 10 min. We then added 1 mL of ethyl acetate, incubated with shaking at 1500 rpm for 30 min, and then centrifuged for 15 min. After repeating the above operation twice, we dried the EP tube to constant weight, and calculated the difference after weighing to obtain the biomass.

For high-performance liquid chromatography (HPLC), samples were dissolved in 2 mL ethyl acetate and were transferred into a clean tube. The extraction operation previously described was repeated five times, and the ethyl acetate extracts were combined. We concentrated the resulting extract with a nitrogen blower at room temperature and then added 4 mL of acetonitrile to re-dissolve it, followed by filtration through 0.22 μm of microporous membrane. HPLC was performed on an Agilent TC-C18 column (4.6 × 250 mm, 30 °C) using acetonitrile/water (7:3, v/v) as the mobile phase at a flow rate of 0.5 mL/min with ultraviolet detection at 254 nm and an injection volume of 10 μL. Peaks were compared to an internal standard of 9α-OH-AD to determine the exact amount of product during the fermentation [20].

The product yield was calculated using the following formula:

$$C_c = \frac{P_c \times C_1 \times D}{P_1}$$

$$\text{Yield}\% = \frac{C_c \times M_d}{C_d \times M_c} \times 100\%;$$

C_1 standard concentration (g/L); C_d phytosterols concentration (g/L); C_c product concentration (g/L); P_1 the peak area of standard analyzed by HPLC; P_c the peak area of product analyzed by HPLC; D sample dilution multiple; M_c product molar mass; M_d phytosterols molar mass [20].

Results and discussion

Genome sequence analysis of *Mycobacterium* sp. LY-1

Based on comparative genomic analysis, preliminary genome sequence analysis of *Mycobacterium* sp. LY-1 was conducted (Table 2). The genome size of *Mycobacterium* sp. LY-1 was 6,320,069 bp. Through gene prediction analysis, the total length of *Mycobacterium* sp. LY-1 coding gene sequence was 5,877,051 bp, which contained 6129 genes, 54 tRNAs, 25 rRNAs, and 33 other RNAs. The coding sequence accounted for 93.0% of the entire genome, the average length of the gene was 959 bp, and the GC content was 66.5%.

Figure 1 is a phylogenetic analysis of *Mycobacterium* sp. LY-1. Nine *Mycobacterium* sp. genome sequences published on National Coalition Building Institute (NCBI) were selected to construct a phylogenetic tree based on the whole-genome sequence. The phylogenetic tree indicated that the genome features of *Mycobacterium* sp. LY-1 were close to two reference strains. One was *M. neoaurum* VKM Ac-1817D which had been reported to transform phytosterols to 9α-OH-AD and the other was *M. tuberculosis* H37Rv [26] which had been extensively studied.

Protein functions and sequences encoded by *Mycobacterium* sp. LY-1 genes were annotated based on the Cluster of Orthologous Groups (COG) database (Fig. 2). All open-reading frames in *Mycobacterium* sp. LY-1 could be divided into 1 unknown functional group and 20 functional groups, resulting in a total of 21 groups. The results indicated that the group with the most protein was Function unknown S (Function unknown), which accounted for 19.0% (Fig. S1). This result also illustrated the exploitability of this strain. In addition, the comprehensive analysis showed that the number of *Mycobacterium* sp. LY-1 in the Q, M, K, C, and E categories was more than

Table 2 General prediction of genome of *Mycobacterium* sp. LY-1

Total bases (bp)	6,320,069
Gene sequence length (bp)	5,877,051
Total number of genes	6129
Number of tRNA	54
Number of rRNA	25
Number of other RNAs	33
Gene coverage ratio (%)	93.0%
Average gene length (bp)	959
GC content (%)	66.5

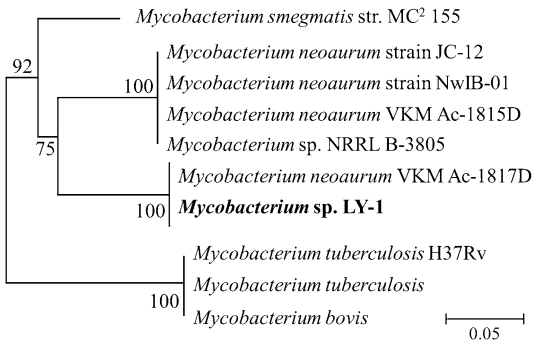

Fig. 1 Phylogenetic analysis of *Mycobacterium* sp. strains

the number in the A, B, W, and Z categories. The gene sequence prediction and functional annotation indicated that the functional proteins of *Mycobacterium* sp. LY-1 were mainly distributed in functional categories such as [Q] (secondary metabolism synthesis and metabolism), [C] (energy production and transport), and [K] (transcription). This result may indicate that strong secondary metabolism was accompanied by a large amount of energy production. This is consistent with the observation that there is a large amount of ATP during the transformation of phytosterols by mycobacteria [27]. At the same time, the functional proteins of *Mycobacterium* sp. LY-1 also accounted for a large proportion of [E] (amino acid transport and metabolism) and [K] (transcription). This may illustrate that *Mycobacterium* sp. LY-1 is a suitable host for expressing protein.

The side-chain degradation pathway of phytosterols of in *Mycobacterium* sp. LY-1

The protein sequences of *Mycobacterium* sp. LY-1 were compared and analyzed with the gene clusters of the sterol metabolic pathway in *M. neoaurum* VKM Ac-1817D and *M. tuberculosis* H37Rv (Table S2). According to the four stages of sterol metabolism, multiple sterol metabolism-related enzyme genes were found in *Mycobacterium* sp. LY-1, including one cholesterol oxidase (Cho) and three 3-hydroxyacyl-CoA dehydrogenases (Hsd) which were responsible for preliminary oxidation of sterols, one enoyl CoA thiolase (FadA5) and five acyl-CoA dehydrogenases (FadE26/27/28/29/30) which were responsible for side-chain degradation of sterols, one 2-Enoyl acyl-CoA hydratase (HsdB), five 3-sterone-9α-hydroxyoxygenase (KshA), one 3-sterone-9α-hydroxyreductase (KshB), and five 3-steroid Ketone-Δ1-dehydrogenases/3-Oxosteroid Δ1-dehydrogenases (KstD1/2/3/4/5) which were responsible for nucleus oxidation of sterols. Compared with *M. tuberculosis* H37Rv, *Mycobacterium* sp. LY-1 had more homologues of the genes responsible for nucleus oxidation of sterols, especially more copies of *KshA*. That may be the reason why *Mycobacterium* sp. LY-1 accumulated 9α-OH-AD. In addition, five 2,3-Dihydroxybiphenyl/1,2-Dioxygenases (HsaC1/2/3/4/5) which were responsible for nucleus cleavage of sterols were also identified (Table 3).

Based on the above results, combined with the previous research work, a putative phytosterol metabolic pathway in the *Mycobacterium* sp. LY-1 was predicted (Fig. 3). The results showed that phytosterols first underwent oxidative

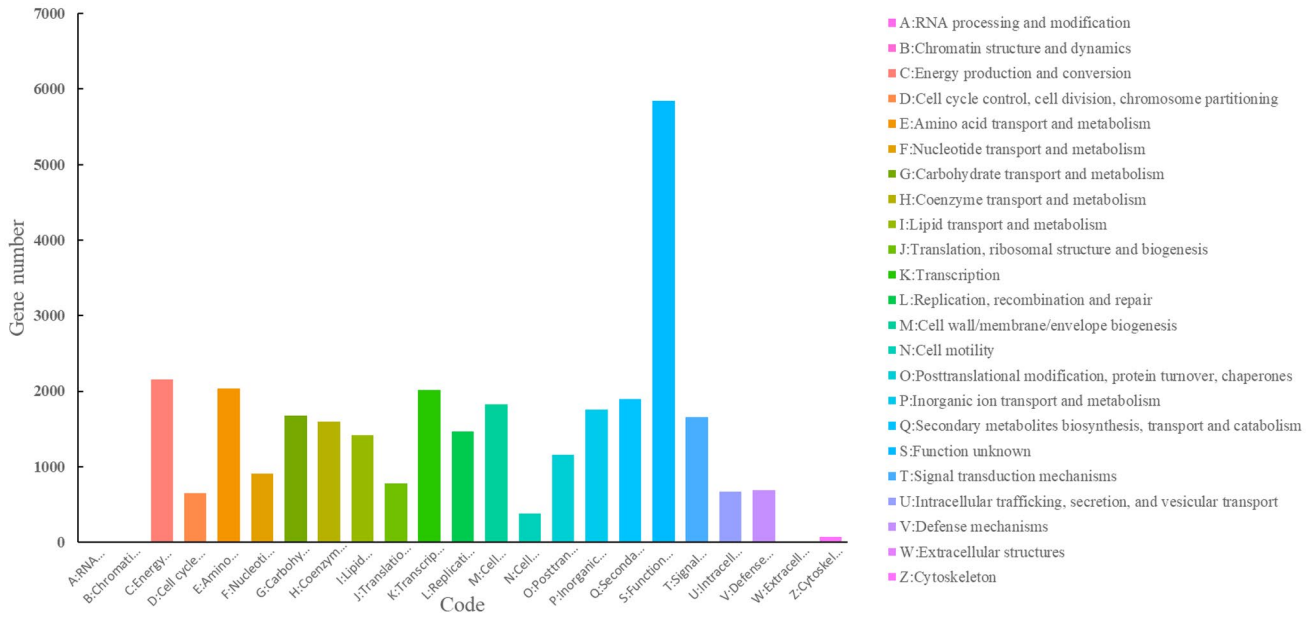

Fig. 2 COG function classification of *Mycobacterium* sp. LY-1

dehydrogenation by cholesterol oxidase, and then entered the complex sterol degradation pathway, and finally resulted in the target product 9α-OH-AD. Through the identification of the product composition and structure, we found that most products (> 98.8%) were C9 hydroxylation ones, which demonstrated that *Mycobacterium* sp. LY-1 had a strong C9 hydroxylation ability for sterols (Fig.S2). According to the previous reports, the enzyme for sterol C9 hydroxylation is 3-ketosteroid-9α-hydroxylase, and the process of hydroxylation requires KshA and KshB to act synergistically [28]. On the basis of the genome sequence analysis, there were five *kshA* genes and one *kshB* gene of 9 hydroxylation oxygen isoenzymes in *Mycobacterium* sp. LY-1. Therefore, we speculated that there may be a special 9α-OH-AD accumulation pathway. During the sterol metabolism of *Mycobacterium* sp. LY-1 (Fig. 3, in blue), after the oxidation of phytosterols, the cholestenone obtained is first hydroxylated by KshA and KshB, introducing a high metabolic flux into this pathway. 9α-OH-cholestenone was obtained by KshA and KshB co-catalysis, and then, 9α-OH-AD was directly obtained by side-chain degradation. Even if a small amount of metabolic flux was introduced into the production pathway of AD and ADD, they will be metabolized by the more active KshA and KshB into the corresponding 9-position hydroxylated products. Therefore, it may be one of the reasons why *Mycobacterium* sp. LY-1 accumulated a large amount of 9α-OH-AD.

Identification of the related key enzymes in sterol metabolic pathway of *Mycobacterium* sp. LY-1

In the conversion process of *Mycobacterium* sp. LY-1, adding soybean oil could improve the solubility of sterols and the vitality of the bacteria [29, 30]. Our previous studies indicated that in the presence of soybean oil, the ability of *Mycobacterium* sp. LY-1 to transform phytosterols was significantly improved. Figure 4 shows the fermentation curve of the strain with and without soybean oil. The maximum yield of 9α-OH-AD was increased from 10.03% to 35.42% and the biomass was also improved. To explore whether this result was related to the difference in the expression of partial enzyme genes in the sterol metabolic pathway, the difference in transcription level of sterol metabolism-related enzyme genes with and without oil were analyzed. 16S rDNA was used as the housekeeping gene, and the primers are shown in Table S1. cDNA at three time points, 60, 84, and 168 h during the transformation process was used as a template for qRT-PCR experiments.

The results indicated that at the three time points with oil conditions, gene expression levels were basically up-regulated (red) (Fig. 5). Among them, in the early stage of transformation (60 h), the nine enzyme genes whose expression levels were down-regulated were *kshA1*, *kshA5*, *rdc1*, *kstd1*, *kstd5*, *hsdC*, *hsd2*, *hsd3*, and *hasC2*, and the rest were all up-regulated; in the middle stage of transformation (84 h) and the later stage of transformation (168 h), all enzyme genes were up-regulated, and nine of them (*cho, fadA5, fadE30, hsdB, kshA2, kshB, hsd1, kstD2,* and *kstD3*) were up-regulated significantly. Among these nine up-regulated genes, there were seven genes (*cho, fadA5, fadE30, hsdB, kshA2, kshB, and hsd1)* that are associated with 9α-OH-AD synthesis and two genes (kstD2 and kstD3) associated with 9α-OH-AD degradation. The result indicated that as the fermentation process progressed, the expression level of enzymes related to the biotransformation of 9α-OH-AD

Table 3 The enzymes of sterol metabolism pathways in *Mycobacterium* sp. LY-1

Sterol metabolic pathway	Enzymes	Abbreviation
Preliminary oxidation of sterols	Cholesterol oxidase	Cho
	3-Hydroxyacyl-CoA dehydrogenase	Hsd1/ Hsd2/ Hsd3
Side-chain degradation of sterols	Acyl-CoA thiolase	FadA5
	Acyl-CoA dehydrogenase	FadE26/27/28/29/30
	Acyl-CoA synthetase	FadD1/ FadD2/ FadD3
	Ferredoxin	FdxD
Nucleus oxidation of sterols	17β-Hydroxysteroid dehydrogenase	HsdA
	2-Enoyl acyl-CoA hydratase	HsdB
	17β-Hydroxysteroid dehydrogenase	HsdC
	3-Ketosteroid 9 α-hydroxylase oxygenase	KshA1/2/3/4/5
	3-Ketosteroid 9 α-hydroxylase reductase subunit	KshB
	Rieske (2Fe-2S) domain-containing protein	Rdc1/2/3
	3-Ketosteroid Δ^1-dehydrogenase	KstD3/4/5
	3-Oxosteroid Δ^1-dehydrogenase	KstD1/2
Nucleus cleavage of sterols	monooxygenase, oxygenase	HsaA
	2,3-Dihydroxybiphenyl/1, 2-Dioxygenase	HsaC1/2/3/4/5

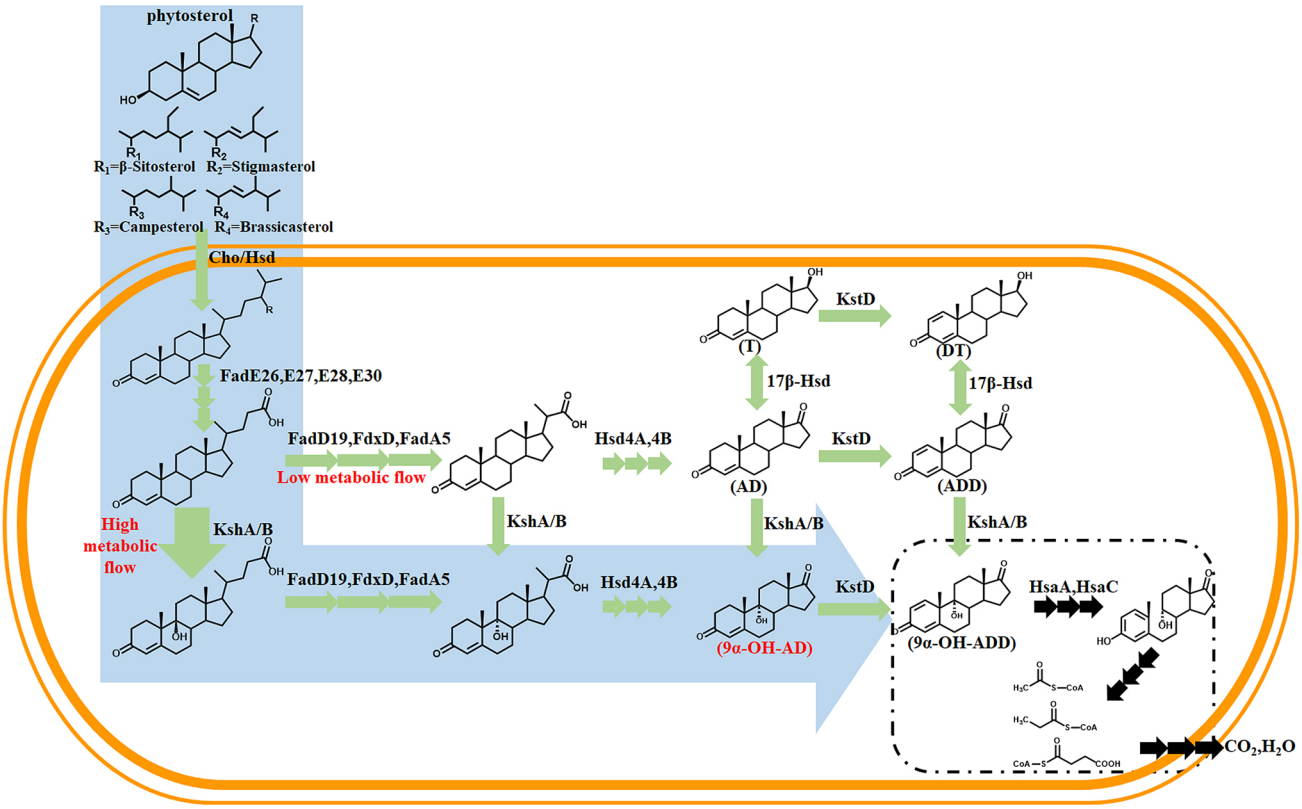

Fig. 3 Putative phytosterols degradation pathways of *Mycobacterium* sp. LY-1

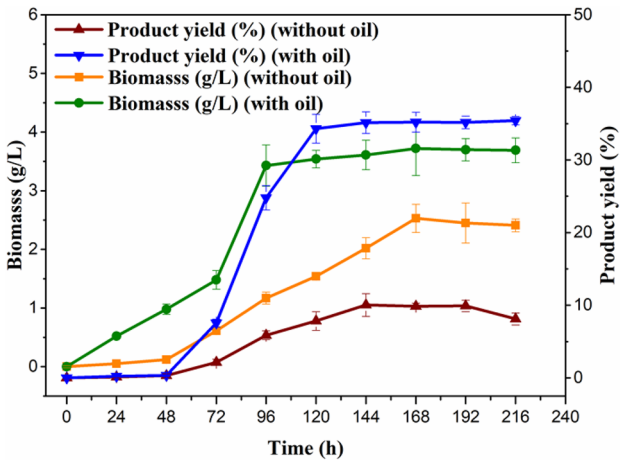

Fig. 4 The growth and 9α-OH-AD yield curves of *Mycobacterium* sp. LY-1 culturing without and with soybean oil. Data represent the mean values from three independent experiments

increased, especially the above-mentioned up-regulated seven genes. Based on the analysis of the differential transcription levels of the enzymes related to the sterol metabolic pathway under the above-mentioned different culture conditions, seven key enzymes for the synthesis of 9α-OH-AD were mainly involved in the accumulation of 9α-OH-AD including Cho, Hsd1, FadA5, FadE30, HsdB, KshA2, and KshB. To verify the above hypothesis, single overexpression and combined overexpression of the proposed 9α-OH-AD synthetic key enzymes were conducted in the following experiments.

Overexpression of key enzymes in *Mycobacterium* sp. LY-1 to increased 9α-OH-AD production

Single overexpressing and co-overexpressing the key genes with significant up-regulation in the sterol metabolism was an effective strategy to improve 9α-OH-AD production [21, 28]. Seven key genes (*cho, fadA5, fadE30, hsdB, kshA2, kshB, and hsd1*) related to 9α-OH-AD synthesis were overexpressed separately to up-regulate the sterol metabolism (Fig. 6). Compared with the control, the yield of 9α-OH-AD in all recombinant strains showed a slight increase. Among them, the 9α-OH-AD production of the LY-1-*kshA2* was the highest, and the yield of 9α-OH-AD was increased to 40.8%. At the same time, the biomass of these recombinant strains decreased significantly. These results showed that although the single overexpression of key enzyme genes could promote the accumulation of 9α-OH-AD to some extent, the degree of improvement was relatively weak (within 10%).

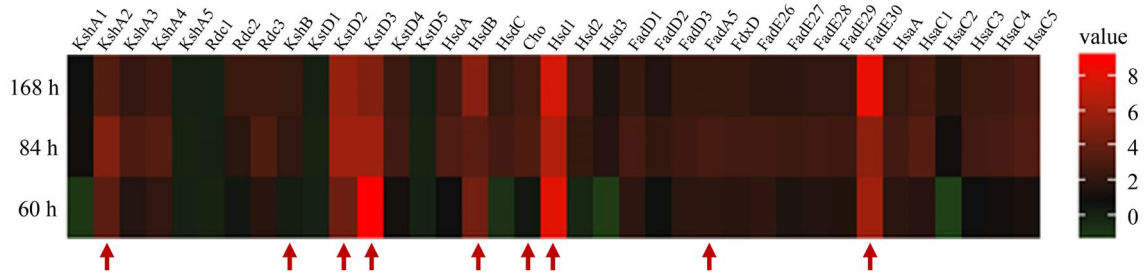

Fig. 5 The transcription-level analysis diagram of sterol metabolic pathway

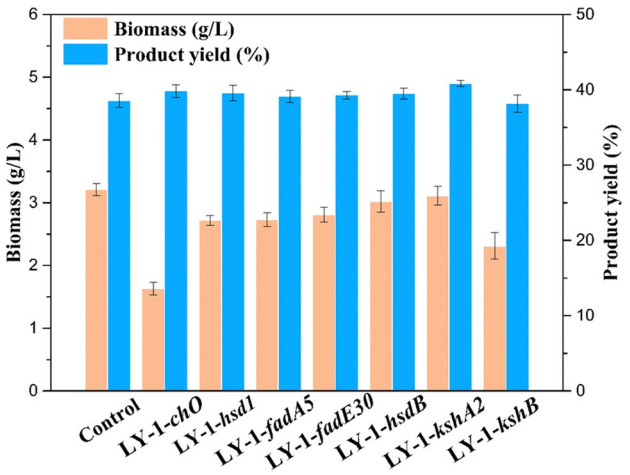

Fig. 6 α-OH-AD yield of *Mycobacterium* sp. LY-1 control strain and single overexpressing strains in shake flask culture. The *Mycobacterium* sp. LY-1 of the electroporated plasmid pMV261 was used as a control, the culture conditions were the same, the substrate concentration was 15 g/L, and the fermentation time was 7 days. All data represent the mean values from three independent experiments

Therefore, co-expression of key enzyme genes involved in sterol metabolism was further investigated.

Based on single overexpression responses, genes of four enzymes including *kshA2*, *kshB*, *hsdB*, and *hsd1* which enabled a higher increase in the 9α-OH-AD production were selected to co-express in *Mycobacterium* sp. LY-1. To reduce the impact of multiple plasmids on the growth of *Mycobacterium* sp. LY-1, a single plasmid was used to overexpress all genes. As shown in Fig. 7, the yield of 9α-OH-AD of co-expressed strains further improved compared with the single overexpression recombinant strains, while the biomass also increased. The 9α-OH-AD yield of LY-1-*kshA2/hsdB* reached 42.07%, an increase of 9.0% from the wild *Mycobacterium* sp. LY-1 (38.58%).

The metabolic pathway from AD to 9α-OH-AD involved the key enzyme 3-ketosteroid-9α-dehydrogenase (Ksh). The Ksh activity consisted of two components: terminal oxygenase (KshA) and ferrooxidase (KshB) [31] and the existence of these two components at the same time could help the biotransformation of 9α-OH-AD. To better promote the

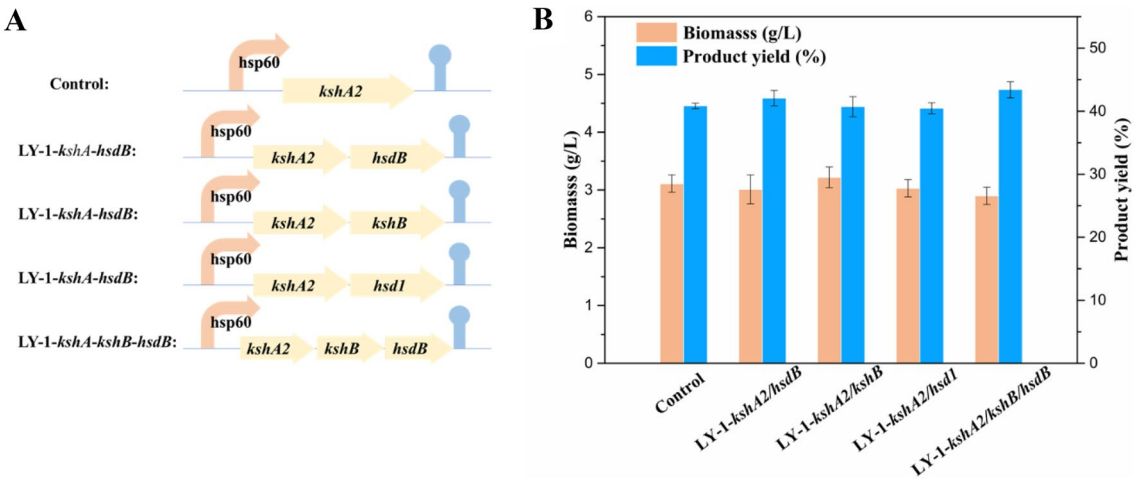

Fig. 7 9α-OH-AD yield of *Mycobacterium* sp. LY-1 control strain and co-expressing strains in shake flask culture. **A** Schematic of the co-overexpression in *Mycobacterium* sp. LY-1. **B** 9α-OH-AD production by the control strain and co-expressing strains at the 168th hours' fermentation. The culture conditions were the same, and the substrate concentration was 15 g/L. All data represent the mean values from three independent experiments

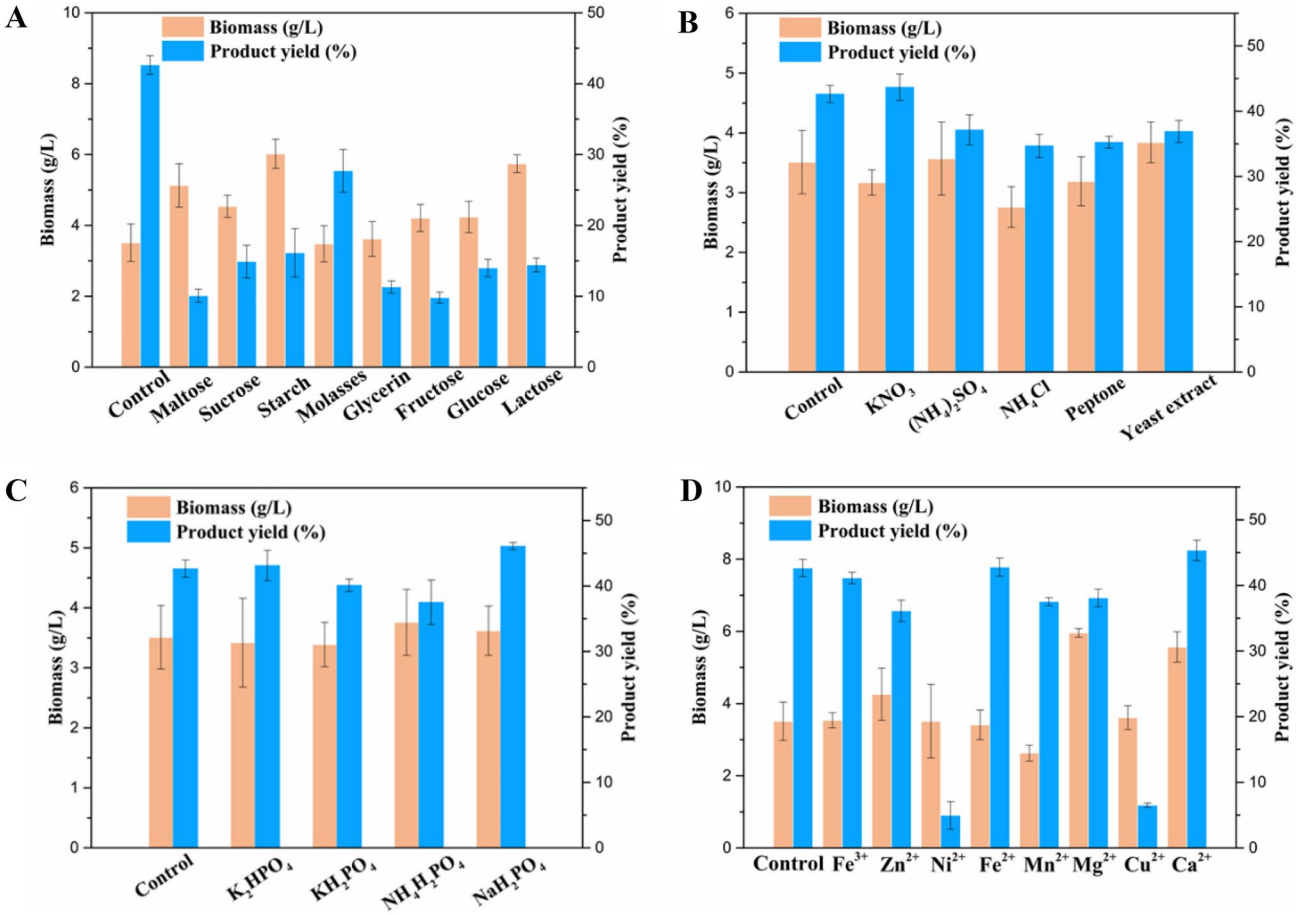

Fig. 8 Optimization of the medium. **A** Effects of different carbon sources on 9α-OH-AD yield. The control used corn steep liquor as the initial carbon source. **B** Effects of different nitrogen sources on 9α-OH-AD yield. The control used $NaNO_3$ as the initial nitrogen source. **C** Effects of different phosphates on 9α-OH-AD yield. The control used $(NH_4)_2HPO_4$ as the initial phosphates. **D** Effects of different metal ions on 9α-OH-AD yield. The control contained no added metal ions. The culture conditions were the same, and the substrate concentration was 15 g/L. All data represent the mean values from three independent experiments

C9 hydroxylation, LY-1-*kshA2/kshB/hsdB* was constructed based on LY-1-*kshA2/hsdB* overexpressing the third gene of *kshB*. The yield of 9α-OH-AD could reach 43.4% (Fig. 7) and the 9α-OH-AD yield of the LY-1-*kshA2/kshB/hsdB* increased by 12.7% compared with the wild *Mycobacterium* sp. LY-1. The results showed that the yield of 9α-OH-AD could be increased by co-overexpressing endogenous genes. In the future, the catalytic properties of the enzymes should be specifically investigated to further improve the yield of 9α-OH-AD by heterologous expression of the relevant key enzymes.

Optimization of biotransformation process of recombinant strain

In the microbial transformation process of phytosterols, the growth of microorganisms is closely related to their growth environment. Various components in the medium and culture conditions had great influences on the growth of strains and the expression of intracellular enzymes [32]. Therefore, the medium and culture conditions were further optimized based on the LY-1-*kshA2/kshB/hsdB*. The effects of carbon sources, nitrogen sources, phosphates, and metal ions in the culture medium on product yield were studied, and the optimum concentration of components in the medium was optimized through orthogonal experiment (Tables S3 and S4). Thus, the optimal medium was 30 g/L, KNO_3 4 g/L, NaH_2PO_4 1.2 g/L, $FeSO_4$ 0.075 g/L, and $CaCl_2$ 0.1 g/L. After the optimization of the medium, the product yield of 9α-OH-AD reached 48.9% (Fig. 8). In addition, the culture conditions were optimized under the conditions of the above-mentioned optimal medium. Finally, it was determined that the optimal fermentation conditions were pH 8.0, inoculum amount 2%, and growth temperature of 30 °C (Fig. 9). Under the optimal conditions, the

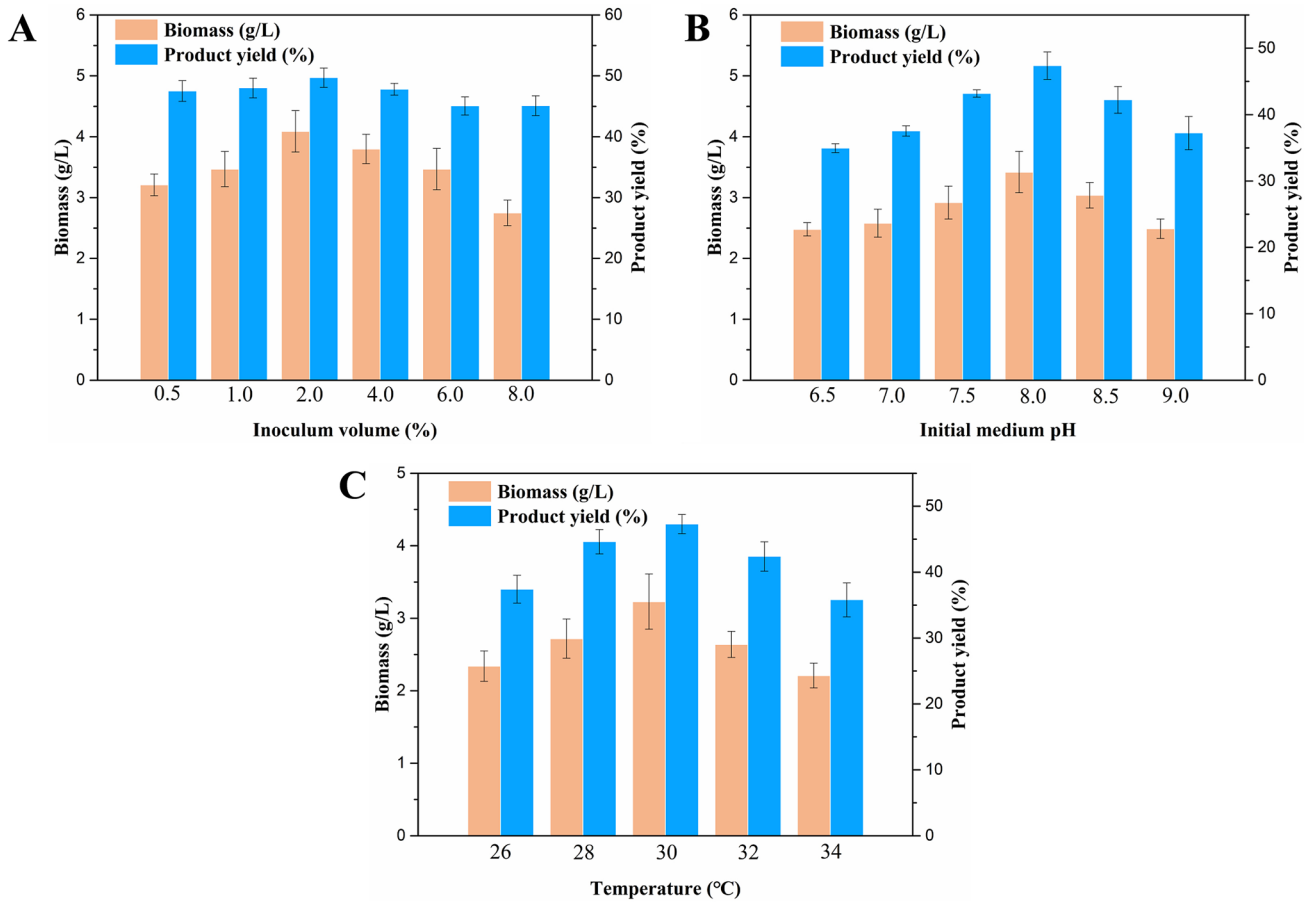

Fig. 9 Optimization of culture conditions. **A** Effects of cell concentration on 9α-OH-AD yield. **B** Effects of Initial medium pH on 9α-OH-AD yield. **C** Effects of temperature on 9α-OH-AD yield. The culture mediums were the same, and the substrate concentration was 15 g/L. All data represent the mean values from three independent experiments

maximum concentration of 9α-OH-AD reached 5.59 g/L and 9α-OH-AD yield reached 50.4%.

Conclusions

In this study, the metabolic pathway of phytosterols was predicted through the analysis of the whole genome of *Mycobacterium* sp. LY-1. Our key observation that the yield of 9α-OH-AD can be significantly improved by adding soybean oil led to the identification of the key enzymes related to the synthesis of 9α-OH-AD in the sterol metabolism pathway. Strain LY-1-*kshA2/kshB/hsdB* with high yield of 9α-OH-AD was obtained by co-expressing three key enzymes. This metabolic engineering strategy when combined with the optimization of the transformation process resulted in further yield of 9α-OH-AD improvement, reaching 50.4% in the presence of 15 g/L phytosterols. This study provides a promising biotransformation method for 9α-OH-AD production in industrial application.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s43393-022-00103-w.

Acknowledgements This work was supported by the National Key R & D Program of China (No. 2019YFA0905300), Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (No. TSBICIP-KJGG-001-14), the National Natural Science Foundation of China (No. 22078126), Qing Lan Project in Jiangsu Province and Fundamental Research Funds for the Central Universities (JUSRP221025).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Donova MV. Steroid bioconversions. In: Barredo J-L, Herráiz I, editors. Microbial steroids. New York: Springer; 2017. p. 1–13.

2. Fernández-Cabezón L, Galán B, García JL. New insights on steroid biotechnology. Front Microbiol. 2018;9:958–73. https://doi.org/10.3389/fmicb.2018.00958.
3. Teixeira MP, Passos EF, Haddad NF, et al. In vitro antitumoral effects of the steroid ouabain on human thyroid papillary carcinoma cell lines. Environ Toxicol. 2021;36(7):1338–48. https://doi.org/10.1002/tox.23130.
4. Liu Z, Liu T, Li W, et al. Insights into the antitumor mechanism of ginsenosides Rg3. Mol Biol Rep. 2021;48(3):2639–52. https://doi.org/10.1007/s11033-021-06187-2.
5. Frye CA. Steroids, reproductive endocrine function, and affect. A review. Minerva Ginecol. 2009;61(6):541–62.
6. Tong WY, Dong X. Microbial biotransformation: recent developments on steroid drugs. Recent Pat Biotechnol. 2009;3(2):141–53. https://doi.org/10.2174/187220809788700157.
7. Lednicer D. Steroid chemistry at a glance. Chichester: John Wiley & Sons; 2011.
8. Javid M, Nickavar B, Vahidi H, et al. Baeyer-Villiger oxidation of progesterone by *Aspergillus sojae* PTCC 5196. Steroids. 2018;140:52–7. https://doi.org/10.1016/j.steroids.2018.07.008.
9. Herráiz I. Chemical pathways of corticosteroids, industrial synthesis from sapogenins. Methods Mol Biol. 2017;1645:15–27.
10. Capyk JK, Kalscheuer R, Stewart GR, et al. Mycobacterial cytochrome P450 125 (Cyp125) catalyzes the terminal hydroxylation of C27 Steroids. J Biol Chem. 2009;284(51):35534–42. https://doi.org/10.1074/jbc.M109.072132.
11. Karpova NV, Andryushina VA, Stytsenko TS, et al. A search for microscopic fungi with directed hydroxylase activity for the synthesis of steroid drugs. Appl Biochem Microbiol. 2016;52(3):316–23. https://doi.org/10.1134/S000368381603008X.
12. Vidal M, Becerra J, Mondaca M, et al. Selection of *Mycobacterium* sp. strains with capacity to biotransform high concentrations of beta-sitosterol. Appl Microbiol Biotechnol. 2001;57(3):385–9. https://doi.org/10.1007/s002530100693.
13. Sukhodolskaya GV, Nikolayeva VM, Khomutov SM, et al. Steroid-1-dehydrogenase of *Mycobacterium* sp. VKM Ac-1817D strain producing 9alpha-hydroxy-androst-4-ene-3,17-dione from sitosterol. Appl Microbiol Biotechnol. 2007;74(4):867–73. https://doi.org/10.1007/s00253-006-0728-4.
14. Galan B, Uhía I, García-Fernández E, et al. *Mycobacterium smegmatis* is a suitable cell factory for the production of steroidic synthons. Microb Biotechnol. 2017;10(1):138–50. https://doi.org/10.1111/1751-7915.12429.
15. Yao K, Xu LQ, Wang FQ, et al. Characterization and engineering of 3-ketosteroid-Δ 1-dehydrogenase and 3-ketosteroid-9α-hydroxylase in *Mycobacterium neoaurum* ATCC 25795 to produce 9α-hydroxy-4-androstene-3, 17-dione through the catabolism of sterols. Metab Eng. 2014;24:181–91. https://doi.org/10.1016/j.ymben.2014.05.005.
16. Shao M, Zhang X, Rao Z, et al. Identification of steroid C27 monooxygenase isoenzymes involved in sterol catabolism and stepwise pathway engineering of *Mycobacterium neoaurum* for improved androst-1,4-diene-3,17-dione production. J Ind Microbiol Biotechnol. 2019;46(5):635–47. https://doi.org/10.1007/s10295-018-02135-5.
17. Szentirmai A. Microbial physiology of sidechain degradation of sterols. J Ind Microbiol Biotechnol. 1990;6(2):101–15. https://doi.org/10.1007/BF01576429.
18. Shtratnikova VY, Schelkunov MI, Dovbnya DV, et al. Complete genome sequence of *Mycobacterium* sp. strain VKM Ac-1817D, capable of producing 9α-Hydroxy-androst-4-ene-3,17-dione from phytosterol. Genome Announcements. 2015;3(1):e01447-14. https://doi.org/10.1128/genomeA.01447-14.
19. Luthra U, Bhosle V, Singh NK, et al. Media Optimization for 9α-hydroxyandrost-4-ene-3,17-dione Production by *Mycobacterium* spp. using Stat Des. 2016.
20. Zhou L, Li H, Xu Y, et al. Effects of a nonionic surfactant TX-40 on 9α-hydroxyandrost-4-ene-3,17-dione biosynthesis and physiological properties of *Mycobacterium* sp. LY-1. Process Biochem. 2019;87:89–94. https://doi.org/10.1016/j.procbio.2019.09.018.
21. Sun H, Yang J, He K, et al. Enhancing production of 9α-hydroxyandrost-4-ene-3,17-dione (9-OHAD) from phytosterols by metabolic pathway engineering of mycobacteria. Chem Eng Sci. 2021;230(47): 116195. https://doi.org/10.1016/j.ces.2020.116195.
22. Chang H, Zhang H, Zhu L, et al. A combined strategy of metabolic pathway regulation and two-step bioprocess for improved 4-androstene-3,17-dione production with an engineered *Mycobacterium neoaurum* - ScienceDirect. Biochem Eng J. 2020;164:107789. https://doi.org/10.1016/j.bej.2020.107789.
23. Gao XQ, Feng JX, Wang XD, et al. Enhanced steroid metabolites production by resting cell phytosterol bioconversion. Chem Biochem Eng Quart. 2015;29(4):567–73. https://doi.org/10.15255/CABEQ.2014.2098.
24. Gao X, Feng J, Hua Q, et al. Investigation of factors affecting biotransformation of phytosterols to 9-hydroxyandrost-4-ene-3,-17-dione based on the HP-β-CD-resting cells reaction system. Biocatal Biotransform. 2014;32(5–6):343–7. https://doi.org/10.3109/10242422.2014.976633.
25. Malaviya A, Gomes J. Androstenedione production by biotransformation of phytosterols. Biores Technol. 2008;99(15):6725–37. https://doi.org/10.1016/j.biortech.2008.01.039.
26. O'Toole RF, Gautam SS. Limitations of the *Mycobacterium tuberculosis* reference genome H37Rv in the detection of virulence-related loci. Genomics. 2017;109(5–6):471–4. https://doi.org/10.1016/j.ygeno.2017.07.004.
27. Zhou X, Zhang Y, Shen Y, et al. Efficient repeated batch production of androstenedione using untreated cane molasses by *Mycobacterium neoaurum* driven by ATP futile cycle. Bioresource Technol. 2020;309:123307. https://doi.org/10.1016/j.biortech.2020.123307.
28. Liu HH, Xu LQ, Yao K, et al. Engineered 3-ketosteroid 9α-hydroxylases in *Mycobacterium neoaurum*: an efficient platform for production of steroid drugs. Appl Environ Microbiol. 2018;84(14):e02777-e2817. https://doi.org/10.1128/AEM.02777-17.
29. Ceen EG, Herrmann JPR, Dunnill P. Solvent damage during immobilised cell catalysis and its avoidance: studies of 11α-hydroxylation of progesterone by *Aspergillus ochraceus*. Enzyme Microbial Technol. 1987;25(6):491–4. https://doi.org/10.1016/0141-0229(87)90061-5.
30. Phase N, Patil S. Natural oils are better than organic solvents for the conversion of soybean sterols to 17-ketosteroids by *Mycobacterium fortuitum*. World J Microbiol Biotechnol. 1994;10(2):228–9. https://doi.org/10.1007/BF00360894.
31. Van der Geize R, Yam K, Heuser T, et al. A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages. Proc Natl Acad Sci. 2007;104(6):1947–52. https://doi.org/10.1073/pnas.0605728104.
32. Yin Y. Effects of different carbon sources on growth, membrane permeability, β-sitosterol consumption, androstadienedione and androstenedione production by *Mycobacterium neoaurum*. Interdisc Sci Comput Life Sci. 2016;8(1):102–7. https://doi.org/10.1007/s12539-015-0116-9.