(86) Date de dépôt PCT/PCT Filing Date: 2008/06/16
(87) Date publication PCT/PCT Publication Date: 2008/12/24
(45) Date de délivrance/Issue Date: 2016/06/07
(85) Entrée phase nationale/National Entry: 2009/12/17
(86) N° demande PCT/PCT Application No.: EP 2008/057522
(87) N° publication PCT/PCT Publication No.: 2008/155302
(30) Priorité/Priority: 2007/06/20 (EP07110670.2)

(54) Titre : PROCEDE DE PRODUCTION D'ALCOOLS OPTIQUEMENT ACTIFS A L'AIDE D'UNE DESHYDROGENASE TIREE DE AZOARCUS SP. EbN1

(54) Title: METHOD FOR PRODUCING OPTICALLY ACTIVE ALCOHOL USING AZOARCUS SP EbN1 DEHYDROGENASE

\[
\begin{align*}
la & : \text{R}_2 & & \text{R}_1 \\
R & : & \text{alkyl, alkenyl, aryl, or alkylaryl groups,} & \\
& & \text{which in turn can be mono-substituted or polysubstituted with alkyl, halogen, SH, SR, OH OR, NO, CN, CO, COOR, NR, or} \\
\text{R}_2 & & \text{R}_3 \\
\end{align*}
\]

(57) Abrégé/Abstract:
A method for producing optically active alcohols of formula (la) or (lb), wherein R, R are alkyl, alkenyl, aryl, or alkylaryl groups, which in turn can be mono-substituted or polysubstituted with alkyl, halogen, SH, SR, OH OR, NO, CN, CO, COOR, NR, or R3.
(57) **Abrégé(suite)/Abstract(continued):**

NR²R³R⁴X, wherein R², R³ and R⁴ independently of each other are H or a low-alkyl or low-alkoxy radical, and X is a counterion, with the stipulation that R¹ is not identical to R², by reduction of the corresponding ketone, the reduction being performed with a dehydrogenase having the polypeptide sequence SEQ ID NO:2 or NO:4, or having a polypeptide sequence in which up to 25% of the amino acid radicals are modified relative to SEQ ID NO:2 or NO:4 by deletion, insertion, substitution, or a combination thereof.
(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum

(51) Internationale Patentklassifikation:
C12P 7/02 (2006.01)

(21) Internationales Anmeldenummer:
PCT/EP2008/057522

(22) Internationales Anmeldedatum:

(25) Einreichungssprache:
Deutsch

(26) Veröffentlichungssprache:
Deutsch

(30) Angaben zur Priorität:
07116760.2


(72) Erfinder; und


(74) Gemeinsamer Vertreter: BASF SE; Gvx/b - C 6, Carl-Bosch-Str. 38, 67056 Ludwigshafen (DE).

(54) Titel: METHOD FOR PRODUCING OPTICALLY ACTIVE ALCOHOLS USING AN AZOARCUS SP. EBNI DEHYDROGENASE

(54) Bezeichnung: VERFAHREN ZUR HERSTELLUNG OPTISCH AKTIVER ALKOHOLEN UNTER VERWENDUNG EINER AZOARCUS SP. EBNI DEHYDROGENASE

(57) Abstract: A method for producing optically active alcohols of formula (Ia) or (Ib), wherein R₁, R₂ are alkyl, alkenyl, aryl, or alkylaryl groups, which in turn can be mono-substituted or poly-substituted with alkyl, halogen, SH, SR, OR, OH, OR₂, NO₂, CN, CO, COOR, NR₂ or NR₂ R X', wherein R₁, R₂ and X independently of each other are H or a low-alkyl or low-alkoxy radical, and X' is a counterion, with the stipulation that R₁ is not identical to R₂, by reduction of the corresponding ketone, the reduction being performed with a dehydrogenase having the polypeptide sequence SEQ ID NO:2 or NO:4, or having a polypeptide sequence in which up to 25% of the amino acid radicals are modified relative to SEQ ID NO:2 or NO:4 by deletion, insertion, substitution, or a combination thereof.

(57) Zusammenfassung: Verfahren zur Herstellung von optisch aktiven Alkoholen der Formel (Ia) oder (Ib) worin R₁, R₂ für Alkyl, Alkenyl, Aryl oder Alkylarylgruppen stehen, die wiederum ein- oder mehrfach substituiert sein können, durch Alkyl, Halogen, SH, SR, OH, OR, OR₂, NO₂, CN, CO, COOR, NR₂ oder NR₂ R X', wobei R₁, R₂ und X unabhängig voneinander für H oder einen Niedrigalkyl- oder niedrigalkoxy-Rest stehen und X für ein Gegenion steht, mit der Maßgabe, dass R₁ ungleich R₂ ist, durch Reduktion des entsprechenden Ketons, wobei die Reduktion mit einer Dehydrogenase mit der Polypeptiddsequenz SEQ ID NO:2 oder NO:4, oder mit einer Polypeptiddsequenz, bei der bis zu 25% der Aminosäurereste gegenüber SEQ ID NO:2 oder NO:4 durch Deletion, Inversion, Substitution oder einer Kombination davon verändert sind, durchgeführt wird.
Method for producing optically active alcohol using Azoarcus Sp EbN1 dehydrogenase

The present invention relates to a process for preparing optically active alcohols of the formula 1a or 1b

\[
\begin{align*}
\text{OH} & \quad \text{OH} \\
\text{R}^2 & \quad \text{R}^2 \\
\text{R}^1 & \quad \text{R}^1 \\
\text{R}^2 & \quad \text{R}^1 \\
\end{align*}
\]

Formula 1a       Formula 1b


Thus, the biocatalysts described for preparing optically active phenylethanol derivatives are chiefly those exhibiting 'Prelog’ selectivity, enzymes exhibiting the contrary enantioselectivity are rarer, although not unknown [Trends Biotechnol (1999) 17:487, J. Org. Chem. (1992) 57:1532].

The present invention relates to a process for preparing optically active alcohols of the formula 1a or 1b

\[
\begin{align*}
\text{OH} & \quad \text{OH} \\
\text{R}^2 & \quad \text{R}^2 \\
\text{R}^1 & \quad \text{R}^1 \\
\text{R}^2 & \quad \text{R}^1 \\
\end{align*}
\]

Formula 1a       Formula 1b

in which
R<sup>1</sup>, R<sup>2</sup> are alkyl, alkenyl, aryl, or alkylaryl groups which may in turn be substituted one or more times by alkyl, halogen, SH, SR<sup>2</sup>, OH, OR<sup>2</sup>, NO<sub>2</sub>, CN, CO, COOR<sup>2</sup>, NR<sup>2</sup>R<sup>3</sup> or NR<sup>2</sup>R<sup>3</sup>R<sup>4+X</sup>, where R<sup>2</sup>, R<sup>3</sup> and R<sup>4</sup> are independently of one another H or a lower alkyl or lower alkoxy radical, and X<sup>-</sup> is a counter ion, with the proviso that R<sup>1</sup> is not equal to R<sup>2</sup>, by reducing the corresponding ketone, where the reduction is carried out with a dehydrogenase having the polypeptide sequence of SEQ ID NO: 2 or NO: 4, or with a polypeptide sequence in which up to 25% of the amino acid residues are altered by comparison with SEQ ID NO: 2 or NO: 4 by deletion, insertion, substitution or a combination thereof.

The present invention relates to a process for preparing an optically active alcohol of formula la,

![Formula la](image)

in which R<sup>1</sup> and R<sup>2</sup> independently are alkyl, alkenyl, aryl, or alkylaryl groups, each of which may be substituted one or more times by alkyl, halogen, SH, SR<sup>2</sup>, OH, OR<sup>2</sup>, NO<sub>2</sub>, CN, CO, COOR<sup>2</sup>, NR<sup>2</sup>R<sup>3</sup> or NR<sup>2</sup>R<sup>3</sup>R<sup>4+X</sup>, where R<sup>2</sup>, R<sup>3</sup> and R<sup>4</sup> independently are H or a lower alkyl or lower alkoxy radical, and X<sup>-</sup> is a counter ion, with the proviso that R<sup>1</sup> is not equal to R<sup>2</sup> and R<sup>1</sup> is less bulky than R<sup>2</sup>, by reducing a corresponding ketone with a dehydrogenase having the polypeptide sequence of SEQ ID NO: 2 or having a polypeptide sequence having at least 95% sequence identity to the polypeptide sequence of SEQ ID NO: 2.
A particularly good embodiment of the invention consists of a process for preparing optically active alcohols of the formula Ia or Ib in which $R^1$ is C1-C10-alkyl and $R^2$ is phenyl, where the radicals $R_1$ and/or $R_2$ are optionally monosubstituted by halogen.

The present invention relates in particular to a process for preparing optically active alcohols of the formula Ia, where the radical $R^1$ is less bulky than $R^2$.

$$\text{Formula Ia}$$

If the $R^2$ radical is more bulky than the $R^1$ radical, the alcohol is allocated, in accordance with Prelog, V., Pure and Applied Chemistry, (1964), 6, 119-130, to the 'anti'-Prelog category.

Chiral alcohols can be differentiated on the basis of their configuration into so-called 'Prelog' and 'anti-Prelog' enantiomers. The assignment to one of the two categories takes place according to the size (bulk) of the two groups which are adjacent to the alcohol group, and the alignment of the hydroxy function in relation to these two groups. Optically active alcohols with 'anti-Prelog' configuration are important precursors for various active ingredients.
General terms and definitions

Unless indicated otherwise, the following general meanings apply:

"Halogen" stands for fluorine, chlorine, bromine, or iodine, in particular fluorine or chlorine.

"Lower alkyl" stands for straight-chain or branched alkyl radicals having 1 to 6 C atoms, such as methyl, ethyl, i- or n-propyl, n-, i-, sec- or tert-butyl, n-pentyl or 2-methylbutyl, n-hexyl, 2-methylpentyl, 3-methylpentyl, 2-ethylbutyl.

"Lower alkenyl" stands for the mono- or polyunsaturated, preferably mono- or diunsaturated, analogs of the abovementioned alkyl radicals having 2 to 6 carbon atoms, where the double bond may be present in any position of the carbon chain.

"Lower alkoxy" stands for the oxygen-terminated analogs of the above alkyl radicals.

"Aryl" stands for a mono- or polynuclear, preferably mono- or dinuclear, optionally substituted aromatic radical, in particular for phenyl or for a naphthyl bonded via any ring position, such as 1- or 2-naphthyl. These aryl radicals may optionally have 1 or 2 identical or different substituents, for example halogen, lower alkyl, lower alkoxy as defined above, or trifluoromethyl.

"Enantioselectivity" in the context of the present invention means that the enantiomeric excess ee (in %) of one of the two possible enantiomers is at least 50\%, preferably at least 80\%, in particular at least 90\% and specifically at least 95\%. The ee is calculated as follows:

\[
\text{ee (\%)} = \frac{\text{enantiomer A} - \text{enantiomer B}}{\text{enantiomer A} + \text{enantiomer B}} \times 100
\]

Biochemical embodiments

Particularly suitable dehydrogenases (EC 1.1.X.X) are especially NAD- or NADP-dependent dehydrogenases (E.C. 1.1.1.1), in particular alcohol dehydrogenases (E.C.1.1.1.1 or E.C.1.1.1.2) which bring about selective reduction of the ketone to the 'anti-Prelog' alcohol. The dehydrogenase is preferably obtained from a microorganism, particularly preferably from a bacterium, a fungus, in particular a yeast, in each case deposited in collections of
strains or obtainable from isolates of natural sources, such as soil samples, biomass samples and the like or by \textit{de novo}-gene synthesis.

The dehydrogenase can be used in purified or partially purified form or in the form of the original microorganism or of a recombinant host organism which expresses the dehydrogenase. Processes for obtaining and purifying dehydrogenases from microorganisms are sufficiently well known to the skilled worker, e.g. from K. Nakamura & T. Matsuda, "Reduction of Ketones" in K. Drauz and H. Waldmann, Enzyme Catalysis in Organic Synthesis 2002, Vol.III, 991-1032, Wiley-VCH, Weinheim. Recombinant processes for generating dehydrogenases are likewise known, for example from W. Hummel, K. Abokitse, K. Drauz, C. Rollmann and H. Gröger, Adv. Synth. Catal. 2003, 345, No. 1 + 2, pp. 153-159.

Suitable bacteria are for example those of the orders of \textit{Burkholderiales}, \textit{Hydrogenophilales}, \textit{Methylphilales}, \textit{Neisseriales}, \textit{Nitrosomonadales}, \textit{Procabacteriales} or \textit{Rhodocycales}.

Particularly preferred dehydrogenases are those from the family of \textit{Rhodocyclaceae}.

Particularly preferred dehydrogenases are from the genera \textit{Azoarcus} \textit{Azoneus}, \textit{Azospira}, \textit{Azovibrio}, \textit{Dechloromonas}, \textit{Ferribacterium}, \textit{Petrobacter}, \textit{Propionivibrio}, \textit{Quadricoccus}, \textit{Rhodocyclus}, \textit{Sterolibacterium}, \textit{Thauera} and \textit{Zoogloea}.

Especially preferred dehydrogenases are from species of the genus \textit{Azoarcus}.

The reduction with the dehydrogenase normally takes place in the presence of a suitable cofactor (also referred to as cosubstrate). The cofactor normally used for reducing the ketone is NADH and/or NADPH. It is possible besides to employ dehydrogenases as cellular systems which intrinsically comprise cofactor, or alternative redox mediators can be added (A. Schmidt, F. Hollmann and B. Bühler "Oxidation of Alcohols" in K. Drauz and H. Waldmann, Enzyme Catalysis in Organic Synthesis 2002, Vol.III, 991-1032, Wiley-VCH, Weinheim).

The reduction with the dehydrogenase normally additionally takes place in the presence of a suitable reducing agent which regenerates the cofactor oxidized during the reduction.
Examples of suitable reducing agents are sugars, especially the hexoses such as glucose, mannose, fructose, and/or oxidizable alcohols, especially ethanol, propanol, butanol, pentanol or isopropanol, and formate, phosphite or molecular hydrogen. To oxidize the reducing agent and, connected therewith, to regenerate the coenzyme it is possible to add a second dehydrogenase such as, for example, glucose dehydrogenase when glucose is used as reducing agent, phosphite dehydrogenase when phosphite is used as reducing agent or formate dehydrogenase when formate is used as reducing agent. This dehydrogenase can be employed as free or immobilized enzyme or in the form of free or immobilized cells. Preparation thereof is possible either separately or by coexpression in a (recombinant) dehydrogenase strain.

The dehydrogenases used according to the invention can be employed free or immobilized. An immobilized enzyme means an enzyme which is fixed to an inert carrier. Suitable carrier materials, and the enzymes immobilized thereon, are disclosed in EP-A-1149849, EP-A-1069183 and DE-A 100193773, and the references cited therein. The disclosure of these publications in this regard is incorporated in its entirety herein by reference. Suitable carrier materials include for example clays, clay minerals such as kaolinite, diatomaceous earth, perlite, silicon dioxide, aluminum oxide, sodium carbonate, calcium carbonate, cellulose powder, anion exchanger materials, synthetic polymers such as polystyrene, acrylic resins, phenol-formaldehyde resins, polyurethanes and polyolefins such as polyethylene and polypropylene. The carrier materials are normally employed in a finely divided particulate form to prepare the carrier-bound enzymes, with preference for porous forms. The particle size of the carrier material is normally not more than 5 mm, in particular not more than 2 mm (grading curve). It is possible analogously to choose a free or immobilized form on use of the dehydrogenase as whole-cell catalyst. Examples of carrier materials are Ca alginate and carrageenan. Both enzymes and cells can also be crosslinked directly with glutaraldehyde (crosslinking to give CLEAs). Corresponding and further immobilization methods are described for example in J. Lalonde and A. Margolin “Immobilization of Enzymes” in K. Drauz and H. Waldmann, Enzyme Catalysis in Organic Synthesis 2002, Vol.III, 991-1032, Wiley-VCH, Weinheim.

The reaction can be carried out in aqueous or nonaqueous reaction media or in 2-phase systems or (micro)emulsions. The aqueous reaction media are preferably buffered solutions which ordinarily have a pH of from 4 to 8, preferably from 5 to 8. The aqueous solvent may,
besides water, additionally comprise at least one alcohol, e.g. ethanol or isopropanol, or dimethyl sulfoxide.

Nonaqueous reaction media mean reaction media which comprise less than 1% by weight, preferably less than 0.5% by weight, of water based on the total weight of the reaction medium. The reaction is preferably carried out in an organic solvent. Examples of suitable solvents are aliphatic hydrocarbons, preferably having 5 to 8 carbon atoms, such as pentane, cyclopentane, hexane, cyclohexane, heptane, octane or cyclooctane, halogenated aliphatic hydrocarbons, preferably having one or two carbon atoms, such as dichloromethane, chloroform, tetrachloromethane, dichloroethane or tetrachloroethane, aromatic hydrocarbons, such as benzene, toluene, the xylenes, chlorobenzene or dichlorobenzene, aliphatic acyclic and cyclic ethers or alcohols, preferably having 4 to 8 carbon atoms, such as diethyl ether, methyl tert-butyl ether, ethyl tert-butyl ether, dipropyl ether, diisopropyl ether, dibutyl ether, tetrahydrofuran or esters such as ethyl acetate or n-butyl acetate or ketones such as methyl isobutyl ketone or dioxane or mixtures thereof.

The reduction with the dehydrogenase is carried out for example in an aqueous-organic, in particular aqueous, reaction medium.

The ketone to be reduced is preferably employed in a concentration of from 0.1 g/l to 500 g/l, particularly preferably from 1 g/l to 50 g/l, in the enzymatic reduction and can be fed in continuously or discontinuously.

The enzymatic reduction ordinarily takes place at a reaction temperature below the deactivation temperature of the dehydrogenase employed and preferably at −10°C at least. It is particularly preferably in the range from 0 to 100°C, in particular from 15 to 60°C and specifically from 20 to 40°C, e.g. at about 30°C.

A possible procedure is for example to mix the ketone with the dehydrogenase, the solvent and, if desired, the coenzymes, if desired a second dehydrogenase to regenerate the coenzyme and/or further reducing agents, thoroughly, e.g. by stirring or shaking. However, it is also possible to immobilize the dehydrogenase(s) in a reactor, for example in a column, and to pass a mixture comprising the ketone and, if desired, coenzymes and/or cosubstrates through the reactor. For this purpose, the mixture can be circulated through the reactor until
the desired conversion is reached. In this case, the keto group of the ketone is reduced to an OH group, resulting in substantially one of the two enantiomers of the alcohol. The reduction is ordinarily managed until the conversion is at least 70%, particularly preferably at least 85% and especially at least 95%, based on the ketone present in the mixture. The progress of the reaction, i.e. the sequential reduction of the ketone, can in this case be followed by conventional methods such as gas chromatography or high-pressure liquid chromatography.

The dehydrogenases employed in the process of the invention are particularly preferably alcohol dehydrogenases having the following properties:

Alcohol dehydrogenase from Azoarcus having an amino acid sequence shown in SEQ ID 2 or SEQ ID 4, and alcohol dehydrogenases having amino acid sequences in which up to 25%, preferably up to 15%, particularly preferably up to 10, especially up to 5%, of the amino acid residues are altered by comparison with SEQ ID NO: 2 or NO: 4 by deletion, insertion, substitution or a combination thereof.

Oxidation of simple alcohols such as, for example, isopropanol, butan-2-ol, pentan-2-ol or cyclohexanol to the corresponding carbonyl with simultaneous reduction of NAD⁺ or NADP⁺.

Alcohol dehydrogenases which catalyze the reduction in an enantiomeric purity of at least 95% ee (in the presence of NADH and/or NADPH; at 30°C and pH 7.0).

The present invention further relates also to an 'anti-Prelog' dehydrogenase having at least one of the aforementioned properties.

The alcohol dehydrogenases exhibit activity in the presence of the following solvents: heptane, hexane, MtBE, n-butanol, butan-2-ol, n-pentanol, pentan-2-ol, pentan-3-ol, DMSO i-propanol, n-propanol, ethanol.

They preferably have a molecular weight in the region of 26 ± 2 kdaltons.

Further modifications of dehydrogenases of the invention: the invention likewise comprises "functional equivalents" of the specifically disclosed
enzymes having dehydrogenase activity and the use thereof in the processes of the invention.

“Functional equivalents” or analogs of the specifically disclosed enzymes are in the context of the present invention polypeptides which differ therefrom and which still have the desired biological activity such as, for example, substrate specificity. Thus, “functional equivalents” mean for example enzymes which reduce from the ketone to the corresponding ‘anti-Prelog’ alcohol and which have at least 20%, preferably 50%, particularly preferably 75%, very particularly preferably 90% of the activity of an enzyme comprising one of the amino acid sequences listed under SEQ ID 2 or SEQ ID 4. Functional equivalents are additionally preferably stable between pH 4 to 10 and advantageousy have a pH optimum between pH 5 and 8 and a temperature optimum in the range from 20°C to 80°C.

“Functional equivalents” also mean according to the invention in particular mutants which have an amino acid other than that specifically mentioned in at least one sequence position of the abovementioned amino acid sequences but nevertheless have one of the abovementioned biological activities. “Functional equivalents” thus comprise the mutants obtainable by one or more amino acid additions, substitutions, deletions and/or inversions, it being possible for said modifications to occur in any sequence position as long as they lead to a mutant having the property profile according to the invention. Functional equivalence also exists in particular when the reactivity patterns agree qualitatively between mutant and unmodified polypeptide, i.e. for example identical substrates are converted at a different rate.

Examples of suitable amino acid substitutions are to be found in the following table:

<table>
<thead>
<tr>
<th>Original residue</th>
<th>Examples of the substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>Ser</td>
</tr>
<tr>
<td>Arg</td>
<td>Lys</td>
</tr>
<tr>
<td>Asn</td>
<td>Gln; His</td>
</tr>
<tr>
<td>Asp</td>
<td>Glu</td>
</tr>
<tr>
<td>Cys</td>
<td>Ser</td>
</tr>
<tr>
<td>Gln</td>
<td>Asn</td>
</tr>
<tr>
<td>Glu</td>
<td>Asp</td>
</tr>
<tr>
<td>Gly</td>
<td>Pro</td>
</tr>
<tr>
<td>Residue</td>
<td>Functional Equivalents</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
</tr>
<tr>
<td>His</td>
<td>Asn; Gln</td>
</tr>
<tr>
<td>Ile</td>
<td>Leu; Val</td>
</tr>
<tr>
<td>Leu</td>
<td>Ile; Val</td>
</tr>
<tr>
<td>Lys</td>
<td>Arg; Gln; Glu</td>
</tr>
<tr>
<td>Met</td>
<td>Leu; Ile</td>
</tr>
<tr>
<td>Phe</td>
<td>Met; Leu; Tyr</td>
</tr>
<tr>
<td>Ser</td>
<td>Thr</td>
</tr>
<tr>
<td>Thr</td>
<td>Ser</td>
</tr>
<tr>
<td>Trp</td>
<td>Tyr</td>
</tr>
<tr>
<td>Tyr</td>
<td>Trp; Phe</td>
</tr>
<tr>
<td>Val</td>
<td>Ile; Leu</td>
</tr>
</tbody>
</table>

“Functional equivalents” in the above sense are also “precursors” of the described polypeptides, and “functional derivatives” and “salts” of the polypeptides.

“Precursors” are in this connection natural or synthetic precursors of the polypeptides with or without the desired biological activity.

The term “salts” means both salts of carboxyl groups and acid addition salts of amino groups of the protein molecules of the invention. Salts of carboxyl groups can be prepared in a manner known per se and comprise inorganic salts such as, for example, sodium, calcium, ammonium, iron and zinc salts, and salts with organic bases such as, for example, amines, such as triethanolamine, arginine, lysine, piperidine and the like. The invention also relates to acid addition salts such as, for example, salts with mineral acids such as hydrochloric acid or sulfuric acid, and salts with organic acids such as acetic acid and oxalic acid.

“Functional derivatives” of polypeptides of the invention can likewise be prepared on functional amino acid side groups or on their N- or C-terminal end with the aid of known techniques. Such derivatives comprise for example aliphatic esters of carboxylic acid groups, amides of carboxylic acid groups, obtainable by reaction with ammonia or with a primary or secondary amine; N-acyl derivatives of free amino groups prepared by reaction with acyl groups; or O-acyl derivatives of free hydroxy groups prepared by reaction with acyl groups.
In the case where protein glycosylation is possible, "functional equivalents" of the invention comprise proteins of the type designated above in deglycosylated or glycosylated form, and modified forms obtainable by altering the glycosylation pattern.

"Functional equivalents" of course also comprise polypeptides which are obtainable from other organisms, and naturally occurring variants. For example, it is possible to establish ranges of homologous sequence regions by comparison of sequences, and to ascertain equivalent enzymes based on the specific requirements of the invention.

"Functional equivalents" likewise comprise fragments, preferably individual domains or sequence motifs, of the polypeptides of the invention, which have, for example, the desired biological function.

"Functional equivalents" are additionally fusion proteins which comprise one of the abovementioned polypeptide sequences or functional equivalents derived therefrom and at least one further, heterologous sequence which is functionally different therefrom and is in functional N- or C-terminal linkage (i.e. with negligible mutual functional impairment of the parts of the fusion protein). Nonlimiting examples of such heterologous sequences are, for example, signal peptides or enzymes.

"Functional equivalents" also comprised in the invention are homologues of the specifically disclosed proteins. These have at least 75%, in particular at least 85%, such as, for example, 90%, 95%, 97% or 99%, homology to one of the specifically disclosed amino acid sequences, calculated by the algorithm of Pearson and Lipman, Proc. Natl. Acad. Sci. (USA) 85(8), 1988, 2444-2448. A percentage homology of a homologous polypeptide of the invention means in particular percentage identity of the amino acid residues based on the total length of one of the amino acid sequences specifically described herein.

Homologues of the proteins or polypeptides of the invention can be generated by mutagenesis, e.g. by point mutation or truncation of the protein.

Homologues of the proteins of the invention can be identified by screening combinatorial libraries of mutants, such as, for example, truncation mutants. For example, a variegated library of protein variants can be generated by combinatorial mutagenesis at the nucleic acid
level, such as, for example, by enzymatic ligation of a mixture of synthetic oligonucleotides. There is a large number of methods which can be used to prepare libraries of potential homologues from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be carried out in an automatic DNA synthesizer, and the synthetic gene can then be ligated into a suitable expression vector. Use of a degenerate set of genes makes it possible to provide all the sequences which encode the desired set of potential protein sequences in one mixture. Methods for synthesizing degenerate oligonucleotides are known to the skilled worker (e.g. Narang, S.A. (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al., (1984) Science 198:1056; lke et al. (1983) Nucleic Acids Res. 11:477).

Several techniques are known in the art for screening gene products of combinatorial libraries which have been prepared by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. These techniques can be adapted to the rapid screening of gene libraries which have been generated by combinatorial mutagenesis of homologues of the invention. The most commonly used techniques for screening large gene libraries, which are subject to high-throughput analysis, comprise the cloning of the gene library into replicable expression vectors, transformation of suitable cells with the resulting vector library and expression of the combinatorial genes under conditions under which detection of the desired activity facilitates isolation of the vector which encodes the gene whose product has been detected. Recursive ensemble mutagenesis (REM), a technique which increases the frequency of functional mutants in the libraries, can be used in combination with the screening tests to identify homologues (Arkin and Yourvan (1992) PNAS 89:7811-7815; Delgraves et al. (1993) Protein Engineering 6(3):327-331).

Further configuration of coding nucleic acid sequences of the invention

The invention relates to the use of nucleic acid sequences (single- and double-stranded DNA and RNA sequences, such as, for example, cDNA and mRNA) which code for an enzyme having dehydrogenase activity according to the invention. Preferred nucleic acid sequences code for example for amino acid sequences shown in SEQ ID 2 or SEQ ID 4 or characteristic partial sequences thereof, or comprise nucleic acid sequences shown in SEQ ID 1 or SEQ ID 3 or characteristic partial sequences thereof.
All nucleic acid sequences mentioned herein can be prepared in a manner known per se by chemical synthesis from the nucleotide building blocks, such as, for example, by fragment condensation of individual overlapping complementary nucleic acid building blocks of the double helix. Chemical synthesis of oligonucleotides can take place for example in a known manner by the phosphoamidite method (Voet, Voet, 2nd edition, Wiley Press New York, pages 896–897). Addition of synthetic oligonucleotides and filling in of gaps using the Klenow fragment of DNA polymerase and ligation reactions, and general cloning methods are described in Sambrook et al. (1989), Molecular Cloning: A laboratory manual, Cold Spring Harbor Laboratory Press.

The invention also relates to nucleic acid sequences (single- and double-stranded DNA and RNA sequences, such as, for example, cDNA and mRNA), coding for one of the above polypeptides and their functional equivalents which are obtainable for example by using artificial nucleotide analogs.

The invention relates both to isolated nucleic acid molecules which code for polypeptides or proteins of the invention or biologically active sections thereof, and to nucleic acid fragments which can be used for example for use as hybridization probes or primers for identifying or amplifying coding nucleic acids of the invention.

The nucleic acid molecules of the invention may additionally comprise untranslated sequences from the 3’ and/or 5’ end of the coding gene region.

The invention further comprises the nucleic acid molecules which are complementary to the specifically described nucleotide sequences, or a section thereof.

The nucleotide sequences of the invention make it possible to generate probes and primers which can be used to identify and/or clone homologous sequences in other cell types and organisms. Such probes or primers usually comprise a nucleotide sequence region which hybridizes under "stringent" conditions (see below) to at least about 12, preferably at least about 25, such as, for example, about 40, 50 or 75, consecutive nucleotides of a sense strand of a nucleic acid sequence of the invention or a corresponding antisense strand.
An "isolated" nucleic acid molecule is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid, and may moreover be substantially free of other cellular material or culture medium if it is prepared by recombinant techniques, or free of chemical precursors or other chemicals if it is synthesized chemically.

A nucleic acid molecule of the invention can be isolated by standard techniques of molecular biology and with the aid of the sequence information provided according to the invention. For example, cDNA can be isolated from a suitable cDNA library by using one of the specifically disclosed complete sequences or a segment thereof as hybridization probe and standard hybridization techniques (as described for example in Sambrook, J., Fritsch, E.F. and Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd edition, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). In addition, a nucleic acid molecule comprising one of the disclosed sequences or a segment thereof can be isolated by polymerase chain reaction, using the oligonucleotide primers constructed on the basis of this sequence. The nucleic acid amplified in this way can be cloned into a suitable vector and be characterized by DNA sequence analysis. The oligonucleotides of the invention can also be prepared by standard synthetic methods, e.g. using an automatic DNA synthesizer.

The nucleic acid sequences of the invention can in principle be identified and isolated from all organisms. The nucleic acid sequences of the invention or the homologues thereof can advantageously be isolated from fungi, yeasts, archaea or bacteria. Bacteria which may be mentioned are gram-negative and gram-positive bacteria. The nucleic acids of the invention are preferably from gram-negative bacteria, advantageously from α-proteobacteria, β-proteobacteria or γ-proteobacteria, particularly preferably from bacteria of the orders of Burkholderiales, Hydrogenophilales, Methylophilales, Neisseriales, Nitrosomonadales, Procabacterales or Rhodocyclales. Very particularly preferably from bacteria of the family of Rhodocyclaceae. Particularly preferably from the genus Azoarcus. Particularly preferably from species Azoarcus anaerobius, Azoarcus buckelii, Azoarcus communis, Azoarcus evansii, Azoarcus indicens, Azoarcus toluesticus, Azoarcus tolulyticus, Azoarcus toluovarans, Azoarcus sp., Azoarcus sp. 22Lin, Azoarcus sp. BH72, Azoarcus sp. CC-11, Azoarcus sp. CIB, Azoarcus sp. CR23, Azoarcus sp. EB1, Azoarcus sp. EbN1, Azoarcus sp. FL05, Azoarcus sp. HA, Azoarcus sp. HxN1, Azoarcus sp. mXyN1, Azoarcus sp. PbN1, Azoarcus sp. PH002, Azoarcus sp. T and Azcarcus sp. ToN1.
Dehydrogenases from *Azoarcus* sp EbN1 are particularly preferably used.

Nucleic acid sequences of the invention can be isolated for example by conventional hybridization processes or the PCR technique from other organisms, e.g. through genomic or cDNA libraries. These DNA sequences hybridize under standard conditions with the sequences of the invention. It is advantageous to use for the hybridization short oligonucleotides of the conserved regions, for example from the active center, which can be ascertained by comparisons with a dehydrogenase of the invention in a manner known to the skilled worker. However, longer fragments of the nucleic acids of the invention, or the complete sequences, can also be used for the hybridization. These standard conditions vary depending on the nucleic acid used (oligonucleotide, longer fragment or complete sequence) or depending on which type of nucleic acid, DNA or RNA, are used for the hybridization. Thus, for example, the melting temperatures for DNA:DNA hybrids are about 10°C lower than those of DNA:RNA hybrids of the same length.

Standard conditions mean for example, depending on the nucleic acid, temperatures between 42 and 58°C in an aqueous buffer solution with a concentration of between 0.1 to 5 x SSC (1 X SSC = 0.15 M NaCl, 15 mM sodium citrate, pH 7.2) or additionally in the presence of 50% formamide such as, for example, 42°C in 5 x SSC, 50% formamide. The hybridization conditions for DNA:DNA hybrids are advantageously 0.1 x SSC with temperatures between about 20°C to 45°C, preferably between about 30°C to 45°C. The hybridization conditions for DNA:RNA hybrids are advantageously 0.1 x SSC with temperatures between about 30°C to 55°C, preferably between about 45°C to 55°C. These stated temperatures for the hybridization are calculated values for the melting temperature by way of example for a nucleic acid with a length of about 100 nucleotides and a G + C content of 50% in the absence of formamide. The experimental conditions for DNA hybridization are described in relevant text books of genetics, such as, for example, Sambrook et al., *Molecular Cloning*, Cold Spring Harbor Laboratory, 1989, and can be calculated by formulae known to the skilled worker, for example depending on the length of the nucleic acids, the type of hybrids or the G + C content. Further information on hybridization can be found by the skilled worker in the following text books: Ausubel et al. (eds), 1985, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York; Hames and Higgins (eds), 1985, *Nucleic Acids Hybridization: A Practical Approach*, IRL Press at

The invention also relates to derivatives of the specifically disclosed or derivable nucleic acid sequences.

Thus, further nucleic acid sequences of the invention can be derived from SEQ ID 1 or SEQ ID 3 and differ therefrom by addition, substitution, insertion or deletion of single or multiple nucleotides, but still code for polypeptides having the desired property profile.

The invention also comprises nucleic acid sequences which comprise so-called silent mutations or are modified according to the codon usage of a specific original or host organism by comparison with a specifically mentioned sequence, as well as naturally occurring variants such as, for example, splice variants or allelic variants thereof.

It likewise relates to sequences obtainable by conservative nucleotide substitutions (i.e. the relevant amino acid is replaced by an amino acid of the same charge, size, polarity and/or solubility).

The invention also relates to the molecules derived from the specifically disclosed nucleic acids through sequence polymorphisms. These genetic polymorphisms may exist between individuals within a population owing to natural variation. These natural variations normally result in a variance of from 1 to 5% in the nucleotide sequence of a gene.

Derivatives of a nucleic acid sequence of the invention mean for example allelic variants which exhibit at least 40% homology at the derived amino acid level, preferably at least 60% homology, very particularly preferably at least 80, 85, 90, 93, 95 or 98% homology over the entire sequence region (concerning homology at the amino acid level, reference may be made to the above statements about the polypeptides). The homologies may advantageously be higher over partial regions of the sequences.

Derivatives also mean in addition homologues of the nucleic acid sequences of the invention, for example fungal or bacterial homologues, truncated sequences, single-stranded DNA or RNA of the coding and noncoding DNA sequence. They possess for
example at the DNA level a homology of at least 40%, preferably of at least 60%, particularly preferably of at least 70%, very particularly preferably of at least 80% over the entire DNA region indicated.

5 Derivatives additionally mean for example fusions with promoters. The promoters which are upstream of the indicated nucleotide sequences may be modified by one or more nucleotide exchanges, insertions, inversions and/or deletions but without impairing the functionality or activity of the promoters. The promoters may in addition have their activity increased through modification of their sequence, or be completely replaced by more efficient promoters even of organisms of different species.

10 Derivatives also mean variants whose nucleotide sequence in the region from -1 to -1000 bases upstream of the start codon or 0 to 1000 bases downstream of the stop codon have been modified in such a way that gene expression and/or protein expression is altered, preferably increased.

The invention further comprises also nucleic acid sequences which hybridize with the abovementioned coding sequences under “stringent conditions”. These polynucleotides can be found by screening genomic or cDNA libraries and if appropriate be amplified therefrom using suitable primers by means of PCR and then isolated for example using suitable probes. It is additionally possible to synthesize polynucleotides of the invention also by a chemical route. By this property is meant the ability of a poly- or oligonucleotide to bind under stringent conditions to a virtually complementary sequence, whereas nonspecific bindings between non-complementary partners do not occur under these conditions. For this purpose, the sequences should be 70-100%, preferably 90-100%, complementary. The property of complementary sequences being able to bind specifically to one another is made use of, for example, in the Northern or Southern blotting technique or in the case of primer binding in PCR or RT-PCR. Oligonucleotides with a length of 30 base pairs or more are normally employed for this purpose. Stringent conditions mean for example in the Northern blotting technique the use of a washing solution at 50-70°C, preferably 60-65°C, for example 0.1 x SSC buffer with 0.1% SDS (20 x SSC: 3M NaCl, 0.3M Na citrate, pH 7.0) for eluting nonspecifically hybridized cDNA probes or oligonucleotides. In this case, as mentioned above, only nucleic acids with a high degree of complementarity remain bound to one another. The setting up of stringent conditions is known to the skilled worker and is
described, for example, in Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6.

Configurations of constructs of the invention:

The invention additionally relates to expression constructs comprising a nucleic acid sequence which codes for a polypeptide of the invention and is under the genetic control of regulatory nucleic acid sequences; and to vectors comprising at least one of these expression constructs.

Such constructs of the invention preferably comprise a promoter 5'-upstream from the respective coding sequence and a terminator sequence 3'-downstream and, if desired, further conventional regulatory elements, in particular in each case operatively linked to the coding sequence.

An "operative linkage" means the sequential arrangement of promoter, coding sequence, terminator and, if desired, further regulatory elements in such a way that each of the regulatory elements is able to perform its function as intended in the expression of the coding sequence. Examples of sequences which can be operatively linked are targeting sequences, and enhancers, polyadenylation signals and the like. Further regulatory elements comprise selectable markers, amplification signals, origins of replication and the like. Suitable regulatory sequences are described for example in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990).

A nucleic acid construct of the invention means in particular those in which the gene for a dehydrogenase of the invention has been operatively or functionally linked to one or more regulatory signals to control, e.g. increase, gene expression.

In addition to these regulatory sequences it is possible for the natural regulation of these sequences still to be present in front of the actual structural genes and, if desired, to have been genetically modified so that the natural regulation has been switched off and the expression of the genes has been increased. The nucleic acid construct may, however, also have a simpler structure, i.e. no additional regulatory signals have been inserted in front of the coding sequence, and the natural promoter with its regulation has not been deleted.
Instead, the natural regulatory sequence is mutated in such a way that regulation no longer takes place and gene expression is increased.

A preferred nucleic acid construct advantageously also comprises one or more of the previously mentioned "enhancer" sequences functionally linked to the promoter, which make increased expression of the nucleic acid sequence possible. Additional advantageous sequences may also be inserted at the 3' end of the DNA sequences, such as further regulatory elements or terminators. The nucleic acids of the invention may be present in one or more copies in the construct. The construct may also comprise further markers such as antibiotic resistances or auxotrophy-complementing genes, if desired for selecting for the construct.

Advantageous regulatory sequences for the process of the invention are for example present in promoters such as cos, tac, trp, tet, trp-tet, lpp, lac, lpp-lac, lacI, T7, T5, T3, gal, trc, ara, rhaP (rhaPbad) SP6, lambda-P_r or in the lambda-P_l promoter, which are advantageously used in gram-negative bacteria. Further advantageous regulatory sequences are present for example in the gram-positive promoters amy and SPO2, in the yeast or fungal promoters ADC1, MFalpha, AC, P-60, CYC1, GAPDH, TEF, rpm28, ADH. Also advantageous in this connection are the promoters of pyruvate decarboxylase and methanol oxidase, for example from *Hansenula*. It is also possible to use artificial promoters for the regulation.

For expression in a host organism, the nucleic acid construct is advantageously inserted into a vector such as, for example, a plasmid or a phage which makes optimal expression of the genes in the host possible. Vectors mean apart from plasmids and phages also all other vectors known to the skilled worker, e.g. viruses such as SV40, CMV, baculovirus and adenovirus, transposons, IS elements, phasmids, cosmids, and linear or circular DNA. These vectors may undergo autonomous replication in the host organism, or chromosomal replication. These vectors represent a further configuration of the invention. Suitable plasmids are for example in *Escherichia coli* pLG338, pACYC184, pBR322, pUC18, pUC19, pKC30, pRep4, pHIS1, pKK223-3, pDHE19.2, pHIS2, pPLc236, pMBL24, pLG200, pUR290, pIN-II^1^11^3^-B1, lgt11 or pBdCl, in *Streptomyces* plJ101, plJ364, plJ702 or plJ361, in *Bacillus* pUB110, pC194 or pBD214, in *Corynebacterium* pSA77 or pAJ667, in fungi pALS1, pIL2 or pBB116, in yeasts 2alphaM, pAG-1, YEp6, YEp13 or pEMBLYe23 or in plants pLGV23,

The nucleic acid construct advantageously comprises, for expression of the further genes present, additionally 3'- and/or 5'-terminal regulatory sequences to increase the expression, which are selected for optimal expression depending on the gene or genes and on the host organism selected.

These regulatory sequences are intended to make specific expression of the genes and of protein expression possible. This may mean, for example depending on the host organism, that the gene is expressed or overexpressed only after induction, or that it is immediately expressed and/or overexpressed.

The regulatory sequences or factors may moreover preferably influence positively, and thus increase, expression of the introduced genes. Thus, enhancement of the regulatory elements can take place advantageously at the level of transcription, by using strong transcription signals such as promoters and/or "enhancers". However, it is also possible in addition to enhance translation by, for example, improving the stability of the mRNA.

In a further embodiment of the vector, the vector comprising the nucleic acid construct of the invention or the nucleic acid of the invention may also advantageously be introduced in the form of a linear DNA into the microorganisms and be integrated by heterologous or homologous recombination into the genome of the host organism. This linear DNA may consist of a linearized vector such as a plasmid or only of the nucleic acid construct or of the nucleic acid of the invention.

For optimal expression of heterologous genes in organisms, it is advantageous to modify the nucleic acid sequences to accord with the specific codon usage used in the organism. The codon usage can easily be ascertained on the basis of computer analyses of other known genes in the relevant organism.

An expression cassette of the invention is prepared by fusing a suitable promoter to a

For expression in a suitable host organism, the recombinant nucleic acid construct or gene construct is advantageously inserted into a host-specific vector which makes optimal expression of the genes in the host possible. Vectors are well known to the skilled worker and can be found for example in "Cloning Vectors" (Pouwels P. H. et al., editors, Elsevier, Amsterdam-New York-Oxford, 1985).

Host organisms which can be used according to the invention
It is possible with the aid of the vectors or constructs of the invention to prepare recombinant microorganisms which are for example transformed with at least one vector of the invention and can be employed to produce the polypeptides of the invention. The recombinant constructs of the invention described above are advantageously introduced into a suitable host system and expressed. Common cloning and transfection methods familiar to the skilled worker, such as, for example, coprecipitation, protoplast fusion, electroporation, retroviral transfection and the like, are preferably used to bring about expression of said nucleic acids in the particular expression system. Suitable systems are described for example in Current Protocols in Molecular Biology, F. Ausubel et al., editors, Wiley Interscience, New York 1997, or Sambrook et al. Molecular Cloning: A Laboratory Manual. 2nd edition, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

It is also possible according to the invention to prepare homologously recombined microorganisms. This entails preparation of a vector which comprises at least one segment of a gene of the invention or a coding sequence in which, if desired, at least one amino acid deletion, addition or substitution has been introduced in order to modify, e.g. functionally disrupt, the sequence of the invention (knockout vector). The introduced sequence may, for example, also be a homologue from a related microorganism or be derived from a
mammalian, yeast or insect source. The vector used for homologous recombination may alternatively be designed so that the endogenous gene is mutated or otherwise modified during the homologous recombination but still encodes the functional protein (e.g. the regulatory region located upstream may be modified in such a way that this modifies expression of the endogenous protein). The modified segment of the gene of the invention is in the homologous recombination vector. The construction of suitable vectors for homologous recombination is described for example in Thomas, K.R. and Capecchi, M.R. (1987) Cell 51:503.

Suitable recombinant host organisms for the nucleic acid of the invention or the nucleic acid construct are in principle all prokaryotic or eukaryotic organisms. It is advantageous to use as host organisms microorganisms such as bacteria, fungi or yeasts. It is advantageous to use gram-positive or gram-negative bacteria, preferably bacteria of the families Enterobacteriaceae, Pseudomonadaceae, Rhizobiaceae, Streptomycetaceae or Nocardiaceae, particularly preferably bacteria of the genera Escherichia, Pseudomonas, Streptomyces, Nocardia, Burkholderia, Salmonella, Agrobacterium or Rhodococcus. The genus and species Escherichia coli is very particularly preferred. Further advantageous bacteria are additionally to be found in the group of α-proteobacteria, β-proteobacteria or γ-proteobacteria.

The host organism or the host organisms of the invention moreover comprise preferably at least one of the nucleic acid sequences, nucleic acid constructs or vectors which are described in this invention and which code for an enzyme having dehydrogenase activity according to the invention.

The organisms used in the process of the invention are grown or cultured in a manner known to the skilled worker depending on the host organism. Microorganisms are ordinarily grown in a liquid medium which comprises a carbon source, usually in the form of sugars, a nitrogen source, usually in the form of organic nitrogen sources such as yeast extract or salts such as ammonium sulfate, trace elements such as iron, manganese, magnesium salts and, if desired, vitamins, at temperatures between 0°C and 100°C, preferably between 10°C to 60°C, while passing in oxygen. The pH of the nutrient liquid can be kept at a fixed value during this, i.e. controlled during the culturing, or not. The culturing can be carried out batchwise, semibatchwise or continuously. Nutrients can be introduced at the start of the
fermentation or be subsequently fed in semicontinuously or continuously. The ketone can be added directly for culturing or advantageously after culturing. The enzymes can be isolated from the organisms by the processes described in the examples, or be used for the reaction as crude extract.

Recombinant preparation of the polypeptides of the invention

The invention further relates to processes for the recombinant preparation of polypeptides of the invention or functional, biologically active fragments thereof, where a polypeptide-producing microorganism is cultivated, if desired, expression of the polypeptides is induced, and the latter are isolated from the culture. The polypeptides can also be produced in this way on the industrial scale if this is desired.

The recombinant microorganism can be cultivated and fermented by known processes. Bacteria can be grown for example in TB or LB medium and at a temperature of 20 to 40°C and a pH of 6 to 9. Suitable cultivation conditions are described for example in detail in T. Maniatis, E.F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989).

If the polypeptides are not secreted into the culture medium, the cells are then disrupted and the product is obtained from the lysate by known protein isolation methods. The cells may alternatively be disrupted by high-frequency ultrasound, by high pressure such as, for example, in a French pressure cell, by osmolysis, by the action of detergents, lytic enzymes or organic solvents, by homogenizers or by a combination of a plurality of the processes mentioned.

The polypeptides can be purified by known chromatographic methods such as molecular sieve chromatography (gel filtration), such as Q-Sepharose chromatography, ion exchange chromatography and hydrophobic chromatography, and by other usual methods such as ultrafiltration, crystallization, salting out, dialysis and native gel electrophoresis. Suitable methods are described for example in Cooper, F. G., Biochemische Arbeitsmethoden, Verlag Walter de Gruyter, Berlin, New York or in Scopes, R., Protein Purification, Springer Verlag, New York, Heidelberg, Berlin.
It may be advantageous for isolation of the recombinant protein to use vector systems or oligonucleotides which extend the cDNA by particular nucleotide sequences and thus code for modified polypeptides or fusion proteins which serve, for example, for simpler purification. Suitable modifications of these types are for example so-called tags which act as anchors, such as, for example, the modification known as hexa-histidine anchor, or epitopes which can be recognized as antigens by antibodies (described for example in Harlow, E. and Lane, D., 1988, Antibodies: A Laboratory Manual. Cold Spring Harbor (N.Y.) Press). These anchors can be used to attach the proteins to a solid support, such as, for example, a polymer matrix, which can for example be packed in a chromatography column, or can be used on a microtiter plate or another support.

These anchors can at the same time also be used for recognition of proteins. It is also possible to use for recognition of the proteins conventional markers such as fluorescent dyes, enzyme markers which form a detectable reaction product after reaction with a substrate, or radioactive markers, alone or in combination with the anchors for derivatizing the proteins.

Further configurations for carrying out the enzymatic reduction process of the invention

The dehydrogenases can be used in the process of the invention as free or immobilized enzyme or as catalyst still present in the recombinant producer organism.

The process of the invention is advantageously carried out at a temperature between 0°C to 95°C, preferably between 10°C to 85°C, particularly preferably between 15°C to 75°C.

The pH in the process of the invention is advantageously kept between pH 4 and 12, preferably between pH 4.5 and 9, particularly preferably between pH 5 and 8.

Enantiopure or chiral products mean in the process of the invention enantiomers which exhibit an enantiomer enrichment. Enantiopurities of at least 70% ee, preferably of at least 80% ee, particularly preferably of at least 90% ee, very particularly preferably of at least 98% ee, are preferably achieved in the process.

It is possible to use for the process of the invention growing cells which comprise the nucleic
acids, nucleic acid constructs or vectors of the invention. Resting or disrupted cells can also be used. Disrupted cells mean for example cells which have been made permeable by a treatment with, for example, solvents, or cells which have been disintegrated by an enzymic treatment, by a mechanical treatment (e.g. French press or ultrasound) or by another method. The crude extracts obtained in this way are advantageously suitable for the process according to the invention. Purified or partly purified enzymes can also be used for the process. Likewise suitable are immobilized microorganisms or enzymes which can be advantageously used in the reaction.

The process of the invention can be carried out batchwise, semibatchwise or continuously.

The process can advantageously be carried out in bioreactors as described for example in Biotechnology, Volume 3, 2nd edition, Rehm et al. editors, (1993), in particular Chapter II.

The following examples are intended to illustrate the invention but without restricting it.

Experimental section
Example 1: Cloning of the alcohol dehydrogenase EbN2 from Azoarcus sp. EbN1.

The sequence of the EbN2 dehydrogenase gene from Azoarcus sp. EbN1 is deposited in databases (SEQ ID 1, [Genbank ID 56475432, Region: 2797788..2798528]). Oligonucleotides were derived from the nucleic acid sequence of the gene and were used to amplify by known methods the gene from genomic DNA of Azoarcus sp. EbN1. The resulting sequence corresponds to the published sequence. The DNA sequence of the oligonucleotides is summarized in Table 1.

PCR conditions:

2 µl of 10*Pfu ultra buffer (Stratagene)
100 ng of primer #1 (cf. Table 1)
100 ng of primer #2 (cf. Table 1)
1 µl of dNTP (10 mM each)
ca. 30 ng of chromosomal DNA from Azoarcus sp. EbN1
1 U of Pfu ultra DNA polymerase
ad 20 µl H₂O
Temperature program:
5 min, 94°C,
60 sec, 50°C,
2 min, 72°C,
\{ (35 cycles)
5
60 sec, 94°C,
10 min, 72°C,
∞, 10°C

The PCR product (approx. 751 bp) was digested with the restriction endonucleases NdeI and BamHI and cloned into correspondingly digested pDHE19.2 vector (DE19848129). The ligation mixtures were transformed into E. coli XL1 Blue (Stratagene). The resulting plasmid pDHE-PDH-L was transformed into the strain E. coli TG10 pAgro4 pHSG575 (TG10: an RhaA derivative of E. coli TG1 (Stratagene); pAgro4: Takeshita, S; Sato, M; Toba, M; Masahashi, W; Hashimoto-Gotoh, T (1987) Gene 61, 63-74; pHSG575: T.

The recombinant E. coli are referred to as LU 13151.

Example 2: Cloning of the alcohol dehydrogenase ChnA from Azoarcus sp. EbN1.
The sequence of the dehydrogenase gene ChnA from Azoarcus sp. EbN1 is deposited in databases ([Genbank ID 56475432, Region: (complement) 192247..192993]). Oligonucleotides were derived from the nucleic acid sequence of the gene and were used to amplify by known methods the gene from genomic DNA of Azoarcus sp. EbN1. The resulting sequence corresponds to the published sequence. The DNA sequence of the oligonucleotides is summarized in Table 2.

PCR conditions:
2 μl of 10*Pfu ultra buffer (Stratagene)
100 ng of primer #3 (cf. Table 2)
100 ng of primer #4 (cf. Table 2)
1 μl of dNTP (10 mM each)
ca. 30 ng of chromosomal DNA from Azoarcus sp. EbN1
1 U of Pfu ultra DNA polymerase
ad 20 μl H₂O
Temperature program:
5 min, 94°C,
60 sec, 50°C,
2 min, 72°C,
60 sec, 94°C,
10 min, 72°C,
∞, 10°C

(35 cycles)

The PCR product (approx. 743 bp) was digested with the restriction endonucleases NdeI and BglII and cloned into a pDHE19.2 vector (DE19848129) restricted with NdeI and BamHI. The ligation mixtures were transformed into E. coli XL1 Blue (Stratagene). The resulting plasmid pDHE-PDH-L was transformed into the strain E. coli TG10 pAgro4 pHSRG75 (TG10: an RhaA- derivative of E. coli TG1(Stratagene); pAgro4: Takeshita, S; Sato, M; Toba, M; Masahashi, W; Hashimoto-Gotoh, T (1987) Gene 61, 63-74; pHSRG75: T.


The recombinant E. coli are referred to as LU 13283.

Example 3: Provision of recombinant 'anti-Prelog' dehydrogenases
LU 13151 or LU 13283 were grown in 20 ml of LB-Amp/Spec/Cm (100 μg/l ampicillin;
100 μg/l spectinomycin; 20 μg/l chloramphenicol), 0.1 mM IPTG, 0.5 g/l rhamnose in 100 ml
Erlenmeyer flasks (baffles) at 37°C for 18 h, centrifuged at 5000×g/10 min, washed once
with 10 mM TRIS*HCl, pH 7.0, and resuspended in 2 ml of the same buffer.

Cell-free crude extract was prepared by disrupting LU 13151 or LU 13283 cell paste
0.7 ml glass beads (d = 0.5 mm) in a vibratory mill (3 x 5 min with intermediate cooling on
ice).

Example 4: Determination of the activity of the recombinant 'anti-Prelog' dehydrogenases
from Azoarcus sp. EbN1
6 transformants in each case were grown in 20 ml of LB Amp/Spec/Cm (100 μg/l amp;
100 mg/l lac; 20 μg/lcm) 0.1 mM IPTG 0.5 g/l rhamnose in 100 ml Erlenmeyer flasks
(baffles) at 37°C for 18 h, centrifuged at 5000×g/10 min, washed once with 10 mM Tris/HCl
pH 7.0, and resuspended in 2 ml of the same buffer.

Cell-free crude extract of the recombinant E. coli which comprised the dehydrogenase
genes was obtained by cell disruption with 0.7 ml of glass beads (d = 0.5 mm) in a vibratory mill (3 x 5 min with intermediate cooling on ice). The consumption of reduced cosubstrates can be followed during the reduction of ketones in a photometer at 340 nm. 10 µl of diluted cell-free crude extract (≈ 10 µg of protein), 10 µmol of ketone and 250 nmol of NADH or NADPH were incubated in 1 ml of 50 mM KP\textsubscript{i}, 1 mM MgCl\textsubscript{2}, pH 6.5, at 30°C. 1 Unit (1U) corresponds to the amount of enzyme which reduces 1 µmol of ketone in 1 min.

Example 5: Phenylethanol analyses

The concentration of acetophenone and phenylethanol can be determined by HPLC. Depending on the choice of the stationary and mobile phases it is possible to determine the ee in addition to the concentration.

Stationary phase: Hydrodex β-6-TBDM (Macherey\&Nagel), length: 25 m, Ø: 250 µm,
Mobile phase: helium, split 100:1, total flow rate: 92 ml/min, pressure: 17 psi
Flow rate: 1.0 ml/min
Detection: FID
Temperature gradient: t = 0 min: 90°C, heating at 3°C/min to 140°C
Detector temperature: 250°C

Retention times: acetophenone: approx. 7.5 min
(1S)-phenylethanol: approx. 12.5 min
(1R)-phenylethanol: approx. 12.1 min

A calibration series is constructed using authentic material and allows the concentration of unknown samples to be determined.

Example 6: Provision of glucose dehydrogenase for cofactor regeneration and cofactor regeneration with glucose dehydrogenase (enzyme coupling)

Glucose dehydrogenase can be used for cofactor regeneration. The enzyme can be obtained from commercial (e.g. Jüllich Fine Chemicals Order No. 22.10 or 19.10) or own sources. The latter is an E. coli XL10 Gold clone which comprises the glucose dehydrogenase gene from Bacillus subtilis (Genbank Acc.No. M12276) in the plasmid pUC19 (this construct is called E. coli LU11293).
The following medium was made up to ferment *E. coli* LU11293:

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>560 g</td>
<td>Yeast extract (65%)</td>
</tr>
<tr>
<td>448 g</td>
<td>Tryptone (Difco)</td>
</tr>
<tr>
<td>42 g</td>
<td>KH₂PO₄</td>
</tr>
<tr>
<td>84 g</td>
<td>Na₂HPO₄</td>
</tr>
<tr>
<td>644 g</td>
<td>Glycerol (99%)</td>
</tr>
<tr>
<td>100 ml</td>
<td>SL4 solution (5x)</td>
</tr>
<tr>
<td>1 g</td>
<td>Tegosipon 3062</td>
</tr>
</tbody>
</table>

Make up medium to 13.5 l with water, adjust pH to 7.0, remove about 300 ml for preculture, then sterilize at 122°C for 30 min.

Add sterile salt solution* (remove the salt solution for the shaken flasks beforehand; see report).

*Salt solution: 2.1 g of CaCl₂ * 2 H₂O
3.5 g of MgSO₄ * 7 H₂O
14 g of NH₄Cl
14 ml of ampicillin solution (100 mg/ml)
dissolve in 500 ml of water and sterilize by filtration

100 ml portions of medium were sterilized in two 1 l Erlenmeyer flasks and completed with 5 ml of sterile salt solution. Inoculation from an LB-ampicillin agar plate was followed by incubation of the precultures at 37°C and 200 rpm for 12 hours, and addition to the fermentation medium. The fermentation was started at 37°C, 0.1 bar internal pressure, pH 7.0 (controlling with 20% phosphoric acid and 25% NaOH) with an aeration rate of 7.5 l/min and 300 rpm (controlling pO₂ at between 20 and 50% with 10-20 l/min inlet air and 500-1500 rpm). After 2 h, 0.1 mM IPTG was added for induction and, after a total of 13 h, the fermentation was terminated. After harvesting and washing of the cells (1.3 kg), they were stored at −20°C until used (2-20 g/l in the mixture).

Equimolar amounts of glucose and ketone are dissolved with 1-30 U/ml glucose
dehydrogenase crude extract and 1-30 U/mL, alcohol dehydrogenase crude extract, 0.02-1 mmol/l NAD or NADP, or NADH or NADPH, in buffer, and incubated at 10-60°C. The pH was kept constant by automatic addition of base.

Example 7: Cofactor regeneration with substrate coupling
Regeneration of the cofactor can also be carried out by the two alcohol dehydrogenases themselves. In this case, addition of a separate regenerating enzyme is unnecessary. The alcohol dehydrogenases ChnA and Ebn2 accept various simple alcohols as reducing agents. They are oxidized to the corresponding carbonyl compounds. Simple alcohols which are suitable for regenerating NADH or NADPH are iso-propanol, butan-2-ol and pentan-2-ol.

Example 8: Cofactor regeneration with formate dehydrogenase (enzyme coupling)
Formate dehydrogenase can be used for cofactor regeneration. The enzyme is obtainable from commercial (e.g. Jülich Fine Chemicals: Order No. 09.11, 24.11 or 25.10) or from own sources. Regeneration of the cofactors can thus take place also with formate dehydrogenase in analogy to Example 6. In this case, equimolar amounts of formate and ketone are dissolved with 1-30 U/ml formate dehydrogenase crude extract and 1-30 U/ml alcohol dehydrogenase crude extract, 0.02-1 mmol/l NAD or NADP, or NADH or NADPH, in buffer and incubated at 10-60°C. The pH was kept constant by automatic addition of acid.

Example 9: Preparation of R-phenylethanol with recombinant anti-Prelog dehydrogenases
E. coli LU 13151 or LU 13283 were grown, harvested and disrupted in accordance with Example 3.
Per liter of reaction volume, 0.2 mmol of NAD, 500 U of glucose dehydrogenase, 1 mol of D-glucose, 1 mol of acetophenone, 100-1000 U of alcohol dehydrogenase from LU 13283 or LU 13151 are dissolved in KP buffer (50 mM KP, 1 mM MgCl2, pH 6.5) and incubated at 30°C. The pH was kept constant by automatic addition of 2M NaOH.
It is perfectly plausible that higher final concentrations of R-phenylethanol can be reached if acetophenone is metered in during the course of the reaction (fed-batch procedure).
Likewise, reaction in the presence of organic, water-insoluble solvents is possible.
We claim:

1. A process for preparing an optically active alcohol of formula la,

   \[
   \text{OH} \\
   R^2 \begin{array}{c}
   \text{R}^1 \end{array}
   \]

   **Formula la**

   in which R₁ and R² independently are alkyl, alkenyl, aryl, or alkylaryl groups, each of which may be substituted one or more times by alkyl, halogen, SH, SR², OH, OR², NO₂, CN, CO, COOR², NR²R³ or NR²R³R⁴X, where R², R³ and R⁴ independently are H or a lower alkyl or lower alkoxy radical, and X⁻ is a counter ion, with the proviso that R¹ is not equal to R² and R¹ is less bulky than R²,

   by reducing a corresponding ketone with a dehydrogenase having the polypeptide sequence of SEQ ID NO: 2 or having a polypeptide sequence having at least 95% sequence identity to the polypeptide sequence of SEQ ID NO: 2.

2. The process of claim 1, wherein the dehydrogenase is expressed recombinantly in a host organism, and the host organism is incubated in a solution in which the process is carried out.

3. The process of claim 1, wherein the dehydrogenase is expressed recombinantly in a host organism and isolated from said host organism prior to incubating in a solution in which the process is carried out.

4. The process of claim 1, wherein the reduction is carried out at a temperature of 20°C to 40°C.

5. The process of claim 1, further comprising generating a reduced cofactor by the dehydrogenase, or wherein glucose dehydrogenase, phosphite
dehydrogenase, formate dehydrogenase or another alcohol dehydrogenase is used as a cofactor-regenerating system.

6. The process of claim 1, wherein the dehydrogenase is expressed recombinantly in a host organism and a crude extract of said host organism is incubated in a solution in which the process is carried out.

7. The process of claim 6, wherein the crude extract is obtained by disrupting the host organism by high-frequency ultrasound, high pressure, osmolysis, homogenization, by action of detergents, lytic enzymes or organic solvents, or any combination thereof.