Advanced Synthesis & Catalysis

Accepted Article

Title: Insights into the Substrate Promiscuity of Novel Hydroxysteroid Dehydrogenases

Authors: Susanna Bertuletti, Erica Elisa Ferrandi, Stefano Marzorati, Marta Vanoni, Sergio Riva, and Daniela Monti

To be cited as: *Adv. Synth. Catal.* 10.1002/adsc.202000120

Link to VoR: https://doi.org/10.1002/adsc.202000120

FULL PAPER

DOI: 10.1002/adsc.202000120

Insights into the Substrate Promiscuity of Novel Hydroxysteroid Dehydrogenases

Susanna Bertuletti,[+a,b] Erica Elisa Ferrandi,[+a] Stefano Marzorati,[a] Marta Vanoni,[a] Sergio Riva,[a,*] and Daniela Monti [a,*]

[a] Istituto di Scienze e Tecnologie Chimiche "G. Natta" (SCITEC), CNR, Via Mario Bianco 9, 20131 Milano, Italy
E-mail: sergio.riva@scitec.cnr.it; daniela.monti@scitec.cnr.it
[b] Università degli Studi di Milano, Via Giuseppe Colombo 60, 20133 Milano, Italy
[+] Equal author contribution

Received:((will be filled in by the editorial staff))

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/adsc.201######.

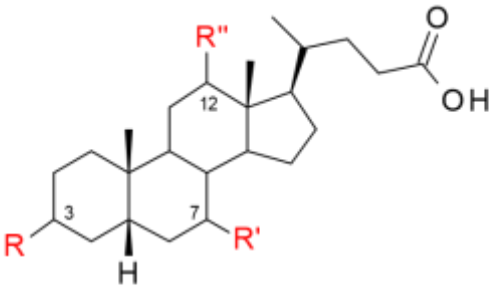
Abstract. Hydroxysteroid dehydrogenase (HSDHs) are valuable biocatalysts for the regio- and stereoselective modification of steroids, bile acids and other steroid derivatives. In this work, we investigated the substrate promiscuity of this highly selective class of enzymes. In order to reach this goal, a preliminary search of HSDH homologues in in-house or public available (meta)genomes was carried out. Eight novel NAD(H)-dependent HSDHs, showing either 7α-, 7β-, or 12α-HSDH activity, and including, for the first time, enzymes from extremophilic microorganisms, were identified, recombinantly produced, and characterized. Among the novel HSDHs, four highly active (up to 92 U mg^{-1}) NAD(H)-dependent 7β-HSDHs showing negligible similarity towards previously described 7β-HSDHs, were discovered. These enzymes, along with previously characterized HSDHs, were tested as biocatalysts for the stereoselective reduction of a panel of substrates including two α-ketoesters of pharmaceutical interest and selected ketones that partially resemble the structural features of steroids. All the reactions were coupled with a suitable cofactor regeneration system. Regarding the α-ketoesters, nearly all of the tested HSDHs showed a good activity toward the selected substrates, yielding the reduced α-hydroxyester with up to 99% conversions and enantiomeric excesses. On the other hand, only the 7β-HSDHs from *Collinsella aerofaciens* and *Clostridium absonum* showed appreciable activity toward more complex ketones, i.e., (±)-trans-1-decalone, but with interesting as well as different selectivity.

Keywords: Hydroxysteroid dehydrogenase; Bile acids; Substrate promiscuity; Enzyme discovery; Stereoselectivity

Introduction

Hydroxysteroid dehydrogenases (HSDHs) are a group of NAD(P)(H)-dependent oxidoreductases that are characterized by their ability to catalyze the oxidation/reduction of the hydroxyl/keto groups of bile acids and steroids.[1–3] Most of the HSDHs described so far show an exceptionally high selectivity that enables their application in the straightforward modification of their natural complex substrates, which is instead quite challenging by conventional chemical synthesis.

Specifically, HSDHs have been shown to oxidize the hydroxyl groups at different positions, e.g., at C-3, C-7, and C-12 of bile acids in a very high regiospecific manner (Scheme 1). Moreover, for each one of these positions, HSDHs usually show a practically absolute stereoselectivity by oxidizing only either the hydroxyl group above (β configuration) or below (α configuration) the plane of the steroid molecule.[4,5] The reactions catalyzed by HSDHs are reversible, thus these enzymes can be applied in the regioselective reduction of the corresponding keto derivatives as well.[4,6,7]



Cholic acid	R = R' = R" = α-OH
Chenodeoxycholic acid	R = R' = α-OH; R" = H
Deoxycholic acid	R = R" = α-OH; R' = H
Ursocholic acid	R = R" = α-OH; R' = β-OH
Ursodeoxycholic acid	R = α-OH; R' = β-OH; R" = H

Scheme 1. Chemical structure of the most common bile acids, natural substrates of hydroxysteroid dehydrogenases.

Thanks to these interesting features, HSDHs have been widely studied during the last years for their

exploitation in the biocatalyzed synthesis of key intermediates of the drug ursodeoxycholic acid (UDCA), a largely applied therapeutic agent used for the dissolution of cholesterol gallstones and for the treatment of different hepatic diseases. These efforts have led to the characterization of different enzymes showing either 7α-, 7β- or 12α-HSDH activity[2,3], as well as to the development of biocatalyzed processes for the preparation of UDCA intermediates.[8–11]

The ability of enzymes to act on substrates possessing chemical structures significantly different from their natural ones is a quite common feature.[12–14] This "substrate promiscuity" was also observed with HSDHs a long time ago by Davies and coworkers who reported the activity of a 3α,20β–HSDH from *Mortierella ramanniana* toward 3-hydroxybicyclo-heptan-6-ones.[15]

However, to date only few additional examples of HSDH promiscuous activities towards non-steroidal substrates have been described. For example, *Comamonas testosteroni* 3α-HSDH[16] and 11β–HSDH[17] were found able to reduce *p*-nitrobenzaldehyde, *p*-nitroacetophenone and metyrapone, while 17β-HSDH from the filamentous fungus *Cochliobolus lunatus*[18] was active toward quinones, menadione, *p*-nitrobenzaldehyde. Moreover, 3β,17β-HSDH from *Comamonas testosteroni* and 3α,20β-HSDH from *Streptomyces hydrogenans* showed activity toward muscarine precursors.[19] More recently, the catalytic promiscuous property of the 7α-HSDH from *Bacteroides fragilis* has been investigated by studying the enzyme substrate specificity and stereoselectivity in the reduction of different aromatic and aliphatic α-ketoesters to the corresponding α-hydroxyesters.[20] To the best of our knowledge, only these six papers have been published so far, describing, as just summarized, the results obtained with unrelated HSDHs (3α,20β–HSDH, 3α-HSDH, 11β–HSDH, 17β-HSDH, 3β,17β-HSDH, 3α,20β-HSDH and 7α-HSDH) on very different substrates (3-hydroxybicyclo-heptan-6-ones, metyrapone, *p*-nitroacetophenone, *p*-nitrobenzaldehyde, quinones, menadione, muscarine precursors, aromatic and aliphatic α-ketoesters).

In this work, we aimed at filling this gap by performing a more systematic investigation on HSDHs substrate promiscuity. This result was achieved by i) increasing the diversity of the available HSDHs by discovering novel enzyme homologues, ii) establishing their regio- and stereoselectivity in the respect of their natural substrates, i.e., bile acids, and iii) evaluating their possible substrate promiscuity toward a set of different non-steroidal substrates.

Results and Discussion

Discovery of novel HSDHs

To the best of our knowledge, all the HSDHs characterized so far belong to the "short-chain dehydrogenases/reductases" (SDR) superfamily, a very large and highly divergent protein family with generally low sequence identity (20-30%), thus making quite challenging their functional annotation if a biochemical characterization is not performed.[21] On the other side, most of the known HSDHs are produced by gut or soil bacteria belonging to the phyla of Proteobacteria, e.g., *Escherichia*[22] and *Stenotrophomonas*[5] strains, Actinobacteria, e.g., *Collinsella*[23] and *Eggerthella*[8,24,25] strains, and Firmicutes, e.g., *Clostridium*,[26–30] *Ruminococcus*,[31,32] and *Lactobacillus*[5] strains, thus resulting in an overall limited diversity.

To improve the diversity of the HSDHs library already available in our lab, a bioinformatic screening was performed in search of HSDH homologues in in-house or public available (meta)genomes, with a special focus for those from extreme environments. Indeed, no HSDHs have been characterized so far from extremophilic microorganisms, e.g., thermophiles and halophiles. However, enzymes from these sources may have practical interest exceeding the scope of the present work, since, thanks to their higher robustness, they could be more suitable than mesophilic enzymes to industrial applications.[33–36]

The search of novel HSDH sequences was carried out by using either the Blastp tool (https://blast.ncbi.nlm.nih.gov/) for mining into the NCBI database (https://www.ncbi.nlm.nih.gov/) or the program LAST (http://last.cbrc.jp/) for the analysis of metagenomes of samples collected in hot terrestrial environments.[33,37,38] In both cases, only sequences of functionally characterized HSDHs (Table S1 in the Supporting Information) were used as queries during the multiple sequence alignments.

At the end of these analyses, ten SDR sequences were selected for cloning and characterization (Table 1).

Two sequences (Dm7α-HSDH and Hh7α-HSDH, entries 1-2 in Table 1) showing about 75% similarity with known 7α-HSDHs (see also Fig. S1 in the Supporting Information) were discovered for the first time in genomes from extremophilic bacteria. Dm7α-HSDH was annotated as a putative glucose 1-dehydrogenase in the genome of *Deinococcus marmoris* strain PAMC 26562, a radiation-resistant and psychro- and draught-tolerant bacterium isolated from an Antarctic rock sample.[39,40] Instead, Hh7α-HSDH was already annotated as a 7α-HSDH in the genome of the halophilic bacterium *Halomonas halodenitrificans*, but not yet characterized.

Since genes coding for HSDHs are often clustered in bacterial genomes, e.g., the gene coding for *Clostridium absonum* 7β-HSDH was found upstream that coding for a co-expressed 7α-HSDH,[29] the genome regions contiguous to Dm7α-HSDH and Hh7α-HSDH genes were checked. No putative dehydrogenase was found in proximity of the gene coding for Dm7α-HSDH.

Instead, a gene annotated as a SDR family oxidoreductase (Genbank: WP_027961749.1) was found upstream the Hh7α-HSDH gene. Interestingly,

this sequence showed only a modest similarity to 7α-HSDHs (Table 1, entry 3, 28-38% identity at the deduced amino acid level), and no significative similarity to other HSDSs, including known 7β-HSDHs.

Even more curiously, when searching for homologues of this new sequence into the NCBI database, we found several uncharacterized SDRs from diverse sources with up to 79% identity. Three SDRs, i.e., one from plasmid material of the photosynthetic cyanobacterium *Stanieria cyanosphaera* PCC 7437 (Genbank: WP_015212061.1), one from a *Brucella* strain (Genbank: WP_004684107.1), and one from a *Rhodobacter sphaeroides* strain (Genbank: WP_011911126.1) were then selected for further studies (Table 1, entries 4-6, see also Fig. S2 in the Supporting Information).

As far as the metagenomes mining concerns, three sequences were found, two of them from the metagenomes of samples collected at drilling wells in the Norwegian continental shelf (Ngi1_7α-HSDH and Ngi7-SDR, entries 7 and 9, Table 1) and one from the metagenome obtained from Icelandic hot spring sediments (Is2-SDR, entry 8, Table 1).

Ngi1_7α-HSDH (entry 7) showed about 70% similarity with known 7α-HSDHs. By performing a Blast analysis in the NCBI database, a putative 7α-HSDH with 99.7% identity in the respect of Ngi1_7α-HSDH was found in the genome of a *Psychrobacter* sp. strain (Genbank: HAR74729.1).

It was more difficult to make hypotheses about the possible catalytic activity of the other two selected metagenomic sequences. In fact, Is2-SDR showed a quite low similarity with respect to both known HSDHs and sequences in the NCBI database [38% and 53% identity with the 7α-HSDH from *Bosea* sp. and with a putative D-threitol dehydrogenase from an *Aerophobetes* strain (Genbank: TKJ47585.1), respectively]. In the case of Ngi7-SDR, it showed a 35% identity with the 12α-HSDH from *Clostridium* sp., and was identical at the amino acidic level to an uncharacterized SDR family oxidoreductase from a *Pseudomonas pelagia* strain (Genbank: QFY56536).

Finally, a sequence from *Lysinibacillus* (formerly *Bacillus*) *sphaericus* showing about 47% identity with known 12α-HSDH was discovered and included in further studies (Ls12α-HSDH, entry 10, Table 1). In fact, from the same source, a commercially available NAD(H)-dependent 12α-HSDH (in the past from Genzyme Biochemicals Ltd., today from Creative Enzymes®) was widely used in previous studies,[11,41] but never cloned and produced in recombinant form.

Recombinant production and characterization of novel HSDHs

The codon-optimized synthetic genes coding for the selected putative HSDHs were cloned into the pETite expression vector (Lucigen) in frame with a C-term His-Tag sequence. Protein over-expression in the corresponding recombinant *Escherichia coli* BL21(DE3) strains was achieved by induction with isopropyl-β-D-thiogalactopyranoside (IPTG) (see Experimental section for details).

As shown in Table 1, all the target enzymes could be produced in recombinant form in *E. coli* and homogeneous protein samples were obtained from the cell extracts with up to >300 mg L^{-1} yields by nickel-nitriloacetic acid (Ni-NTA) chromatography, as confirmed by SDS-PAGE analysis (Fig. S3, Supporting Information).

The assessment of the catalytic activity of these novel SDRs toward different bile acids (Scheme 1) was first carried out by spectrophometric assays in oxido-reduction reactions in the presence of NAD(P)(H) cofactors. Subsequently, the respective HSDH activity was confirmed by setting up small-scale oxido-reduction reactions with suitable bile acids (see Experimental section, Fig. S4 and Scheme S1 in the Supporting Information for details).

As expected from the previously reported sequence analysis, Dm7α-HSDH (Table 1, entry 1), Hh7α-HSDH (entry 2), and Ngi1_7α-HSDH (entry 7) showed NAD(H)-dependent 7α-HSDH activity, thus demonstrating for the first time the occurrence of HSDH activities in extremophilic microorganisms.

Interestingly, Dm7α-HSDH (entry 1) showed a thermophilic behavior by displaying an optimum temperature at 60°C in the regioselective oxidation of cholic acid (see Fig. S5 in the Supporting Information), which is, to our knowledge, the highest value reported so far for HSDHs. It is also noteworthy the broad temperature range at which this enzyme showed activity, since about 15% of the maximal activity recorded at 60°C was retained at both 20°C and 80°C. The finding of a thermostable enzyme in a cold adapted microorganism looks contradictory at a first glance. However, despite their psychrophilic origin, the occurrence of thermophilic and thermostable enzymes in Antarctic microbes has been previously reported.[42–45] To a lower extent, also Ngi1_7α-HSDH (entry 7), whose origin, as previously mentioned, is likely close to psychrophilic bacteria, showed a moderate thermophilic behavior, with an optimum temperature at 40°C (see Fig. S5 in the Supporting Information). Remarkably, Dm7α-HSDH and Ngi1_7α-HSDH are phylogenetically related (Fig. 1), and close also to the recently discovered 7α-HSDH from *Stenotrophomonas maltophilia* (Sm7α-HSDH), which showed a marked thermophilicity as well,[5] thus constituting a distinct clade from the other (mesophilic) 7α-HSDHs.

Table 1. Expression yields and activity of the SDRs selected in this work.

Entry	Enzyme	Source	Identity with known HSDHs[a] (%)	HSDH activity[b]	Cofactor	Yield (U)[b,c]	Yield (mg)[c]	Specific activity (U mg^{-1})
1	Dm7α-HSDH	*Deinococcus marmoris*	75.40 (Sm7α-HSDH)	7α-HSDH	NAD(H)	788	14	56
2	Hh7α-HSDH	*Halomonas halodenitrificans*	74.51 (Sh7α-HSDH)	7α-HSDH	NAD(H)	1986	64	31
3	Hh7β-HSDH	*Halomonas halodenitrificans*	38.89 (Bs7α-HSDH)	7β-HSDH	NAD(H)	955	95	10
4	Sc7β-HSDH	*Stanieria cyanosphaera*	38.40 (Gri7α-HSDH)	7β-HSDH	NAD(H)	261	7	38
5	Bsp7β-HSDH	*Brucella* sp.	36.00 (Pa7α-HSDH)	7β-HSDH	NAD(H)	8400	91	92
6	Rs7β-HSDH	*Rhodobacter sphaeroides*	39.68 (Bs7α-HSDH)	7β-HSDH	NAD(H)	2348	80	29
7	Ngi1_7α-HSDH	metagenome	69.05 (Sm7α-HSDH)	7α-HSDH	NAD(H)	24780	174	142
8	Is2-SDR	metagenome	38.49 (Bsp7α-HSDH)	n.a.[d]	NADP(H)	-	107	-
9	Ngi7-SDR	metagenome	35.50 (Csp12α-HSDH)	n.a.[d]	NAD(H)	-	326	-
10	Ls12α-HSDH	*Lysinibacillus sphaericus*	47.39 (El12α-HSDH)	12α-HSDH	NAD(H)	348	135	2.6

[a] For details about known HSDHs (source and Genbank accession numbers), see Table S1 in the Supporting Information. [b] Activity assays were performed in the presence of NAD^+ as cofactor and respective substrates: 7α-HSDH, cholic acid; 12α-HSDH, deoxycholic acid; 7β-HSDH, ursodeoxycholic acid (see Experimental section for details). [c] Yields refer to 1 L-scale enzyme production under optimized conditions (see Experimental section for details). [d] n.a.: not active towards any of the tested bile acids (see Experimental section for details).

The third novel 7α-HSDH, Hh7α-HSDH (entry 2), showed, in agreement with the phylogenetic analysis (Fig. 1), a mesophilic behavior (optimal temperature at 30°C). Moreover, consistently with its origin, it showed a slight halophilic character. In fact, an about 40% decrease of activity was observed when removing by dialysis the salts contained in the elution buffer of the Ni-NTA chromatography.

Coming to the next four sequences (entries 3-6, Table 1), these enzymes showed unequivocally NAD(H)-dependent 7β-HSDH activity and were consequently named Hh7β-HSDH, Sc7β-HSDH, Bsp7β-HSDH, and Rs7β-HSDH, respectively. As shown in Fig. 1, these 7β-HSDHs constitute a separate clade from that including other known 7β-HSDHs, which are all NADP(H)-dependent enzymes with the only exclusion of the recently discovered NAD(H)-dependent 7β-HSDH from *Lactobacillus spicheri* (Ls-7β-HSDH).[5] Moreover, the novel 7β-HSDHs showed a significantly higher specific activity (up to 92 U mg^{-1}) than Ls-7β-HSDH (3.1 U mg^{-1}),[5] thus suggesting a highly interesting potential for further biocatalytic applications in bile acid modifications.

Subsequent characterization of Hh7β-HSDH (entry 3) showed that the presence of salts had a stronger influence on its activity than that observed with its clustered enzyme Hh7α-HSDH (entry 2). In fact, dialyzed samples of Hh7β-HSDH showed about 2 orders of magnitude lower specific activity (0.07-0.16 U mg^{-1}) than that observed after the affinity chromatography step (10 U mg^{-1}). Interestingly, Hh7β-HSDH fully recovered its starting activity by addition of NaCl (1 M final concentration) to the dialyzed solution.

Sc7β-HSDH (entry 4) showed a rather limited solubility during the recombinant production with a detrimental effect on the expression yields, even under optimized conditions (Table S3, Supporting Information). However, this enzyme is quite thermophilic, showing an optimum of temperature at around 50°C (Figure S4, Supporting Information).

Excellent yields and specific activity were obtained instead with Bsp7β-HSDH (entry 5) and Rs7β-HSDH (entry 6), both enzymes showing a mesophilic character.

As far as the metagenomic sequences Is2-SDR and Ngi7-SDR concern, both proteins were obtained with very good yields in soluble form (entries 8 and 9, Table 1). However, neither of the two showed catalytic activity toward any of the tested bile acids (see Experimental section for details). Indeed, both enzymes demonstrated to be functional dehydrogenases in the presence of NADPH (Is2-SDR) or NADH (Ngi7-SDR) as cofactors and different ketone substrates (see in the following).

10.1002/adsc.202000120

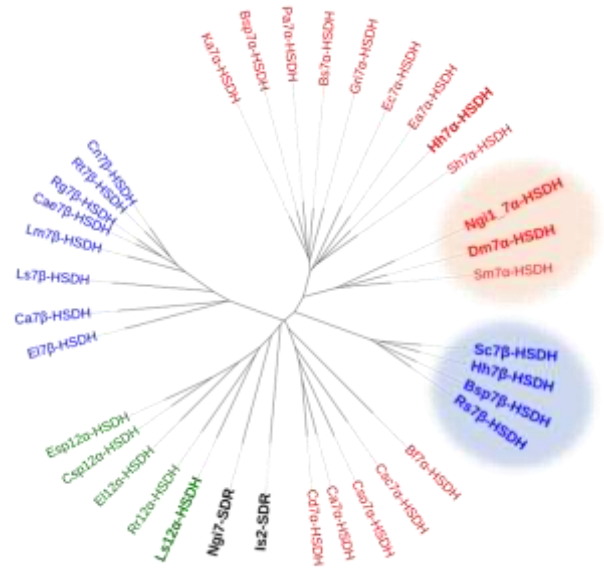


Figure 1. Unrooted phylogenetic tree of HSDHs. 7α-HSDHs are indicated in red, 7β-HSDHs in blue, 12α-HSDHs in green. Novel enzymes discovered in this work are in bold. Thermophilic 7α-HSDHs and the novel NAD(H)-dependent 7β-HSDHs are in light red and light blue background, respectively. The tree was generated with Clustal Omega[46] and visualized with iTOL.[47]

Finally, Ls12α-HSDH (entry 10, Table 1) was obtained with excellent yield from recombinant production and confirmed as a NAD(H)-dependent 12α-HSDH, thus included, together with the other seven novel HSDHs, in the following study of the substrate promiscuity of this group of enzymes.

Substrate scope of HSDHs

The demonstration of the catalytic activity of the eight novel HSDHs in oxidoreduction reactions of different bile acids paves the way to their exploitation as biocatalysts for the stereo- and regioselective preparation of steroid derivatives. Our group is currently performing specific investigations on this topic and the results will be reported in due time.

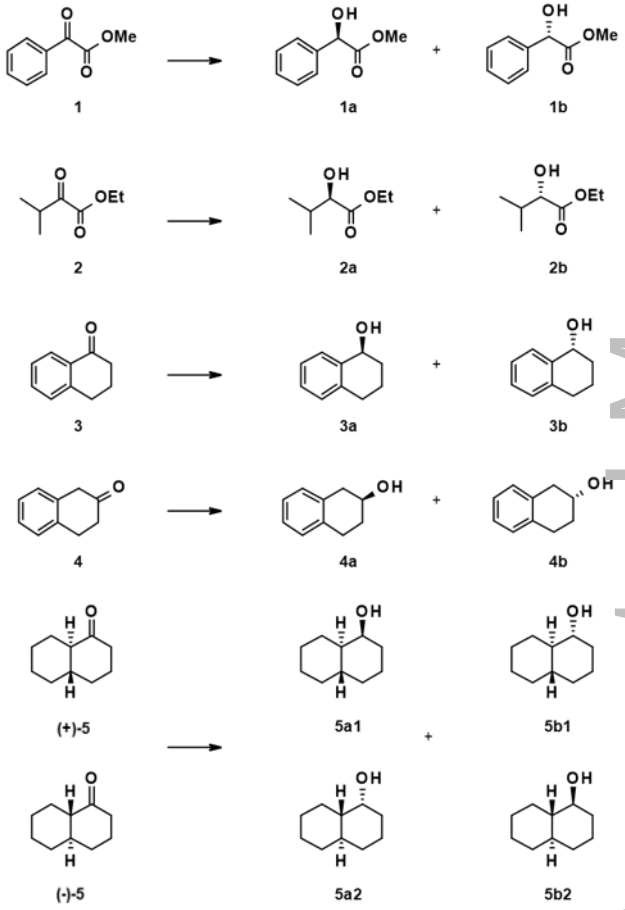
Additionally, we considered of synthetic interest a more detailed evaluation of the substrate promiscuity of this group of enzymes. Accordingly, the newly discovered 7α-, 7β-, and 12α-HSDHs were tested, along with our in-house collection of previously characterized HSDHs (Table 2, for details see also Table S1, Supporting Information), in the stereoselective reduction of a panel of substrates (**1**-**5**) (Scheme 2).

Specifically, the tested substrates included two α-ketoesters (**1**-**2**) of pharmaceutical interest,[48] chosen among the panel of ketoesters that, as previously discussed,[20] were selected by Zhu and co-workers to investigate the catalytic promiscuous properties of the *B. fragilis* 7α-HSDH, and three bicyclic ketones (**3**-**5**) partially resembling the structural features of steroids.

To allow the use of catalytic amounts of the cofactors and to provide the necessary driving force to product formation, the reactions were coupled with a suitable cofactor regeneration protocol exploiting a formate/formate dehydrogenase (FDH) system (see Experimental section and Scheme S2 in Supporting information).

Reaction conversions were estimated after 24-48 h by GC-MS (**1, 3**-**5**) or by HPLC (**2**) analysis and enantiomeric excesses (ees) were evaluated by chiral GC (**1** and **5**) or chiral HPLC analysis (**2**).

For the sake of comparison, compounds **1**-**5** were also submitted to the action of the two new identified SDRs not showing HSDH activity (Is2-SDR and Ngi7-SDR, lines 8 and 9 of Table 1) and the results are reported in the Supplementary Materials (Table S4).



Scheme 2. Substrates **1**-**5** and their possible reduction products.

Regarding the α-ketoesters **1**-**2**, nearly all the tested HSDHs showed a good activity towards these substrates, yielding the corresponding reduced α-hydroxyesters, with high conversions and high enantiomeric excesses (Table 2).

Specifically, quantitative conversions of **1** were obtained with Ca7α-HSDH, Dm7α-HSDH, Ca7β-HSDH, Cae7β-HSDH and Bsp7β-HSDH (Table 2,

10.1002/adsc.202000120

entries 2, 3, 6, 7, and 10), while complete conversions of **2** were obtained with Dm7α-HSDH, Ca7β-HSDH, Cae7β-HSDH, Bsp7β-HSDH, Rs7β-HSDH and Ls12α-HSDH (Table 2, entries 3, 6, 7, 10, 11 and 13).

On the other hand, the most enantioselective HSDHs in the reduction of **1** were Ngi1_7α-HSDH, Rs7β-HSDH and Csp12α-HSDH (Table 2, entries 5, 11, and 12), while Ec7α-HSDH and Hh7α-HSDH (Table 2, entries 1 and 4) showed the best enantioselectivity in the reduction of **2**.

Interestingly, the stereoselectivity in the reduction of **1** was significantly lower than that observed with **2** when testing the NADP(H)-dependent Ca7β-HSDH and Cae7β-HSDH (entries 6 and 7, respectively), and, to a minor extent, the novel Ls12α-HSDH enzyme (entry 13). These results suggest that in these selected HSDHs the aromatic substituent in **1** may promote alternative substrate binding modes in either pro-*R* or pro-*S* fashion, whereas in the case of the aliphatic α-keto ester **2** only the pro-*R* binding mode is productive.

It is also worth of mentioning that reactions catalyzed by Hh7α-HSDH and Hh7β-HSDH (Table 2, entries 4 and 9) were carried out in parallel in the presence and in the absence of NaCl (0.4 M). As expected, conversions were very low in the absence of NaCl (<5%, data not shown), while they were up to about 98% in the presence of salts, thus confirming the halophile behavior of these novel HSDHs.

The thermophilic Dm7α-HSDH (Table 2, entry 3) demonstrated to perform well also at sub-optimal temperatures, giving complete conversions of **1** and **2** at 25°C.

The stereochemistry of the products obtained from the enzymatic reduction of **1** (the enantiomers **1a** or **1b**) was determined *via* analytical comparison with standards. All the HSDHs showed to be (*R*)-selective, compound **1a** being the preferred product. Instead, in order to assign the absolute configurations of the products obtained by the reduction of **2** (the enantiomers **2a** or **2b**), the reaction catalyzed by 7β-HSDH from *Collinsella aerofaciens* (Cae7β-HSDH, Table 2, entry 7), which combined quantitative conversion of **2** and high enantioselectivity, was scaled up (100 mg). Optical rotation measurements were carried out on the purified product (63% isolated yield), the resulting $[\alpha]_D^{22}$ value (-2.70°) being consistent with the value reported in literature for **2a**.[49] Thus, also in the case of the reduction of **2**, all the HSDHs demonstrated to be (*R*)-selective.

Table 2. Conversions and enantioselectivity of HSDH-catalyzed reduction of substrates **1** and **2**.

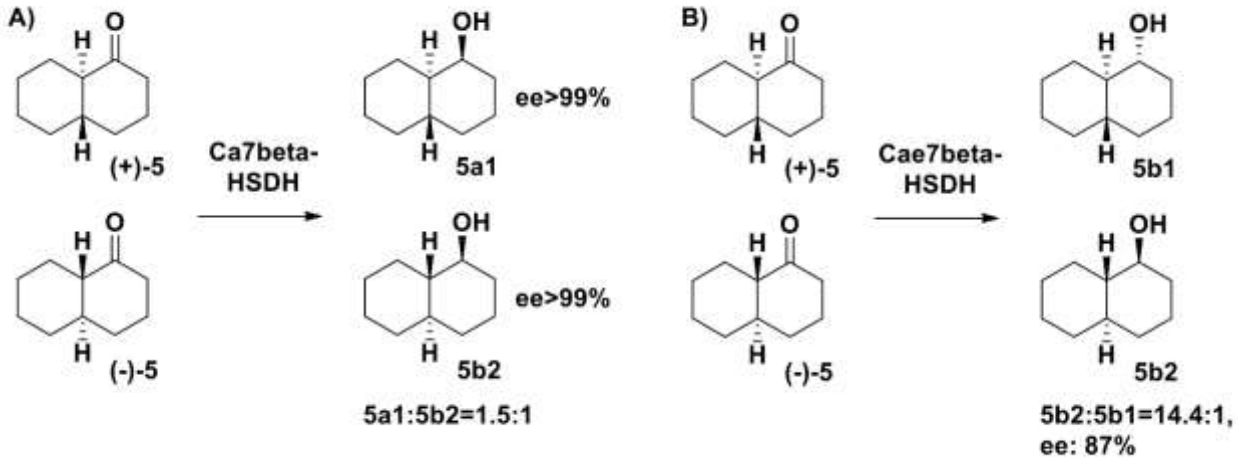
Entry	Enzyme[a)]	Compound **1**		Compound **2**	
		c[b)] (%)	ee_R[b)] (%)	c[b)] (%)	ee_R[b)] (%)
1	Ec7α-HSDH	89.7	90.2	81.6	>99
2	Ca7α-HSDH	>99	88.3	42.3	91.1
3	Dm7α-HSDH	>99	82.9	>99	93.9
4	Hh7α-HSDH	95.9	97.9	81.0	>99
5	Ngi1_7αHSDH	26.1	>99	64.9	92.4
6	Ca7β-HSDH	>99	45.2	>99	89.3
7	Cae7β-HSDH	>99	23.3	>99	97.2
8	Sc7β-HSDH	61.7	92.5	92.0	94.6
9	Hh7β-HSDH	97.4	95.2	65.8	64.9
10	Bsp7β-HSDH	>99	89.6	>99	96.0
11	Rs7β-HSDH	75.4	>99	>99	95.5
12	Csp12α-HSDH	94.5	>99	93.9	85.7
13	Ls12α-HSDH	94.1	64.7	>99	93.6

[a)] Sources of previously characterized HSDHs: Ec7α-HSDH (entry 1): *Escherichia coli*; Ca7α-HSDH (entry 2) and Ca7β-HSDH (entry 6): *Clostridium absonum*; Cae7β-HSDH (entry 7): *Collinsella aerofaciens*; Csp12α-HSDH (entry 12): *Clostridium* sp. (see Table S1, Supporting Information, for details). [b)] Conversions and ees were assigned via chiral HPLC for reduction of **1**, while chiral GC was used for reduction of **2** (see Experimental section for details).

Concerning the reduction of the bicyclic substrates **3**-**5**, none of the tested enzymes showed activity towards the tetralones **3** and **4**, whereas only Ca7β-HSDH and Cae7β-HSDH were found able to catalyze the reduction of (±)-trans-1-decalone ((±)-**5**). These two enzymes, working under unoptimized reaction conditions, gave moderate conversions (18% and 35% after 48 h for Ca7β-HSDH and Cae7β-HSDH, respectively), but accompanied by interesting stereoselectivity (Scheme 3).

10.1002/adsc.202000120

Considering that commercially available **5** is a racemate of the enantiomers (+)-**5** (4a*S*,8a*R*) and (-)-**5** (4a*R*,8a*S*), its chemical reduction can yield four different stereoisomers, i.e., **5a1** (1*S*,4a*S*,8a*R*), **5a2** (1*R*,4a*R*,8a*S*), **5b1** (1*R*,4a*S*,8a*R*), and **5b2** (1*S*,4a*R*,8a*S*) (Scheme 2). The four isomers were prepared by chemical reduction of **5** with $NaBH_4$, isolated and analyzed by NMR spectroscopy. In this way, it was possible to assign the relative stereochemistry of the diastereomers **5a** and **5b** (see Supporting Information for further details).Moreover, the four isomeric products (**5a1**, **5a2**, **5b1**, **5b2**) could be separated by chiral GC.



Scheme 3. Enzymatic reductions of substrate (±)-**5** with A) Ca7β-HSDH and B) Cae7β-HSDH.

The same analysis of the samples obtained from the two enzymatic reductions indicated that Ca7β-HSDH catalyzed the formation of one enantiomer of each diastereomer (Scheme 3A, see also Scheme S3 in the Supporting Information for details). On the contrary, Cae7β-HSDH performed a kinetic resolution of the starting ketone. The unreacted substrate **5** showed an ee of 72%, whereas the product **5b** showed an ee of 87% (Scheme 3B). Therefore, Ca7β-HSDH and Cae7β-HSDH showed a different stereopreference for the enantiomers of (±)-**5**.

To assign the absolute configurations of the residual substrate and of the formed products, the reduction of (±)-**5** catalyzed by Cae7β-HSDH was scaled up to 100 mg.

The remaining substrate and the product (**5b**, 15% isolated yield; product **5a** was present in traces and therefore it was not isolated) were purified *via* flash chromatography, then submitted to NMR analysis in order to identify the relative stereochemistry and to optical rotation measurements to assess the absolute configurations (see Supporting Information for further details). The resulting $[\alpha]_D^{25}$ values (+44.45° and +2,75° for produced **5b** and residual **5**, respectively) were compared with the data reported in literature,[50] indicating that the residual ketone was enriched in the enantiomer **(+)-5** and the product was enriched in the enantiomer **5b2** (1*S*,4a*R*,8a*S*). These data are summarized in Table 3.

Table 3. Conversion and enantioselectivity of reduction of substrate (±)-**5** catalyzed by Cae7β-HSDH and Ca7β-HSDH.

Enzyme	c[a)]	ee[a)]		
		residual substrate	product **5b**	product **5a**
	(%)	(%)	(%)	(%)
Cae7β-HSDH	35	72 (4a*S*,8a*R*)	87[b)]	-
Ca7β-HSDH	18	9 (4a*R*,8a*S*)	> 99[b)]	> 99[c)]

[a)] Conversions and ees were assigned via chiral GC (see Experimental section for details). [b)] Major product stereochemistry: (1*S*,4a*R*,8a*S*). [c)] Major product stereochemistry: (1*S*,4a*S*,8a*R*).

Our previous investigations on the 3D structure of Cae7β-HSDH[51] showed that its substrate binding site, including a mobile substrate loop, is substantially different from that of other HSDHs, e.g., *E. coli* 7α-HSDH. Although the overall active site architecture results quite similar in this family of oxidoreductases, it is likely that substrate recognition and, consequently, stereoselectivity, are ruled by a network of very subtle interactions between the substrate and the active site residues. Further investigations in this respect are currently under considerations in our lab.

Conclusion

As far as the mining of (meta)genomes for novel HSDHs is concerned, the results obtained in the first part of this work show that the occurrence of redox enzymes acting on steroids largely exceeds the range of microbes usually considered as possible HSDHs sources so far, i.e., gut and soil bacteria. In fact, functionally active HSDHs have been found for the first time in extremophilic bacteria as well as in photosynthetic cyanobacteria. This finding poses several interesting questions that were clearly out of the scope of this investigation, e.g., about the physiological role of these HSDHs in these microbes, as well as about how these enzymatic functions were acquired or evolved from common ancestral genes. However, since extremophilic enzymes are very often more robust than their mesophilic counterparts,[52] the discovery of these novel HSDHs could also be useful in the development of industrial biocatalytic processes for the stereo- and regioselective modification of steroid derivatives. We are currently working on this specific topic, as well as on the synthetic exploitation of the novel and highly active NAD(H)-dependent 7β-HSDHs for the preparation of bile acid derivatives.

Coming to the study of the activity of HSDHs toward non-steroidal substrates, we have confirmed that enzymes evolved to accommodate structurally complex and bulky substrates in their active sites, maintain their selectivity towards much simpler molecules, as observed with compounds **1**, **2** and **5**. These results corroborate the few scant previous literature reports, point out the importance of having in hands libraries of enzymes and encourage us to continue these studies in order both to acquire more information on the general substrate scope of HSDHs and to exploit molecular modeling approaches[53] to find suitable rationales for the observed selectivity.

Experimental Section

General

Methyl benzoylformate (purity 98%), ethyl 3-methyl-2-oxobutyrate (purity 97%), 1-tetralone (purity 97%), 2-tetralone (purity 98%) and (±)-trans-1-decalone (purity 98%) were purchased from Merck (Darmstadt, Germany) (catalog numbers M30507, 218456, T19003, T19208 and 156655, respectively) and used with no further purification. The analytical standards methyl D-mandelate (purity≥99%, $[\alpha]_D^{20}$=144°±2, c=2 in MeOH) and methyl L-mandelate (purity>99%) were from Fluka Chemika (catalog numbers 63456 and 63466, respectively). Formate dehydrogenase (FDH) from *Candida boidinii* and lactate dehydrogenase (LDH) from rabbit muscle were from Merck. The NADP(H) dependent FDH from methylotrophic bacterium *Pseudomonas* sp. 101 was a kind gift from Prof. Tishkov (M.V. Lomonosov Moscow State University). Unless otherwise stated, all other chemicals were of analytical grade and were purchased from Merck.

Reactions were monitored via TLC (thin-layer chromatography) on pre-coated glass plates silica gel 60 with fluorescent indicator UV_{254} and treated with an oxidizing solution [4-hydroxybenzaldehyde (6.3 g), H_2SO_4(50% v/v in H_2O, 40 mL), MeOH (400 mL)]. The isolation of pure products was allowed via extraction with EtOAc and subsequent flash chromatography on silica gel 60 (70-320 mesh, Merck, eluent CH_2Cl_2).

Environmental sample collection, DNA extraction from samples, DNA sequencing, and generation of databases of metagenomic sequences were carried out as previously described.[54] The metagenomic sample "ngi" was collected in drilling wells in Svalbard island and DNA extraction and sequencing from this sample was performed as described previously for sample It3.[54]

Bacterial strains.

E. coli BL21(DE3) and *E. coli* HI-Control 10G were from Lucigen (Wisconsin, USA).

E. coli BL21(DE3)/pET24b-Ec7αHSDH producing the 7α-HSDH from *E. coli*, *E. coli* BL21(DE3)/pET24b-Cae7βHSDH producing the 7β-HSDH from *Collinsella aerofaciens*, *E. coli* BL21(DE3)/pETite-Ca7βHSDH producing the 7β-HSDH from *Clostridium absonum*, *E. coli* TOP10/pBAD-Ca7αHSDH producing the 7α-HSDH from *Clostridium absonum* were part of our in-house collection and were expressed as fusion proteins with a (6x)His-tag at the C-terminus.

Analytical methods

At scheduled times, reaction samples (50 µl) were extracted with EtOAc and dried over Na_2SO_4 to obtain a 10 mM final concentration sample suitable to chiral GC (and/or GC-MS) analysis, or evaporated, resuspended in CH_3CN to 10 mM final concentration and analyzed by chiral RP-HPLC.

GC-MS analyses were performed using an Agilent HP 5MS column (30 m × 0.25 mm × 0.25 µm) on a Finnigan TRACE DSQ GC/MS instrument (ThermoQuest, San Jose, CA). Elution conditions: 60°C, 1 min; +10°C min^{-1} until 300°C; hold 1 min; flow rate: 1.0 mL min^{-1}; inlet temperature: 250°C; ion source temperature: 250°C; MS transfer line temperature: 250°C. Retention times: (**1**): 9.53 min; (**1a**, **1b**): 9.38 min; ((-)-**5** and(+)-**5**): 9.57 min; (**5b1**, **5b2**): 9.25 min.

Chiral GC analyses were carried out on an Agilent Technologies 6850 Network GC system gas chromatograph equipped with split/splitless injector, FID detector and MEGA-DEX DAC Beta chiral capillary column (25 m x 0.25 mm x 0.25 µm). Elution conditions for **2** and **5:** 80°C, 1 min; + 1°C min^{-1} until 90°C; +10°C min^{-1} until 180°C; hold 5 min. Retention times: (**2**): 9.1 min; (**2a**, **2b**): 7.9 min, 8.8 min ((-)-**5** and(+)-**5**, respectively): 17.5 min, 17.6 min; (**5b1**, **5b2**): 17.7 min, 18.6 min. Elution conditions for **3** and **4:** 120°C, 1 min; + 5°C min^{-1} until 180°C; +10°C min^{-1} until 200°C; hold 5 min. Retention times: (**3**) 8.8 min; (**3a**, **3b**): 9.1 min, 9.3 min; (**4**) 10.5 min; (**4a, 4b**): 10.0 min. Flow rate: 1.5 mL min^{-1}; detector temperature 200°C; inlet temperature 250°C.

Chiral HPLC analyses were performed on a Shimadzu LC-20AD high performance liquid chromatography system equipped with a Shimadzu SPD-20A UV detector and a Phenomenex Lux Cellulose-1 5µ chiral column. HPLC conditions: injection volume 10 µL; mobile phase H_2O + 0.1% trifluoroacetic acid : CH_3CN = 75 : 25 (isocratic elution); flow rate: 1.0 mL min^{-1}; detection 254 nm; temperature 30°C. Retention times: (**1**): 4.0 min; (**1b**): 5.5 min; (**1a**): 7.0 min. Prior to performing HPLC analyses, the molar extinction coefficients of **1** and **1a**, **1b** were determined (see Supporting information for details), to be able to assess conversions while considering the different molar absorbivity of substrate and products.

The NMR spectra (^{1}H and ^{13}C) were acquired in $CDCl_3$ or in DMSO-d_6 at rt on a Bruker AV 400 MHz spectrometer with a z gradient at 400 MHz for ^{1}H-NMR analysis and 101 MHz for ^{13}C-NMR.

ESI-MS spectra were recorded on a Bruker Esquire 3000 PLUS instrument (ESI Ion Trap LC/MSn System), equipped with an ESI source and a quadrupole ion trap detector (QIT). The samples were dissolved in methanol to 10-2 g L^{-1} and then directly syringed in the ESI-MS at 4 µL min^{-1} rate. The analyses were performed in positive mode. The acquisition parameters were optimized as such: 4.5 kV needle voltage, 10 L h^{-1} N_2 flow rate, 40 V cone voltage, trap drive set to 46, 115.8 V capillary exit, 13000 (m/z) s^{-1} scan resolution over the 35-900 m/z mass/charge range, source temperature 250°C.

Optical rotations were measured on a Jasco P-2000 polarimeter. The specific rotation was calculated as the $[\alpha]_\lambda^T=\alpha 100/cd$, where α represents the recorded optical rotation, c the analyte concentration (mg mL^{-1}), d the cuvette length (dm). As such, the specific rotation is expressed as (10^{-1} deg cm^{-2} g^{-1}). λ is reported in nm and T in °C. λ usually corresponds to sodium D line (589 nm), thus the optical rotation is referred to as $[\alpha]_D$. T, c and the solvent were chosen according to references reported in literature: c = 1 in $CHCl_3$ at 22°C for **2**; c = 0.5 in EtOH at 25°C for **5**; c = 0.75 in $CHCl_3$ at 25°C for **5a** and **5b**.[49,50]

In silico screening for novel HSDHs and bioinformatic analysis

Bioinformatic search for new HSDHs was performed by aligning query sequences (entry 1, 2, 4, 14-21, 33 Table S1 Supporting Information) with database metagenomic sequences using the program LAST (http://last.cbrc.jp/) with default settings[55] or with GenBank database sequences using standard protein Blast tool (https://blast.ncbi.nlm.nih.gov/).

Phylogenetic trees were created using the Clustal Omega tool (http://www.ebi.ac.uk/Tools/msa/clustalo/)[46] and visualized with iTOL webserver (http://itol.embl.de/).[47]

Gene cloning and generation of recombinant bacterial strains

The codon-optimized genes coding for putative HSDHs (Table S1, Supporting Information, entries 11-13, 26-29, 31, 43, and 44,) and for Csp12α-HSDH (entry 33) were synthesized and cloned into the pUC57 vector by BaseClear (Leiden, The Netherlands).

Putative HSDH genes were amplified by PCR using primers reported in Table S2 (Supporting Information), for the subsequent cloning in the pETite C-His Kan vector *in frame* with C-term His Tag sequence. PCR amplifications were carried out in 50 µl reaction mixtures containing 20 ng of pUC57 vector including the desired gene, primers (1 µM each), dNTPs (0.2 mM each), 2 U of Xtra Taq polymerase and 5 µl of buffer containing $MgCl_2$. All PCR reagents were from Genespin (Milan, Italy). PCR conditions were as follows: 95 °C for 2 min, followed by 35 cycles at 94 °C for 30 s, primers T_m -5°C for 30 s, 72 °C for 1 min, and then 72 °C for 10 min. PCR products were then purified from agarose gel (0.7% (w/v)) using the Wizard® SV Gel and PCR Clean-Up System (Promega, Wisconsin, USA) before cloning. Purified sequences were subsequently cloned in the pETite C-His Kan plasmid using the Expresso T7 Cloning and Expression kit from Lucigen (Wisconsin, USA). Specifically, purified sequences were mixed with the pETite C-his linear plasmid and transformed in chemically competent *E. coli* HI-Control 10G cells following the manufacturer's instructions.

Csp12α-HSDH cloning into the pETite N-His Kan vector *in frame* with N-term His-Tag sequence was carried out by gene amplification with primers F10/R10 (Table S2, Supporting Information), using the PCR protocol described above. The amplified sequence was subsequently cloned in the pETite N-His Kan vector using the same kit described for the cloning in the pETite C-His Kan plasmid.

The resulting plasmids were purified from *E. coli* HI-Control 10G cells using the E.Z.N.A. Plasmid Mini kit II (Omega/VWR) and plasmid inserts were sequenced on both strands by Biofab Research (Rome, Italy) using primers T7 promoter and pETite reverse (Table S2, Supporting Information). Recombinant pETite plasmids were finally transformed in *E. coli* BL21(DE3) chemically competent cells (Lucigen) for the expression of the corresponding genes.

HSDH/SDR and FDH expression and purification

Recombinant *E. coli* strains from our in-house collection (see bacterial strain section) and recombinant *E. coli* BL21(DE3) strains prepared as described above, were inoculated overnight in LB medium supplemented with the opportune antibiotic (Table S3, Supporting Information) (100 mL) and grown at 37°C, 220 rpm. 25 mL of precultures were subsequently inoculated in 500 mL of LB medium containing the corresponding antibiotic and incubated at 37°C, 220 rpm till the OD_{600} cell density reached 0.5–1. Gene expression was induced by the addition of IPTG (for IPTG final concentrations, see Table S3, Supporting Information) and the culture was transferred to 17-37°C (see Table S3, Supporting Information) with shaking at 220 rpm and grown for 4-72h (see Table S3, Supporting Information). In case of *E. coli* TOP10/pBAD-Ca7αHSDH, precultures were inoculated in TB medium containing rhamnose as protein expression inducer.

After recovery by centrifugation (5000 rpm, 30 min, 4°C), cell pellet was resuspended in 20 mL wash buffer (20 mM potassium phosphate (KP) buffer, pH 7.0, 500 mM NaCl, 20 mM imidazole) and cells were disrupted by sonication. Soluble protein fraction was separated from the cell debris by centrifugation (11,000 rpm, 30 min) and clear lysates were assessed for the presence of soluble protein by SDS-PAGE (12% T, 2.6% C).

For protein purification, Ni Sepharose 6 Fast Flow agarose resin (Ni-NTA) (GE Healthcare, Italy) was incubated with clear cell lysates containing soluble protein for 90 min at 4 °C under mild shaking. The mixture was then loaded onto a glass column (10 × 110 mm) and the resin was washed with 20 mL of wash buffer (20 mM imidazole, 500 mM NaCl, 20 mM KP buffer). His tagged proteins were eluted using a 3-step gradient (10 ml washing buffer with 100, 200, and 300 mM imidazole, respectively) and, if not stated otherwise, dialyzed against 5 L of a suitable buffer (Table S3, Supporting Information), at 4°C for 16 h and stored at −80°C. Protein content was measured using the Bio-Rad Protein Assay according to the method of Bradford and protein purity was verified by SDS-PAGE analysis (12% T, 2.6% C). The molecular weight protein standard mixture from Bio-Rad (Karlsruhe, Germany) was used as reference. Gels were stained for protein detection with Coomassie Brilliant Blue.

Activity assays

Dehydrogenase activity of HSDHs and FDHs was determined spectrophotometrically by measuring the reduction of $NAD(P)^+$ at 340 nm (ε: 6.22 $mM^{-1}cm^{-1}$) in the presence of the opportune substrate. If not stated otherwise, cholic acid was used for 7α- and 12α-HSDHs, ursodeoxycholic acid was used for 7β-HSDHs, and ammonium formate was used for FDHs.
Specifically, assays were carried out in polyethylene cuvettes at 25°C by adding the opportune purified

dehydrogenase (1-20 µL) to the following assay mixtures (1 mL final volume):
HSDH assay: 2.5 mM substrate; 50 mM potassium phosphate buffer at pH 9.0; 0.20 mM $NAD(P)^+$.
FDH assay: 20 mM substrate; 50 mM potassium phosphate buffer at pH 7.0; 0.20 mM $NAD(P)^+$.
One unit of activity is defined as the enzyme activity that reduces 1 µmol of $NAD(P)^+$ per min under the assay conditions described above.
HSDHs optimal temperatures were determined by heating the assay solution in cuvettes in a water bath in the range 20–90 °C for 10 min before adding the enzyme. Results were compared to blanks.

Small-scale biotransformations of bile acids

Oxidation of cholic Acid to 3α,12α-dihydroxy-7-oxo-5β-cholanoic acid or to 3α,7α-dihydroxy-12-oxo-5β-cholanoic acid. Cholic acid oxidation reactions were coupled with a pyruvate/LDH system to regenerate NAD^+. Specifically, reactions were carried out in a 1 mL solution containing 50 mM potassium phosphate buffer, pH 8.0, 10 mM cholic acid, 50 mM sodium pyruvate, 0.4 mM NAD^+, 0.5 U LDH from rabbit muscle, 2 U of opportune 7α-HSDH (entry 11-13, Table S1, Supporting Information) to obtain 3α,12α-dihydroxy-7-oxo-5β-cholanoic acid (Scheme S1, A) or Ls12α-HSDH to obtain 3α,7α-dihydroxy-12-oxo-5β-cholanoic acid (Scheme S1, B) at 25°C for 24 h.
Reaction progress was monitored by TLC using chloroform-methanol-acetic acid, 10:1:0.08, as eluting system.

Reduction of 3α,12α-dihydroxy-7-oxo-5β-cholanoic acid to ursocholic acid. 3α,12α-diHydroxy-7-oxo-5β-cholanoic acid reduction was coupled with a formate/FDH system to regenerate NADH. Specifically, reactions were performed in a 1 mL solution containing 50 mM potassium phosphate buffer, pH 8.0, 10 mM cholic acid, 50 mM NH_4HCO_2, 0.4 mM NADH, 0.5 U FDH from *Candida boidinii*, 2 U of opportune 7β-HSDH (entry 26-29, Table S1, Supporting Information) (Scheme S1, C) at 25°C for 24 h.
Reaction progress was monitored by TLC using the same eluting system described for cholic acid oxidation (chloroform-methanol-acetic acid, 10:1:0.08), as eluting system.

Preparation of standard racemate mixtures

The reductions of substrates **1-5** were performed following a standard protocol of reduction with $NaBH_4$.[56] To a stirred solution of 0.13 M substrate (1 eq, 100 mg) in MeOH (5 mL) at 0°C, $NaBH_4$ (1 eq, 13 mg) was added. When the reaction mixture became clear, it was brought to rt and stirred for 2-4 h. The reaction was quenched with a saturated aqueous solution of NH_4Cl, then it was extracted with EtOAc (3x). The organic phase was dried over Na_2SO_4, filtrated and the solvent was evaporated under reduced pressure to yield the desired products (**1a**+**1b**, **2a**+**2b**, **3a**+**3b**, **4a+4b, 5a1**+**5a2**+**5b1**+**5b2**) in quantitative yields. The products were characterized via ^{1}H-NMR analysis and chiral GC or HPLC analyses (see Supporting information, and analytical methods).

1a+1b: ^{1}H NMR (400 MHz, CDCl3) δ 7.49 – 7.30 (m, 1H), 5.20 (s, 1H), 3.79 (s, 1H).

2a+2b: ^{1}H NMR (400 MHz, CDCl3) δ 4.42 – 4.15 (m, 1H), 4.04 (s, 1H), 2.72 (s, 1H), 2.19 – 2.00 (m, J = 13.7, 6.9, 3.5 Hz, 1H), 1.33 (t, J = 7.1 Hz, 1H), 1.05 (d, J = 6.9 Hz, 1H), 0.89 (d, J = 6.9 Hz, 1H).

3a+3b: ^{1}H NMR (400 MHz, CDCl3) δ 7.48 – 7.41 (m, 1H), 7.26 – 7.18 (m, 2H), 7.15 – 7.08 (m, 1H), 4.88 – 4.72 (m, 1H), 3.04 – 2.52 (m, 2H), 2.21 – 1.74 (m, 4H).

4a+4b: ^{1}H NMR (400 MHz, CDCl3) δ 7.24 – 7.02 (m, 4H), 4.30 – 3.98 (m, 1H), 3.12 (dd, J = 15.9, 4.5 Hz, 1H), 3.07 – 3.06 (m, 1H), 2.99 (dt, J = 17.0, 5.8 Hz, 1H), 2.92 – 2.84 (m, 1H), 2.80 (dd, J = 16.3, 7.9 Hz, 1H), 2.15 – 2.04 (m, 1H), 1.92 – 1.79 (m, 1H).

5a1+**5a2**: ^{1}H NMR (400 MHz, CDCl3) δ 3.76 (q, J = 2.5 Hz, 1H), 1.92 – 1.74 (m, 2H), 1.74 – 1.55 (m, 4H), 1.56 – 1.43 (m, 3H), 1.44 – 1.16 (m, 5H), 1.15 – 0.85 (m, 3H).

5b1+**5b2:** ^{1}H NMR (400 MHz, CDCl3) δ 3.20 (ddd, J = 10.2, 9.1, 4.3 Hz, 1H), 2.18 – 2.07 (m, J = 6.1, 3.0 Hz, 1H), 2.04 – 1.91 (m, 1H), 1.87 – 1.59 (m, 4H), 1.61 – 1.17 (m, 7H), 1.12 – 0.81 (m, 5H).

Enzymatic reduction of substrates 1-5 catalyzed by HSDHs

Reactions catalyzed by HSDHs were coupled with a formate/FDH system to regenerate NAD(P)H. For the initial activity screening the general reaction protocol was as follows: 76 mM NH_4HCO_2; 0.2 U mL^{-1} FDH; 0.4 mM $NAD(P)^+$; HSDH, 3.4 U mL^{-1}; 12.5 mM substrate; 5% v/v DMSO; 50 mM potassium phosphate buffer, pH 7.0 (total volume: 1 mL). Reaction catalyzed by Hh7α-HSDH and Hh7β-HSDH were performed in the presence and in the absence of 0.4M NaCl. The mixtures were shaken at 25°C and 100 rpm for 24 to 48 h and monitored over time via TLC (eluent CH_2Cl_2). Reaction conversions and enantiomeric excesses were evaluated by GC-MS, chiral GC or chiral HPLC analyses. The absolute configurations were assigned via analytical comparison with standards (products **1a** and **1b**) or via optical rotation measurements and confrontation with values reported in literature (products **2a** and **2b**, **(+)-5** and **(-)-5**, **5b1** and **5b2**).[49,50]

Reactions on substrates **2** and **5** catalyzed by *Cae*7β-HSDH were subsequently scaled up to semi-preparative scale (100 mg, 0.7 mmol in 50 mL total volume) following the protocol described above. The isolated products were characterized via ^{1}H- and ^{13}C-NMR, ESI-MS, chiral GC and $[\alpha]_D^{25}$ (see Supporting information, and analytical methods).

2a+2b: ^{1}H NMR (400 MHz, DMSO) δ 5.05 (d, *J* = 5.9 Hz, 1H), 4.04 – 3.89 (m, 2H), 3.64 (t, *J* = 5.4 Hz, 1H), 1.83 – 1.70 (m, 1H), 1.05 (t, *J* = 7.1 Hz, 3H), 0.74 (d, *J* = 6.9 Hz, 3H), 0.69 (d, *J* = 6.8 Hz, 3H); ^{13}C NMR (101 MHz, CDCl3) δ 174.90, 77.35, 77.04, 76.72, 74.99, 61.48, 41.00, 32.14, 18.74, 15.97, 14.23; $[M+Na]^+$: 167.1; conversion: 99.9%; isolated yield: 63.2%; ee (chiral GC): 97.1%; $[\alpha]_D^{22}$: -2.70°.

5: ^{1}H NMR (400 MHz, CDCl3) δ 2.41 – 2.25 (m, J = 13.5, 5.0, 1.8 Hz, 2H), 2.10 – 2.00 (m, J = 16.3, 5.9, 2.8 Hz, 1H), 2.00 – 1.85 (m, 2H), 1.85 – 1.59 (m, 5H), 1.49 – 1.30 (m, 2H), 1.30 – 1.09 (m, 4H); ^{13}C NMR (101 MHz, CDCl3) δ 212.72, 77.35, 77.03, 76.71, 55.08, 44.96, 41.80, 34.38, 33.04, 26.48, 25.75, 25.43, 25.10; $[M+Na]^+$: 175.1; ee (chiral GC): 62.2%; $[\alpha]_D^{25}$: +2,75°.

5b1+**5b2:** ^{1}H NMR (400 MHz, CDCl3) δ 3.29 – 3.11 (m, 1H), 2.23 – 2.06 (m, J = 6.0, 3.0 Hz, 1H), 2.02 – 1.91 (m, 1H), 1.84 – 1.58 (m, 4H), 1.58 – 1.43 (m, 2H), 1.41 – 1.13 (m, 4H), 1.11 – 0.77 (m, 5H); ^{13}C NMR (101 MHz, CDCl3) δ 77.33, 77.01, 76.70, 75.04, 50.46, 41.13, 35.82, 33.59, 33.43, 29.00, 26.34, 26.15, 24.03; $[M+Na]^+$: 177.1; conversion (GC-MS): 35.8%; isolated yield: 15%; ee (chiral GC): 89.1%; $[\alpha]_D^{25}$: +44.45°.

Acknowledgements

We would like to kindly acknowledge Prof. Vladimir I. Tishkov and Dr. Anastasia A. Pometun (M.V. Lomonosov Moscow State University) for their support in the preparation of the

Pseudomonas sp. FDH. We warmly thank also Laura Cucinotta for valuable technical assistance in enzyme production and characterization and Walter Panzeri (SCITEC-CNR) for analytical support.

References

[1] D. Monti, G. Ottolina, G. Carrea, S. Riva, *Chem. Rev.* **2011**, *111*, 4111–4140.

[2] T. Eggert, D. Bakonyi, W. Hummel, *J. Biotechnol.* **2014**, *191*, 11–21.

[3] F. Tonin, I. W. C. E. Arends, *Beilstein J. Org. Chem.* **2018**, *14*, 470–483.

[4] S. Riva, R. Bovara, P. Pasta, G. Carrea, *J. Org. Chem.* **1986**, *51*, 2902–2906.

[5] F. Tonin, L. G. Otten, I. W. C. E. Arends, *ChemSusChem* **2019**, *12*, 3192–3203.

[6] D. Bakonyi, A. Wirtz, W. Hummel, *Zeitschrift fur Naturforsch. - Sect. B J. Chem. Sci.* **2012**, *67*, 1037–1044.

[7] M. Braun, B. Sun, B. Anselment, D. Weuster-Botz, *Appl. Microbiol. Biotechnol.* **2012**, *95*, 1457–68.

[8] F. Tonin, N. Alvarenga, J. Z. Ye, I. W. C. E. Arends, U. Hanefeld, *Adv. Synth. Catal.* **2019**, *361*, 2448–2455.

[9] S. Shi, Z. You, K. Zhou, Q. Chen, J. Pan, X. Qian, J. Xu, C. Li, *Adv. Synth. Catal.* **2019**, *361*, 4661–4668.

[10] Z. N. You, Q. Chen, S. C. Shi, M. M. Zheng, J. Pan, X. L. Qian, C. X. Li, J. H. Xu, *ACS Catal.* **2019**, *9*, 466–473.

[11] D. Monti, E. E. Ferrandi, L. Zanellato, L. Hua, F. Polentini, G. Carrea, S. Riva, *Adv. Synth. Catal.* **2009**, *351*, 1303–1311.

[12] U. T. Bornscheuer, R. J. Kazlauskas, in *Enzym. Catal. Org. Synth. Third Ed.*, **2012**, pp. 1695–1733.

[13] U. T. Bornscheuer, R. J. Kazlauskas, *Angew. Chemie - Int. Ed.* **2004**, *43*, 6032–6040.

[14] K. Hult, P. Berglund, *Trends Biotechnol.* **2007**, *25*, 231–238.

[15] S. Butt, H. G. Davies, M. J. Dawson, G. C. Lawrence, J. Leaver, S. M. Roberts, M. K. Turner, B. J. Wakefield, W. F. Wall, J. A. Winders, *J. Chem. Soc. Perkin Trans. 1* **1987**, 903–907.

[16] U. C. Oppermann, E. Maser, *Eur. J. Biochem.* **1996**, *241*, 744–749.

[17] E. Maser, G. Bannenberg, *Biochem. Pharmacol.* **1994**, *47*, 1805–1812.

[18] M. Zorko, H. E. Gottlieb, M. Žakelj-Mavric, *Steroids* **2000**, *65*, 46–53.

[19] M. De Amici, C. De Micheli, G. Molteni, D. Pitre, G. Carrea, S. Riva, S. Spezia, L. Zetta, *J. Org. Chem.* **1991**, *56*, 67–72.

[20] D. Zhu, J. E. Stearns, M. Ramirez, L. Hua, *Tetrahedron* **2006**, *62*, 4535–4539.

[21] Y. Kallberg, U. Oppermann, B. Persson, *FEBS J.* **2010**, *277*, 2375–2386.

[22] T. Yoshimoto, H. Higashi, A. Kanatani, X. S. Lin, H. Nagai, H. Oyama, K. Kurazono, D. Tsuru, *J. Bacteriol.* **1991**, *173*, 2173–2179.

[23] L. Liu, A. Aigner, R. D. Schmid, *Appl. Microbiol. Biotechnol.* **2011**, *90*, 127–35.

[24] S. M. Mythen, S. Devendran, C. Méndez-García, I. Cann, J. M. Ridlon, *Appl. Environ. Microbiol.* **2018**, *84*, e02475-17.

[25] A. S. Devlin, M. A. Fischbach, *Nat. Chem. Biol.* **2015**, *11*, 685–690.

[26] X. Chen, Y. Cui, J. Feng, Y. Wang, X. Liu, Q. Wu, D. Zhu, Y. Ma, *Adv. Synth. Catal.* **2019**, *361*, 2497–2504.

[27] J. P. Coleman, L. L. Hudson, M. J. Adams, *J. Bacteriol.* **1994**, *176*, 4865–4874.

[28] S. F. Baron, C. V. Franklund, P. B. Hylemon, *J. Bacteriol.* **1991**, *173*, 4558–4569.

[29] E. E. Ferrandi, G. M. Bertolesi, F. Polentini, A. Negri, S. Riva, D. Monti, *Appl. Microbiol. Biotechnol.* **2012**, *95*, 1221–1233.

[30] D. Bakonyi, W. Hummel, *Enzyme Microb. Technol.* **2017**, *99*, 16–24.

[31] M. M. Zheng, R. F. Wang, C. X. Li, J. H. Xu, *Process Biochem.* **2015**, *50*, 598–604.

[32] J. Y. Lee, H. Arai, Y. Nakamura, S. Fukiya, M. Wada, A. Yokota, *J. Lipid Res.* **2013**, *54*, 3062–3069.

[33] R. Wohlgemuth, J. Littlechild, D. Monti, K. Schnorr, T. van Rossum, B. Siebers, P. Menzel, I. V. Kublanov, A. G. Rike, G. Skretas, et al., *Biotechnol. Adv.* **2018**, *36*, 2077–2100.

[34] S. Patel, M. Saraf, in *Halophiles Biodivers. Sustain. Exploit.* (Eds.: D.K. Maheshwari, M. Saraf), Springer International Publishing, Cham, **2015**, pp. 403–419.

[35] T. Satyanarayana, Y. Kawarabayasi, J. Littlechild, *Thermophilic Microbes in Environmental and Industrial Biotechnology*, Springer Netherlands, Dordrecht, **2013**.

[36] J. A. Littlechild, *Front. Bioeng. Biotechnol.* **2015**, *3*, 1–9.

[37] E. E. Ferrandi, C. Sayer, M. N. Isupov, C.

Annovazzi, C. Marchesi, G. Iacobone, X. Peng, E. Bonch-Osmolovskaya, R. Wohlgemuth, J. A. Littlechild, et al., *FEBS J.* **2015**, *282*, 2879–2894.

[38] E. E. Ferrandi, C. Sayer, S. A. De Rose, E. Guazzelli, C. Marchesi, V. Saneei, M. N. Isupov, J. A. Littlechild, D. Monti, *Front. Bioeng. Biotechnol.* **2018**, *6*, 1–16.

[39] P. Hirsch, C. A. Gallikowski, J. Siebert, K. Peissl, R. Kroppenstedt, P. Schumann, E. Stackebrandt, R. Anderson, *Syst. Appl. Microbiol.* **2004**, *27*, 636–645.

[40] Y. M. Lee, E. H. Kim, H. K. Lee, S. G. Hong, *World J. Microbiol. Biotechnol.* **2014**, *30*, 2711–2721.

[41] N. Tamasawa, M. Yoneda, I. Makino, K. Takebe, S. Ueda, H. Misaki, *Gastroenterol. Jpn.* **1988**, *23*, 646–651.

[42] T. Oikawa, T. Kazuoka, K. Soda, *J. Mol. Catal. B Enzym.* **2003**, *23*, 65–70.

[43] H. R. Novak, C. Sayer, J. Panning, J. A. Littlechild, *Mar. Biotechnol.* **2013**, *15*, 695–705.

[44] L. Chrast, K. Tratsiak, J. Planas-Iglesias, L. Daniel, T. Prudnikova, J. Brezovsky, D. Bednar, I. Kuta Smatanova, R. Chaloupkova, J. Damborsky, *Microorganisms* **2019**, *7*, 498.

[45] O. Kirk, M. W. Christensen, *Org. Process Res. Dev.* **2002**, *6*, 446–451.

[46] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, et al., *Mol. Syst. Biol.* **2011**, *7*, 539.

[47] I. Letunic, P. Bork, *Nucleic Acids Res.* **2011**, *39*, 475–478.

[48] C. Mellon, R. Aspiotis, C. W. Black, C. I. Bayly, E. L. Grimm, A. Giroux, Y. Han, E. Isabel, D. J. McKay, D. W. Nicholson, et al., *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3886–3890.

[49] D. Zhu, Y. Yang, L. Hua, *J. Org. Chem.* **2006**, *71*, 4202–4205.

[50] F. Fernandez, D. N. Kirk, M. Scopes, *J. Chem. Soc. Perkin Trans. 1* **1974**, 18–21.

[51] S. Savino, E. E. Ferrandi, F. Forneris, S. Rovida, S. Riva, D. Monti, A. Mattevi, *Proteins Struct. Funct. Bioinforma.* **2016**, *84*, 859–865.

[52] J. A. Littlechild, *J. Ind. Microbiol. Biotechnol.* **2017**, *44*, 711–720.

[53] S. A. Serapian, M. W. Van Der Kamp, *ACS Catal.* **2019**, *9*, 2381–2394.

[54] P. Menzel, S. R. Gudbergsdóttir, A. G. Rike, L. Lin, Q. Zhang, P. Contursi, M. Moracci, J. K. Kristjansson, B. Bolduc, S. Gavrilov, et al., *Microb. Ecol.* **2015**, *70*, 411–424.

[55] S. M. Kielbasa, R. Wan, K. Sato, M. C. Frith, P. Horton, *Genome Res.* **2011**, *21*, 487–493.

[56] D. E. Ward, C. K. Rhee, *Can. J. Chem.* **1989**, *67*, 1206–1211.

10.1002/adsc.202000120

FULL PAPER

Insights into the Substrate Promiscuity of Novel Hydroxysteroid Dehydrogenases

Adv. Synth. Catal. **Year**, *Volume*, Page – Page

Susanna Bertuletti, Erica Elisa Ferrandi, Stefano Marzorati, Marta Vanoni, Sergio Riva,* Daniela Monti*