High Throughput Discovery of Alcohol Dehydrogenases for Industrial Biocatalysis

Gilles Ravot, Denis Wahler, Olivier Favre-Bulle, Valérie Cilia, Fabrice Lefevre*

Proteus SA, Parc Scientifique George Besse, 70 allée Graham Bell, 30000 Nîmes, France
Phone: (+33)-4-66-70-64-64, Fax: (+33)-4-66-70-64-60, e-mail: flefevre@proteus.fr

Received: January 24, 2003; Accepted: April 11, 2003

Abstract: To speed up the discovery of novel biocatalysts for industrial synthesis, an efficient screening approach has been set up. The huge potential of natural biodiversity has been taken advantage of in a discovery protocol using bioinformatics analysis, polymerase chain reaction (PCR) techniques and direct *in vitro* expression of enzymes. This approach led to the direct characterization of 19 novel thermoactive alcohol dehydrogenases with a broad substrate range, that set up a useful enzymatic "toolbox" for fine chemistry synthesis.

Keywords: alcohol dehydrogenase; biocatalysis; biodiversity; enzyme screening; gene expression; high throughput screening; ketones; oxidation

The production of enantiopure chiral intermediates is a key step in the synthesis of optically active drugs and agrochemicals, and, as enzymes can display very high stereoselectivities, the use of biotechnological processes has been developed rapidly in industrial organic synthesis.[1]

Among several groups of chiral building blocks, optically active hydroxy compounds represent one of the most important classes of synthons, and the difficulty to prepare them with classical chemistry has long led the manufacturers to consider oxidoreductases to prepare various useful asymmetric alcohols from prochiral ketones.[2] However, an important limitation to obtain the right biocatalyst is due to the fact that industrial substrates are usually artificial, and that in general no suitable enzyme is known. Therefore the best way to find dedicated enzymes for the asymmetric reduction of prochiral carbonyl compounds is to screen large number of microorganisms or mutant enzymes.[3]

A first solution is the exploitation of whole cells of microorganisms, which possess a high occurrence of alcohol dehydrogenases with differing substrate specificity and stereoselectivity.[4] But each candidate enzyme has to be independently purified and characterized for an industrial application, leading to very time-consuming procedures.

An efficient way to isolate these enzymes is to clone genes encoding for known alcohol dehydrogenases in heterologous laboratory hosts.[5] Nevertheless, expression cloning techniques in heterologous hosts are also very time-consuming, and this approach does not always lead to active and properly folded alcohol dehydrogenases.[6]

Here we report on a new method to rapidly detect and characterize desired or novel alcohol dehydrogenase activities in a high-throughput format, without the need for enzyme purification.

To discover novel enzymes, one way is to screen microorganisms originating from diverse environments using high-throughput screening (HTS) technologies, and a second way is to adapt existing enzymes by shuffling the corresponding genes.[7] The key point in the discovery of these enzymes is to be compatible with the time scales of industry. To fit with these timelines, we are using these two approaches, but when some sequences are available, we can also combine the extensive potential of natural biodiversity with bio-informatics and the use of our *in vitro* expression technology called Phenomics™ (see Figure 1).[8]

This technology is fully acellular. Amplified genomic DNA fragments are isolated and placed *in vitro* under the control of specific transcription controlling sequences, such as T7 transcription promoter. This approach therefore escapes all technical pitfalls linked to regulation of gene transcription. *In vitro* transcribed mRNAs are then translated into proteins using *in vitro* expression extracts. These extracts are selected to optimize gene expression according to the source of genomic DNA. These extracts allow a perfect match between the specific requirement linked to the origin of the DNA and the acellular translation machinery used. This eliminates all identified obstacles that interfere with the compatibility between expression vectors and DNA fragments on the one hand and cellular host selected for expression on the other. Standardized *in vitro* expression extracts enable us to translate extremophilic genes from a simple polymerase chain reaction (PCR) product with high efficiency and to achieve protein concentrations in the range of 30 to 300 μg/mL.

We used this approach to discover and characterize a pool of new thermoactive alcohol dehydrogenases

 DOI: 10.1002/adsc.200303019

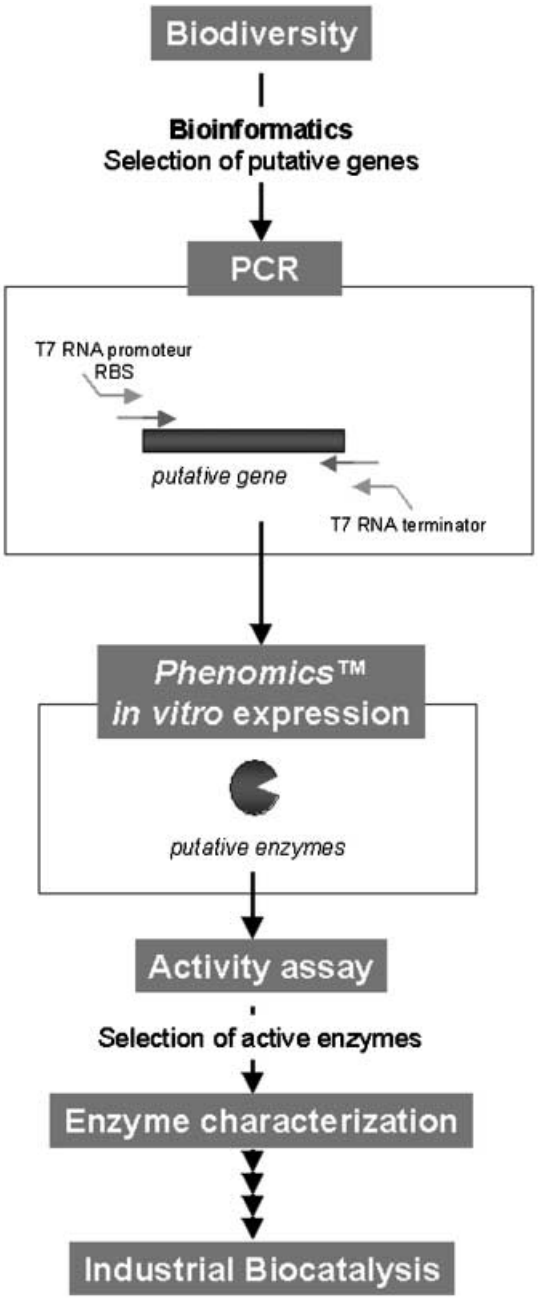

Figure 1. Discovery procedure.

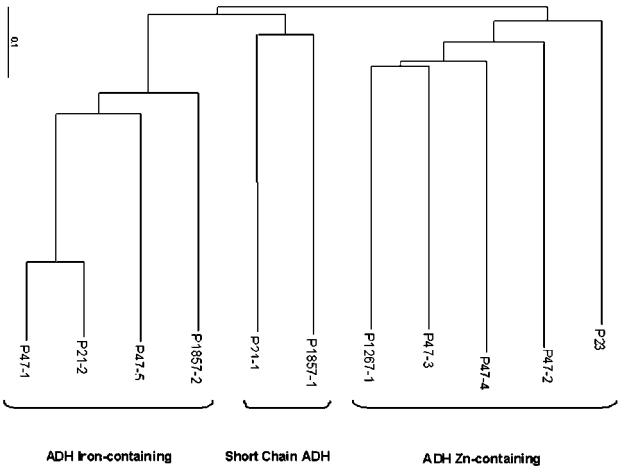

Figure 2. Phylogenetic positions of some of the selected ADH. Dendrograms were constructed using the neighbor-joining method[9] available on PhyloWin program.[11]

(ADH, also called ketone reductases). Eighteen putative alcohol dehydrogenase genes were selected by processing information amongst sequences that have been published in databases (see Table 1).

These sequences have been selected on the basis of the growth temperature of the corresponding microorganisms (to retrieve thermophilic enzymes).The maximum level of identity between two selected sequences is 57% (see Figure 2).

Corresponding microorganisms were cultivated from our culture collection, and genomic DNAs were prepared and purified. For each ADH gene, a specific set of primers was designed. These primers were used in PCR reactions to amplify each gene. Each PCR product was used in a second PCR reaction to introduce a T7 RNA polymerase promoter, a ribosome binding site at the 5′-end of each gene, and a T7 RNA polymerase terminator at the 3′-end of each gene (the second set of primers has been designed on the basis of the sequences of T7 promoter, RBS and T7 terminator of pET26b+ vector (Novagen).

Then each product of this second PCR reaction was purified and incubated in a transcription/translation reaction mix in order to express the corresponding putative ADH. After a short denaturation step at 80 °C for 30 seconds of the transcription/translation mix (to inactivate potential endogenous oxidoreductases that are present in the translation extracts, and then to lower the potential background), the enzymatic activity of each putative ADH was tested in an oxidation reaction (alcohol dehydrogenation assay). This assay is based on a formazan precipitation reaction as described by Mayer et al.,[10] with a mixture of NAD and NADP as cofactors, and a pool of seven different alcohols (primary and secondary alcohols with short or long chains) as substrates.

Reaction mixture with an active ADH becomes dark blue within a few minutes at 80 °C, while the negative control (without enzyme) stays clear.

The activity test at 80 °C of the 18 putative ADHs was achieved, and revealed the presence of active ADH in each reaction, demonstrating that all of these putative genes do encode for thermoactive ADHs.

In order to qualify these ADHs for their potential use in industrial biotransformations, we further characterized cofactor dependence and substrate range of each ADH. First, an activity test was assayed with either NAD or NADP and a pool of seven different alcohols as substrates. Activity was monitored following UV ab-

Table 1. Microbial origin of the selected putative ADH genes.

ADH putative gene number	Strain number	Phylogenetic position	Optimal growth temperature [°C]	Environment
P21–1	P21	*Pyrococcus* sp.	100	Heated marine sediments (Vulcano, Italy)
P21–2	P21	*Pyrococcus* sp.	100	Heated marine sediments (Vulcano, Italy)
P23-T	P23	*Thermoanaerobacter* sp.	65–70	Thermal spring (Yellowstone National Park, USA)
P23-ten	P23	*Thermoanaerobacter* sp.	65–70	Thermal spring (Yellowstone National Park, USA)
P47–1	P47	*Thermotoga* sp.	80	Geothermal heated seafloor (Vulcano, Italy)
P47–2	P47	*Thermotoga* sp.	80	Geothermal heated seafloor (Vulcano, Italy)
P47–3	P47	*Thermotoga* sp.	80	Geothermal heated seafloor (Vulcano, Italy)
P47–4	P47	*Thermotoga* sp.	80	Geothermal heated seafloor (Vulcano, Italy)
P47–5	P47	*Thermotoga* sp.	80	Geothermal heated seafloor (Vulcano, Italy)
P48–5	P48	*Thermotoga* sp.	80	Shallow submarine hot spring (Naples, Italy)
P376–1	P376	*Thermotoga* sp.	80	Oil well
P376–2	P376	*Thermotoga* sp.	80	Oil well
P376–3	P376	*Thermotoga* sp.	80	Oil well
P376–5	P376	*Thermotoga* sp.	80	Oil well
P1267–1	P1267	*Pyrococcus* sp.	98	Marine hydrothermal vent (Pacific Ocean)
P1267–2	P1267	*Pyrococcus* sp.	98	Marine hydrothermal vent (Pacific Ocean)
P1857–1	P1857	*Pyrococcus* sp.	96	Hot fluid from an active chimney (Pacific Ocean)
P1857–2	P1857	*Pyrococcus* sp.	96	Hot fluid from an active chimney (Pacific Ocean)

Table 2. Characteristics of the novel alcohol dehydrogenases.

ADH gene	Cofactors[a]	Volumetric activity on various substrates (nmol min^{-1} μL^{-1})[b]					
		Methanol	Ethanol	2-Propanol	Butanol	Pentanol	Decanol
P21–1	NADP > NAD	5.25	6.86	5.83	5.30	ND	4.50
P21–2	**NADP**	**ND**	**1.88**	**0.49**	**2.18**	**ND**	**0.19**
P23-T	NAD=NADP	2.09	2.02	2.15	5.09	ND	0.58
P23-ten	NADP > NAD	1.08	1.21	0.90	0.61	ND	1.64
P47–1	NADP ≫ NAD	1.14	ND	1.24	2.12	ND	1.34
P47–2	NAD > NADP	3.54	3.54	4.66	4.24	ND	4.77
P47–3	NAD=NADP	5.83	4.88	3.60	5.30	ND	4.56
P47–4	NADP > NAD	0.50	0.85	0.59	0.63	ND	0.77
P47–5	NAD > NADP	4.45	4.98	4.24	7.31	ND	4.17
P48–5	**NADP**	**2.45**	**2.52**	**0.10**	**0**	**ND**	**0**
P376–1	**NAD=NADP**	**6.78**	**5.77**	**5.41**	**6.86**	**6.36**	**5.09**
P376–2	NADP#NAD	3.46	2.76	3.60	5.51	ND	5.51
P376–3	NADP > NAD	0.71	0.60	0.83	0.81	ND	1.04
P376–5	NADP > NAD	0.24	0.18	0.28	0.25	ND	0.54
P1267–1	NAD ≫ NADP	4.35	5.19	4.03	0.36	ND	1.02
P1267–2	**NADP**	**0.72**	**0.18**	**0.16**	**6.10**	**ND**	**0.77**
P1857–1	NADP > NAD	3.96	2.91	4.44	4.85	ND	3.34
P1857–2	NAD > NADP	3.39	2.49	3.47	3.47	ND	3.60

[a] Cofactor used : preferred cofactor.
[b] Volumetric activity: nmol of alcohol oxidized per min at 70 °C per μL of translation mix (mU/μL).

sorbance increase at 340 nm, and the results of cofactor dependence is reported in Table 2. In a second step, to set up a rapid characterization of substrate profile of these novel enzymes, every ADH was incubated with its preferred cofactor in a reaction containing only one alcohol. As can be seen in Table 2, these alcohol dehydrogenases show various substrate specificities depending on chain length, and on substitution degree of the alcohol function (primary or secondary alcohols).

ADH P48–5 seems to be highly specific of very short chain alcohols, while P1267–2 has a narrow range of substrates, butanol being the preferred one. On the other hand, other ADHs such as P376–1, P1857–1, P1857–2 have a broader substrate range and could be used in several different biotransformations. These data provide basic information for the selection of a suitable ADH in a specific synthesis, and further substrate profiles will be run with industrial substrates.

Biocatalysis has been seen as an area with great promise for chemical synthesis but industrial applications are modest. The success of biocatalysis depends ultimately on the economics of specific processes and among other things on the rapid discovery of highly efficient biocatalysts. Among biocatalysts, alcohol dehydrogenases (ADH) are of particular interest since these enzymes catalyze the reduction of a ketone into an enantiomerically pure alcohol. As this paper shows, the use of proper technologies and strategies allows the rapid screening of the natural biodiversity that results in the isolation of 18 ADH having a broad substrate range in less than a week. This set of enzymes can be used as a "toolbox" to rapidly screen for activity on substrates of industrial interest. This method of direct expression of putative genes without purification can be applied for other catalytic activities such as dioxygenases, esterases, epoxide hydrolases, nitrilases or amidases.

In some cases, there are no natural enzymes to synthesize industrial compounds and the performance of the selected ADH could not fulfill all the industrial requirements. Then, *in vitro* directed evolution techniques, like L-Shuffling™,[7] provides the means to "breed" and optimize new, non-natural enzymes. Therefore, the availability of 18 enzymes with a broad substrate range, allows us to tailor the specificity and the activity of alcohol dehydrogenases and by this way, to fulfill all the specifications of the desired process.

Experimental Section

Bioinformatics

Protein sequences data were imported into the sequence editor SeaView.[11] The full sequences were aligned using the ClustalW program available on EBI site. The alignment sequence was then adjusted manually.

Gene Expression

Genomic DNAs have been extracted and purified using QiaAmp DNA kit following the manufacturer's recommendations. PCR reaction was realized in Perkin Elmer 9700 apparatus as described in Mc Pherson et al.[12] *In vitro* expression was achieved as described in Patent WO 00/09747[8] with slight modifications.

Enzyme Assay

Five μL of translated proteins were incubated for a few minutes at 80 °C with 30 μL of buffer A (Tris/HCl 50 mM, pH 8.0, 0.13% gelatin), 20 μL of an isomix of methanol, ethanol, 2-propanol, butanol, decanol, benzyl alcohol and octanol, 5 μL of 20 mM NAD, 5 μL of 20 mM NADP and 140 μL of buffer B (30 μM of methosulfate phenazine and 300 μM of nitroblue tetrazolium).

Enzyme Characterization

Cofactor specificity and substrate profiles were determined by incubating at 70 °C for a few minutes 5 μL (or 10 μL for P376 – 3 and P 376 – 5 enzymes) of translated ADH with 250 μL of buffer C (100 mM glycine, 0.9 M NaCl pH 10.0), 5μL of 100 mM cofactor (NAD or NADP), 20 μL of alcohol substrate and 250 μL of water (or 245 μL for P376 – 3 and P 376 – 5 enzymes). The reaction was monitored at 340 nm ($\varepsilon_{340} = 6.25\ mM^{-1}\ cm^{-1}$) with a Uvikon 922 (Kontron instrument) recording spectrophotometer. One unit of enzyme represents the amount of enzyme that catalyzes the oxidation of 1 μmol of alcohol per minute at 70 °C.

Acknowledgements

We thank V. Duvivier, M. Mazel, and M. Gilly for their useful contributions to this work.

References

[1] a) A. J. J. Straathof, S. Panke, A. Schmid, *Curr. Opin. Biotechnol.* **2002**, *13*, 548; b) C. R. Johnson, G. W. Wells, *Curr. Opin. Chem. Biol.* **1998**, *2*, 70; c) A. Liese, M. V. Filho, *Curr. Opin. Biotechnol.* **1999**, *10*, 595; d) B. G. Davis, V. Boyer, *Nat. Prod. Rep.* **2001**, *18*, 618; e) O. Kirk, T. V. Borchert, C. C. Fuglsang, *Curr. Opin. Biotechnol.* **2002**, *13*, 345; f) J. B. van Beilen, Z. Li, *Curr. Opin. Biotechnol.* **2002**, *13*, 338.
[2] J. D. Stewart, *Curr. Opin. Chem. Biol.* **2001**, *5*, 120.
[3] J. Ogawa, S. Shimizu, *Curr. Opin. Biotechnol.* **2002**, *13*, 367.
[4] J. D. Stewart, *Curr. Opin. Biotechnol.* **2000**, *11*, 363.
[5] S. Ammendola, C. A. Raia, C. Caruso, L. Camardella, S. D'Auria, M. De Rosa, M. Rossi, *Biochemistry* **1992**, *31*, 12514.
[6] J. van der Oost, W. G. Voorhorst, S. W. Kengen, A. C. Geerling, V. Wittenhorst, Y. Gueguen, W. M. de Vos, *Eur J Biochem.* **2001**, 268, 3062.
[7] *Patent* WO00/09679.
[8] a) *Patent* WO 00/09747; b) *Patent* WO00/34512; c) *U. S. Patent* 6,514,703.
[9] N. Saitou, M. Nei, *Molecular Biology and Evolution* **1987**, *4*, 406.
[10] K. M. Meyer, F. H. Arnold, *Journal of Biomolecular Screening* **2002**, *7*, 135.
[11] N. Galtier, M. Gouy, C. Gautier, *Comput. Applic. Biosci.* **1996**, *12*, 543.
[12] M. J. Mc Pherson, P. Quirke, G. R. Taylor, in *PCR – A practical approach*, Vol. 1, (Eds.: D. Rickwood, B. D. Hames), IRL Press, Oxford, **1991**.

