Rational development of mycobacteria cell factory for advancing the steroid biomanufacturing

Xin-Xin Wang1,2 · Xia Ke1,2 · Zhi-Qiang Liu1,2 · Yu-Guo Zheng1,2

Received: 23 June 2022 / Accepted: 28 July 2022 / Published online: 17 August 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract
Steroidal resource occupies a vital proportion in the pharmaceutical industry attributing to their important therapeutic effects on fertility, anti-inflammatory and antiviral activities. Currently, microbial transformation from phytosterol has become the dominant strategy of steroidal drug intermediate synthesis that bypasses the traditional chemical route. Mycobacterium sp. serve as the main industrial microbial strains that are capable of introducing selective functional modifications of steroidal intermediate, which has become an indispensable platform for steroid biomanufacturing. By reviewing the progress in past two decades, the present paper concentrates mainly on the microbial rational modification aspects that include metabolic pathway editing, key enzymes engineering, material transport pathway reinforcement, toxic metabolic intermediates removal and byproduct reconciliation. In addition, progress on omics analysis and direct genetic manipulation are summarized and classified that may help reform the industrial hosts with more efficiency. The paper provides an insightful present for steroid biomanufacturing especially on the current trends and prospects of mycobacteria.

Keywords Mycobacterium · Steroid biomanufacturing · Substrate availability · Omics analysis · Genetic manipulation

Introduction
Steroidal compounds (steroids) are terpenoid lipids with four cycloalkane rings (A-D) as a unique nucleus, which usually harbors a side chain at C-17 (Fig. 1A). They are widespread in nature as important cell membrane components and hormone precursors, which also gives them precious biological resources value. The common structural and stereochimical skeleton of these compounds can be further modified or substituted with different side chain groups, resulting in multiple biological functions (Wollam and Antebi 2011). About three hundred steroid-based drugs have been approved for clinical application, which makes them the second largest category of pharmaceuticals (Donova 2017). Most recently, the cheap and common steroid dexamethasone was found to use in COVID-19 severely infected patients significantly reduced the mortality by about one-third of patients under respiratory assistance (Ledford 2020). Steroid drugs are mainly divided into the categories of sex hormone and adrenocortical hormone analogs, due to their wide range of therapeutic applications, the steroidal drug market demand has become a strong driver for the fast-growing production scale (Fernandez-Cabezon et al. 2018). Steroidal drugs are derived from a variety of steroid intermediates, which contain a gonane core with different side-chains at C-17 as illustrated in Fig. 1A. Examples include androst-4-ene-3,17-dione (AD), androsta-1,4-diene-3,17-dione (ADD), 9α-hydroxy-4-androstene-3,17-dione (9-OH-AD), 22-hydroxy-23,24-bisnorchole-4-ene-3-one (4-HBC), 22-hydroxy-23,24-bisnorchole-1,4-diene-3-one (1,4-HBC), progesterone (PG), and testosterone (TS), among many others.

The preparation of steroid intermediates mainly includes chemical methods and microbial transformation. The main raw materials for chemical conversion are diosgenin, a key raw material extracted from Dioscorea species, and the chemically produced 16-dehydroprogrenolone acetate (Fig. 1B), as the central intermediates for...
Phytosterols are better sustainable resources for the conventional route of cholesterol-side chain degradation due to their low price and comparatively higher conversion rate (Sripalakit et al. 2006). The mapping and analysis of the complicated sterol catabolic pathway has always been one of the research hotspots in the field, involving a variety of enzymes and reactions, some of which show discrepancies among different species, which precluded the elucidation of the entire metabolic pathway (Wilbrink 2011).

With the development of bioinformatics, the genetic information of Actinomycetes was annotated, and the mystery on the functional genes involved in the general sterol catabolism route has been gradually unveiled (Donova 2017; Giorgi et al. 2019). As shown in Fig. 2, cholesterol is firstly converted to 4-cholesten-3-one by cholesterol oxidase (chox, ChOx) or 3β-hydroxysteroid oxidase (hsd, 3β-HSD) that was responsible for the initiation of bacterial cholesterol catabolism (Fig. 2M1). The side-chain degradation process was confirmed to be consistent with the fatty acid oxidation pathway (Kendall et al. 2007). Steroid C26-monoxygenase (cyp125, cyp142) is proposed to mainly catalyze side-chain to form a terminal carboxyl group (Fig. 2M2). After the terminal acylation of C-27, the cholesterol side-chain is activated by the terminal CoA thioesterification, and then the carboxyacyl-CoA of C-27 enters the β-oxidation reaction, consisting of acyl-CoA dehydrogenation, alkenyl-CoA hydration, β-hydroxyacyl-CoA dehydrogenation and β-ketoacyl-CoA sulfuration sequentially (Niu et al. 2019) (Fig. 2M3–M4). Module 5 shows the key intermediates synthesis in which aldolases and propionyl-CoA reductase are applied to generate 4-HBC, which can be further synthesized into PG. In the others pathway, the side-chain is completely truncated to produce AD, ADD and 9-OH-AD, and 3-ketosteroid Δ1-dehydrogenases (KstD) and 3-Ketosteroid-9α-hydroxylase (KSH) that play a crucial role in initiating the opening of the steroid nucleus (Li et al. 2019; Rohman and Dijkstra 2021).

The complete microbial degradation of a sterol requires a large array of steroid-degrading enzymes with different specificities. Through the combination of metabolic modification and genetic engineering technology, researchers hope to obtain excellent strains that can efficiently transform phytosterols into high-value steroidal intermediates. The corresponding products and intermediates of steroids obtained via cholesterol or phytosterol side chain degradation are listed in Table 1. The recently reported yield indicated the current research level and industrialization potential in this field. Since most of the classical enzymes...
involved in sterol catabolism have been summarized in previous studies (Donova and Egorova 2012; Zhao et al. 2021), we focused on the most recent advances on the key enzymes involved in the initiation of sterol catabolism in this paper.

**Significance of monooxygenase resources for steroid modification**

**Steroid C27 monooxygenase: initiation of side-chain degradation**

Microbial degradation of steroids encompasses three processes, including the oxidation and isomerization of the sterane core by 3β-hydroxysteroid dehydrogenases or cholesterol oxidase, degradation and oxidation of the sterol side chain, and finally decomposition of the steroid nucleus (Fernandez-Cabezon et al. 2017). The degradation of the side chain is crucial for sterol metabolism. Although many enzymes and cofactors participate in the process, this step is mainly driven by sterol C27 monooxygenases belonging to different cytochrome P450 families (Rosloniec et al. 2009).

CYP125 is demonstrated to be crucial both in hydroxylating cholest-4-en-3-one at C27 and oxidizing it into cholest-4-en-3-one-27-carboxylic acid. Therefore, reinforcement of CYP125 activity can enhance the transformation ability of steroids. The CYP 125 family has a “letterbox” active center, which can accommodate sterol molecules with a polycyclic structure, and a funnel-shaped entrance near the active-site heme. The alkyl chain of the substrate extends downward along the narrow junction funnel, and the terminal methyl carbon of the side chain is brought in proximity of the heme iron for oxidation (McLean et al. 2010). Accordingly, the CYP 125 level was upregulated with the increase of cholesterol (Brengel et al. 2016). Deletion of the gene greatly affected the growth of mycobacteria utilizing cholesterol as carbon source (Capyk et al. 2009). Overexpression...
of CYP 125–3 in *M. neoaurum* M3 increased the specific activity by 22%, and the NAD+/NADH ratio was increased by 31%, the yield of AD was improved approximately 18% in the biotransformation of phytosterol (Su et al. 2018).

In addition, three isoenzymes (SMO1, SMO2, SMO3) of steroid C27 monooxygenase were identified in *M. neoaurum* JC-12. The strongest catalytic activity was exhibited by steroid C27 monooxygenase 2, and when it was co-expressed with ChOx and KstD, the production of ADD reached 20 g/L (Shao et al. 2019a). A whole-cell catalyst for the production of AD overexpressing cholesterol oxidase and steroid C27 monooxygenase exhibited 88.6% capacity improvement, and the highest reported AD yield of 25.8 g/L was achieved in a two-step bioprocess (Chang et al. 2020). Hence, in the process of transformation of steroid by mycobacteria, C27 hydroxylation of the side chain by CYP 125 is both the initial step of side chain degradation and the rate-limiting step of cell growth, which plays an important role in metabolism (Ouellet et al. 2010).

### Table 1 Steroidal drug intermediates obtained by phytosterols biotransformation with *Mycobacterium neoaurum*

| Strain          | Strain information                              | Phytosterols | Main product | Molar yield (%) | References            |
|-----------------|-------------------------------------------------|--------------|-------------|----------------|-----------------------|
| JC-12           | From soil                                       | 40 g/L       | 18.6 g/L ADD | 67.8           | Shao et al. (2015)    |
| JC-12 s2-chOM-ksdΔkshA | JC-12 overexpressing cyp125A2, chOM-ksd, deleting kshA | 30 g/L | 20.1 g/L ADD | 91.6 | Shao et al. (2019a)   |
| NRRL-B3683      | DSMZ-German Collection of Microorganisms and Cell Cultures | 25 g/L | 12.0 g/L ADD | 70.0 | Wang et al. (2006)    |
| NwIB-01MS       | UV mutant of NwIB-01                            | 15 g/L       | 5.57 g/L ADD | 54.1           | Yao et al. (2013)     |
| NwIB-yV (Resting cell) | NwIB-01 deleting kstD1&2&3, overexpressing kshA | 70 g/L | 36.4 g/L 9-OH-AD | 71.3 | Gao (2016)            |
| W1embc (Resting cell) | *M. neoaurum* deleting kstD1, kstD2, kstD3 and embC | 20 g/L | 9.90 g/L 9-OH-AD | 67.9 | Xiong et al. (2020a)  |
| W1kasb (Resting cell) | *M. neoaurum* deleting kstD1, kstD2, kstD3 and kasB | 20 g/L | 10.90 g/L 9-OH-AD | 74.7 | Xiong et al. (2020b)  |
| Ac-1815D ChoD-SMO-Hsd4A | Ac-1815D overexpressing choD, cyp125A, hsd4A | 50 g/L | 25.8 g/L AD | 74.7 | Chang et al. (2020)   |
| B3805           | DSMZ-German Collection of Microorganisms and Cell Cultures | 20 g/L | 7.4 g/L AD | 53.6 | Mancilla et al. (2018) |
| NwIB-R10        | UV mutant of NwIB-01                            | 15 g/L       | 6.85 g/L AD | 66.1           | Yao et al. (2013)     |
| NdhF            | TCCC 11978 overexpressing ndh                   | 5 g/L        | 3.28 g/L AD | 95.0           | Zhou et al. (2019a)   |
| WII             | ATCC 25795 deleting kshA1.2, kstD1.2,3, hsd4A | 80 g/L | 57.4 g/L 4-HBC | 85.8 | Hu et al. (2020)      |
| DSM 138I ΔkstD1 | DSM 1381 deleting kstD1                         | 20 g/L | 14.18 g/L 4-HBC | 89.0 | Zhang et al. (2018)   |
| WII             | WII overexpressing choD, cyp125A and hsd4A      | 20 g/L | 7.59 g/L 4-HBC | 47.6 | Xiong et al. (2017)   |
| mjTU8-nsOpccR   | CCTCC AB2019054 deleting kstD, hsd4A and overexpressing nsOpccR | 10 g/L | 7.41 g/L 4-HBC | 93.0 | Peng et al. (2021)    |

*a*Phytosterols including β-sitosterol (414.7 g/mol), stigmasterol (412.7 g/mol), campesterol (400.7 g/mol), brassicasterol (398.7 g/mol) and other sterols. The average molecular weight of phytosterols is calculated as 414.7 g/mol (National Center for Biotechnology Information)

*b*The main final products are shown and other by-products are not enumerated

*c*The molar yields are calculated by the following formula: \( Y_m = \frac{M_S \times C_S \times 100\%}{M_P \times C_P} \), \( Y_m \), molar yield of the steroidal product, %, \( M_S \), the molecular weight of phytosterols, g/mol, \( C_S \) the working concentration of phytosterols, g/L; \( M_P \): the molecular weight of steroidal intermediates, g/mol \( C_P \) the end concentration of steroidal products, g/L.

**Nucleus hydroxylation for the activation of steroid molecules**

Sequencing analysis of *Mycobacterium tuberculosis* (*M. tuberculosis*) revealed that there were twenty P450s in its genome. These cytochrome P450s have been designated with systematic “CYP” numbers, such as CYP 125, CYP 142, CYP 106A, CYP 109, CYP 154, CYP 260 and other families (Szaleniec et al. 2018). Many CYPs have also been used in the reaction of steroid modification, and the obtained new steroid resources have more medicinal value. The important sites of the steroidal nucleus, such as C6, C7, C11, C12,
C14, C15, C16, C17 and C25, are also a focus of continuous research (Fig. 3). For example, 25-hydroxycholesterol possesses a significant antiviral and counteract inflammation activity (Diczfalusy, 2013; Tricarico et al. 2019), and could be produced from cholesterol by CYP109E1 from *Bacillus megaterium* (Putkaradze et al. 2019). CYP11A1 can biosynthesize pregnenolone by diametrically catalyzing the side-chain hydroxylation and cleavage of cholesterol or phytosterols, which avoids multiple rounds of step-by-step degradation (Zhang et al. 2019a). CYP154C5, CYP109E1 were found to convert testosterone into 16α-hydroxytestosterone, which is an important precursor of steroidal drugs in the pharmaceutical industry (Bracco et al. 2013; Jóźwik et al. 2016). Similarly, 17α-hydroxyprogesterone is an important pharmaceutical intermediate in the production of megestrol, hydrocortisone acetate, dexamethasone and other drugs (Gonzalez et al. 2018). In recent years, 17α-hydroxylases (CYP17A) were used to convert 17α-hydroxyprogesterone. At present, CYP17A2 had been successfully expressed in *Pichia pastoris* and used for the synthesis of 7α-hydroxyprogesterone, but the production was only 120.9 mg/L (Liu et al. 2022a).

The hydroxylation of C7, C11, C14 has great significance in the pharmaceutical industry (Chen et al. 2019; Milecka-Tronina et al. 2014; Schiffer et al. 2015). To extend the steroidal bioconversion pathway, exploration and screening of hydroxylase with higher activity and selectivity towards more active steroids and their functional expression in *Mycobacterium* is regarded as novel direction for the rational development of mycobacteria cells in the future (Donova 2021). Therefore, it is of great significance that we carry out more research on steroid oxidation by CYP enzymes, which will help to expand the application of cholesterol derivatives.

**Enhancement of substrate availability**

Reaction engineering for improving the accessibility of sterol substrates

The metabolic pathway is only one of the limiting factors for sterol transformation, low substrate solubility and transport barriers are also important constraints (Xiao et al. 2020). The problems of sterol emulsion have not been completely resolved, since the particles disperse poorly and require a complicated biocompatibility balance (Gao 2016). To improve the bioavailability of sterols by the application of emulsion technologies so that the mass transfer interface area is amplified (McClements 2012), researchers mainly add materials such as surfactants (Tween 80, Triton X-100), cyclodextrins (CDs) and their modified variants (2-hydroxylpropyl-β-cyclodextrin, HP-β-CD, methyl-β-cyclodextrin) (Shao et al. 2015). There are many

---

*Fig. 3* Cytochrome P450s with major target sites on steroid substrates. The steroidal precursors are modified by hydroxylation for different functions. C17 and C11 of hydroxylation were particularly concerned owing to its significance for steroids activation. Atoms in red represent the first critical hydroxyl group, and atoms in blue represent the second critical hydroxyl group. C6-hydroxylation: CYP4A21, CYP260B1; C7-hydroxylation: CYP7A2, CYP2A2; C11-hydroxylase: CYP106A2, CYP11B; C12-hydroxylase: CYP8B1; C14-hydroxylation: P450lun; C15-hydroxylation: CYP106A1, CYP106A2; C16-hydroxylation: CYP154C2, CYP109A1; C17-hydroxylation: CYP17A1, CYP17A2; C25-hydroxylase: CYP27A1, CYP109E1, CYP46A1; C26/C27-hydroxylase: CYP125A1, CYP142A1
shortcomings in these strategies, since surfactants interfere with cell growth and make wastewater treatment more challenging, while the cost of cyclodextrins remains prohibitive for industrial application (Mancilla et al. 2018).

In the process of mycobacterial growth and metabolism, the addition of vancomycin, glycine, ethambutol, polyethylenimine and cyclodextrin was found to change the permeability of the cell membrane or cell wall to enhance the trans-membrane delivery of sterols (Wang et al. 2006; Malaviya and Gomes 2008). For example, glycine was found to interfere with the formation of the peptidoglycan network structure. The resulting incomplete cell wall is conducive to breaking the osmotic barrier to promote the transport of sterols and other lipid substances, and thereby increase the accumulation of steroidal intermediates. Notably, most of these studies enhance the permeability of hydrophobic substances by altering the phenotype of cell wall, while the genetic regulation mechanism of sterol transport was seldomly addressed. This aspect had been effectively supplemented by omics analysis and molecular biology technology.

**Enhancement of transport channels in the cell membrane**

The structure of envelope in mycobacteria is another key factor for improving the accessibility of sterol substrates, not only maintains the life activities, but also represents an osmotic barrier for material transfer. The mycobacterial cell wall is mainly composed of peptidoglycan, arabinogalactan and the very-long-chain mycolic acids (Fig. 4). In these unusual core complexes, mycolic acid endows the cell wall with hydrophobic properties, while the crosslinking of peptidoglycan gives the cell wall its density, these structures play different functions in the process of transport (Kieser and Rubin 2014). With the development of omics technology, many genes regulating substrate transport and the osmotic barrier function of the membrane have been discovered. An analysis of the genomes of *M. tuberculosis* and *M. smegmatis* also found similar gene clusters with *mce* homologs. Their genome respectively accommodated 4 and 6 gene clusters corresponding to *mce* operons, and each of *mce* loci contained 2 *yrbE* and 6 *mce* genes, which encoded two integral membrane proteins. The membrane proteins have numerous similarities with the ATP-binding cassette transporters, which are closely related to cell membrane remodeling and cholesterol assimilation (Kumar et al. 2005).

The *mce4* genes have been confirmed to form part of the operon of *M. tuberculosis* for the uptake and utilization of cholesterol. Similarly, the *mce4* operon is also involved in the cholesterol transport of *M. smegmatis* (Klepp et al. 2012). Moreover, it was found that any of the *mce4* mutants

![Fig. 4](image_url) The structure of cell envelope in mycobacteria refer to the model of Jackson's cell capsule (Jackson et al., 2020). The cell envelope is made up of periplasm, outer membrane and capsule. The component of periplasm is peptidoglycan and arabinose and their biosynthesis is regulated by *pbpA/B, araR, embR* genes respectively, which further affects the efficiency of cholesterol uptake. *kasB, ephD, aceE* and *MmpL3* are relevant to the extension of mycolic acid and trehalose. Lipoarabinomannan and lipomannan are important backbones in the capsule, which are elaborated and modified covalently by succinate, methythioxylose, mannan and glucan. To provide detailed information, the corresponding schematic modules and structures were not drawn to scale.
could not grow on cholesterol. Hence, these genes are necessary for cholesterol uptake, while the homologous genes of other *mce* operons did not have the ability to replace them (Garcia-Fernandez et al. 2017). The Mce4 transporters were recently found to form a multicomponent complex that is essential for the mycobacterial cells to take up cholesterol. At the same time, it was found that Mce4A might bind to Mce systems, including Mce1 (Rank et al. 2021). When *mceG*, *yrbE4A* and *yrbE4B* from *M. tuberculosis*, which respectively encode energy-related factors and two integral membrane proteins, were expressed in *Mycobacterium* sp. MS136-GAB, they increased the product yield by 20% compared with the wild-type strain, indicated that the system promoted the mass transfer of phytosterols (He et al. 2018).

### Regulation of cell envelope

In recent years, the mycobacterial cell wall structure has also been widely investigated because of its unique function in sterol uptake. The *mmpl3* gene encodes a trehalose transmembrane transporter, and is also an important gene affecting mycobacteria cell envelope assembly (Tahlan et al. 2012). When the expression of *mmpl3* was reduced in *M. neoaurum* ATCC 25795, the cell permeability was increased by 23.4%, and the uptake of sterols was enhanced by 15.6% (Xiong et al. 2017). This is only one of the strategies that can be used to change the cell permeability by modifying the mycobacterial membrane and membrane proteins to affect the uptake and transport of substrates. The *pbpAB* genes were modified to change the structure of the peptidoglycan layer, and the yield of 4-HBC could be increased by 28% compared with the wild type (Sun et al. 2019a). What’s more, the mycolic acid layer can also be changed. The deletion of the nonessential gene *kasB*, which encodes β-ketoacyl acyl carrier synthetase, resulted in the disorder of the proportion of mycolic acids in *M. neoaurum* ATCC 25795. Furthermore, the cell permeability was increased by about twofold, and the yield was significantly increased by 2.38-fold, greatly reducing the conversion time (Xiong et al. 2020b). *embC* converts lipomannan to lipoarabinomannan in *M. tuberculosis*. The knockout of the *embC* resulted in a dysfunction of the cell envelope and increase of cell permeability, increasing sterol uptake by 52.4% over the parental strain (Xiong et al. 2020a). Recent reports indicate that *LmeA* and *SucT* mediate the maturation of lipoarabinomannan and can change the hydrophobicity and stiffness of the capsule (Palcekova et al. 2019; Rahlwes et al. 2020). Modulation of the mycobacterial cell envelope had not only become a new research hotspot, but also strengthened the sterol production performance of the strains from another aspect.

### Coordination strategy of toxic metabolic intermediates

#### Rational modulation of propionyl-CoA

*Mycobacterium* sp. can use cholesterol as a carbon source for growth, or convert sterols into steroids, what mainly involves the degradation of the sterol side chain and the decomposition of the core rings, whereby propionyl-CoA and acetyl-CoA are produced in the metabolic pathway (Uhia et al. 2012). The propionyl-CoA can maintain metabolic balance and provide energy in *Mycobacterium* sp., such as acylation of amino acids, metabolism of lipids (Crowe et al. 2017; Xu et al. 2020). However, studies have shown that when *M. smegmatis* and *M. tuberculosis* grow on propionate, the cells will accumulate excessive amounts of propionyl-CoA, which acts as a toxic intermediate that triggers metabolic stress and inhibits growth (Lee et al. 2013). Therefore, the detoxification of the propionyl-CoA pool is essential. There are three pathways for the degradation of toxic propionyl-CoA, including the methyl citrate cycle (MCC), the vitamin B12 dependent methylmalonyl pathway (MMP) and the synthesis of methyl-branched chain lipids in the cell envelope (Liu et al. 2018c). MCC is characterized by the initial condensation of propionyl-CoA and oxaloacetate to form methyl citrate, which is converted into pyruvate and succinate to achieve the goal of detoxification after several reactions (Otzen et al. 2014).

Based on these findings, in order to improve the ability of microbial transformation of AD, it is necessary to relieve the toxicity of propionyl-CoA. In *M. neoaurum*, there is a complete MCC pathway encoded by a *prp* gene cluster including *prpB*, *prpC*, and *prpD* encoding key enzymes, as well as a transcriptional regulator encoded by *prpR*. It was demonstrated that the expression of *prpR* and related genes (*prpDBC*) could be enhanced by the phosphate regulator RegX3 in MCC (Pei et al. 2021). Overexpression of *prpB-CDR* significantly increased the transcription level of the *prp* gene cluster and conversion of AD. Compared with the wild type, *prpR* could effectively reduce the level of propionyl-CoA and improve the cell viability. Furthermore, the nitrogen transcription regulator GlnR inhibited the transcription of the *prp* operon in nitrogen- limited medium. The deletion of *glnR* brought about the enhancement of the *prp* gene cluster transcription and expression, and promoted propionate detoxification (Zhang et al. 2020).

MMP is another detoxification strategy for propionyl-CoA. Firstly, propionyl-CoA is carboxylated into methyl malonyl-CoA, and then isomerized to non-toxic succinyl-CoA. The key enzyme in this metabolic process is biotin dependent propionyl-CoA carboxylase (PCC) (Savvi et al. 2008). For example, increasing the expression of *pccB* enhanced the biotransformation ability of phytosterols...
(Zhou et al. 2019b). Furthermore, the pccB and prpR genes were overexpressed to construct a citrate-based ATP futile cycle (AFC) and pyruvate-based AFC. These cycles could reduce the accumulation of ATP and propionyl-CoA in cells, enhancing cell viability and increasing the AD conversion rate (Zhou et al. 2020). In conclusion, propionyl-CoA is one of the key factors restricting the growth and production efficiency of host cells in the process of microbial transformation of phytosterols into steroids. The research and application of detoxification pathways will help improve the conversion rate of sterol substrates.

Metabolic detoxication of reactive oxygen species

Intracellular reactive oxygen species (ROS), such as superoxide radical, hydrogen peroxide (H2O2), hydroxyl radical and lipid peroxides are mainly produced by NADPH oxidase, xanthine oxidase, or during electron transfer in the respiratory process (Shao et al. 2017; Yeware et al. 2017). ROS has traditionally been regarded as a regulator of many cellular functions, such as transcription, direct oxidative modification, protein turnover, protein interaction and enzyme modification (Finkel 2011). However, excessive accumulation of ROS leads to oxidative stress and damages a variety of biological molecules, including DNA, proteins, and lipids (Tyagi et al. 2015).

Most mycobacteria have evolved antioxidant systems to resist the damage of ROS. There are two frequently-used detoxification mechanisms, based either on oxygen scavenging enzymes, such as catalase (CAT), catalase peroxidase and superoxide dismutase, or on low molecular weight mercaptans, such as glutathione (GSH), fungal mercaptan (MSH) and ergothioneine (EGT) (Richard-Greenblatt et al. 2015). When Mycobacterium sp. utilizes steroids, a series of redox reactions will occur inside the cell, which requires a large number of NAD+ and FAD molecules, producing ROS beyond the normal threshold (Shao et al. 2017). Excessive ROS brings about oxidative stress on the metabolism and subsequently damages various biological molecules, and ultimately injures cells critically (Imlay 2003), thus both the cell growth and steroidal biosynthesis capability of mycobacteria was severely inhibited. Therefore, it is an important strategy to reduce ROS level for enhancing steroid transformation.

Enhancing type II NADH dehydrogenase in M. neoaurum effectively increased the NAD+/NADH ratio and ATP level, and also effectively reduced the oxidative stress, resulting in greater cell viability. After enhancing NADH dehydrogenase, ROS was decreased by 42.3%, resulting in a 54.2% increase of viability; thus substantially raising the productivity of AD (Zhou et al. 2019a). NADH oxidase and catalase genes from Bacillus subtilis were overexpressed to eliminate the toxic H2O2 accumulated during steriol bioconversion, which improved the sterol transformation efficiency by 80% compared to the initial strain (Shao et al. 2019b). Mycobacterial endogenous antioxidants, ROS scavenging enzymes and low molecular weight mercaptans were used to construct detoxification pathways. The combination of CAT, MSH and EGT could boost cell viability by 54.2% and increase 4-HBC production by 47.5% (Sun et al. 2019b). Thus, normal ROS levels were restored by skillful metabolic adjustment, which may be employed as an effective strategy for enhancing sterol bioconversion.

Omics-guided studies provide new insights

Multi-omics strategies are valuable tools in the search for key genes and proteins of metabolic pathways. In the wake of the continuous development and convergence of internet and biotechnology, the application of bioinformatics is becoming increasingly widespread (Shtratnikova et al. 2021). The mycobacteria research can be divided into two categories. One is the research on pathogenic mycobacteria such as M. tuberculosis (Malone et al. 2018). Through the means of bioinformatics, the mechanism of infection, therapeutic drugs and metabolic pathways are continuously being explored (Goff et al. 2020). Another branch focuses on non-pathogenic mycobacteria (such as M. smegmatis and M. neoaurum) which have industrial potential and can be used to produce steroids (Uhia et al. 2012) (Table 2). Multi-omics data analysis of samples from different growth conditions or life cycles is not only help understand the biology of the strains, but is also conducive to the mining of key genes and new metabolic pathways, aiming to achieve the goal of industrialization (Li et al. 2016).

Genomics and transcriptomics assist the rational improvement of mycobacteria

Genome and transcriptome sequencing can reveal the gene expression changes that may be responsible for steroid catabolism of M. neoaurum. On the basis of significant differences in omics data, combined with single nucleotide polymorphisms and dynamic sRNA data, we can understand gene associations and interactions, as well as the regulatory mechanism of small molecules in various pathways, such as lipid transport and metabolism, amino acid metabolism, signal transduction, cell envelope function and ATP biosynthesis (Liu et al. 2016a, b). For example, a high yield and substrate resistant mutation of M. neoaurum MN4 was obtained by physical and chemical mutagenesis and screening. Experimental data from the combined omics analysis between the wild type and mutation revealed that both Chox and enzymes involved in β-oxidation were mutated (Xu et al. 2017). By comparing and analyzing the genome
of *Mycobacterium* sp. VKM, researchers found several key enzymes of steroid catabolism, such as KSH and KstD (Bragin et al. 2013). These results not only confirmed the hypotheses raised by previous studies, but also guided the design of experiments to increase production (Rohman and Dijkstra 2021; Zhang et al. 2019b).

Furthermore, transcriptome analysis was applied more frequently in sterol biodegradation research. The high-throughput transcriptome sequencing of *Mycobacterium* sp. VKM Ac-1817D grown on phytosterols showed that core genes encoding KSH and KstD exhibited tenfold of upregulation (Bragin et al. 2019). The changes in the transcriptome and proteome of *M. neoaurum* ATCC 25795 were used to panoramically analyze the sterol transformation process, substrate uptake, carbon metabolism, cell wall and membrane synthesis, regulatory changes of efflux protein and ATP at different levels, which provided targets for the removal of by-products and increase of sterol uptake (Liu et al. 2018b, c).

### Profiling of biotransformation intermediates using metabolomics

Many studies investigated the metabolome of mycobacteria, but they mainly focused on the physiological networks, pathogenic mechanisms and novel drug targets of *M. tuberculosis* (Rego et al. 2021). By comparison, the research on the metabolome of *M. neoaurum* is sparse. Tween 80, a detergent that promotes the dissolution of sterols and reduces the aggregation of bacterial cells, is conventionally added to the medium for steroid transformation. Metabolomics revealed that Tween 80 induced the upregulation of multiple pathways in central carbon metabolism and affected the synthesis of triacylglycerols, proteinogenic amino acids and nucleotide precursors (Pietersen et al. 2020). The in-depth influence of detergents on the synthesis of steroids at the metabolic level is worth exploring. In addition to these results, propionyl-CoA, ROS, and 4-ene-3-ketosteroids were identified as toxic molecules that inhibit the synthesis of 9α-OH-AD, and an in-situ adsorption scheme improved the production efficiency by 23.15% (Wang et al. 2021). Development of omics not only rapidly promoted the technological development of the biosynthesis of steroids, but also provided genetic data for constructing efficient pathways and designing biological cell factories.

### Advances in genetic manipulation approaches of mycobacteria

#### General genetic manipulation

Non-pathogenic *M. smegmatis* and *M. neoaurum* are widely applied in the development of steroid biotransformation, but wild-type bacteria usually have a long growth period, large accumulation of by-products and low product yield, so it is necessary to conduct extensive genetic engineering to obtain high-yield strains (Loraine and Smith 2017). Genetic manipulations can be divided into three broad strategies: (I) Gene replacement based on homologous recombination and resistance marker screening; (II) Gene silencing based on homologous sequences and controllable promoters; (III) CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) editing technology, which was recently also introduced in *Mycobacterium* sp. (Choudhary et al. 2016).

It is difficult and inefficient to eliminate resistance markers to obtain label-free mutants. To solve this problem, two strategies have been developed, based on suicide plasmids carrying screening gene cassettes and mutually dependent or incompatible plasmids (Pashley et al. 2003). These vectors can be easily eliminated after silencing the target gene. However, there are still many problems such as multiple transformation steps and low efficiency of screening. In addition to the influence of resistance markers, the introduction of exogenous genes into mycobacterial hosts will cause a greater burden on growth and metabolism, which is not conducive to the implementation of genetic manipulation (Brown and Parish, 2006). Genetic manipulation for activation or silencing...
at appropriate times can effectively relieve the growth burden of the host.

**Precise genome editing**

The emergence of CRISPR greatly simplified gene editing on the chromosome. CRISPR-Cas is a natural defense system found in many bacteria and archaea, which can selectively cleave exogenous invading DNA (Han and She 2017; Makarova and Koonin 2015). It has become the “No. 1 scissors” of genetic engineering, and its inventors were awarded the Nobel Prize in physiology and medicine in 2020. Before plasmids can be introduced into the acceptor strain, it needs alkali or ultraviolet treatment to improve the success rates of transformation and recombination (Song et al. 2015). The time-consuming traditional tools limited the research on mycobacteria. CRISPR-Cas can shorten the time of genetic operation, reduce the difficulty of marker selection, and remove the restriction of homologous arm length (Doudna and Charpentier 2014). Consequently, the application of the CRISPR system in mycobacteria is becoming increasingly popular.

Many successful cases have been demonstrated in various mycobacteria. For example, the introduction of codon optimized dCas9 and sgRNA from *Streptococcus pyogenes* into *M. tuberculosis* resulted in complete inhibition of single or multiple target genes (Tang and Fu 2018). Using the CRISPR-Cas9 system, the expression of *sepF* was successfully inhibited in *M. smegmatis*. The inhibition rate is as high as 98%, and no off-target effects were found (Xiao et al. 2019). Moreover, CRISPR-Cas9 combined with non-homologous end joining (NHEJ) could achieve efficient frameshift mutation of genes in *M. marinum* and *M. tuberculosis*, which accelerated the research on these lethal pathogens (Meijers et al. 2020). However, CRISPR-Cas9 has many limitations, such as toxicity, single strand cleavage and low efficiency in GC rich recognition regions. CRISPR-Cas12a (Cpf1) edits bacterial genomes more effectively, providing a new tool for special hosts that cannot accept Cas9. It has different recognition target sites and the ability to cut double stranded DNA, as well as lower host toxicity (Tang and Fu 2018; Yan et al. 2017).

When different Cas effector proteins were tested in *M. smegmatis*, the expression and editing effect of FnCpf1-cg was significantly better than that of Cas9. The efficiency of FnCpf1 combined with the NHEJ system was as high as 70%, which greatly reduced the genetic operation time (Sun et al. 2018). CRISPR-Cas12a assisted recombination technology was also applied to steroidal biosynthesis. *M. smegmatis* strains with inactivated *kstD1*, *kstD2* and *kstD3* were successfully constructed, and accumulated 2.86 g/L 9α-OH-AD (Meng et al. 2020). Unfortunately, there are few reports on the application of CRISPR-Cas systems in *M. neoaurum*. In 2021, it was reported that *M. neoaurum* could be edited using Cas12 in combination with NHEJ or HR-mediated DSB repair. CRISPR-Cas12a was able to delete target fragments of 1–24 kb with a maximum efficiency of 70%. Combined with a suicide plasmid system, the probability of single crossover increased to 100% (Liu et al. 2022b). CRISPR-Cas has many advantages in genetic operation, and there are still many expectations on how to help reform industrial hosts for the production of steroids.

**Conclusions and future prospects**

The research on the transformation of sterols into steroids by *Mycobacterium* sp. has been ongoing for decades. Although the direction and content of these studies are very extensive, and many breakthroughs have been made, there are still many problems that have not been solved. For example, (I) the low solubility of the substrates is one of the most important factors limiting the microbial uptake and transformation of sterols; (II) More than ten enzymes are involved in the degradation of the sterol side chain, which lack unbridged analysis of their catalytic mechanisms and isozyme relationships; (III) There are still no efficient, systematic and simple genetic operation tools; (V) There is no ideal host that can overcome the low material transport efficiency and slow growth. Accordingly, many new research strategies are needed in the future.

Synthetic biology and bioinformatics are currently very popular disciplines. Combined with emerging technologies, they will not only encourage technological innovation in the research of steroids, but also promote the efficient development of industrialization processes for steroid biosynthesis (Fernandez-Cabezon et al. 2018). By taking the high-speed trains of bioinformatics, researchers can deeply analyze the interaction relationship of genes, proteins, regulatory factors, and metabolic networks at the individuality or global level using massive data, and point out the direction for accurate mining and rational improvement of strains. Using the technology of synthetic biology, sterol conversion systems will be modularized and componentized, skillfully transplanted and assembled into a more efficient host. This will enable researchers to reconstruct the metabolic pathway, build an efficient cell factory, and break through the traditional boundaries. Artificial cell synthesis instead of natural extraction will achieve great economic benefits and contribute to industrial upgrading tremendously.

Existing mycobacteria cell factories cannot meet the requirements of the steroid industry revolution thoroughly. Nowadays, even though the biocatalytic potential of mycobacteria towards steroids is gradually disclosed, the development of an ideal mycobacterial chassis strain is still a pressing problem. An industrial chassis strain must have
numerous robust features, such as low nutritional requirements and short growth cycle, efficient substrate uptake and product efflux, high copy and correct translation of genetic information, knock out or silencing of by-product genes, flexible and modular gene circuit design, as well as strong environmental adaptability and tolerance to substrates and products.

The metabolism of steroids in mycobacteria is highly complex. The optimization of metabolic pathways will greatly increase the flux of steroid synthesis and reduce the difficulty of regulation. Related strategies include the mining and heterologous expression of aldolases or high-performance hydroxylases. It is not only beneficial to improve the biosynthesis efficiency of steroids, but also to achieve the synthesis of related derivatives. What’s more, enzymes that are difficult to adapt can be directed or artificially created to modify steroids. Recent research on the change of catalytic sites in the directed evolution of enzymes can serve as inspiration (Grobe et al. 2021).

Cholesterol is a model substrate for the study of sterol metabolism, but it is too expensive for large-scale industrial production of steroids. At present, phytosterols are the main raw materials for steroid production, but they also have many disadvantages, such as complex composition, limited sources and many by-products. There were many reports on the synthesis of lanosterol, ergosterol, campesterol and cholesterol from glucose by yeast (Xu and Li 2020; Zhang et al. 2017). Therefore, combining the advantages of yeasts in the synthesis of polycyclic substances with the mycobacterial characteristics in the degradation and utilization of sterols holds promise to completely solve the problem of substrate availability in the future.

Author contributions All the authors were involved in the concept and design of this review. XXW: Conceptualization and original draft editing. XK: Review and editing. ZQL: Review, editing and supervision. YGZ: Review, editing and supervision.

Funding This research was supported by the National Key Research and Development Project of China (2019YFA0905300).

Declarations

Conflict of interest The authors have declared no conflict of interest.

References

Baruah D, Das RN, Konwar D (2015) Facile green synthesis of 16-dehydropregnenolone acetate (16-DPA) from diosgenin. Synthetic Commun 46:79–84. https://doi.org/10.1080/00397911.2015.1121280

Behra PRK, Peterson BMF, Ramesh M et al (2019) Insight into the biology of Mycobacterium mucogenicum and Mycobacterium neoaurum clade members. Sci Rep 9:19259. https://doi.org/10.1038/s41598-019-55464-5

Bracco P, Janssen DB, Schallmey A (2013) Selective steroid oxygenation by CYP154C5, a bacterial cytochrome P450. Microb Cell Fact 12:95. https://doi.org/10.1186/1475-2859-12-95

Bragin EY, Shratnikova VY, Dovbnya DV 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. https://doi.org/10.1016/j.jsbmb.2013.02.016

Bragin EY, Shratnikova VY, Schelkunov MI et al (2019) Genomewide response on phytosterol in 9-hydroxysteroido-dione-producing strain of Mycobacterium sp. VKM Ac-1817D. BMC Biotechnol 19:39. https://doi.org/10.1186/s12896-019-0533-7

Brengel C, Thomann A, Schirfin A et al (2016) Discovery and biophysical evaluation of first low nanomolar hits targeting CYP125 of M. tuberculosis. ChemMedChem 11:2385–2391. https://doi.org/10.1002/cmdc.201600361

Brown AC, Parish T (2006) Instability of the acetamide-inducible expression vector pAM2 in Mycobacterium tuberculosis. Plasmid 55:81–86. https://doi.org/10.1016/j.plasmid.2005.06.005

Capyk JK, Kalscheuer R, Stewart GR et al (2009) Mycobacterial cytochrome p450 125 (CYP125) catalyzes the terminal hydroxylation of C27 steroids. J Biol Chem 284:35534–35542. https://doi.org/10.1074/jbc.M109.072132

Chang H, Zhang H, Zhu L et al (2020) A combined strategy of metabolic pathway regulation and two-step bioprocess for improved 4-androstene-3,17-dione production with an engineered Mycobacterium neoaurum. Biochem Eng J. https://doi.org/10.1016/j.bej.2020.107789

Chen J, Tang JL, Xi YY et al (2019) Production of 14 alpha-hydroxy steroids by a recombinant Saccharomyces cerevisiae biocatalyst expressing of a fungal steroid 14 alpha-hydroxylation system. Appl Microbiol Biot 103:8363–8374. https://doi.org/10.1007/s00253-019-10076-x

Choudhary E, Lunge A, Garwal N (2016) Strategies of genome editing in mycobacteria: achievements and challenges. Tuberculosis (edinn) 98:132–138. https://doi.org/10.1016/j.tube.2016.03.005

Crowe AM, Casabon I, Brown KL et al (2017) Catabolism of the last two steroid rings in Mycobacterium tuberculosis and other bacteria. Mbio. https://doi.org/10.1128/mBio.00321-17

Diczfalusy U (2013) On the formation and possible biological role of 25-hydroxycholesterol. Biochimie 95:455–460. https://doi.org/10.1016/j.biochi.2012.06.016

Donova MV (2017) Steroid bioconversions. Methods Mol Biol 1645:1–13. https://doi.org/10.1007/978-1-4939-7183-1_1

Donova MV (2021) Microbial steroid production technologies: current trends and prospects. Microorganisms. https://doi.org/10.3390/microorganisms10010053

Donova MV, Egorova OV (2012) Microbial steroid transformations: current state and prospects. Appl Microbiol Biotechnol 94:1423–1447. https://doi.org/10.1007/s00253-012-4078-0

Doudna JA, Charpentier E (2014) Genome editing. The new frontier of molecular genetics. Science 346:1258096. https://doi.org/10.1126/science.1258096

Fernandes P, Cabral JMS (2007) Phytosterols: applications and recovery methods. Bioresource Technol 98:2335–2350. https://doi.org/10.1016/j.biortech.2006.10.006

Fernandez-Cabezon L, Garcia-Fernandez E, Galan B et al (2017) Molecular characterization of a new gene cluster for steroid degradation in Mycobacterium smegmatis. Environ Microbiol 19:2546–2563. https://doi.org/10.1111/1462-2920.13704

Fernandez-Cabezon L, Galan B, Garcia JL (2018) New insights on steroid biotechnology. Front Microbiol 9:958. https://doi.org/10.3389/fmicb.2018.00958
Zhang R, Zhang Y, Wang Y et al (2019a) Pregnenolone overproduction in *Yarrowia lipolytica* by integrative components pairing of the cytochrome p450scc system. Acs Synth Biol 8:2666–2678. https://doi.org/10.1021/acssynbio.9b00018

Zhang X, Zhu M, Han R et al (2019b) A novel 3-phytosterone-9alpha-hydroxylase oxygenation component and its application in bioconversion of 4-androstene-3,17-dione to 9alpha-hydroxy-4-androstene-3,17-dione coupling with a nadh regeneration formate dehydrogenase. Molecules. https://doi.org/10.3390/molecules24142534

Zhang Y, Zhou X, Wang X et al (2020) Improving phytosterol biotransformation at low nitrogen levels by enhancing the methylcitrate cycle with transcriptional regulators *PrpR* and *GlnR* of *Mycobacterium neoaurum*. Microb Cell Fact 19:13. https://doi.org/10.1186/s12934-020-1285-8

Zhao A, Zhang X, Li Y et al (2021) *Mycolicibacterium* cell factory for the production of steroid-based drug intermediates. Biotechnol Adv 53:107860. https://doi.org/10.1016/j.biotechadv.2021.107860

Zhou X, Zhang Y, Shen Y et al (2019a) Efficient production of androstenedione by repeated batch fermentation in waste cooking oil media through regulating NAD(+)/NADH ratio and strengthening cell vitality of *Mycobacterium neoaurum*. Biore sour Technol 279:209–217. https://doi.org/10.1016/j.biortech.2019.01.144

Zhou X, Zhang Y, Shen Y et al (2019b) Economical production of androstenedione and 9alpha-hydroxyandrostenedione using untreated cane molasses by recombinant mycobacteria. Biore sour Technol 290:121750. https://doi.org/10.1016/j.biortech.2019.121750

Zhou X, Zhang Y, Shen Y et al (2020) Efficient repeated batch production of androstenedione using untreated cane molasses by *Mycobacterium neoaurum* driven by ATP futile cycle. Biore sour Technol 309:123307. https://doi.org/10.1016/j.biortech.2020.123307

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.