Latest development in the synthesis of ursodeoxycholic acid (UDCA): a critical review

Fabio Tonin[1] and Isabel W.C.E. Arends*[1]

[1]Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, The Netherlands.

Email: I.W.C.E.Arends@tudelft.nl

* Corresponding author

Abstract:

Ursodeoxycholic acid (UDCA) is a pharmaceutical ingredient widely used in clinics. As bile acid it solubilizes cholesterol gallstones and improves liver function in cholestatic diseases. UDCA can be obtained from cholic acid (CA), which is the most abundant and least expensive bile acid available. This transformation requires several protection and de-protection steps and the use of toxic and dangerous reagents: this results in an overall yield of about 30% and a series of waste products. The organic synthesis for the preparation of CDCA and UDCA starting from taurinated and glycinated cholic acid, is a long process. It is complicated and risky due to the nature and toxicity of the reagents used, and the purification processes involved.

For that reason, studies have been performed towards the development of microbial transformations or chemo-enzymatic procedures for the synthesis of UDCA employing CA or chenodeoxycholic acid (CDCA). This promising approach led several research groups to focus their attention on the development of biotransformations with non-pathogenic, easy-to-manage microorganisms, and their enzymes. In particular, the reactions that can achieve this goal are

selective hydrolysis, epimerization of hydroxyl functions (by oxidation and subsequent reduction), the specific hydroxylation and dehydroxylation of suitable positions in the steroid rings.

In this mini-review we critically analyze the state of art on the production of UDCA by several chemical, chemo-enzymatic and enzymatic routes reported, highlighting the bottlenecks of each production step. Particular attention is placed on the precursors availability as well as the substrate loading in the process.

Potential new routes and recently developments are discussed, in particular on the employment of flow-reactors. The latter technology allows to develop processes characterized by shorter reaction times and lower costs for the chemical and enzymatic reactions involved.

Keywords: Bile acids; Biotransformation; Hydroxysteroid dehydrogenases; Production process; UDCA

Abbreviations: CA, Cholic acid; CDCA, Chenodeoxycholic acid; DCA, Deoxycholic acid; DHCA, Dehydrocholic acid; LCA, Lithocholic acid; UCA, Ursocholic acid; UDCA, Ursodeoxycholic acid; HSDH, Hydroxysteroid dehydrogenase

1 Introduction

Ursodeoxycholic acid (UDCA), is applied in the pharmaceutical industry (Figure 1) [1]. As reported in several ’90 papers, UDCA solubilizes cholesterol gallstones [2,3], it improves the liver function in cholestatic diseases [4-8] and it significantly decreases cholesterol saturation in the bile [8-10]. In terms of pharmacology, it is considered to be better than chenodeoxycholic acid (CDCA) in the treatment against biliary calculus, since it possesses high efficacy and total absence of side effects [11].

UDCA is commonly produced by transformation of cholic acid (CA), which is the most abundant and least expensive bile acid available. Because of the molecular complexity of bile acids, the chemical modification requires several protection and de-protection steps, resulting in an overall yield of about 30% [12-15]. For that reason, research has been performed on the development of more selective procedures which involve less reaction steps, for the synthesis of UDCA. In particular microbial transformations [16-19] or chemo-enzymatic procedures [20,21] employing CA, CDCA or lithocholic acid (LCA) as starting material have been studied.

This mini-review summarises the different aspects to be addressed and hurdles to be taken in the development of a selective and sustainable process for the production of UDCA. Different chemical, chemo-enzymatic and enzymatic routes will be considered. In addition, the precursors availability as well as the substrate loading in the process and the requisites for potential new routes will be discussed. Furthermore, the potential benefits of a flow reactor set-up for this multi-step synthesis will be discussed.

2 Precursor availability

2.1 Bile acids

The most important active ingredients of bile are the bile acids. They are responsible for absorption, emulsification, and digestion of lipids. Bile acids are 24- carbon containing 5β-steroids. Their structure contains multiple hydroxyl substituents: the position and the stereochemistry of these OH- groups influence the solubility and biochemical properties of the compounds. In CA, for example, the OH- groups on the steroidal ring are all in position α with respect to the ring plane, defining a structure in which the acid has a polar and an apolar surface. For this reason, this molecule and its derivatives are defined as amphipathic. Because of this, bile acids are considered very important molecules for their ability to form micelles in aqueous environment [22].

Bile acid synthesis takes place in the liver starting from cholesterol: 17 enzymes are involved in the production of these molecules. The final products are the primary bile acids: CDCA and CA [23]. Subsequently, bile acids can be modified by intestinal bacteria to form the secondary bile acids as, for example, deoxycholic acid, LCA and UDCA. Secondary bile acids can be subsequently resorbed and returned to the liver where they are re-secreted in a process known as enterohepatic circulation.

In mammals, bile acids are secreted as conjugated molecules with glycin or taurine (Figure 2), forming the so called bile salts, with slightly different properties (pKa, solubility) in comparison to the corresponding free acid [24,25]. These bile salts also leads to an increased retention in the intestine.

The only economically viable resource of bile acids is the bovine bile, which must be extracted at the time of slaughter.

In slaughterhouses, the bovine gallbladder is recovered during the processing of the meat: from a single cow, around 230 mL of bile can be obtained. The commercial prize of bile is in the range of 0.1-0.4 $/L. Bile acids represent roughly 0.7% (w/w) of the bile [26].

In order to extract and purify the different bile acids, bile is freeze and lyophilized: from 100 mL of bile 8 gr of apparently dry powder can be obtained. From this about 6.9 g of 90% pure bile acids can be obtained [27]: cholesterol, cholesterol esters, triglycerides and free fatty acids are selectively extracted with organic solvents from aqueous buffers at different pHs. Then, bile acids in the neutralized aqueous fraction were freed of inorganic salts by extraction of dry residue with absolute ethanol.

The major components of the obtained mixture are the primary bile acids (CA and CDCA), secondary bile acids (deoxycholic acid and LCA) and bile salts like taurocholic acid and glycocholic acid (derivatives of CA), taurochenodeoxycholic acid and glycochenodeoxycholic acid (derivatives of CDCA) and other conjugated salts of their 7-α-dehydroxylated derivatives [27].

2.2 CA

3α,7α,12α-trihydroxy-5β-cholan-24-oic acid, also named cholic acid (Figure 3, A), is one the primary bile acids. It is almost insoluble in water, but soluble in methanol, ethanol and acetic acid. Salts of CA are called cholates. From the 3D structure of CA is it possible to observe an hydrophilic and an hydrophobic face, giving to CA its characteristic surfactant properties. CA is sold as a treatment for children and adults with bile acid synthesis disorders. Because of its abundance in bovine bile, it is the precursor for UDCA.

2.3 CDCA

3α,7α-dihydroxy-5β-cholan-24-oic acid, also known as CDCA (Figure 3, B), is a primary bile acid in human but only traces can be found in bovine bile: this is the main reason why it is not used as precursor for the preparation of UDCA.

CDCA acid can be used to treat gallstones avoiding, unlike CA, the downregulation of the cholesterol-7-α-hydroxylase, that represent the rate-limiting step in bile acid synthesis [28].

It can be metabolised by bacteria in the colon to form the secondary bile acid known as lithocholic acid [29].

2.4 LCA and other bile acids

LCA, also known as 3α-hydroxy-5β-cholan-24-oic acid (Figure 3, C), is a secondary bile acid. It is produced by bacteria in the colon from CDCA trough the reduction of the hydroxyl functional group at C-7 in the "B" ring of the steroid framework.

Low percentages of other secondary bile acids and related keto derivatives can be found in the bile. The solubility properties, interactions and metabolisms are related to the position and stereochemistry of the hydroxyl groups attached to the steroid ring. A general structure with the names of several bile acids is reported in Figure 3, C.

2.5 Deconjugation

Bile acids can be obtained from bile salts through a deconjugation step.

Chemically, it is an hydrolysis of the amide derivatives, that can be carried out at high temperature in alkaline environment. This reaction requires large amounts of sodium hydroxide (30%) and high temperatures (120 ° C) for extremely long times (8-12 hours). Few enzymes (acylases, EC: 3.5.1) have been reported to hydrolyse glycinates and taurinates to the corresponding carboxylic acid. Recently, Pedrini et al. [30] have isolated and characterized a

cholylglycine hydrolase from *Xanthomonas maltophilia* CBS 827.97: this enzyme completely hydrolyses glycine and taurine conjugates in 20 minutes at 50 °C. Unfortunately the protein sequence of this enzyme is not reported making recombinant expression and industrial use impossible. A second enzyme, isolated in *Lactobacillus plantarum* and recombinantly expressed in *E. coli*, was reported by Christiaens et al. [31]: this enzyme shows almost the same properties of the one described above but with lower activities (eg. 100 U/mg vs. 3.42 U/mg on glycocholic acid as substrate).

From a biocatalytic point of view, other acylases and the well-known lipases can be used to achieve the same reaction. Few literature reports can be found on the promiscuous amidase activity of wild-type or engineered lipases [32,33], but no one have tested their activities on bile salts.

2.6 Conclusions on the precursor availability

UDCA is a very expensive active ingredient since, to date, it can only be obtained by isolation from bear bile (practice used in traditional Chinese medicine) or by chemical transformation of CA and CDCA from bovine bile leading to high costs and high environmental impact.

CA is the main constituent of bovine bile and is the main precursor for the synthesis of ursodeoxycholic acid.

One of the main problems regarding the availability of precursors for the synthesis of UDCA is the direct relation with meat industries. The major manufacturers of bovine meat are in developing countries, in particular south America and India, where the processes of food production are still managed in a traditional manner. In these slaughterhouses there is often a lack of adequate technical conditions and hygienic protocols leading to environmental pollution and the need to include sanitary procedures in the processing of these bile acids.

Alternative sources of sterols can be found in eukaryotic microorganisms like yeast and algea [34]. However, technological and scientific knowledge on these metabolic pathways area are still in an early stage, and will not be included in this review.

3 C-12 dehydroxylation

3.1 Chemical dehydroxylation

UDCA can be obtained by a multistep chemical synthesis starting from CA. Two main steps are involved: the dehydroxylation at C-12 and the epimerization of the 7-OH group.

In order to achieve chemical dehydroxylation, firstly CA has to be oxidized to the corresponding ketone, after which Wolff-Kishner reduction can be applied. This whole sequence comprises 5 steps [13]: after the protection of the carboxylic group by acid catalyzed esterification (quantitative yield), the 3- and 7-OH groups are protected selectively with acetic anhydride and pyridine (yield 92%). The 12-OH group is oxidized with CrO_3 (yield 98%) and, after a deprotection step in alkaline environment, the formed ketone group can be removed by a Wolff-Kisner reaction yielding CDCA (yield 82%). The overall yield of the dehydroxylation step is around 65%.

3.2 Wolff–Kishner reduction

The Wolff–Kishner reaction is widely used by chemists to remove carbonyl moieties from unsubstituted alkyl-chains. The reaction requires hydrazine to form a hydrazone from the ketone. Under highly basic conditions and heat, N_2 is cleaved and the desired alkyl-chain is obtained (Figure 4). This reaction is applied to the synthesis of UDCA in order to remove the carbonyl group at C-12.

Despite its simplicity, the use of hydrazine represent an high-importance factor for chemical industries, in terms of safety (explosive risk) and economic cost (it should be used in equimolar

amount to the desired product). In order to reduce the explosive risk, different hydrazine-derivatives have been discovered and tested (e.g. methyl hydrazinocarboxylate, [35]). Other reports have demonstrated the feasibility of this reaction in a flow-system, avoiding the large excess of hydrazine, giving high yields in a more rapid reaction [36]. In addition, a microwave-assisted Wolff–Kishner reduction reaction has been examined with good results in a 30 seconds reaction [37]. Another option is the production of hydrazine *in situ*, using chemical methods. Also enzymatic activities towards hydrazine have been discoved [38], however not enough optimized to be applicable.

3.3 A Wolff-Kishner alternative: the Mozingo reduction

Mozingo reduction catalyses the same reaction as Wolff-Kishner, albeit under neutral condition (Figure 4). It involves two steps: Firstly, the carbonyl compound is converted into a thioketal by adding a dithiol. The mechanism for this step is analogous to the mechanism for ketal or acetal formation except sulfur replaces oxygen as the nucleophile attacking the carbonyl.

In a second step, the thioketal is reduced to the corresponding methylene compound by hydrogenolysis in presence of Raney nickel (actually used for the hydrogenation of fatty acids). In comparison to Wolff-Kishner, the use of hydrazine is replaced by the use of hydrogen gas that can be seen as a double-bladed knife. The reduction step can also be performed with $NaBH_4$ or other reductants. At the moment, a complete and clear reaction mechanism has not been well identified yet.

There is one report that suggests the application of this reaction for the synthesis of UDCA (yield 95%) [39]. The major problem related to this reaction is the very characteristic odor of ehtanedithiol which is compared by many people to rotten cabbage. Ideally, there is the possibility of using other types of less volatile compounds but no reports thereof have been found.

3.4 12α-Hydroxysteroid dehydrogenase

12α-hydroxysteroid dehydrogenases (12α-HSDH) are particularly interesting for the selective oxidation of the 12-hydroxyl group of CA (Figure 5). These enzymes belong to the family of oxidoreductases with NAD^+ or $NADP^+$ as electron acceptor. This oxidation is a mandatory step for removing the OH functionality at C-12. In all the chemo-enzymatic routes reported by Eggert et al., [20] the carbonyl group resulting from the oxidation of 12-OH was subsequentially reduced by the Wolff–Kishner reaction (see above). The few cases showing a dehydroxylation in position C-12 by bacteria are reported in the next paragraph. The 12α-HSDHs activity (Table 1, [40-43]) is abundant among the strains of the genus *Clostridium*: two $NADP^+$-dependent 12α-HSDHs have been detected in *Clostridium leptum* [40] and in *Clostridium* group P strain C 48–50 [41]. Up to now, the 12α-HSDH from *Eubacterium lentum* and *Clostridium perfringens* are the only two NAD^+-dependent enzymes reported [42,43].

3.5 Biocatalytic C-12 dehydroxylation

In contrast to the reports on the epimerization of CA and CDCA with enzymes (see paragraph 4), the dehydroxylation of CA remains a undiscovered field for microbiologists and biochemists. Up to our knowledge, the only evidence of bacterial catalysed 12α-dehydroxylation was reported by Edenharder in 1983 [44]. He found eight strains of the *Bacteroides* genus that specifically dehydroxylate CA to CDCA. However, no other studies have been carried out concerning this topic. A method for the production of 12-dehydro steroids which comprises subjecting a 12-hydroxysteroid to the action of 12-dehydroxylase producing microorganism (*Clostridium perfringens* ATCC 19574) was patented in 1976 [45]. However, the presence and the expression of a protein that can catalyse this reaction were never confirmed in other papers.

The putative molecular mechanism for the C-12 dehydroxylation is still unknown: it can resemble the dehydroxylation mechanism described for position C-7 [46,47] (Supporting

information, Figure 1). Interestingly all the genes catalysing that reaction where clustered in the BAI operon and, by analogy, it can be possible to design a biochemical pathway that specifically acts on C-12 (See postulated sequence of steps in Supporting information, Figure 2). The reaction sequence can be divided in 3 steps: firstly, the substrate is oxidized by a specific alcohol dehydrogenase and an ene-reductase-like enzyme. Then dehydration occurs, catalysed by a specific dehydratase. The dehydrated product is then reduced through a 3-step-cascade reaction (catalysed by 3 different enzymes) giving the final dehydroxylated product.

4 7-OH epimerization: shift the equilibrium

4.1 Chemical epimerization of CDCA into UDCA

The second step of UDCA synthesis from CDCA, is the epimerization of the 7-OH group. Chemically, the 7α-OH group of CDCA is selectively oxidized in the presence of sodium bromate [48] (yield 88%), N-Bromosuccinimide [13,15] (ungiven yield) or 1-hydroxy-1,2-benziodoxol-3(1H)-one 1-oxide [49] (yield 90%) and subsequently reduced with metallic sodium in presence of imidazole and 1-propanol (yield 80%) yielding the 7β-OH epimer (UDCA) as imidazole salt. Notably, the regio-specific oxidoreduction of 7α-OH group is achieved using weak oxidants: this behavior can be explained by the peculiar conformation of CDCA (the 7α-OH group is surrounded by alkyl chains, generating an hydrophobic environment that favors oxidation to the ketone, which is not the case for the other epimer). These data are supported by the density functional calculation or rather the differential change in electron density due to an infinitesimal change in the number of electrons [50]. The overall yield of the epimerization step is around 70% [12,15,51].

A further purification step is necessary for preparation of free UDCA: it can be easily obtained with sequential esterification, extraction and hydrolysis (yield 91%). The theoretical yield of the whole process fluctuates around 30 to 40%.

4.2 Enzymes for the production of UDCA from CDCA

There are different combinations of enzymes that can be used for the transformation of CDCA into UDCA. The enzymatic activities, the equilibrium of the reaction, the inhibition of enzymes by substrate and products, as well as their stabilities have to be evaluated in order to find the optimum reaction conditions for high yields of the desired product.

A list of enzymes that can be used for the transformation of CDCA is presented in the next paragraphs.

4.3 7α-Hydroxysteroid dehydrogenases (7α-HSDH)

These enzymes are able to oxidise specifically the α-hydroxyl group at C-7 together with the concomitant reduction of NAD^+ or $NADP^+$ (Figure 6, A). All of them are part of the group of the short chain dehydrogenases/reductases (SDR), showing a molecular weight around 30 kDa and a homotetrameric quaternary structure. Reported 7α-HSDHs were isolated from both aerobic and anaerobic bacteria: the state of art, together with the cofactor dependence and the specific activities, are summarized in Table 2 [30,52-59].

In addition to these reported enzymes biotransformations, many additional 7α-HSDHs have been discovered and reported over the past years. About 500 entries can be found in NCBI database typing “7 alpha hydroxysteroid dehydrogenase”.

4.4 7β-Hydroxysteroid dehydrogenases (7β-HSDH)

Unlike their homologues, only few examples of bioconversion with 7β-HSDH have been reported in literature (Table 3, [30,54,60-63]) (Figure 6, B): one $NADP^+$ dependent dehydrogenase from *Clostridium absonum* [54] was used in two different processes for the production of UDCA [64,65].The $NADP^+$-dependent enzyme from *Eubacterium aerofaciens* shows a significantly lower specific activity [60] and another $NADP^+$-dependent enzyme was isolated from *Ruminococcus gnavus* [61]. In order to increase the activity and stability of 7β-

HSDHs, protein engineering studies were carried out, as described in literature by Weuster-Botz et al. [63] and Zheng et al. [62]. Up till now, *Xanthomonas maltophilia* 7β-HSDH (33 U/mg) represents the only isolated NAD^+-dependent enzyme [30]. Unfortunately its protein sequence has not been reported.

4.5 Biocatalytic processes

Both microorganisms and purified enzymes have been applied for the fully biocatalytic epimerization of the 7-OH group. Several examples reported in literature are summarized in Table 4 [66-71].

The use of whole-cell conversion offers both advantages and disadvantages: wild-type microorganisms are normally difficult to grow, especially if the enzyme expression is related to anaerobic conditions. In addition the pathogenicity of these microorganisms represents a problem for their use in the pharmaceutical industry; additional steps of purification and control of sterility are necessary to obtain a safe product for the market. Otherwise, the circumvention of protein isolation and production makes it cheaper than their free-enzyme analogues.

In this way, the use of lyophilized whole-cell containing recombinant HSDHs can represent a solution in the reduction of catalyst costs, maintaining a reasonable safety. This approach was followed by Braun et al. and Sun et al. obtaining the 12-keto-UDCA with a yield of 99.5% using engineered *E. coli* cells [72,73].

Several enzymatic systems have been proposed in literature, together with cofactor regeneration systems. As general rule, the oxidative and reductive steps are coupled with a different cofactor dependence and related regeneration system. In this way, the equilibrium of the reaction can be pushed to the production of UDCA. An overview of reported enzymatic and chemo-enzymatic cascades is summarized in Table 5 and Figure 7 [12,30,64,72-80].

The decoupling of the 2 reactions is an elegant way to spin the equilibrium but, in every catalytic cycle, the co-substrates used to regenerate the cofactor have to be added in great surplus, leading to additional costs and additional problems in the downstream process. The most used enzymes for the cofactor regeneration are glucose dehydrogenase (glucose to glucuronic acid), lactate dehydrogenase (pyruvate to lactate), glutamate dehydrogenase (α-ketoglutarate to glutamate) and formate dehydrogenase (formate to CO_2). In particular, the last enzyme is interesting because formate is cheap and, because of the gaseous nature of CO_2 as product, the equilibrium of the reaction is entropically favoured.

Pedrini et al., in 2006, [30] reported the successful epimerization of CDCA in UDCA using a redox-neutral cascade reaction, with two NAD^+ dependent dehydrogenases. In this way the requirement of external systems for cofactor regeneration was circumvented and UDCA was obtained with a final yield of 75%. Interestingly, the addition of 2-hexanol led to an increase of NADH available for the reduction of 7-keto-LCA and a final yield of 82% was observed: according to the authors, the presence of another alcohol dehydrogenase in the partially-purified enzyme preparation increases the amount of NADH for the 7-keto-LCA reduction.

To conclude, it is difficult to denote the “best” route for 7-OH epimerization: all the processes mentioned, demonstrate reasonable yield and high selectivity. A redox-neutral cascade seems most elegant, but in order to fully understand and push the equilibrium of the reaction, a full biochemical characterization and a deep knowledge of the kinetics and stability of the involved enzymes is required.

4.6 Other ways to obtain 7-OH epimerization

Other chemical routes for the production of UDCA have been patented and published: for example, Dangate et al. [49] proposed a chemical route where the order of the two steps is

reverted and the specific oxidation of 7- and 12- OH group can be achieved without any protection step (yield 53%).

Another interesting chemo-enzymatic way to obtain the epimerization of the 7-OH group consists is the removal of the functionality and the subsequent re-hydroxylation with a specific final chiral conformation. Both steps can be performed by enzymes and/or microorganisms: Sawada et al. [81] reported that a fungal strain (*Fusarium equiseti* M41) was able to introduce a 7β-hydroxyl group into LCA by hydroxylation forming UDCA directly.

Later, many other microorganisms with a 7β-hydroxylating activity were discovered in strains of actinobacteria and filamentous fungi [82,83]. The key-enzyme in that pathway is a P450-like enzyme that catalyses the specific and irreversible 7β-hydroxylation.

The possibility to access that kind of chemical and chemo-enzymatic reactions pave the way for the design of other unexplored routes for the production of UDCA (example in Figure 8).

In addition, other reported enzymes can eventually play a role in the cascade reaction synthesis of UDCA. For example the 3α-HSDHs [43,84], catalyze the oxidoreduction of the 3α-OH groups to the correspecive ketones and the well-known laccase-TEMPO system [85], can be used for the unselective oxidation of CA to dehydrocholic acid (DHCA),.

5 Solvent and substrate loading considerations in processing

For an economically and environmentally sustainable process volumetric productivities have to be considered. In other words substrate loadings cannot be too low. While it does not represent a problem in chemical synthesis (UDCA, CDCA and CA are pretty soluble in alcohols like methanol and ethanol), the water-based environment required by enzymes is an obstacle in the development of a biocatalytic process.

The solubility of CDCA and UDCA at pH 8.0 (typically used for HSDHs) is around 10 mM, and it could be increased up to 20 mM when adding methanol or ethanol as co-solvent. Notably, HSDHs are relatively stable and active in 10-20% methanol. Moreover, the immobilization of the enzyme can provide an higher stability to the protein and make the system work also at higher concentrations of co-solvent. However, working with a diluted solution, produce a big amount of wastewater that had to be treated before the reintroduction into the environment.

Another option is represented by biphasic systems: in these cases, the organic phase works as reservoir of reagents and products. This methodology is widely used in biocatalysis to solve solubility issues. Unfortunately, the solubility of hydroxysteroids in non-alcoholic organic solvents (eg. ethylacetate, ethers, alkanes, dichloromethane, chloroform) is even lower than in water.

Of no lesser importance, the increased amount of substrates and products up to relevant concentrations for industrial application, can inhibit the enzymes used in the biocatalytic process. Several examples are reported in literature about substrate or product inhibition of HSDHs. Protein engineering could help to solve or lowering the effect of this issues, leading to the optimization of the bio-catalyst for the industrial applications. In addition, the use of flow-reactors can be beneficial to diminish substrate and product inhibition by controlling the contact time.

In conclusion, increase of the substrate loading is one of the main challenges in the development of an efficient biocatalytic system for the production of UDCA form CA. More research is needed to address this aspects.

6 Conclusion

The organic synthesis of CDCA and UDCA starting from taurinated and glycinated cholic acid is a long process, complicated and risky due to the nature and toxicity of the reagents used, (the costs of disposal of large amounts of NaOH, chromium salts and organic solvents) and the purification processes (necessary to eliminate by-products formed at each step of reaction) involved. All this extends the time, increases costs and decreases production yields. Therefore, research nowadays is geared towards more economical synthesis methods, that are waste-free and safe to operate.

An approach that bears great promise is the biotransformation with non-pathogenic, easy-to-manage microorganisms, and their enzymes. Several chemical, chemo-enzymatic and enzymatic routes have been proposed for the production of UDCA. In view of sustainability, instead of pursuing a step-wise approach, an integrated one-pot or one-flow reaction, involving highly selective enzymatic steps would be preferred.

When a multi-enzyme system is employed, the different enzyme activities, pH optima, cross reactions and inhibitions have to be taken into account in order to reach high product yields [86-88]. Furthermore, when a combination of chemical and enzymatic steps is employed special attention has to be paid to the compatibility. The challenge is to find a suitable combination of biotransformation and chemical steps to reach high yields of UDCA.

Nowadays, the most promising system for the biocatalytic production of UDCA are flow-reactors: they can be used for the set-up of continuous working systems, lowering the quantity of catalyst needed and the time of each reaction. This technology was recently employed by Zheng et al. [79] with great results, in terms of yields (99%) and productivity (88.5 g L^{-1} d^{-1}) for the epimerization of CDCA to UDCA. However, the employed enzymes have different cofactor specificities, leading to the consumption of stoichiometric amounts of sacrificial

substrates (pyruvate and glucose). In addition, substrate loadings in the latter process are still modest (10 mM). Therefore, there is much room improvement and further studies are needed to design a truly sustainable integrated process for the production of UDCA.

7 Acknowledgments:

This work was done as part of the ONE-FLOW project (https://one-flow.org) that has received funding from the European Union's FET-Open research and innovation actions (proposal 737266 - ONE-FLOW). We thank Linda Otten, Duncan McMillan and all the BOC group of TU Delft for the scientific discussions.

8 Supporting information

Supporting information features Figure 1 relative to the C-7 dehydroxylation mechanism of hydroxysteroids and Figure 2 relative to the postulated biochemical pathway for the C-12 dehydroxylation.

Supporting Information File 1:

File Name: S1

File Format: .pdf

Title: Supporting Figure 1 and 2

9 References:

1. Ikegami, T.; Matsuzaki, Y. *Hepatol. Res.* **2008**, *38*, 123.
2. Crosignani, A.; Battezzati, P. M.; Setchell, K. D.; Invernizzi, P.; Covini, G.; Zuin, M.; Podda, M. *Dig. Dis. Sci.* **1996**, *41*, 809.

3. Salen, G.; Colalillo, A.; Verga, D.; Bagan, E.; Tint, G.; Shefer, S. *Gastroenterology* **1980**, *78*, 1412.

4. Colombo, C.; Setchell, K. D.; Podda, M.; Crosignani, A.; Roda, A.; Curcio, L.; Ronchi, M.; Giunta, A. *J. Pediatr.* **1990**, *117*, 482.

5. Colombo, C.; Crosignani, A.; Assaisso, M.; Battezzati, P. M.; Podda, M.; Giunta, A.; Zimmer-Nechemias, L.; Setchell, K. D. *Hepatology* **1992**, *16*, 924.

6. Colombo, C.; Battezzati, P. M.; Podda, M.; Bettinardi, N.; Giunta, A. *Hepatology* **1996**, *23*, 1484.

7. Combes, B.; Carithers, R. L.; Maddrey, W. C.; Lin, D.; Mcdonald, M. F.; Wheeler, D. E.; Eigenbrodt, E. H.; Muñoz, S. J.; Rubin, R.; Garcia-Tsao, G. *Hepatology* **1995**, *22*, 759.

8. Podda, M.; Zuin, M.; Battezzati, P.; Ghezzi, C.; De Fazio, C.; Dioguardi, M. L. *Gastroenterology* **1989**, *96*, 222.

9. Thistle, J. L. In *Semin. Liver Dis.*; © 1983 by Thieme Medical Publishers, Inc.: 1983; Vol. 3, p 146.

10. Ward, A.; Brogden, R.; Heel, R.; Speight, T.; Avery, G. *Drugs* **1984**, *27*, 95.

11. Roda, E.; Bazzoli, F.; Labate, A. M. M.; Mazzella, G.; Roda, A.; Sama, C.; Festi, D.; Aldini, R.; Taroni, F.; Barbara, L. *Hepatology* **1982**, *2*, 804.

12. Carrea, G.; Pilotti, A.; Riva, S.; Canzi, E.; Ferrari, A. *Biotechnol. Lett.* **1992**, *14*, 1131.

13. Hofmann, A. F. *Acta Chem. Scand.* **1963**, *17*, 173.

14. Kanazawa, T.; Shimazaki, A.; Sato, T.; Hoshino, T. *Nippon Kagaku Zasshi* **1955**, *76*, 297.

15. Fieser, L. F.; Rajagopalan, S. *JACS* **1950**, *72*, 5530.

16. Sutherland, J. D.; Macdonald, I. A.; Forrest, T. P. *Prep. Biochem.* **1982**, *12*, 307.

17. Prabha, V.; Ohri, M. *World J. Microbiol. Biotechnol.* **2006**, *22*, 191.

18. Midtvedt, T.; Norman, A. *Acta Pathol. Microbiol. Immunol. Scand. [B]*. **1967**, *71*, 629.

19. Mahato, S. B.; Garai, S. *Steroids* **1997**, *62*, 332.

20. Eggert, T.; Bakonyi, D.; Hummel, W. *J. Biotechnol.* **2014**, *191*, 11.

21. Bortolini, O.; Medici, A.; Poli, S. *Steroids* **1997**, *62*, 564.

22. Mukhopadhyay, S.; Maitra, U. *Curr. Sci.* **2004**, *87*, 1666.

23. Hofmann, A. F. *Arch. Intern. Med.* **1999**, *159*, 2647.

24. Fini, A.; Roda, A. *J. Lipid Res.* **1987**, *28*, 755.

25. Bonar-Law, R. P.; Davis, A. P. *Tetrahedron* **1993**, *49*, 9829.

26. Maldonado Rodríguez, M. E., Università degli Studi di Ferrara, 2013.

27. Ahrens, E. H.; Craig, L. C. *J. Biol. Chem.* **1952**, *195*, 763.

28. Iser, J.; Sali, A. *Drugs* **1981**, *21*, 90.

29. Carey, M. C.; Mass Medical Soc: 1975.

30. Pedrini, P.; Andreotti, E.; Guerrini, A.; Dean, M.; Fantin, G.; Giovannini, P. P. *Steroids* **2006**, *71*, 189.

31. Christiaens, H.; Leer, R.; Pouwels, P.; Verstraete, W. *Appl. Environ. Microbiol.* **1992**, *58*, 3792.

32. Nakagawa, Y.; Hasegawa, A.; Hiratake, J.; Sakata, K. *Protein Eng. Des. Sel.* **2007**, *20*, 339.

33. Torres-Gavilán, A.; Castillo, E.; López-Munguía, A. *J. Mol. Catal. B: Enzym.* **2006**, *41*, 136.

34. Volkman, J. *Appl. Microbiol. Biotechnol.* **2003**, *60*, 495.

35. Cranwell, P. B.; Russell, A. T.; Smith, C. D. *Synlett* **2016**, *27*, 131.

36. Newman, S. G.; Gu, L.; Lesniak, C.; Victor, G.; Meschke, F.; Abahmane, L.; Jensen, K. F. *Green Chem.* **2014**, *16*, 176.

37. Parquet, E.; Lin, Q. *J. Chem. Educ* **1997**, *74*, 1225.

38. Dietl, A.; Ferousi, C.; Maalcke, W. J.; Menzel, A.; de Vries, S.; Keltjens, J. T.; Jetten, M. S.; Kartal, B.; Barends, T. R. *Nature* **2015**, *527*, 394.

39. Sato, Y.; Ikekawa, N. *J. Org. Chem.* **1959**, *24*, 1367.

40. Harris, J.; Hylemon, P. *Biochim. Biophys. Acta* **1978**, *528*, 148.

41. Braun, M.; Lunsdorf, H.; Buckmann, A. F. *Eur. J. Biochem.* **1991**, *196*, 439.

42. Macdonald, I. A.; Jellett, J. F.; Mahony, D. E.; Holdeman, L. V. *Appl. Environ. Microbiol.* **1979**, *37*, 992.

43. Macdonald, I. A.; Meier, E. C.; Mahony, D. E.; Costain, G. A. *Biochim. Biophys. Acta* **1976**, *450*, 142.

44. Edenharder, R. *J. Steroid Biochem.* **1984**, *21*, 413.

45. Saltzman, W. H.; Google Patents: 1976.

46. Wells, J. E.; Hylemon, P. B. *Appl. Environ. Microbiol.* **2000**, *66*, 1107.

47. Ridlon, J. M.; Kang, D.-J.; Hylemon, P. B. *Anaerobe* **2010**, *16*, 137.

48. Ferrari, M.; Zinetti, F.; Google Patents: 2015.

49. Dangate, P. S.; Salunke, C. L.; Akamanchi, K. G. *Steroids* **2011**, *76*, 1397.

50. Parr, R. G.; Yang, W. *JACS* **1984**, *106*, 4049.

51. Samuelsson, B. *Acta Chem. Scand.* **1960**, *14*, 17.

52. Coleman, J. P.; Hudson, L. L.; Adams, M. J. *J. Bacteriol.* **1994**, *176*, 4865.

53. Baron, S.; Franklund, C.; Hylemon, P. *J. Bacteriol.* **1991**, *173*, 4558.

54. Ferrandi, E. E.; Bertolesi, G. M.; Polentini, F.; Negri, A.; Riva, S.; Monti, D. *Appl. Microbiol. Biotechnol.* **2012**, *95*, 1221.

55. Bakonyi, D.; Hummel, W. *Enzyme. Microb. Technol.* **2017**, *99*, 16.

56. Yoshimoto, T.; Higashi, H.; Kanatani, A.; Lin, X. S.; Nagai, H.; Oyama, H.; Kurazono, K.; Tsuru, D. *J. Bacteriol.* **1991**, *173*, 2173.

57. Ueda, S.; Oda, M.; Imamura, S. *Int. J. Biol. Macromol.* **2004**, *4*, 33.

58. Bennett, M. J.; McKnight, S. L.; Coleman, J. P. *Curr. Microbiol.* **2003**, *47*, 475.

59. Ji, W.; Chen, Y.; Zhang, H.; Zhang, X.; Li, Z.; Yu, Y. *Microbiol. Res.* **2014**, *169*, 148.

60. Liu, L.; Aigner, A.; Schmid, R. D. *Appl. Microbiol. Biotechnol.* **2011**, *90*, 127.

61. Lee, J.-Y.; Arai, H.; Nakamura, Y.; Fukiya, S.; Wada, M.; Yokota, A. *J. Lipid Res.* **2013**, *54*, 3062.

62. Zheng, M.-M.; Chen, K.-C.; Wang, R.-F.; Li, H.; Li, C.-X.; Xu, J.-H. *J. Agric. Food. Chem.* **2017**, *65*, 1178.

63. Weuster-Botz, D.; Braun, M.; Aigner, A.; Sun, B.; Kantzow, C.; Bresch, S.; Bakonyi, D.; Hummel, W.; Google Patents: 2011.

64. Bovara, R.; Carrea, G.; Riva, S.; Secundo, F. *Biotechnol. Lett.* **1996**, *18*, 305.

65. Monti, D.; Ottolina, G.; Carrea, G.; Riva, S. *Chem. Rev.* **2011**, *111*, 4111.

66. Macdonald, I. A.; Hutchison, D. M. *J. Steroid Biochem.* **1982**, *17*, 295.

67. Sutherland, J. D.; Macdonald, I. A. *J. Lipid Res.* **1982**, *23*, 726.

68. Macdonald, I.; Rochon, Y.; Hutchison, D.; Holdeman, L. *Appl. Environ. Microbiol.* **1982**, *44*, 1187.

69. Hirano, S.; Masuda, N. *J. Lipid Res.* **1982**, *23*, 1152.

70. Sutherland, J. D.; Holdeman, L. V.; Williams, C. N.; Macdonald, I. A. *J. Lipid Res.* **1984**, *25*, 1084.

71. Medici, A.; Pedrini, P.; Bianchini, E.; Fantin, G.; Guerrini, A.; Natalini, B.; Pellicciari, R. *Steroids* **2002**, *67*, 51.

72. Braun, M.; Sun, B.; Anselment, B.; Weuster-Botz, D. *Appl. Microbiol. Biotechnol.* **2012**, *95*, 1457.

73. Sun, B.; Kantzow, C.; Bresch, S.; Castiglione, K.; Weuster-Botz, D. *Biotechnol. Bioeng.* **2013**, *110*, 68.

74. Bakonyi, D.; Wirtz, A.; Hummel, W. *Z. Naturforsch., B: Chem. Sci.* **2012**, *67*, 1037.

75. Liu, L.; Braun, M.; Gebhardt, G.; Weuster-Botz, D.; Gross, R.; Schmid, R. D. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 633.

76. Giovannini, P. P.; Grandini, A.; Perrone, D.; Pedrini, P.; Fantin, G.; Fogagnolo, M. *Steroids* **2008**, *73*, 1385.

77. Monti, D.; Ferrandi, E. E.; Zanellato, I.; Hua, L.; Polentini, F.; Carrea, G.; Riva, S. *Adv. Synth. Catal.* **2009**, *351*, 1303.

78. Zheng, M.-M.; Wang, R.-F.; Li, C.-X.; Xu, J.-H. *Process Biochem.* **2015**, *50*, 598.

79. Zheng, M.-M.; Chen, F.-F.; Li, H.; Li, C.-X.; Xu, J.-H. *ChemBioChem* **2017**.

80. Ji, Q.; Tan, J.; Zhu, L.; Lou, D.; Wang, B. *Biochem. Eng. J.* **2016**, *105*, 1.

81. Sawada, H.; Kulprecha, S.; Nilubol, N.; Yoshida, T.; Kinoshita, S.; Taguchi, H. *Appl. Environ. Microbiol.* **1982**, *44*, 1249.

82. Carlström, K.; Kirk, D.; Sjövall, J. *J. Lipid Res.* **1981**, *22*, 1225.

83. Kollerov, V.; Monti, D.; Deshcherevskaya, N.; Lobastova, T.; Ferrandi, E.; Larovere, A.; Gulevskaya, S.; Riva, S.; Donova, M. *Steroids* **2013**, *78*, 370.

84. Möbus, E.; Maser, E. *J. Biol. Chem.* **1998**, *273*, 30888.

85. Arends, I. W.; Li, Y.-X.; Ausan, R.; Sheldon, R. A. *Tetrahedron* **2006**, *62*, 6659.

86. Ricca, E.; Brucher, B.; Schrittwieser, J. H. *Adv. Synth. Catal.* **2011**, *353*, 2239.

87. Oroz-Guinea, I.; García-Junceda, E. *Curr. Opin. Chem. Biol.* **2013**, *17*, 236.

88. Riva, S.; Fessner, W.-D. *Cascade Biocatalysis Integrating Stereoselective and Environmentally Friendly Reactions*; John Wiley & Sons, 2014.

Table 1: Summary of reported 12α-HSDH. The given activity is on cholic acid as substrate.

Microbial Source	Ref.	Cofactor	Specific activity	Sequence
Clostidium leptum	[40]	$NADP^+$	3.3 U/mg	/
Clostidium group P.	[41]	$NADP^+$	128 U/mg	GenBank: HC036073.1
Eubacterium lentum	[42]	NAD^+	0.5 U/mg	/
Clostridium perfringens	[43]	NAD^+	/	/

Table 2: Summary of reported 7α-HSDH. The given activity is on chenodeoxycholic acid as substrate.

Microbial Source	Ref.	Cofactor	Specific activity	Sequence
Clostridium sordelii	[52]	$NADP^+$	1.1 U/mg	GenBank: AAA53556.1
Eubacterium scindens	[53]	$NADP^+$	338 U/mg	GenBank: AAB61151.1
Clostidium absonum	[54]	$NADP^+$	59 U/mg	GenBank: JN191345.1
Clostridium difficile	[55]	$NADP^+$	8.5 U/mg	Genbank: YP 001086529
Escherichia coli	[56]	NAD^+	190 U/mg	GenBank: KXH01569.1
Pseudomonas sp. B-0831	[57]	NAD^+	941 U/mg	GenBank: D50325.1
Bacteroides fragilis	[58]	NAD^+	351 U/mg	GenBank: OGX95366.1
Xanthomonas maltophilia	[30]	NAD^+	70 U/mg	/
Comamonas testosteroni	[59]	/	/	/

Table 3: Summary of reported 7β-HSDH The given activity is on 7-keto-LCA as substrate.

Microbial Source	Ref.	Cofactor	Specific activity	Sequence
Clostidium absonum	[54]	$NADP^+$	65 U/mg	GenBank: JN191345.1
Eubacterium aereofaciens	[60]	$NADP^+$	30 U/mg	GenBank: ZP0177306.1
Ruminococcus gnavus	[61]	$NADP^+$	23 U/mg	GenBank: ZP02041813
Collinsella aerofaciens	[63]	$NADP^+$	15 U/mg	GenBank: WP_006236005
Collinsella aerofaciens	[63]	$NADP^+$	21 U/mg	Engineered[a]
Ruminococcus torques	[62]	$NADP^+$	8.6 U/mg	GenBank: WP_015528793
Ruminococcus torques	[62]	$NADP^+$	46.8 U/mg	Engineered[b]
Xanthomonas maltophilia	[30]	NAD^+	33 U/mg	/

[a] G39A variant of the 7β-HSDH from *Collinsella aerofaciens*;
[b] T198V/V207M variant of the 7β-HSDH from *Ruminococcus torques.*

Table 4: Summary of reported whole-cell transformation with wild type microorganisms. Epimerization yields of CDCA to UDCA are given.

Microorganism	Ref.	Yield (%)
Colinsiella aerofaciens	[66]	/
Clostridium absonum	[67]	75%
E. coli + Bacteroides fragilis	[68]	25-30%
Colinsiella aerofaciens + Bacteroides fragilis	[68]	95%
Mixed culture	[69]	/
Clostridium limosum	[70]	55-60% (75-80%[a])
Stenotrophomonas maltophilia	[71]	27% (80%[b])

[a] Reported yield of epimerization of CA to ursocholic acid;
[b] Reported yield of epimerization of 12-keto-CDCA to 12-keto-UDCA.

Table 5: Summary of reported chemo-enzymatic transformations with purified enzymes.

Reaction pathway	Ref.	Conversion yield (%)
DHCA→12-keto-UDCA	[12]	85%
	[72]	95%
	[73]	99%
	[74]	95%
	[75]	99%
CA→UDCA	[76]	70%
CA→12-keto-UDCA	[64]	88%
	[77]	73%
CDCA→UDCA	[30]	82%
	[78]	100%
	[79]	100%
	[80]	63%

Figure legends:

Figure 1: Chemical structure of UDCA.

Figure 2: Chemical structures of bile salts.

Figure 3: Chemical structure of (A) cholic acid, (B) chenodeoxycholic acid and (C) lithocholic acid.

Figure 4: Comparison between Wolff-Kishner and Mozingo reduction. Notably the overall chemical reaction is the same for both cases.

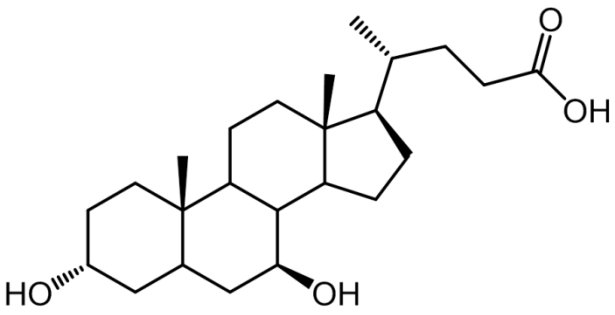
Figure 5: Reaction catalysed by the 12α-HSDH; the 12-OH group is of CA or UCA is oxidized yielding 12-keto-UDCA or 12-keto-CDCA, respectively, with the concomitant reduction of one molecule of $NAD(P)^+$.

Figure 6: Epimerization reaction catalysed by the (A) 7α-HSDH and (B) 7β-HSDH; the 7α-OH group of CA (R=OH) or CDCA (R=H) is firstly oxidized by the 7α-HSDH yielding 7-keto-DCA or 7-keto-LCA, respectively. Subsequently, the keto group is reduced by the 7β-HSDH giving the final products UCA or UDCA.

Figure 7: Overview of the chemo-enzymatic process for the production of UDCA from CA: the oxidation, reduction and dehydroxylation reactions are highlighted with a red, blue and green arrow, respectively. (A) DHCA; (B) 12-keto-UDCA; (C) CA; (D) 7-keto-DCA; (E) 7,12-diketo-LCA; (F) UDCA; (G) 12-keto-CDCA; (H) CDCA; (I) 7-keto-LCA. R=4-pentanoic acid.

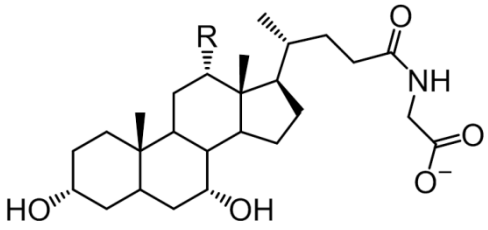
Figure 8: Chemo-enzymatic pathways for the formation of UDCA from CA that profit by the C-7 hydroxylation activity described by Sawada et al. [81]. CA (A) can be transformed to 7,12-diketo-LCA (C) trough specific oxidation of 7α-OH and 12α-OH with a 7α-HSDH and 12α-HSDH, respectively. Alternatively, 7,12-diketo-LCA can be obtained chemically oxidizing (eg. with CrO_3) all the hydroxyl groups, yielding DHCA (B) and then reducing the 3-keto group to 3α-OH by a 3α-HSDH. LCA (D) can be obtained from 7,12-diketo-LCA through dehydroxylation by Wolff-Kishner or Mozingo reduction. Finally, UDCA (E) can be obtained from LCA by 7β-hydroxilation.

Figure 1:

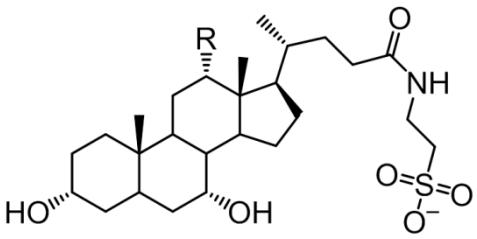


UDCA

Figure 2:

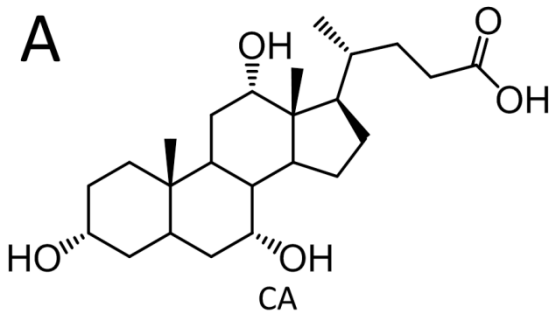


Glycocholate; R=OH
Glycochenodeoxycholate; R=H

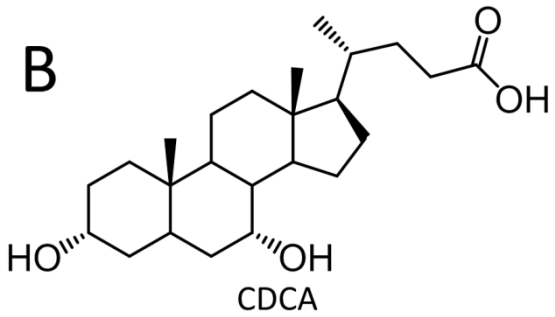


Taurocholate; R=OH
Taurochenodeoxycholate; R=H

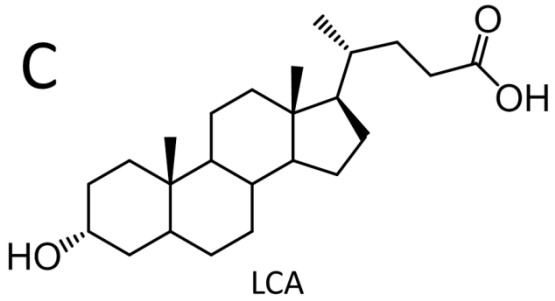
Figure 3:



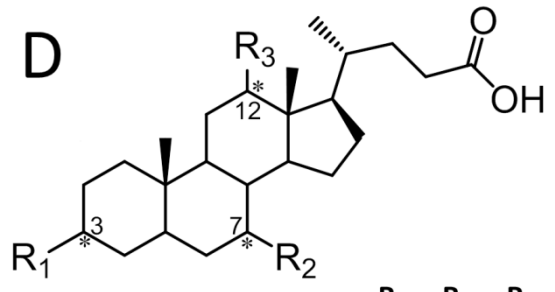
CA



CDCA



LCA



		R_1	R_2	R_3
Cholic Acid	CA	α-OH	α-OH	α-OH
Chenodeoxycholic Acid	CDCA	α-OH	α-OH	H
Deoxycholic Acid	DCA	α-OH	H	α-OH
Dehydrocholic Acid	DHCA	keto	keto	keto
Lithocholic acid	LCA	α-OH	H	H
Ursocholic acid	UCA	α-OH	β-OH	α-OH
Ursodeoxycholic acid	UDCA	α-OH	β-OH	H

Figure 4:

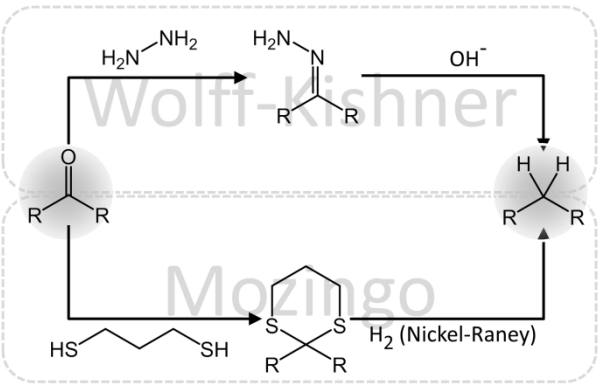


Figure 5:

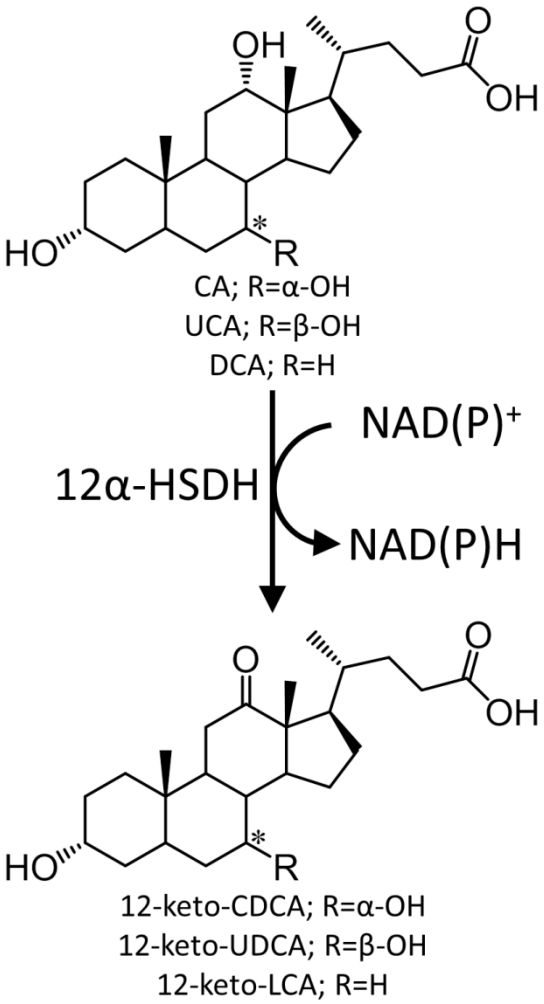


Figure 6:

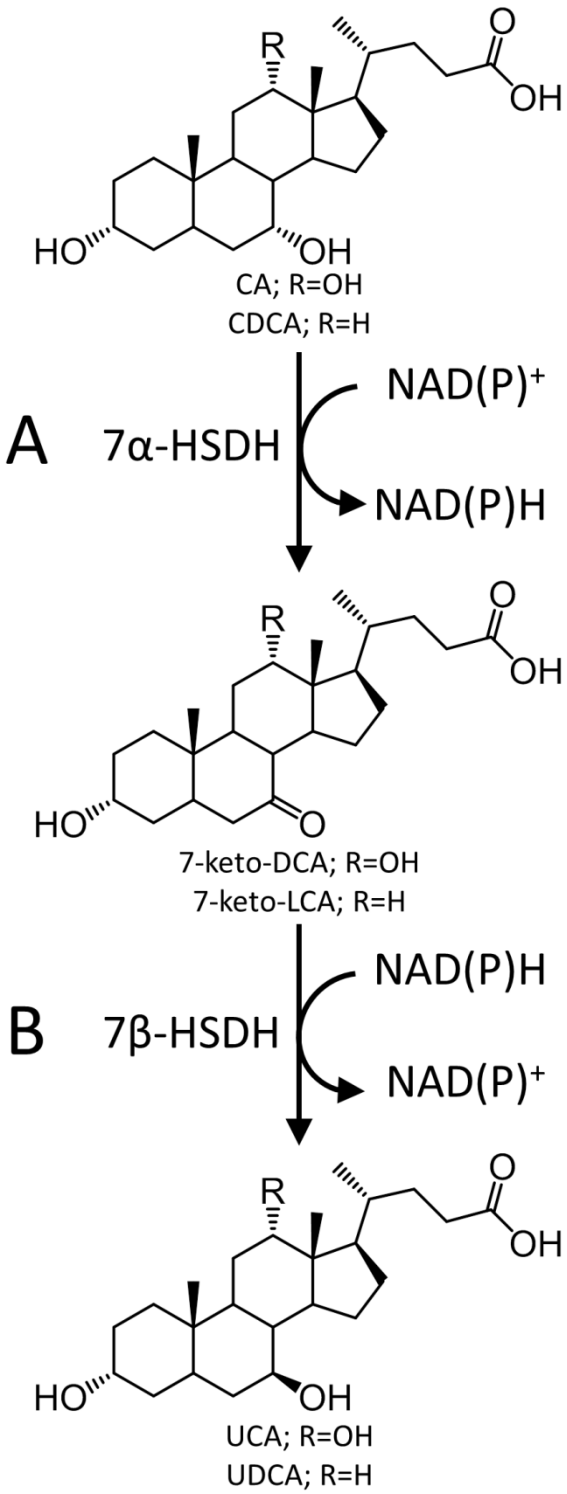

R
O
OH
HO
OH
CA; R=OH
CDCA; R=H
A
7α-HSDH
NAD(P)+
NAD(P)H
R
O
OH
HO
O
7-keto-DCA; R=OH
7-keto-LCA; R=H
B
7β-HSDH
NAD(P)H
NAD(P)+
R
O
OH
HO
OH
UCA; R=OH
UDCA; R=H

Figure 7:

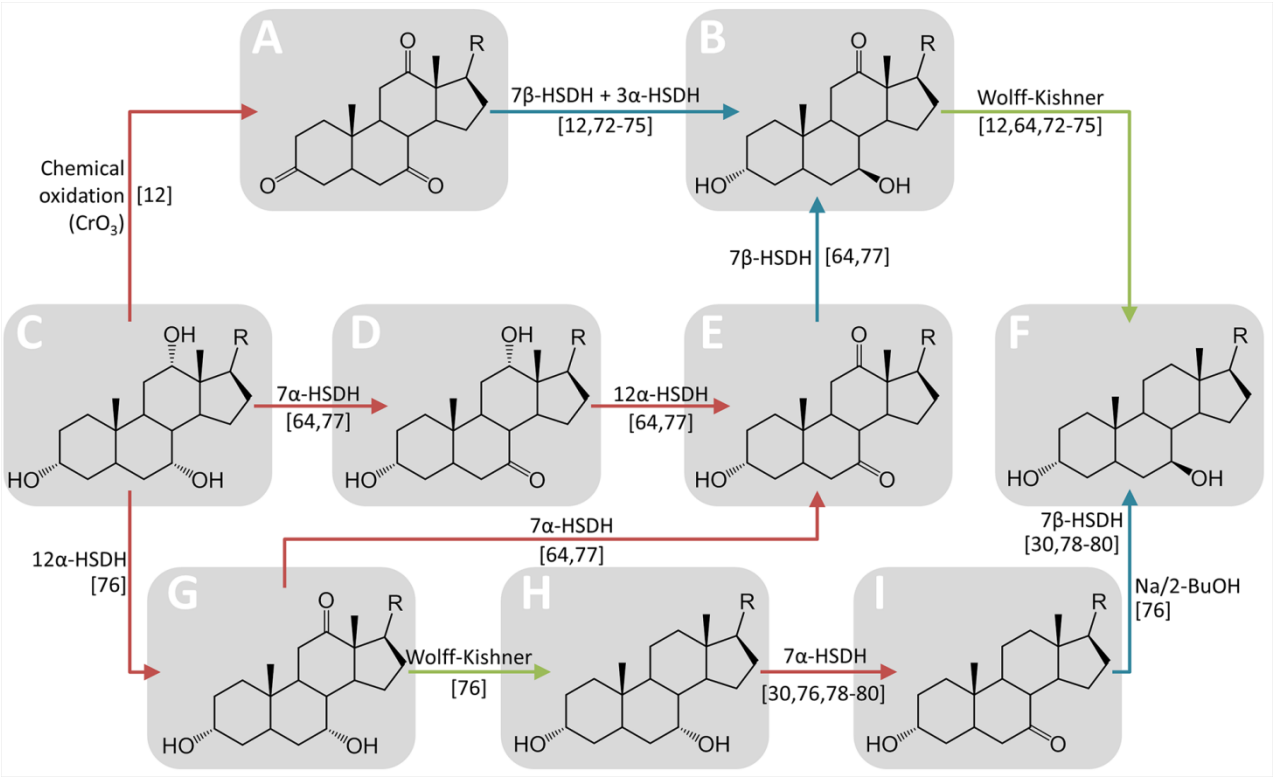
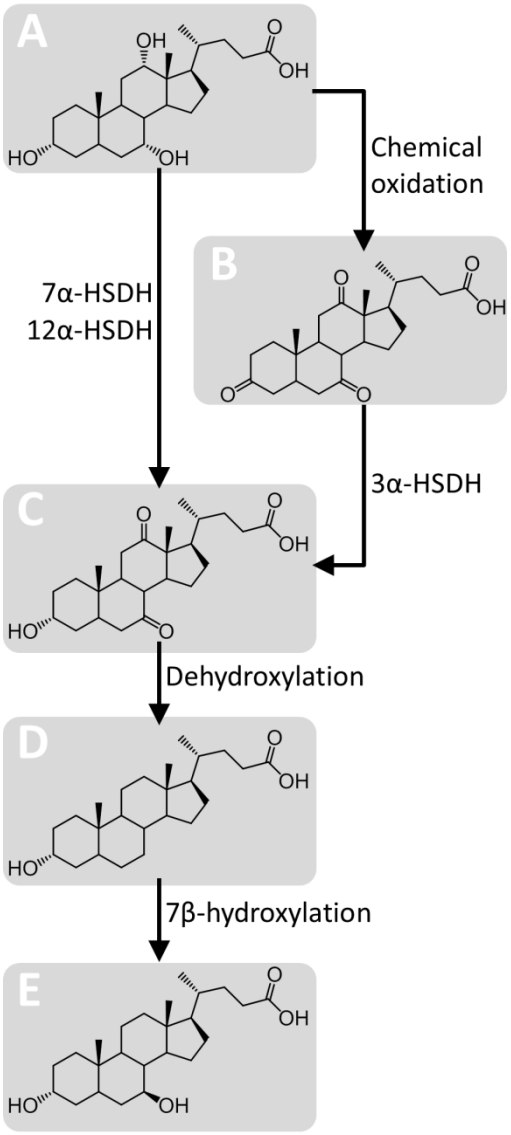

A
B
C
D
E
F
G
H
I
R
O
OH
HO
7β-HSDH + 3α-HSDH
[12,72-75]
Wolff-Kishner
[12,64,72-75]
Chemical
oxidation
(CrO3)
[12]
7β-HSDH [64,77]
7α-HSDH
[64,77]
12α-HSDH
[64,77]
12α-HSDH
[76]
7α-HSDH
[64,77]
7β-HSDH
[30,78-80]
Na/2-BuOH
[76]
Wolff-Kishner
[76]
7α-HSDH
[30,76,78-80]

Figure 8:


A
OH
O
OH
HO
OH
Chemical
oxidation
7α-HSDH
12α-HSDH
B
O
O
OH
O
O
3α-HSDH
C
O
O
OH
HO
O
Dehydroxylation
D
O
OH
HO
7β-hydroxylation
E
O
OH
HO
OH