



CHEMISTRY & SUSTAINABILITY

CHEM **SUS** CHEM

ENERGY & MATERIALS

Accepted Article

Title: NAD⁺-dependent enzymatic route for the epimerization of hydroxysteroids

Authors: Fabio Tonin, Linda G. Otten, and Isabel Arends

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *ChemSusChem* 10.1002/cssc.201801862

Link to VoR: <http://dx.doi.org/10.1002/cssc.201801862>

WILEY-VCH

www.chemsuschem.org

A Journal of



NAD⁺-dependent enzymatic route for the epimerization of hydroxysteroids

Fabio Tonin^[a], Linda G. Otten^[a] and Isabel W.C.E. Arends^{[a,b],*}

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

[b] Present address: Faculty of Science, Utrecht University, Budapestlaan 6, 3584 CD Utrecht, The Netherlands.

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

* Corresponding author

1 Abstract:

The epimerization of cholic and chenodeoxycholic acid (CA and CDCA, respectively) is a notable conversion for the production of ursodeoxycholic acid (UDCA). Two enantiocomplementary hydroxysteroid dehydrogenases (7α - and 7β -HSDHs) can carry out this transformation fully selectively by specific oxidation of the 7α -OH group of the substrate and subsequent reduction of the keto intermediate to the final product (7β -OH). In our quest to develop robust and active biocatalysts, we sought novel NADH-active 7β -HSDH's, allowing us to carry out a solely NAD⁺-dependent redox-neutral cascade for UDCA production. In this research a wild-type NADH dependent 7β -HSDH from *Lactobacillus spicheri* (*Ls7 β -HSDH*) was identified, recombinantly expressed, purified and biochemically characterized. Using this novel

NAD⁺ dependent 7 β -HSDH enzyme in combination with 7 α -HSDH from *Stenotrophomonas maltophilia*, permitted the biotransformations of CA and CDCA in the presence of catalytic amounts of NAD⁺, resulting in high yields (>90%) of UCA and UDCA.

2 Keywords:

Bile acids; Biotransformation; Hydroxysteroid dehydrogenases; UDCA; nicotinamide cofactor switching.

Accepted Manuscript

3 Introduction

Ursodeoxycholic acid (UDCA) is an active pharmaceutical ingredient (API) used in clinical therapy for the treatment of cholestatic diseases^[1]. It is a secondary bile acid that, as reported in several papers and reviews, shows several pharmacological effects^[2, 3]. Particularly, UDCA dissolves cholesterol gallstones, acting as a surfactant^[4, 5]. Additionally, other functions and influences on human physiology have been discovered: for example, UDCA improves the liver function in cholestatic diseases through secretion of glutathione^[6, 7] and interaction with several receptors^[8, 9], and significantly decreases cholesterol saturation in the bile^[10, 11]. Recently, other studies demonstrated anti-tumorigenic effects and a lower risk for some cancers^[12]. In comparison to chenodeoxycholic acid (CDCA) which is also used in therapies, UDCA provides total absence of side effects maintaining a high efficacy, in terms of pharmacological effects^[13, 14].

This API has been known in Chinese traditional medicine for many years, but natural UDCA can only be obtained by isolation from bear bile, making it very expensive^[15].

The current viable route for this API consists of chemical transformation of cholic acid (CA) or CDCA. These bile acids are obtained from the cheaper and more available bovine bile, that represents, in terms of industrial feasibility, the only source of hydroxysteroids. Currently, CA is chemically transformed into UDCA by a seven step chemical synthesis^[16, 17]. Unfortunately this route, first proposed by Hofmann^[16], requires the employment of toxic and dangerous reagents (eg. hydrazine, Cr₂O, pyridine) and results in large amounts of waste. In addition, it only has an overall yield of about 30%. In order to improve the efficiency of UDCA synthesis, alternative chemical routes have been proposed. Dangate^[18] for example achieved the transformation of CA into UDCA (53% yield using *o*-iodobenzoic acid, hydrazine and metallic sodium in *n*-propanol)

without the protection and de-protection steps used in the original route. However, this process is still not optimized in terms of costs and environmental impact. UDCA can also be obtained by epimerization of CDCA (intermediate in the synthesis of CA). In the latter case a viable production route would need to be complemented by a novel synthetic route for CDCA, because of the low natural availability of this compound.^[19]

An emerging technology for the modification of hydroxysteroids is the use of enzymatic transformations^[20, 21]. The regiospecificity of enzymes offers the possibility to avoid protection steps of the hydroxyl groups not employed in the transformation. This approach was followed by several research groups: attention was focused on the identification of enzymes able to play a role in this synthesis. In an enzymatic process for the production of UDCA from its precursors, several steps have to be performed. These are hydrolysis/deconjugation of glycine and taurine CA derivatives (used as raw starting material), the stereoinversion of the hydroxyl functions (by oxidation and subsequent reduction) and the specific hydroxylation and dehydroxylation of desirable positions in the steroid rings^[20].

The epimerization reaction (stereoinversion of 7 α -OH group), can be performed by two enzymes (7 α - and 7 β - hydroxysteroid dehydrogenase (HSDH)): the first enzyme specifically oxidizes the 7 α -OH group to the ketone, that is further reduced by the second enzyme to the corresponding alcohol group with inverted stereochemistry (β -OH). HSDHs are enzymes of the class of alcohol dehydrogenases (ADH) that use NAD(P)⁺ and NAD(P)H as electron acceptor and donor, respectively. In literature, several cascades that use these enzymes are reported^[22-27]: however, since known enzymes have different cofactor specificities, i.e. 7 α -HSDHs (NAD⁺ dependent) and 7 β -HSDHs (NADPH dependent) the oxidative and reductive reaction steps are decoupled in currently reported procedures (Figure 1 – state of art). The cofactors, in this case, are regenerated

by the addition of two other enzymes and two sacrificial substrates^[21-23]. In this specific case, the reaction is driven towards completion by the addition of sacrificial substrate (in great surplus): Although good yields were achieved (>80%), large amounts of side-products and waste were formed that have to be removed during the downstream processing steps, increasing the overall cost and the environmental burden of UDCA production.

Pedrini et al., already in 2006,^[28] reported the successful epimerization of CDCA to UDCA using a catalytic NAD⁺-dependent cascade reaction, with two NAD⁺ dependent dehydrogenases isolated from *Xanthomonas maltophilia* CBS 897.97 (recently reclassified as *Stenotrophomonas maltophilia*). In this way the requirement of external systems for cofactor regeneration was circumvented and UDCA was obtained with a final yield of 75%. However, the cultivation of this class 2 safety level microorganism is complex^[29] (microaerophilic conditions have to be used) and the 7 α -HSDH and 7 β -HSDH enzymes were only produced in low amounts (4 and 0.2 mg/L_{culture}, respectively). In addition, the amino acid or DNA sequences of these enzymes were not reported, making the recombinant expression of these proteins not possible. Since the activity and cofactor specificity is not clear from sequence alone, the univocal identification of the target corresponding gene in genomes databases is not possible. In addition, other works^[23, 24, 30, 31] have been reported on the employment of a NADP⁺ dependent redox neutral cascade for CDCA epimerization: however, in comparison to NAD⁺, the use of NADP⁺ is non-optimal because of the higher price, lower stability^[32] and lower natural availability. Generally, by the employment of a redox neutral cascade, 80% of UDCA is produced.

In this paper we report the development of an efficient enzymatic cascade, using well defined and easy to produce enzymes, for the production of UDCA and UCA from CDCA and CA, respectively. A full NAD⁺ mediated cascade, that circumvents the need for cofactor regeneration

was set-up resulting in conversion above 90%. The key discovery in this paper is the protein engineering of the NADP⁺ dependent 7 β -HSDH from *Clostridium sardiniense* that led to the identification of key residues responsible for NAD/NADP cofactor recognition. With this new sequence at hand, the natural NAD⁺ dependent enzyme from *Lactobacillus spicheri*, was disclosed for the first time as being highly active in the reduction of the 7-oxo-derivatives of CA or CDCA to either UCA or UDCA.

4 Results and discussion

4.1 Protein expression and purification

The genes coding for the *Sm7 α* -HSDH, *Cs7 β* -HSDH and *Ls7 β* -HSDH were cloned into pET24d(+) plasmid, yielding enzymes containing a C-terminal 6x His-tag. The recombinant enzymes were produced in *E. coli* BL21(DE3) host cells grown at 37 °C in LB medium. IPTG was added at the late exponential phase of growth and the cells were collected after another 18 h of incubation at 25 °C while shaking. The expression level under these conditions of the different proteins is reported in Table 1 and Supplementary Table 2. The His-tagged enzymes were purified by HiTrap chelating chromatography: all of the enzymes were isolated with \geq 95% purity, as judged by SDS-PAGE analysis (Supplementary Figure 2).

4.2 Mutagenesis of *Cs7 β* -HSDH

Since the three-dimensional structure of 7 β -HSDH from *C. sardiniense* is unknown, a model of the protein was built by using the comparative homology-modeling server SWISS-MODEL^[33]. 7 β -HSDH from *Collinsella aerofaciens* (PDB code 5GT9) was chosen as template because of the high sequence identity (42%) with *Cs7 β* -HSDH^[34]. The binding mode of the co-substrate NADP⁺ in the model of the *Cs7 β* -HSDH active site was analyzed: the ribose bound phosphate group, interacts with two arginine residues (R40 and R41) (Figure 2a).

On the basis of the *in silico* analysis, site-saturation mutagenesis (SSM)^[35] was performed at positions 39, 40 and 41 using the QuikChange kit and the wild-type *Cs7 β* -HSDH cDNA as template. In the first round of mutagenesis an aspartate was specifically introduced in position 39 while positions 40 and 41 were more randomly changed. The use of a DHK codon in these positions allows the introduction of 13 different amino acids (A, D, E, F, I, K, L, M, N, S, T, V

and Y) with low repetition, representing different chemical functionalities and steric encumbrances^[36]. The use of this codon drastically reduces the number of clones that have to be screened. The activity of *Cs7β*-HSDH variants on NAD⁺ as co-substrate was screened on a microtiter plate using a spectrophotometric method (production of NADH, measured at 340 nm) and an automated liquid-handling system. For the first round of mutagenesis, (G39D, R40X and R41X), 769 clones were screened, giving a probability of 91% that every combination of amino acids is measured. The clones most active on NAD⁺ as identified through the screening procedure were isolated and the substitutions were identified by DNA sequencing.

After the first round of mutagenesis and screening most of the enzyme variants obtained are inactive or have an activity lower than the control activity of wild-type *Cs7β*-HSDH on NAD⁺ (~1 mU mL⁻¹ under these conditions). However, four *Cs7β*-HSDH variants were isolated that showed a clear activity on NAD⁺: G39D, R40L, R41N (DLN variant); G39D, R40F, R41F (DFF variant); G39D, R40K, R41S (DKS variant) and G39D, R40K, R41V (DKV) variant. From the molecular modelling of these variants it could be observed that the hydroxyl group of threonine in position 17 may interfere with aspartate 39 in the hydrogen binding of the free OH-group of ribose (Figure 2b). A second pair of degenerated primers were designed in order to mutate that threonine (T17X). For this second round of mutagenesis, all four variants were used as template for the mutagenesis and 380 clones (96 clones for each variant) were screened (99% coverage). Out of all screened mutants, one variant (T17A, G39D, R40L, R41N – ADLN variant) showed a very high activity towards NAD⁺ as cosubstrate in the microtiter plate assay employing cell extract (Figure 2c).

The ADLN variant was expressed in *E. coli* BL21(DE3) cells and purified by HiTrap chelating chromatography (>90 % purity). This variant shows an expression yield similar to that of the

wild-type *Cs7* β -HSDH (in terms of purified protein/liter of fermentation broth, see Supplementary Table 2). The Michaelis-Menten kinetics of all selected enzyme variants were measured using NAD⁺ or NADP⁺ as cofactor.

As shown in Table 2, among the *Cs7* β -HSDH variants obtained by SSM, the ADLN was identified because of a thirty-fold increase in specific activity compared to wild-type enzyme.

In order to further increase the activity on NAD⁺, the T17A, E18D, G39D, R40L, R41N (ADDLN) variant was designed and recombinantly expressed: the substitution of the glutamate with the smaller aspartate in position 18 may result in the formation of a second hydrogen bond between the cofactor and the protein (Figure 2d). The comparison of the kinetic parameters of the ADLN and ADDLN variants of *Cs7* β -HSDH indicates that the addition of a second hydrogen bond increases the affinity of the ADDLN variant for the NAD⁺ (8-fold decrease in K_m), see Table 2. However, the specific activity on NAD⁺ of this variant decreases (10-fold lower in comparison to previous isolated one). In order to restore the specific activity on NAD⁺ a fourth round of SSM was carried out employing the same primers used in the first round and the ADDLN variant as template. The T17A, E18D, G39D, R40A, R41A variant (ADDAA) was isolated. The specific activity under standard conditions of this variant is 0.3 U mg⁻¹ (Table 2). Interestingly the affinity for the cofactor is lower than the one observed in the previous variants. However, the expression level of this protein is fifteen-fold higher than the wild-type *Cs7* β -HSDH (486 vs. 20 mg/L_{culture}), resulting in the higher activity under the screening conditions. Overall, we conclude that ADLN is the best obtained variant for our purpose due to its highest catalytic efficiency.

4.3 New 7 β -HSDH from *Lactobacillus spicheri*

With the new sequence of NADH dependent 7 β -HSDH (ADLN and ADDLN variants) in hand, a

new alignment search was carried out in order to find potentially better, previously non-identified enzymes, able to convert CA and its derivatives. The *Ls7* β -HSDH sequence was identified using the Basic Local Alignment Search Tool (BLASTp): the predicted sequence analysis showed a 792 bp ORF corresponding to a protein of 264 amino acids residues.

This enzyme was identified as a putative NADH dependent 7 β -HSDH based on the fact that the amino acids relevant to the binding and recognition of NADH are present: specifically, the alanine and aspartate in position 18 and 19 and the stretch DYS in position 40-42 are conserved^[37], analogous to the situation in analogue *Cs7* β -HSDH (Figure 3b). The predicted Mw of 29 kDa and the predicted homodimeric quaternary structure (Figure 3a) put *Ls7* β -HSDH in the short chain dehydrogenase/reductase superfamily.

4.4 Biochemical characterization of HSDHs

Sm7 α -HSDH showed a strict NAD⁺ activity on both CDCA and CA, although the activity on CA is considerably lower (halved). No activity was detected when NADP⁺ was used as electron acceptor. *Sm7* α -HSDH displayed a K_m of 0.22 and 0.96 mM for CDCA and CA, respectively. The *Sm7* α -HSDH did not show any substrate inhibition on CA, but a K_i of 11 mM on CDCA was measured. The K_m value for NAD⁺ is 0.55 mM (Table 3) which is similar to earlier reported enzymes.

The pH and temperature dependence of *Sm7* α -HSDH activity was investigated: both the maximum activity and stability occurred at slightly alkaline pH values (Figure 4a). The enzyme is quite thermophilic, showing an optimum at around 70 °C (Supplementary Figure 3a), and is quite stable: after 24 h incubation at 25 and 37 °C, the enzyme maintained ca. 100 and 70% of its initial activity, respectively. Incubations at higher temperatures (60 °C) resulted in a complete loss of enzymatic activity. The enzymatic activity of *Sm7* α -HSDH was also investigated in the

presence of different concentrations of methanol, that could be used as a potential co-solvent in order to increase the solubility of hydroxysteroids in water. The enzyme shows no loss of activity in 10% methanol and it conserves 90% of its activity in 20% methanol (Figure 5a).

On the other hand, wild-type *Cs7 β* -HSDH showed a strict NADP⁺ activity on UDCA (0.74 U mg⁻¹ under standard conditions). The activity on NAD⁺ is roughly 100-fold lower showing a K_m of 2.6 mM and a specific activity of 0.023 U mg⁻¹. *Cs7 β* -HSDH displayed a K_m of 0.16 mM for UDCA (Table 4). The pH and temperature dependence of *Cs7 β* -HSDH activity was investigated: both the maximum activity and stability occurred at slightly alkaline pH values (Figure 4b). Notably, the pH optimum for the reduction reactions was detected at pH 6-7. In the presence of NADPH, this enzyme is able to reduce both 7-oxo-DCA and 7-oxo-LCA (see Table 4).

In general, the enzyme is less thermophilic than the *Sm7 α* -HSDH, showing an optimum at around 60 °C (Supplementary Figure 3b), and less stable: after 24 h incubation at 25 and 37 °C, *Cs7 β* -HSDH maintained ca. 85 and 62% of its initial activity, respectively. *Cs7 β* -HSDH shows a good tolerance to concentrations of methanol above 10% and 70% of its activity is conserved in presence of 20% methanol.

The ADLN, ADDLN and ADDAA variants showed an increase in activity towards NAD⁺ and NADH as cosubstrate, with little change in specificity for the different substrates (Table 4).

Interestingly, the isolated variants showed higher stability than the wild-type *Cs7 β* -HSDH. From this comparison, it can be observed that the ADLN variant maintains 95% of its activity after incubation for 24h at 25 °C (compared to the 85% for the wild-type under the same conditions).

As predicted by *in silico* analysis, *Ls7 β* -HSDH showed a strict NAD⁺ activity (3.10 U mg⁻¹ in standard condition). A weak activity was detected when NADP⁺ was used as electron acceptor (Table 2). *Ls7 β* -HSDH showed a 0.15, 0.04 and 0.13 mM K_m for UDCA, 7-oxo-LCA and 7-oxo-

DCA, respectively. *Ls7* β -HSDH is inhibited by UDCA showing a K_i of 0.8 mM for this substrate. The K_m value for NAD^+ is 0.08 mM. From all the 7 β -HSDH-variants in this paper, this enzyme shows the highest NAD^+/NADH activities for both the reduction and the oxidation of 7 β -OH hydroxysteroids and its derivatives. All kinetic data are presented in Tables 2 and 4.

From the biochemical characterization it can be observed that this enzyme showed similar features to the other isolated enzymes, with the exception of the pH dependence: the maximum activity for both oxidation and reduction reactions was observed at lower pH (6-7) (Figure 4f).

Ls7 β -HSDH is quite thermophilic, showing an optimum at around 70 °C (Supplementary Fig. 4C), and it is stable at 25 and 37 °C, maintaining, after 24 h of incubation, ca. 98 and 72% of its initial activity, respectively. *Ls7* β -HSDH conserves 60% of its activity in presence of 20% methanol: thus for follow-up bioconversion studies 10% methanol was used, which limits the amount of substrate that can be loaded in a biotransformation (Figure 5f).

4.5 Bioconversion of CDCA and CA

In order to assay the applicability of the enzymes described above as biocatalyst for the production of UDCA and UCA in a batch bioreactor, lab-scale bioconversions were carried out: 10 μmol of CDCA in the presence of 1 μmol of NAD^+ and 10% MeOH was converted into UDCA (92% yield) in 200 min by 1 U of *Sm7* α -HSDH (0.23 μg) and 0.6 U of *Ls7* β -HSDH (190 μg) (Figure 6a). HPLC analysis showed no formation of undesired products, except for the 7-oxo-LCA (0.5 μmol), intermediate of the enzymatic cascade (Supplementary Figure 4a).

Under the same conditions, 10 μmol of CA was converted into UCA (91%) in 120 min (Figure 6b).

Surprisingly the addition of MeOH (10-20%) resulted in an increase (10%) of conversion (Table 5): without cosolvent only 80% of UDCA can be obtained by this biotransformation.

The effect of cosolvents on the equilibrium of the epimerization reaction was already discussed in literature^[38] but never observed for this specific biotransformation.

Additionally, in order to optimize the reaction conditions, a set of bioconversions were carried out: as reported in Table 5, >90% of UDCA was observed employing 1 or 0.5 μmol of NAD^+ . Biocatalytic conversions were also carried out in the presence of 0.2 μmol of NAD^+ using CDCA as substrate: in this case the conversion is slower, yielding 24% of UDCA in 20 h (Figure 6c). Notably, here the detected amount of keto intermediate was lower than 1% (0.1 μmol). Bioconversions were also tested at different pHs (6 and 7) but no improvements were observed (after 150 min in presence of 1 mM NAD^+ , 77%, 81% and 87% conversion was observed at pH 6, 7 and 8, respectively). Unfortunately, the increase of substrate loading to 20 mM negatively affects the bioconversion, yielding 20% of product after 12 h of incubation.

The ADLN variant of *Cs7 β* -HSDH was also tested in the cascade reaction in presence of 1 mM NAD^+ for the epimerization of CDCA and CA: although no conversion was observed when CDCA was used as substrate, 10 μmol of CA was converted into UCA (91% yield) by 790 μg of enzyme after 30 min (Figure 6d and Supplementary Figure 4b). This behavior can partially be explained by the kinetic parameters of this variant: K_i for the intermediate 7-oxo-LCA is 4 times lower than the one for 7-oxo-DCA.

Notably, in the different bioconversions the amount of keto intermediate is limited by the amount of cofactor used. Thus in a typical reaction where 10% cofactor was used, 5% keto intermediate is formed as byproduct.

5 Conclusion

Epimerization of CA and its derivatives is a pivotal step in the synthesis of UDCA (see Figure 1). Enzymatic epimerization circumvents the needs for protection/deprotection steps and can be

carried out with high yield. A solely NAD^+ dependent cascade for CDCA and CA epimerization, in which NAD^+ is present in catalytic amounts, represents a promising way towards a more efficient process in terms of atom economy and number of reagents required.

In order to achieve that goal it is fundamental to employ robust, active and selective enzymes: *Sm7 α -HSDH* shows kinetic parameters better or comparable to its *E. coli* or *Bacteroides* homologues, widely used by researchers in these kind of processes^[22, 27, 39]. This enzyme selectively oxidizes the 7 α -OH group of CA and CDCA. In terms of stability it shows promising tolerance to the presence of methanol and elevated temperatures.

Until now, all sequences of 7 β -HSDHs reported in literature are NADP^+ dependent enzymes^[20]. However, protein engineering provides a straightforward method to change cofactor specificity: this has been the subject of many researchers in both academia and industry for the past decades^[40-42]. Unfortunately, the resulting enzymes usually have a perfect cofactor switch, but in several cases, the enzyme activity and/or substrate specificity decreases tremendously^[43]. A software for the prediction of the specificity-determining residues (CSR-SALAD^[44, 45]) has been developed and is now widely used^[46], but there is still no comprehensive solution for cofactor switching.

Following this approach, an iterative mutagenesis of *Cs7 β -HSDH* was performed. A semi-rational approach was pursued, based on the identification of the residues responsible for the cofactor recognition discovered in the homology model of the enzyme.

In this way, residues G39, R40 and R41 were identified to be responsible for binding of the phosphate group at the 2'-position of the ribose. Interestingly, the importance of arginine in the binding of NAD(P)^+ was shown before on the enantiocomplementary 7 α -HSDHs^[47]. The amino acid in position 39 was mutated to an aspartate. Notably, a similar attempt of cofactor switching,

performed by Bakonyi and Hummel on the *Clostridium difficile* 7 α -HSDH^[48] has led to the identification of similar variants (A37D). In our case, the free hydroxyl group at 2'-of ribose forms a hydrogen bond with the aspartate (length 2.2 Å). R40 and R41 residues were randomly mutagenized in order to identify a couple of amino acids that I) do not interfere with aspartate in NAD⁺ binding and II) create a binding pocket for NAD⁺. SSM with limited codons helped to introduce in the protein a certain degree of variability while generating smart libraries containing <1000 clones for each round.

In the same way, T17 was subjected to SSM. From the 3D structure of the model it was observed that the OH-group of the side chain is 2.5 Å from the aspartate. The obtained ADLN variant shows a reasonable activity on NAD⁺ (in comparison with the wild-type enzyme on NADP⁺). The K_m of this variant for NAD⁺ is 2 mM, which is actually close to the value reported by Pedrini et al., for their 7 β -HSDH from *S. maltophilia*^[28]. Additionally the mutation E18D improved the affinity for NAD⁺ about 10-fold but unfortunately, the specific activity of the protein decreases. We hypothesize that the E18 residue in the *Cs7 β -HSDH* also has a structural role. From this optimization round we thus concluded that the ADLN variant was the best mutant.

These data suggest that the efficiency and the selectivity in cofactor binding depends for a large part, on the molecular dynamics of the *Cs7 β -HSDH*. Although this feature can be reasonably simulated, the engineering of the target enzyme is still complex^[49]. For this reason, the attention was focused on the identification of a wild-type enzyme that conserved both I) the structure of *Cs7 β -HSDH* and II) the residues identified in *Cs7 β -HSDH* to be responsible for cofactor recognition. The new analysis lead us to identify a sequence of a putative NAD⁺ dependent 7 β -HSDH from *Lactobacillus spicheri*. *Ls7 β -HSDH* shows indeed a NAD⁺ dependence activity (K_m 80 μ M) and an activity higher than the *Cs7 β -HSDH* on the respective cofactor. As shown, both the ADLN variant of *Cs7 β -HSDH* and the *Ls7 β -HSDH* are suitable for the epimerization of CA

(to UCA) but, only the *Ls7* β -HSDH is able to epimerize CDCA (to UDCA): this can partially be explained by the comparison of the kinetic parameters of the two enzymes: firstly, the ADLN variant, although it was engineered, shows a K_m for the NADH 35-fold higher than the *Lactobacillus* one; secondly the ADLN variant is inhibited by 7-oxo-LCA ($K_i=2.0$ mM). In general, *Ls7* β -HSDH is a more interesting biocatalyst because of the high activity on both 7-oxo derivatives.

Bioconversions were carried out at the substrate concentration used in previous studies (10 mM). This choice was done in order to I) have a straight comparison between the reported systems and the current redox neutral cascade and II) ensure solubility, because solubility of CDCA in a water environment (pH 8.0) is low (around 10 mM as specify by Zheng^[23]).

The developed cascade shows an epimerization yield >90% for both CA and CDCA under non-optimized conditions. This shows that the currently developed set of enzymes and the epimerization cascade belonging to it, provides a highly promising alternative for both chemical and enzymatic synthetic procedures applied currently^[20, 21]. This could provide a basis for a much more elegant and atom efficient synthesis of UDCA from CDCA.

Since the epimerization of CA and CDCA is carried out with catalytic amounts of NAD^+ , the shift of the equilibrium depends on the difference in physio-chemical properties between reagents and products. Particularly, in a preliminary study using a molecular mechanics approach, we have identified a difference in the ΔG_0 between the two epimers of -1.23 kJ mol⁻¹ (calculated at standard conditions). In addition the differences in solubility and in critical micelles concentration (CMC) of the two isomers can play a role in the shift of the equilibrium to the products. Similar results were shown by Dangate^[18] and Giovannini^[50] in chemical and chemo-enzymatic routes, respectively: Dangate et al. reported that 7-oxo-DCA is preferably reduced to the 7 β epimer (UCA:CA – 80%:10% yield) in presence of sodium in anhydrous *n*-propanol.

Under the same conditions, Giovannini et al. showed similar results for the reduction of 7-oxo-LCA (UDCA:CDCA – 82%:15% yield). Also in our case we observed that the reduction towards the 7 β epimer is thermodynamically favored, but differently from the chemical cascade, the reversibility of the enzymatic catalyzed reactions allows to reach the thermodynamic equilibrium between the two epimers. In comparison to the previously reported NADP⁺-dependent redox neutral systems we observed an higher conversion yield (>90% vs. 80% reported by Zheng^[23]). This increase can be explained by the use of the cosolvent (10% MeOH) that influences the reaction environment (Table 5) without affecting the activity of the employed enzymes. Another advantage of our system is given by the use of NAD⁺ instead of NADP⁺, that lowers the cost of the required cofactor and shows higher stability in solution.

As previously observed by Zheng^[23], in all the epimerization reactions, 5% of keto intermediate (7-oxo-DCA or 7-oxo-LCA) was obtained: For future studies in order to decrease this byproduct, since the amount of intermediate is related to the amount of cofactor in solution, lower NAD⁺ concentrations should be used at the expense of a longer reaction time (e.g. if 0.2 mM of NAD⁺ is used, the amount of intermediate cannot exceed the 2%). In order to evaluate the performance of our system, process metrics were calculated: under the condition described in this study (1 and 0.6 U mL⁻¹ of *Sm7 α* -HSDH and *Ls7 β* -HSDH, respectively) the system shows a TTN of 2.9 million and a space-time yield of 26 g L⁻¹ d⁻¹. Despite a good TTN, the space time yield is not yet high enough for industrial application. Further studies, including the use of flow reactors, will be carried out to address the down-stream processing aspects of this reaction. This, in combination with solvent engineering and the employment of biphasic systems, both targeted to the increase of substrate loading (and decrease the cofactor loading), could optimize the condition for the industrial preparation of UDCA^[51].

In conclusion, we have shown that epimerization reaction of CA and CDCA can be achieved by a

two-enzyme-one-cofactor cascade which benefits from favorable thermodynamics for the 7 β -OH isomer. In particular *Sm7 α -HSDH* and *Ls7 β -HSDH* are a compatible, stable set of enzymes and promising candidates as biocatalyst in the synthesis of hydroxysteroid derivatives. Further investigations will be carried out in order to develop a flow system or a membrane reactor similar to the ones described in literature^[22, 52] for our newly developed biocatalytic UDCA synthesis procedure and to increase the substrate loading in the process.

6 Experimental Section

6.1 Bacterial strains and materials

Escherichia coli TOP 10 [F⁻ mcrA Δ (mrr-hsdRMS-mcrBC) ϕ 80lacZ Δ M15 Δ lacX74 nupG recA1 araD139 Δ (ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ] and BL21(DE3) [F⁻ *ompT gal dcm lon hsdS_B(r_B⁻m_B⁻)* λ (DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*⁺]_{K-12}(λ ^S)] were purchased from Invitrogen (Darmstadt, Germany). Tryptone, yeast extract, cholic acid, ursodeoxycholic acid and nicotinamide adenine dinucleotide cofactors (NAD(P)⁺/H) were from Sigma-Aldrich (St. Louis, US). Ursodeoxycholic acid was from Tokio Chemical Industry (TCI) (Tokyo, Japan). Restriction enzymes and Phusion Q-5 DNA polymerase mastermix were from New England Biolabs (NEB, Ipswich, US). All other reagents were of analytical grade and commercially available.

6.2 Identifying and homology models of 7 α -hydroxysteroid dehydrogenases

The 7 α -hydroxysteroid dehydrogenase from *Stenotrophomonas maltophilia* (*Sm7 α -HSDH*) and 7 β -hydroxysteroid dehydrogenase from *Lactobacillus spicheri* (*Ls7 β -HSDH*) gene sequences were identified by multiple sequence analysis using BLASTp (NCBI, <https://blast.ncbi.nlm.nih.gov>): for *Sm7 α -HSDH*, protein sequences with known activity from

Clostridium sardiniense (GenBank: AET80685.1), *Clostridium difficile* (GenBank: CAJ66880.1, putative), *Clostridium sordelli* (GenBank: L12058.1), *Eubacterium* sp. VPI 12708 (GenBank: M58473.1), *Bacteroides fragilis* (GenBank: AF173833.2) and *Escherichia coli* (Gene ID: ACI83195.1) were used as query, restricting the organisms list to *Stenotrophomonas maltophilia* strains (taxid:40324). A multiple sequence analysis of all 7 α -HSDH is shown in Supplementary Figure 1a.

A 3D structure model of this enzyme was obtained using SWISS-MODEL (<https://swissmodel.expasy.org/interactive>), employing the crystal structure of the 7 α -HSDH from *Escherichia coli* (PDB ID: 1AHI.1) as template.

For *Ls7* β -HSDH, protein sequences from *Clostridium sardiniense* (GenBank: AET80684.1), *Ruminococcus gnavus* (Genbank: WP004843516.1), *Ruminococcus torques* (GenBank: WP015528793.1) and *Collinsella aerofaciens* (GenBank: WP006236005.1) together with the variant of the *Cs7* β -HSDH obtained (see below) were used as query. A multiple sequence analysis of all 7 β -HSDH is shown in Supplementary Figure 1b.

The 3D models of *Cs7* β -HSDH (GenBank: AET80684.1) and *Ls7* β -HSDH were built using the same server employing the entire structure of 7 β -HSDH from *Collinsella aerofaciens* (PDB code 5GT9) as template.

6.3 Cloning and expression of *Sm7* α -HSDH, *Cs7* β -HSDH and *Ls7* β -HSDH

The synthetic cDNAs encoding the *Sm7* α -HSDH, *Cs7* β -HSDH and *Ls7* β -HSDH were designed by *in silico* back translation of the amino acid sequences (GenBank: KRG42928.1, AET80684.1 and WP045806907, respectively). In order to subclone into the pET24d(+) plasmid (Merck Millipore, Burlington, US), sequences corresponding to *Nco*I (CCATGG) and *Xho*I (CTCGAG)

restriction sites were added at the 5'- and 3'-ends of the cDNAs, respectively. The codon usage of the synthetic genes was optimized for expression in *E. coli* and produced by BaseClear B.V. (Leiden, The Netherlands). *Sm7 α* -HSDH, *Cs7 β* -HSDH and *Ls7 β* -HSDH cDNAs were cloned into the pET24d(+) vector using the *Nco*I and *Xho*I sites, resulting in 6.0-, 6.0- and 6.1-kb construct (pET24-*Sm7 α* -HSDH, pET24-*Cs7 β* -HSDH and pET24-*Ls7 β* -HSDH). The genes were cloned in frame with the C-terminal 6-Histidines tag of the vector.

The obtained expression plasmids were then used to transform BL21(DE3) *E. coli* cells.

Starting cultures (100 mL) were prepared from a single recombinant BL21(DE3) *E. coli* colony grown in LB medium containing kanamycin (30 $\mu\text{g mL}^{-1}$), under vigorous shaking (200 rpm) at 37 °C. These cultures were diluted to a starting OD_{600nm} of 0.1 in 1 L of LB medium (LB, 10 g L⁻¹ bacto-tryptone, 10 g L⁻¹ NaCl and 5 g L⁻¹ yeast extract) and then incubated at 37 °C on a rotatory shaker at 200 rpm until an OD_{600nm} of 1.0 was reached. Protein expression was induced by adding 0.25 mM IPTG: cultures were grown for another 12 h at 25 °C with shaking (200 rpm). Cells were harvested by centrifugation at 10,000 x g for 10 min at 4 °C, washed with 50 mM KPi buffer pH 8.0 and stored at -20 °C for at least 1 day before purification.

6.4 Protein purification

E. coli cell pellets were resuspended in lysis buffer (50 mM KPi buffer, 1 M NaCl, 5% glycerol (v/v) and 10 $\mu\text{g mL}^{-1}$ DNase, pH 8.0) and disrupted by French press (Constant Systems Limited, Low March, UK) (2 cycles, 180 psi). The insoluble fraction of the lysates were removed by centrifugation at 39,000 x g for 30 min at 4 °C. Crude extract was loaded onto a HiTrap chelating affinity columns (GE Healthcare, Little Chalfont, UK), previously loaded with Ni²⁺ metal ions and equilibrated with 50 mM KPi buffer, 1 M NaCl and 5% glycerol (v/v) pH 8.0. The columns were washed with this buffer until the absorbance value at 280 nm was that of the buffer and the

bound proteins were eluted with 50 mM KPi buffer, 250 mM imidazole and 5% glycerol (v/v), pH 8.0. The fractions containing the desired activity were dialyzed overnight against 50 mM KPi buffer and 5% glycerol (v/v), pH 8.0, using a 3-kDa dialysis tube.

During the purification procedure, 7 α -HSDH and 7 β -HSDH activities were assayed using the standard activity assay (see below).

6.5 Activity and kinetic measurements

7 α -HSDH's enzymatic activity in the crude extract and of the purified enzyme was determined at 25 °C using 1.0 mM CDCA, 2.0 mM NAD⁺, in 50 mM KPi buffer and 10% methanol (v/v), pH 8.0. 7 β -HSDH's enzymatic activity was determined at room temperature (25 °C) using 1.0 mM UDCA, 2.0 mM NAD(P)⁺, in 50 mM KPi buffer and 10% methanol (v/v), pH 8.0. The production or the consumption of NAD(P)H was followed at 340 nm (extinction coefficient of NAD(P)H is 6,220 M⁻¹·cm⁻¹). One unit (U) was defined as the amount of enzyme producing 1 μ mol of product per minute at 25 °C and at pH 8.0. Blank measurements were performed in absence of CDCA or UDCA, NAD⁺ and enzyme.

The kinetic parameters of the purified samples were determined at room temperature in the presence of: different concentrations of substrates (5–10000 μ M), 2.0 mM NAD(P)⁺ in 50 mM KPi buffer and 10% methanol (v/v), pH 8.0, at 25 °C; different concentrations of NAD(P)⁺ (1–5000 μ M), 2.0 mM CDCA (for 7 α -HSDH) or UDCA (for 7 β -HSDH) in 50 mM KPi buffer and 10% methanol (v/v), pH 8.0, at 25 °C. The specific activity was expressed as unit per mg of protein (determined by spectrophotometric analysis at 280 nm). The kinetic data were fitted to the Michaelis–Menten equation, or to the one modified to account for substrate inhibition when necessary.

The effect of pH on the enzymatic activities was determined using 1.0 mM CDCA (for 7 α -

HSDH) or UDCA (for 7 β -HSDH), 2.0 mM NAD(P)⁺, in 100 mM citrate-phosphate buffer (66 mM citrate, 34 mM Na₂HPO₄) and 10% methanol (v/v), in the 3.0–10.0 pH range.

The effect of methanol concentration on the enzymes activity toward CDCA and UDCA was determined using 1.0 mM CDCA (for 7 α -HSDH) or UDCA (for 7 β -HSDH), 2.0 mM NAD(P)⁺ in 50 mM KPi buffer and different concentration of methanol (0-50% (v/v)), pH 8.0, at 25 °C.

Temperature dependence of the enzymatic activities was determined using 1.0 mM CDCA (for 7 α -HSDH) or UDCA (for 7 β -HSDH), 2.0 mM NAD⁺ in 50 KPi buffer and 10.0% methanol (v/v), pH 8.0 in the 18–95 °C temperature range.

Enzymatic stability was measured by incubating the enzyme solution in 100 mM citrate-phosphate buffer (66 mM citrate, 34 mM Na₂HPO₄) in the 3.0–9.0 pH range at 25 °C^[53], in 50 mM KPi buffer with different concentration of methanol (0-50% (v/v)) at pH 8.0 at 25 °C and in 50 mM KPi buffer, at pH 8.0 at different temperatures: samples were withdrawn at different times and residual activity was determined using the enzymatic activity assay.

6.6 SDS-PAGE

Proteins from crude extract and the purified enzyme fractions were separated by SDS-PAGE on a 12% polyacrylamide resolving gel (BioRAD, Hercules, US): samples were resuspended in an appropriate volume of Laemmli sample buffer and boiled. Proteins were visualized by staining with SimplyBlue safe stain (Novex, Carlsbed, US).

6.7 Site-saturation mutagenesis (SSM) and screening NAD⁺ dependent enzyme variants

SSM was carried out at different amino acid positions of *Cs7* β -HSDH by using whole plasmid PCR^[54]; pET24d-*Cs7* β -HSDH vector as template and a set of degenerated synthetic oligonucleotides was employed to prepare the mutant libraries (Supplementary Table 1). PCRs

were carried out in a final volume of 20 μL : in all cases, 1 μL of template DNA ($\sim 25 \text{ ng } \mu\text{L}^{-1}$), 0.5 μL of each primer (final concentration 0.25 μM), 10 μL of Phusion Q-5 DNA polymerase mastermix and 8 μL of MilliQ water were added in sterile PCR tube. After an initial denaturation step (98 $^{\circ}\text{C}$ for 1 min), reaction was carried out for 30 cycles (denaturation at 98 $^{\circ}\text{C}$ for 30 sec., annealing at 58 $^{\circ}\text{C}$ for 30 sec. and elongation at 72 $^{\circ}\text{C}$ for 7 min). A final elongation step (72 $^{\circ}\text{C}$ for 7 min.) was added. The template DNA was eliminated by enzymatic digestion with 1 μL *DpnI* restriction enzyme at 37 $^{\circ}\text{C}$ for 2 hours; the PCR products were used to transform *E. coli* TOP10 cells. Subsequently, the recombinant plasmids were transferred to *E. coli* BL21(DE3) cells, and these clones were used for the screening procedure. The introduction of the mutations was confirmed by automated DNA sequencing. The mutant libraries obtained from SSM were screened by means of a rapid colorimetric assay based on the reduction on NAD^{+} (as described before) and by means of an automated JANUS Liquid Handler Workstations (Perking Elmer, Waltham, US). To a saturated *E. coli* culture (1 mL, growth in 2mL DeepWell plate) 0.250 mM IPTG were added and the culture was then incubated at 25 $^{\circ}\text{C}$ for 18 h. The culture was centrifuged at 5,000 $\times g$ for 2 min, and the cell pellet was resuspended in 200 μL of 50 mM KPi buffer, pH 8.0 containing of 1 mg mL^{-1} lysozyme. Cell lysis was performed incubating the plate for 30 minutes at 37 $^{\circ}\text{C}$, 200 rpm. The crude extracts were centrifuged at 5,000 $\times g$ for 30 min and then 50 μL of the supernatant was transferred to a 96-well flat-bottom plate. The activity was assayed on the crude extract by adding 150 μL of 1.33X substrate solution (1.33 mM UDCA, 2.66 mM NAD^{+} , 13.3% MeOH in 50 mM KPi buffer, pH 8.0).

The initial activity was determined by measuring the increase of the absorbance at 340 nm for 5 min at 25 $^{\circ}\text{C}$ in a microtiter plate reader (Synergy2, Biotek, Winoosky, US) and compared with cultures expressing the wild-type *Cs7 β -HSDH* and untransformed cells as controls. The selected variants were sequenced and biochemically characterized.

6.8 Preparation of 7-oxo-DCA, 7-oxo-LCA, UCA and UDCA

In a 2 L round-bottom flask 2.1 g of CA, 133 mg of NAD^+ and 1.98 g of oxalacetate were dissolved in 50 mM of KPi buffer, pH 8.0 with 10% MeOH (final volume 1 L). The reaction was initiated adding 15 mg of *Sm7 α* -HSDH (5100 U_{tot}) and 200 U_{tot} of malate dehydrogenase (Sigma-Aldrich, St. Louis, US).

The reaction was gently stirred at 25 °C. Conversion was checked with HPLC. At complete reaction (~1 h) the solution was acidified with HCl till pH 2.0, leading to the formation of white suspension. In order to increase the precipitation 50 g of NaCl was added and dissolved. The product was filtered with a porous glass filter and washed with 20 mL of 0.01 M HCl solution. The powder was then dried and crystallized in MeOH giving 7-oxo-DCA. Yield: 1.9 g (95%); Purity: 98% (assayed by HPLC analysis); ^{13}C NMR (100 MHz, DMSO): δ 211.75 ($\text{C}_7=\text{O}$), 175.15 (COOH), 70.26 ($\text{C}_{12}\alpha\text{-OH}$), 69.33 ($\text{C}_3\alpha\text{-OH}$); full ^1H and ^{13}C NMR spectra are reported in supplementary material.

The same procedure was used for the preparation of 7-oxo-LCA from 1.9 g of CDCA. Yield: 1.6 g (84%); Purity: 98% (assayed by HPLC analysis); ^{13}C NMR (100 MHz, DMSO): δ 211.40 ($\text{C}_7=\text{O}$), 174.86 (COOH), 69.09 ($\text{C}_3\alpha\text{-OH}$); full ^1H and ^{13}C NMR spectra are reported in supplementary material.

UCA and UDCA were obtained by using the enzymatic cascade reported in this paper starting from CA and CDCA, respectively: 0.55 g of CA or 0.51 g of CDCA and 86 mg of NAD^+ were dissolved in 50 mM of KPi buffer, pH 8.0 with 10% MeOH (final volume 130 mL). The reaction was initiated by adding 0.18 mg of *Sm7 α* -HSDH (63 U_{tot}) and 12.54 mg of *Ls7 β* -HSDH (39 U_{tot}). The reaction was gently stirred at 25 °C. After 24 h the reaction was complete (92% and 91% conversion of CA into UCA and of CDCA into UDCA, respectively). The reactions was acidified with HCl (final pH 3.0) and extracted with diethyl ether. UDCA and UCA were purified by

Reveleris X2 apparatus (GRACE, Columbia, US) equipped with C18 cartridge (40 μ m, 12 g) employing Water (TFA 0.1%) : acetonitrile gradient. UCA: Yield: 478 mg (87%); Purity: >99% (assayed by HPLC analysis); ^{13}C NMR (100 MHz, DMSO): δ 174.98 (COOH), 70.66 (C $_7\beta$ -OH), 69.83 (C $_{12\alpha}$ -OH), 69.56 (C $_3\alpha$ -OH). UDCA: Yield: 438 mg (86%); Purity: >99% (assayed by HPLC analysis); ^{13}C NMR (100 MHz, DMSO): δ 174.90 (COOH), 69.74 (C $_7\beta$ -OH), 69.46 (C $_3\alpha$ -OH). Full ^1H and ^{13}C NMR spectra are reported in supplementary material.

^1H and ^{13}C NMR spectra, in agreement with previous results^[18, 29, 50] were acquired with Agilent ProPulse 400 MHz (Santa Clara, US) and are reported in supplementary materials.

Molecular mass of 7-oxo-DCA, 7-oxo-LCA, UCA and UDCA were confirmed using an RP-ESI-Q-TOF-MS system (M-Class and Q-TOF Premier, Waters, UK) which was operated in positive ionization mode (ES+). Mass calibration was performed using [Glu 1]-Fibrinopeptide B. Data were analyzed using the MassLynx 4.1 tool box. The MS analyses are reported in supplementary material.

6.9 Bioconversion of CDCA to UDCA.

All bioconversions were carried out employing 1 U_{tot} of purified *Sm7 α* -HSDS and 0.6 U_{tot} of purified 7 β -HSDS on 10 mM (if not differently specified) of CDCA and CA, NAD $^+$ (0.2, 0.5 or 1.0 mM). As general procedure, 1 mL of reaction mixture containing 10% MeOH and 50 mM of KPi buffer, pH 8.0, was incubated at 25 $^\circ\text{C}$: at fixed times 50 μL of reaction were withdrawn, diluted with 250 μL of MeOH and centrifuged at 14000 $\times g$ for 2 min. 10 μL of the obtained samples were analyzed by HPLC. HPLC analyses were performed on a Shimadzu (Kyoto, Japan) apparatus equipped with a LC20AT pump and an ELSD-LTII detector and fitted with a XTerra RP C18 column (Waters, Milford, US) (length/internal diameter 150/4.6 mm, pore size 5 μm) under the following conditions: for analyses of CA, 7-oxo-DCA and UCA, eluent

H₂O/CH₃CN/TFA (70/30/0.1), flow 1.0 mL min⁻¹. Retention times CA=6.75 min, 7-oxo-DCA=4.82 min, UCA= 4.03 min. For analyses of CDCA, 7-oxo-LCA and UDCA, eluent H₂O/CH₃CN/TFA (65/35/0.1), flow 1.0 mL min⁻¹. Retention times CDCA=4.11 min, 7-OXO-LCA=3.49 min, UDCA=3.21 min.

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 Dr. M. Pabst and P. van Dam for LC-MS measurement and Prof. U. Hanefeld, Dr. F. Hollmann, Dr. D. McMillan and all the members of the biocatalysis section of TU Delft for the scientific discussions.

8 Abbreviations:

API, active pharmaceutical ingredient; CA, Cholic acid; CDCA, Chenodeoxycholic acid; DCA, Deoxycholic acid; DHCA, Dehydrocholic acid; LCA, Lithocholic acid; UCA, Ursocholic acid; UDCA, Ursodeoxycholic acid; HSDH, Hydroxysteroid dehydrogenase; LDH, Lactate dehydrogenase; GDH, Glucose dehydrogenase; SSM, site-saturation mutagenesis.

9 References:

- [1] T. Ikegami, Y. Matsuzaki *Hepatol. Res.* **2008**, 38, 123-131.
- [2] T. Chen, S. Duncan, S. Solanki, K. F. Haq, M. K. M. Abdul, S. Soufleris, S. Shrestha, Z. Khan, F. Kamal, M. A. Khan *Gastroenterology.* **2018**, 154, S-782.
- [3] J. Reardon, T. Hussaini, M. Alsaifi, V. M. Azalgarra, S. R. Erb, N. Partovi, E. M. Yoshida *J. Clin. Transl. Hepatol.* **2016**, 4, 192-205.

- [4] A. Crosignani, P. M. Battezzati, K. D. Setchell, P. Invernizzi, G. Covini, M. Zuin, M. Podda *Dig. Dis. Sci.* **1996**, 41, 809-815.
- [5] G. Salen, A. Colalillo, D. Verga, E. Bagan, G. Tint, S. Shefer *Gastroenterology*. **1980**, 78, 1412-1418.
- [6] S. Arisawa, K. Ishida, N. Kameyama, J. Ueyama, A. Hattori, Y. Tatsumi, H. Hayashi, M. Yano, K. Hayashi, Y. Katano *Biochem. Pharmacol.* **2009**, 77, 858-866.
- [7] P. W. Pemberton, A. Aboutwerat, A. Smith, T. W. Warnes *Redox Rep.* **2006**, 11, 117-123.
- [8] J. D. Amaral, R. J. Viana, R. M. Ramalho, C. J. Steer, C. M. Rodrigues *J. Lipid Res.* **2009**, 50, 1721-1734.
- [9] J. D. Amaral, R. E. Castro, S. Solá, C. J. Steer, C. M. Rodrigues *J. Biol. Chem.* **2007**, 282, 34250-34259.
- [10] A. Stiehl, R. Raedsch, P. Czygan, R. Götz, C. Männer, S. Walker, B. Kommerell *Gastroenterology*. **1980**, 79, 1192-1198.
- [11] A. Stiehl, P. Czygan, B. Kommerell, H. Weis, K. Holtermüller *Gastroenterology*. **1978**, 75, 1016-1020.
- [12] S. Akare, S. Jean - Louis, W. Chen, D. J. Wood, A. A. Powell, J. D. Martinez *Int. J. Cancer*. **2006**, 119, 2958-2969.
- [13] E. Roda, F. Bazzoli, A. M. M. Labate, G. Mazzella, A. Roda, C. Sama, D. Festi, R. Aldini, F. Taroni, L. Barbara *Hepatology*. **1982**, 2, 804-810.
- [14] K. Nilsell, B. Angelin, B. Leijd, K. Einarsson *Gastroenterology*. **1983**, 85, 1248-1256.
- [15] Y. Feng, K. Siu, N. Wang, K.-M. Ng, S.-W. Tsao, T. Nagamatsu, Y. Tong *J. Ethnobiol. Ethnomed.* **2009**, 5, 2.
- [16] A. F. Hofmann *Acta Chem. Scand.* **1963**, 17, 173-186.
- [17] T. Kanazawa, A. Shimazaki, T. Sato, T. Hoshino *Nippon Kagaku Zasshi*. **1955**, 76, 297-301.
- [18] P. S. Dangate, C. L. Salunke, K. G. Akamanchi *Steroids*. **2011**, 76, 1397-1399.
- [19] T. Washizu, I. Tomoda, J. J. Kaneko *J. Vet. Med. Sci.* **1991**, 53, 81-86.
- [20] F. Tonin, I. W. C. E. Arends *Beilstein J. Org. Chem.* **2018**, 14, 470-483.
- [21] T. Eggert, D. Bakonyi, W. Hummel *J. Biotechnol.* **2014**, 191, 11-21.
- [22] M. M. Zheng, F. F. Chen, H. Li, C. X. Li, J. H. Xu *ChemBioChem*. **2018**, 19, 347-353.
- [23] M.-M. Zheng, R.-F. Wang, C.-X. Li, J.-H. Xu *Process Biochem.* **2015**, 50, 598-604.

- [24] Q. Ji, J. Tan, L. Zhu, D. Lou, B. Wang *Biochem. Eng. J.* **2016**, 105, 1-9.
- [25] R. Bovara, G. Carrea, S. Riva, F. Secundo *Biotechnol. Lett.* **1996**, 18, 305-308.
- [26] R. Bovara, E. Canzi, G. Carrea, A. Pilotti, S. Riva *J. Org. Chem.* **1993**, 58, 499-501.
- [27] D. Monti, E. E. Ferrandi, I. Zanellato, L. Hua, F. Polentini, G. Carrea, S. Riva *Adv. Synth. Catal.* **2009**, 351, 1303-1311.
- [28] P. Pedrini, E. Andreotti, A. Guerrini, M. Dean, G. Fantin, P. P. Giovannini *Steroids.* **2006**, 71, 189-198.
- [29] A. Medici, P. Pedrini, E. Bianchini, G. Fantin, A. Guerrini, B. Natalini, R. Pellicciari *Steroids.* **2002**, 67, 51-56.
- [30] J. Shi, J. Wang, L. Yu, L. Yang, S. Zhao, Z. Wang *J. Ind. Microbiol. Biotechnol.* **2017**, 44, 1073-1082.
- [31] Q. Yang, B. Wang, Z. Zhang, D. Lou, J. Tan, L. Zhu *RSC Advances.* **2017**, 7, 38028-38036.
- [32] J. T. Wu, L. H. Wu, J. A. Knight *Clin. Chem.* **1986**, 32, 314-319.
- [33] N. Guex, M. C. Peitsch *Electrophoresis.* **1997**, 18, 2714-2723.
- [34] R. Wang, J. Wu, D. K. Jin, Y. Chen, Z. Lv, Q. Chen, Q. Miao, X. Huo, F. Wang *Acta Crystallogr. Sect. F: Struct. Biol. Commun.* **2017**, 73, 246-252.
- [35] M. T. Reetz, J. D. Carballeira *Nat. Protoc.* **2007**, 2, 891-903.
- [36] S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z.-G. Zhang, D. J. Opperman, M. T. Reetz, J. P. Acevedo *ACS Synth. Biol.* **2012**, 2, 83-92.
- [37] S. Savino, E. E. Ferrandi, F. Forneris, S. Rovida, S. Riva, D. Monti, A. Mattevi *Proteins: Struct. Funct. Bioinform.* **2016**, 84, 859-865.
- [38] C. Held, G. Sadowski *Annu. Rev. Chem. Biomol. Eng.* **2016**, 7, 395-414.
- [39] I. Macdonald, Y. Rochon, D. Hutchison, L. Holdeman *Appl. Environ. Microbiol.* **1982**, 44, 1187-1195.
- [40] R. Machielsen, L. L. Looger, J. Raedts, S. Dijkhuizen, W. Hummel, H. G. Hennemann, T. Daussmann, J. Van der Oost *Eng. Life Sci.* **2009**, 9, 38-44.
- [41] N. Richter, A. Zienert, W. Hummel *Eng. Life Sci.* **2011**, 11, 26-36.
- [42] X. Xu, J. Chen, Q. Wang, C. Duan, Y. Li, R. Wang, S. Yang *ChemBioChem.* **2016**, 17, 56-64.
- [43] A. M. Chánique, L. P. Parra *Front. Microbiol.* **2018**, 9, 194.

- [44] J. K. Cahn, C. A. Werlang, A. Baumschlager, S. Brinkmann-Chen, S. L. Mayo, F. H. Arnold *ACS Synth. Biol.* **2016**, 6, 326-333.
- [45] J. K. Cahn, S. Brinkmann-Chen, F. H. Arnold in *Synthetic Metabolic Pathways: Methods and Protocols* (Eds.: M. K. Jensen, J. D. Keasling), Springer New York, New York, NY, **2018**, pp.15-26.
- [46] N. Borlinghaus, B. M. Nestl *ChemCatChem.* **2018**, 10, 183-187.
- [47] D. Lou, Y. Wang, J. Tan, L. Zhu, S. Ji, B. Wang *Comput. Biol. Chem.* **2017**, 70, 89-95.
- [48] D. Bakonyi, W. Hummel *Enzyme. Microb. Technol.* **2017**, 99, 16-24.
- [49] A. Pabis, V. A. Risso, J. M. Sanchez-Ruiz, S. C. Kamerlin *Curr. Opin. Struct. Biol.* **2018**, 48, 83-92.
- [50] P. P. Giovannini, A. Grandini, D. Perrone, P. Pedrini, G. Fantin, M. Fogagnolo *Steroids.* **2008**, 73, 1385-1390.
- [51] V. Hessel *Chem. Int.* **2018**, 40, 12-16.
- [52] G. Carrea, A. Pilotti, S. Riva, E. Canzi, A. Ferrari *Biotechnol. Lett.* **1992**, 14, 1131-1134.
- [53] F. Tonin, R. Melis, A. Cordes, A. Sanchez-Amat, L. Pollegioni, E. Rosini *New Biotechnol.* **2016**, 33, 387-398.
- [54] R. M. Siloto, R. J. Weselake *Biocatal. Agric. Biotechnol.* **2012**, 1, 181-189.

Table 1

a) Purification of recombinant His-tagged *Sm7 α* -HSDH from *E. coli* BL21(DE3) cells (7 g corresponding to 1.0 L of fermentation broth).

	Protein [mg]	Specific activity [U mg ⁻¹]	Purification [fold]	Yield [%]
Cell extract	1050	233.2 ^[a]	-	100.0
HiTrap chelating	232.3	430.4 ^[a]	1.8	40.8

b) Purification of recombinant His-tagged *Cs7 β* -HSDH from *E. coli* BL21(DE3) cells (3.5 g corresponding to 0.5 L of fermentation broth).

	Protein [mg]	Specific activity [U mg ⁻¹]	Purification [fold]	Yield [%]
Cell extract	177	0.05 ^[b]	-	100.0
HiTrap chelating	10	0.74 ^[b]	14.8	83.3

c) Purification of recombinant His-tagged *Ls7 β* -HSDH from *E. coli* BL21(DE3) cells (2.2 g corresponding to 0.3 L of fermentation broth).

	Protein [mg]	Specific activity [U mg ⁻¹]	Purification [fold]	Yield [%]
Cell extract	122	1.28 ^[c]	-	100.0
HiTrap chelating	29	3.10 ^[c]	2.4	57.9

[a] Activity was assayed on 1.25 mM CDCA and 2.5 mM NAD⁺ as substrate in 50 mM KPi buffer, pH 8.0. [b] Activity was assayed on 1.0 mM UDCA and 2.0 mM NADP⁺ as substrate in 50 mM KPi buffer, pH 8.0. [c] Activity was assayed on 1.0 mM UDCA and 2.0 mM NAD⁺ as substrate in 50 mM KPi buffer, pH 8.0.

Table 2

Kinetic parameters of purified recombinant 7 β -HSDHs on NADP⁺ and NAD⁺. The kinetic parameters were determined in the presence of 1.0 mM UDCA. The preference is calculated as ratio between the catalytic efficiency on the two cofactors.

	NADP ⁺					NAD ⁺				
	$k_{\text{cat}}^{[a]}$ [s ⁻¹]	K_m [mM]	K_i [mM]	k_{cat}/K_m	Preference	$k_{\text{cat}}^{[a]}$ [s ⁻¹]	K_m [mM]	k_{cat}/K_m	Preference	
wt <i>Cs7</i> β -HSDH	2.40 ± 0.16	0.033 ± 0.007	10 ± 3	72.600	8354	0.023 ± 0.001	2.62 ± 0.21	0.009	0.0001	
ADLN variant	0.010 ± 0.002	2.600 ± 0.300	/	0.004	0.01	0.75 ± 0.03	2.83 ± 0.27	0.266	70	
ADDLN variant		below detection				0.08 ± 0.01	0.33 ± 0.02	0.234	/	
ADDAA variant		below detection				0.32 ± 0.01	4.64 ± 0.27	0.068	/	
<i>Ls7</i> β -HSDH	0.025 ± 0.004	10 ± 1	/	0.002	0.0001	3.05 ± 0.08	0.08 ± 0.01	40.637	16587	

[a] k_{cat} values were calculated considering a MW of 59.4 and 58.8 kDa for *Cs7* β -HSDH variants and *Ls7* β -HSDH, respectively.

Table 3

Kinetic parameters of purified recombinant *Sm7* α -HSDH

	$k_{\text{cat}}^{[c]}$ [s ⁻¹]	K_m [mM]	K_i [mM]	k_{cat}/K_m
CA ^[a]	497.1 ± 17.2	0.960 ± 0.110	/	517.8 ± 77.2
CDCA ^[a]	873.1 ± 34.1	0.218 ± 0.024	11 ± 1.6	4005.2 ± 597.5
NAD ^{+[b]}	944.1 ± 9.1	0.560 ± 0.017	/	1685.8 ± 67.4

[a] The kinetic parameters were determined in the presence of 2.5 mM NAD⁺. [b] The kinetic parameters were determined in the presence of 2.0 mM CDCA. [c] k_{cat} values were calculated considering a MW of 107.2 kDa.

Table 4Kinetic parameters of purified recombinant 7 β -HSDHs on different substrates.

	UDCA				7-oxo-LCA				7-oxo-DCA			
	$k_{cat}^{[a]}$ [s ⁻¹]	K_m [mM]	K_i [mM]	k_{cat}/K_m	$k_{cat}^{[b]}$ [s ⁻¹]	K_m [mM]	K_i [mM]	k_{cat}/K_m	$k_{cat}^{[b]}$ [s ⁻¹]	K_m [mM]	K_i [mM]	k_{cat}/K_m
wt <i>Cs7</i> β -HSDH	3.61 ± 0.34	0.160 ± 0.030	1.4 ± 0.3	22.6	1.37 ± 0.07	0.128 ± 0.027	/	10.7	2.27 ± 0.03	0.106 ± 0.005	/	21.3
ADLN variant	1.35 ± 0.35	0.420 ± 0.150	0.6 ± 0.2	3.2	0.24 ± 0.05	0.175 ± 0.080	2.0 ± 1.0	1.4	0.59 ± 0.09	0.630 ± 0.189	8.6 ± 3.6	0.9
ADDLN variant	0.07 ± 0.01	0.075 ± 0.030	2.4 ± 1.0	1.0	0.05 ± 0.01	0.055 ± 0.034	2.4 ± 1.3	0.9	0.08 ± 0.01	0.203 ± 0.034	4.2 ± 0.9	0.4
ADDAA variant	0.30 ± 0.04	0.207 ± 0.040	0.9 ± 0.2	1.4	0.15 ± 0.03	0.246 ± 0.041	/	0.6	0.18 ± 0.01	0.323 ± 0.070	/	0.6
<i>Ls7</i> β -HSDH	11.91 ± 2.28	0.156 ± 0.051	0.8 ± 0.3	76.1	7.55 ± 0.62	0.038 ± 0.014	/	200.8	8.02 ± 0.41	0.131 ± 0.028	/	61.0

[a] The kinetic parameters were determined in the presence of 2.0 mM NAD⁺. [b] The kinetic parameters were determined in the presence of 0.5 mM of NADH. The wt *Cs7* β -HSDH was characterized adding NADP instead of the NAD. In all the cases, k_{cat} values were calculated considering a MW of 59.4 and 58.8 kDa for *Cs7* β -HSDH variants and *Ls7* β -HSDH, respectively.

Accepted Manuscript

Table 5: Bioconversion results employing different initial concentration of substrate, cofactor and MeOH. Incubation (1 mL volume) were carried out in employing 1 U of *Sm7 α* -HSDH (0.23 μ g) and 0.6 U of *Ls7 β* -HSDH (190 μ g) in 50 mM KPi, pH 8.0 at 25 °C for 6 hours.

Bioconversion conditions			Conversion [%]			
[CDCA] [mM]	[NAD ⁺] [mM]	[MeOH] [%]	[UDCA]		[7-oxo-LCA]	
10	1	0	80.4	\pm 1.9	3.5	\pm 0.7
10	1	10	91.3	\pm 0.5	4.7	\pm 0.6
10	1	20	90.2	\pm 0.2	2.5	\pm 1.0
10	0.5	10	92.0	\pm 0.1	1.3	\pm 0.6
10	0.2	10	14.6 (21.1 ^[a])	\pm 2.9	3.0	\pm 1.3
20	1	10	14.0	\pm 0.0	1.1	\pm 0.1
20	0.5	10	19.1	\pm 2.5	1.6	\pm 0.8
50	1	20	0.1	\pm 0.0	0.4	\pm 0.2

[a]12 hours of incubation.

Figure 1: Epimerization of the 7 α -OH group of CA and CDCA. In our work, the NADH produced in the first reaction step (oxidation of 7 α -OH) is reused by a NADH dependent 7 β -HSDH giving the 7 β -OH epimer. E1 and E2 are 2 additional enzymes (i.e. alcohol dehydrogenases) used to regenerate the cofactors converting the sacrificial substrates 1 and 2, respectively, in waste products.

Figure 2: (a) Hypothetical binding mode of the cofactor inside the *Cs7 β -HSDH* active site: (on the left) the two arginines R40 and R41 are the main amino acids responsible for recognition and binding of NADPH. The hypothetical NADH binding mode of the ADLN variant (on the right) shows the putative formation of a hydrogen bond between the aspartate side-chain (D39) and the 2'-OH group of ribose (distance 1.8 - 2.1 Å). (b) Interference of T17 with the D39 side chain in the binding of the 2'-OH group (distance 1.5 - 1.7 Å). (c) Screening results of the second round of SSM on *Cs7 β -HSDH* libraries. Here, four different templates, isolated during the first round were all subjected to SSM at position 17. The activity of the isolated ADLN variant is indicated with an asterisk. (d) Hypothetical binding mode of the ADDLN variant to the NADH: the aspartate in position 18 can form a supplementary hydrogen bond with the 3'-OH group of the ribose.

Figure 3: 3D model of *Ls7 β -HSDH* built employing the entire structure of 7 β -HSDH from *Collinsella aerofaciens* (PDB code 5GT9) as template: the homodimeric quaternary structure (a) is conserved. The hypothetical binding mode of NADH (b) is primary due to the side chains of D19 and D40 to the 3' and 2'-OH groups of ribose, respectively.

Figure 4: Effect of pH on the oxidative (grey bars) and reductive (black bars) activities of purified (a) *Sm7 α -HSDH*, (b) wt *Cs7 β -HSDH*, (c) ADLN, (d) ADDLN and (e) ADDAA variants of *Cs7 β -HSDH*

and (f) wt *Ls7β*-HSDH. The activity at pH 8.0 was taken as 100%. Enzymatic activities were determined by measuring NAD(P)⁺ reduction and NAD(P)H oxidation in the same conditions described in materials and methods. Values represent the mean of three independent experiments (mean ± standard error).

Figure 5: Effect of MeOH concentration on the activities of purified (a) *Sm7α*-HSDH, (b) wt *Cs7β*-HSDH, (c) ADLN, (d) ADDLN and (e) ADDAA variants of *Cs7β*-HSDH and (f) wt *Ls7β*-HSDH.

In panel a, the activity in the absence of MeOH was taken as 100%. Due to the lower solubility of UDCA in water, in all the other panels the activity in presence of 10% of MeOH was taken as 100%. Values represent the mean of three independent experiments (mean ± standard error).

Figure 6: Bioconversion time-courses of: (a) 10 mM CDCA with 1.0 mM NAD⁺, (b) 10 mM CA with 1.0 mM NAD⁺, (c) 10 mM CDCA with 0.2 mM NAD⁺, employing 1 U of *Sm7α*-HSDH (0.23 μg) and 0.6 U of *Ls7β*-HSDH (190 μg) in 50 mM KPi, pH 8.0 and 10% MeOH at 25 °C. (d) Bioconversion time-course of 10 mM CA with 1.0 mM NAD⁺, employing 1 U of *Sm7α*-HSDH (0.23 μg) and 0.6 U of ADLN variant of *Cs7β*-HSDH (790 μg) in 50 mM KPi, pH 8.0 and 10% MeOH at 25 °C. In all cases, the substrates, products and the corresponding 7-oxo derivatives are represented by green triangles, blue diamonds and red squares, respectively.

Figure 1

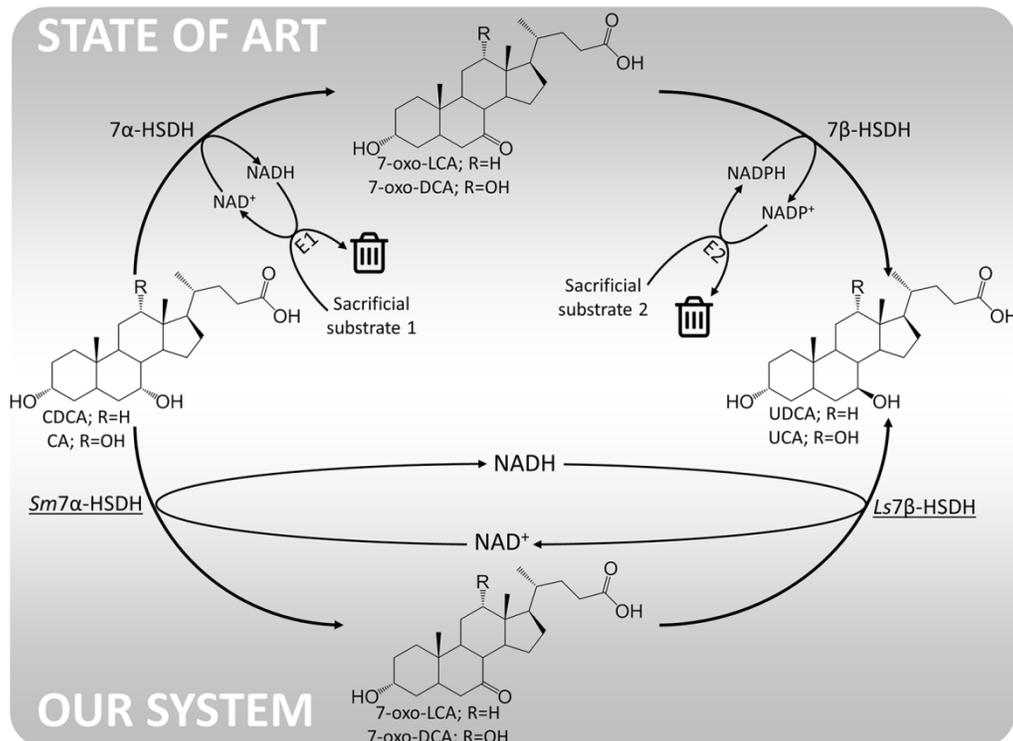


Figure 2

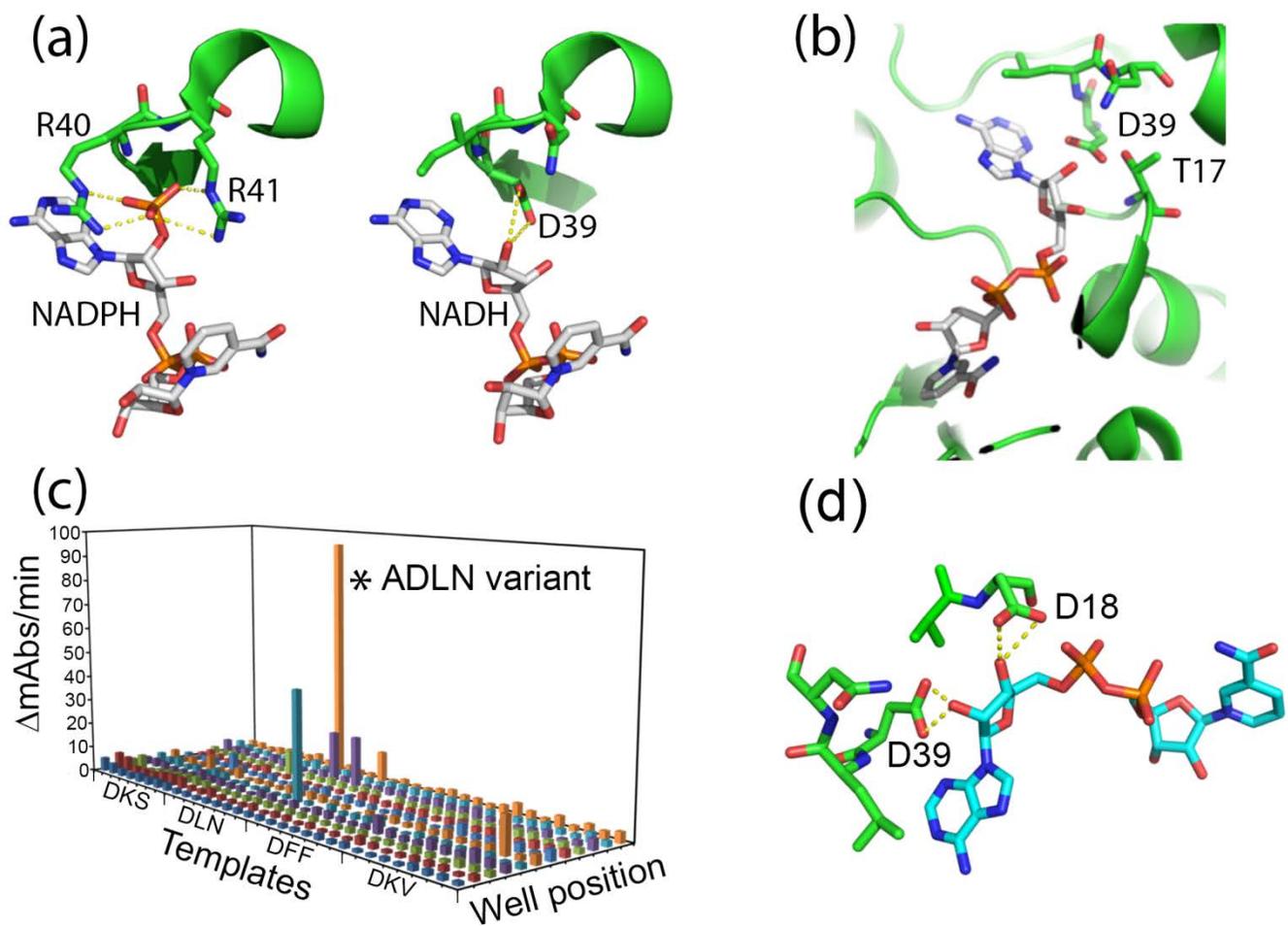


Figure 3

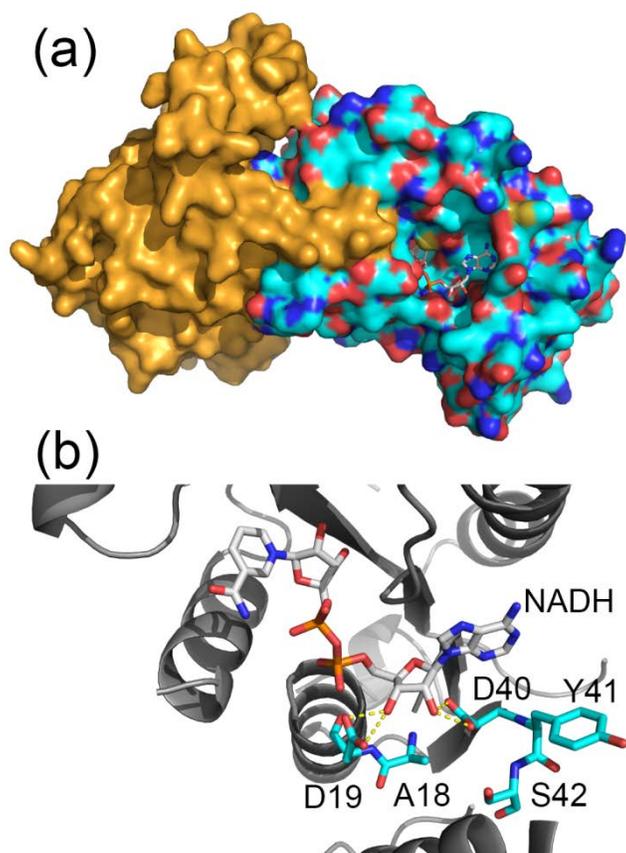


Figure 4

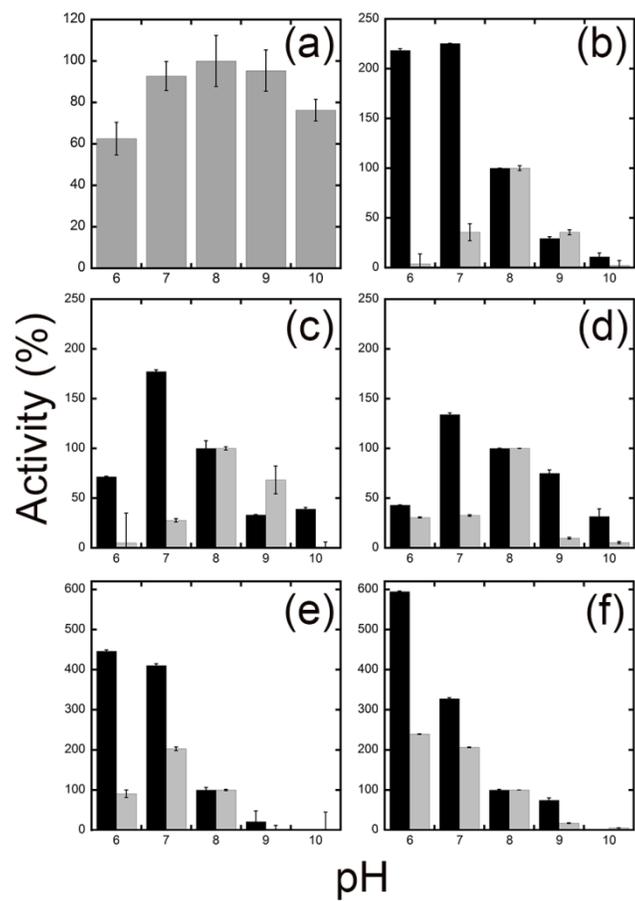


Figure 5

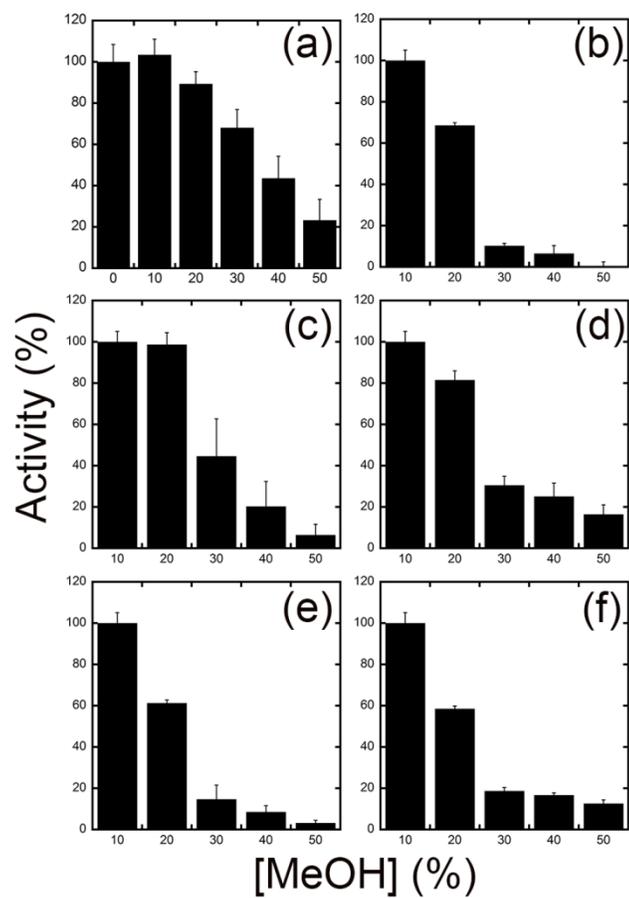
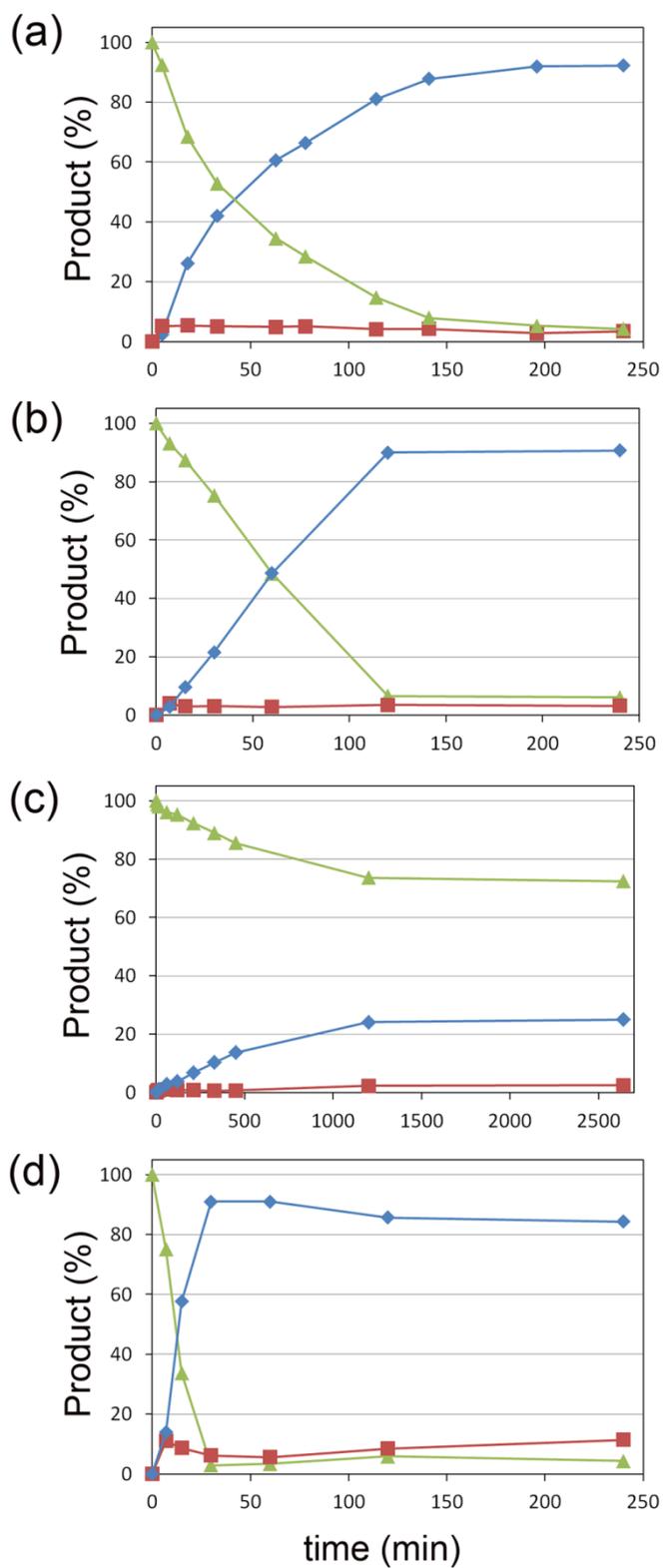
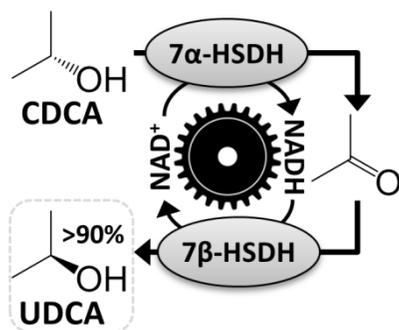


Figure 6:



Graphical abstract:

By employing a compatible set of newly designed and stable enzymes, epimerization reaction of hydroxysteroids was carried out in a redox-neutral environment by using a NAD⁺ dependent cascade. The racemization of CDCA is substantially favoured towards the products, as governed by thermodynamic and solubility properties of the compounds involved, circumventing the requirement of separate nicotinamide cofactors regeneration systems.