Nitrile hydratases from metagenome libraries

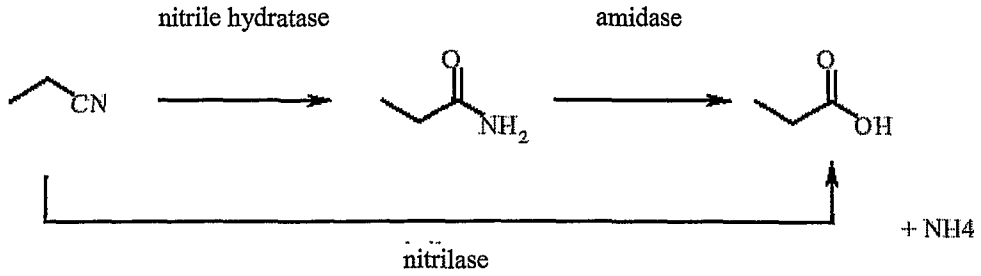
The present invention is directed toward specific degenerate primers. These latter are preferably employed in a process for preparing nitrile hydratases. The present invention therefore also relates to the nitrile hydratases which are prepared by the process which is performed using the primers and to other proteins which are required for the nitrile hydratases to be active. The nucleic acids encoding these protein sequences, and expression systems comprising these nucleic acids, likewise form an additional part of the subject matter of the application. The use of the nitrile hydratases, and of the underlying nucleic acid sequences, constitutes an additional aspect of the present invention.

The amide and carboxylic acid structural classes are becoming ever more important as precursors of fine chemicals. Specific aminoamides and (proteinogenic and nonproteinogenic) amino acids are key intermediates for synthesizing pharmaceutical and agrochemical products as well as in the foodstuffs field. Enantiomerically pure amides and amino acids, in particular, play an ever greater role in the abovementioned areas of application.

Aminonitrile precursors, as are required for preparing the abovementioned compound classes, can be readily obtained in racemic form by way of what is termed Strecker synthesis. The nitriles which have been obtained in this way can then be converted into the corresponding amides and carboxylic acids by means of chemical or enzymic hydrolysis.

Three enzymes which are capable of being involved in the enzymic hydrolysis of nitriles are known. Nitrilases convert a nitrile function directly into the acid whereas nitrile hydratases (E.C. 4.2.1.84) in this case form the corresponding amide. This latter can then finally be

converted into the corresponding carboxylic acid by means of an amidase (E.C. 3.5.1.4) (scheme 1).



Scheme 1:

Using isolated enzymes or whole-cell catalysts to hydrolyze nitriles to give the corresponding amides and acids helps to save large quantities of salt which would otherwise accrue in the neutralization step following the chemical hydrolysis of nitriles. For this reason, the enzymic hydrolysis of nitriles to give, for example, aminoamides and/or amino acids represents a more sustainable production process.

In their active form, nitrile hydratases consist of 2 nonhomologous, α and β subunits. These latter form heterodimers and tetramers, while even decamers have been demonstrated to be present in the case of *Rhodococcus rhodochrous* J1. While the α and β subunits are of approximately the same size, they do not otherwise share any similarities. Nitrile hydratases are metalloproteins which contain $Fe3^+$ or $Co3^+$ ions (Bunch A. W. (1998), Nitriles, in: Biotechnology, Volume 8a, Biotransformations I, Chapter 6, Eds.: Rehm HJ, Reed G, Wiley-VCH, p. 277-324; Shearer J, Kung IY, Lovell S, Kaminsky W, Kovacs JA (2001) Why is there an "inert" metal center in the active site of nitrile hydratase? Reactivity and ligand dissociation from a five-coordinate Co(III) nitrile hydratase model. J Am Chem Soc 123: 463-468; Kobayashi M, Shimizu S (2000) Nitrile hydrolases. Current Opinion in Chemical Biology 4: 95-102).

One of the greatest challenges thus far has been that of heterologously preparing nitrile hydratases in a suitable host, preferably *E. coli*. This Gram-negative bacterium is known to express heterologous proteins at high rates. An additional advantage is the yield of biomass in high cell density fermentations using *E. coli*. In such fermentations, it is possible to achieve productivities of more than 100 g of dry biomass (DBM) in a period of from 24 to 44 hours (Lee SY (1996) High cell-density culture of Escherichia coli. TIBTECH 14:98-105; Riesenberg D, Guthke R (1999) High-cell-density cultivation of microorganisms. Appl Microbiol Biotechnol 51:422-430).

Most nitrile hydratase α and β subunit sequences are known from the genus Rhodococcus. However, it is precisely the nitrile hydratases from this genus which it has thus far only been possible to express in *E. coli* with particular difficulty (Ikehata O, Nishiyama M, Horinouchi S, Beppu T (1989) Primary structure of nitrile hydratase deduced from the nucleotide sequence of a Rhodococcus species and its expression in Escherichia coli. Eur J Biochem 181: 563-570).

The literature describes systems for expressing nitrile hydratases whose specific activities are between 4.2 and 12.2 U/mg of total protein in the case of Co-dependent nitrile hydratases from *R. rhodochrous* J1 (Kobayasjhi M, Nishiyama M, Nagasawa T, Horinouchi S, Beppu T, Yamada H (1991) Cloning, nucleotide sequence and expression in Escherichia coli of two cobalt-containing nitrole hydratase genes from Rhodococcus rhodochrous. Biochim Biophys Acta 1129: 23-33) and 452 U/mg of total protein in the case of an iron-dependent nitrile hydratase from *Rhodococcus spec.* N-771 (Njori M, Yohda M, Odaka M, Matsushita Y, Tsujimura M, Yoshida T, Dohmae N, Takio K Endo I (1999) Functional expression of Nitrile hydratases in E. coli: Requirement of

a nitrile hydratase activator and a post-translational modification of a ligand cysteine. J Biochem 125: 696-704), with this corresponding to about approx. 248 U/mg of DBM (dry biomass) (calculation in accordance with Goodsell DS (1991) Inside a cell. TIBS 16: 203-206). Interestingly, it was not possible to reproduce the latter activity using nitrile hydratases from *R. erythropolis*, which is closely related to *Rhodococcus spec.* N-711, and employing similar vector systems and arrangements of the structural genes. There was, therefore, still a need for processes and systems which enable the enzymes in question to be made available in a manner which is adequate for the industrial scale.

Methods which have thus far been described for screening for nitrile hydratases have been restricted to isolating microorganisms which exhibit a corresponding enzyme activity. These microorganisms have either been taken from existing strain collections or selectively propagated in what are termed enrichment media (Colquhoun JA, Heald SC, Li L, Tamaoka J, Kato C, Horikoshi K and Bull AT (1998a) Taxonomy and biotransformation activities of some deep-sea actinomycetes. Extremophiles 2: 269-277; Colquhoun JA, Mexson J, Goodfellow M, Ward AC, Horikoshi K and Bull AT (1998b) Novel rhodococci and other mycolate actinomycetes from the deep sea. Antonie van Leeuwenhoek 74: 27-40). A disadvantage of these screening methods is that they have in the main found microorganisms of the genera Rhodococcus, Bacillus or Pseudomonas, or the like, and the diversity of the corresponding nitrile hydratases in regard to activity or substrate specificity is also limited as a consequence (Bunch A. W. (1998), Nitriles, in: Biotechnology, Volume 8a, Biotransformations I, Chapter 6, Eds.: Rehm H.J., Reed G., Wiley-VCH, p. 277-324; Cowan D, Cramp R, Pereira R, Graham D, Almathawa Q (1998) Biochemistry and biotechnology of mesophilic and thermophilic nitrile metabolising enzymes. Extremophiles 2: 207-216; Yamada H, Kobayashi M

(1996) Nitrile hydratase and its application to industrial production of acrylamide. Biosci Biotechnol Biochem 60: 1391-141).

Current estimates assume that, as a rule, only 0.01 - 1% of the microorganisms in a habitat can be cultured and thereby made available for screening in accordance with the above-described method (Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. 59: 143-169; Pace NR (1997) A molecular view of microbial diversity and the biosphere. Science 276: 734-740; Cowan DA (2000) Microbial genomes-the untapped resource. Trends Biotechnol. 18: 14-16). "Directly cloning" the genomic DNA of, if at all possible, all the organisms in a soil sample (i.e. the "metagenome"), and making this DNA available, in the form of metagenome gene libraries, for a genetic screening, is therefore increasingly gaining importance for identifying novel industrial enzymes. In this approach, the genetic screening for enzyme-encoding genes can either be effected on the basis of sequence homology, based on conserved sequence motifs, or, when suitable enzyme tests/indicator media are available, on the basis of activity homology (Lorenz P, Köhler B, Wolf M, Eck J, Zinke H (2000) Expression Cloning of Metagenome DNA from Soil. Biotechnol. 2000, Book of Abstr. Vol 2: 306).

While nitrile hydratases have already been amplified from metagenomic DNA by means of a PCR-based screening method using degenerate primers, the nature of the sequences is such that they exhibit very high degrees of similarity (90 - 99%) with those for known nitrile hydratases (Precigou S, Goulas P, Duran R, (2001) Rapid and specific identification of nitrile hydratase encoding genes in soil samples by polymerase chain reaction, FEMS Microbiol. Letters 204: 155-161). It is not possible to make any proper assessment of the primers which are used in this

reference since the authors do not reveal the sequences of the degenerate primers or their degree of degeneracy. The high degree of similarity of the sequences to those for the *Rhodococcus rhodochrous* J1 nitrile hydratases suggests that the substrate specificity is not markedly different from that of this latter enzyme, either.

There was, therefore, still a need for processes and systems which enable additional enzymes of the type in question to be made available in a manner which is adequate for the industrial scale.

The object of the present invention was therefore to specify an additional process for preparing nitrile hydratases. In particular, the process should be able to identify nitrile hydratases which are present in what are termed nonculturable organisms. An additional object of the present invention was to prepare nitrile hydratases which are superior to those in the prior art.

These objects, and additional objects which are not specified in detail but which ensue in an obvious manner from the prior art, are achieved by the specification of specific primer constituents as claimed in claim 1 and their use in a process which has the features of claim 2. Claims 3 to 5 relate to preferred embodiments of the process in question. Claim 6 protects the protein sequences according to the invention while claim 7 is directed toward the nucleic acid sequences which encode them and claim 8 relates to the expression systems which are provided with these nucleic acid sequences. Claim 9 deals with novel nitrile hydratases which have been prepared in accordance with the invention. Claims 10 and 11 are directed toward specific uses.

A prerequisite for finding novel nitrile hydratases is to specify nucleic acid sequences which are able to serve as

probes for nitrile hydratase genes which are present in metagenome DNA libraries. By means of specifying degenerate primer constituents from the group consisting of

A-01f : gcsmrsgcstgg (Seq. ID NO. 1)
B-01f : ggsctsccscc (Seq. ID NO. 2)
B-01r : ggsggsagscc (Seq. ID NO. 3)
C-01r : ggncgcwbsgg (Seq. ID NO. 4)
A-01f : gcnmrrgcntgg (Seq. ID NO. 5)
B-01f : ggnytnccncc (Seq. ID NO. 6)
B-01r : ggnggnarncc (Seq. ID NO. 7)
C-01r : gwngwrtccca (Seq. ID NO. 8)
A-01f : gcntggrynga (Seq. ID NO. 9)
B-01f : ggnytsccncc (Seq. ID NO. 10)
B-01r : ggnggsarncc (Seq. ID NO. 11)
C-01r : swnswrtccca (Seq. ID NO. 12)

the skilled person obtains, completely surprisingly but no less advantageously for that, special nucleic acid sequences which help to construct specific probes for screening for nitrile hydratase genes in metagenome DNA libraries. These nucleic acid sequences are degenerate sequences which, on the one hand, are specific enough for finding only nitrile hydratase genes but which, on the other hand, are so nonspecific that, if at all possible, all the nitrile hydratase genes which are present are detected. At the time of the invention, it was not possible to deduce their preparation from the prior art in an obvious manner.

The present invention accordingly also related to a process for preparing protein sequences which are required for constructing the activity of a nitrile hydratase, such that
a) a metagenome DNA library of a habitat is prepared,
b) this library is contacted with in each case at least one forward(f) primer and one reverse(r) primer exhibiting a

degenerate nucleic acid sequence in accordance with sequences 1 to 12,

c) a PCR is carried out using these primers,

d) the full-length sequences of the nucleic acids encoding protein sequences which are required for constructing the activity of a nitrile hydratase are generated from the part sequences which are obtained, and

e) these full-length sequences are cloned into a host organism and expressed.

As already noted at the outset, nitrile hydratases consist of at least two different subunits (α and β subunits). However, further protein sequences, in addition to these two subunits, may also be required to enable the nitrile hydratases to be active. The presence of particular putative "activators" (e.g. folding proteins, etc.) may sometimes be required in order to enable the corresponding nitrile hydratases to become active. The nucleic acid sequences which encode these activators are frequently located in the immediate vicinity of the nucleic acid sequences which encode the corresponding nitrile hydratase subunits. It is consequently possible, by screening for nucleic acid sequences encoding nitrile hydratases, also at the same time to detect all the protein sequences which are required for the nitrile hydratases to be active. According to the process in accordance with the invention, the first step is that of preparing a metagenome DNA library of a particular habitat. The skilled person is familar with the way in which this library is prepared (Knietsch, AW, Tanja; BS; Henne, ADR (2003) Metagenomes of Complex Microbial Consortia Derived from Different Soils as Sources for Novel Genes Conferring Formation of Carbonyls from Short-Chain Polyols on Escherichia coli. Journal of Molecular Microbiology and Biotechnology 5(1): 46-56; Rondon, MR; August, PR; Bettermann AD; Brady, SF; Grossman, T H; Liles, MR; Loiacono, KA; Lynch, BA; MacNeil, IA; Mino,r C; Tiong, CL; Gilman, M; Osburne, MS; Clardy, J; Handelsman, J;

Goodman, RM (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Applied and environmental microbiology 66(6):2541-7). Primers which exhibit the degenerate nucleic acid sequences according to the invention (Seq. ID Nos. 1-12) are then added to this library. A PC reaction is subsequently carried out, with this reaction giving rise to part sequences of the nucleic acid sequences encoding nitrile hydratase subunits. Employing methods of the prior art, the skilled person can then use these part sequences to identify the corresponding full-length nucleic acid sequences (Schloss, PD; Handelsman, J(2003) Biotechnological prospects from metagenomics. Current Opinion in Biotechnology, 14(3): 303-310; Rondon, MR; August, PR; Bettermann AD; Brady, SF; Grossman, T H; Liles, MR; Loiacono, KA; Lynch, BA; MacNeil, IA; Mino,r C; Tiong, CL; Gilman, M; Osburne, MS; Clardy, J; Handelsman, J; Goodman, RM (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Applied and environmental microbiology 66(6):2541-7). Finally, the nucleic acid sequences which have been found are recombinantly prepared in particular expression systems. The methods for doing this are likewise known to the skilled person (lit. see above).

In a preferred embodiment, the degenerate nucleic acid sequences according to the invention (Seq. ID Nos. 1-12) are employed in the present process such that in each case primer pairs composed of primers exhibiting the nucleic acid sequences A-01f (Seq. ID Nos. 1, 5 and 9) and B-01r (Seq. ID Nos. 3, 7 and 11) or C-01r (Seq. ID Nos. 4, 8 and 12) and also B-01f (Seq. ID Nos. 2, 6 and 10) and C-01r (Seq. ID Nos. 4, 8 and 12) are used in the PCR. When these combinations are employed, the nucleic acid sequences encoding the protein sequences which are required for

nitrile hydratase activity are detected in a preferred and efficient manner.

Preference is furthermore given to locating certain other nucleic acid sequences (e.g. "stabilizing regions") upstream of the above-described degenerate primer constituents (Kwok S, Chang SY, Sninsky JJ, Wang A, 1995, "Design and use of mismatched and degenerate primers" In: "PCR Primer, A laboratory Manual" Dieffenbach CW & Dveksler GS (Editors), Cold Spring Harbor Laboratory Press, pp143-155; Compton T, 1990, "Degenerate Primers for DNA Amplification" In: "PCR Protocols, A Guide to Methods and Applications", Innis MA, Gelfand DH, Sninsky JJ, White TJ (Editors) Academic Press, San Diego, pp39-45). In this case, the primers employed in the PC reaction consist of degenerate nucleic acid sequences of the above-described type (sequences ID Nos. 1 to 12) and the nucleic acid sequences mentioned in the sequences having the ID Nos. 13 to 23. Very particular preference is given, therefore, to a process in which nucleic acid sequences selected from the group consisting of:

GCCAAGGTCGTC	(Seq. ID NO. 13)
GGCCGGTCCTG	(Seq. ID NO. 14)
TCCTTGTACCAGGTC	(Seq. ID NO. 15)
GCCCGCC	(Seq. ID NO. 16)
GGCGCTAATGTTGTT	(Seq. ID NO. 17)
TGGCCGGTTCTG	(Seq. ID NO. 18)
CAAATTCTTTATACCAAGTC	(Seq. ID NO. 19)
CCATATATCGCATTTCAGCT	(Seq. ID NO. 20)
GGTCGTGGCCAAG	(Seq. ID NO. 21)
GGCCGGTCCTG	(Seq. ID NO. 22)
TCCTTGTACCAGGTC	(Seq. ID NO. 23)
GCGCATTTCGGCG	(Seq. ID NO.

are placed upstream of the degenerate nucleic acid sequences (Seq. ID Nos. 1-12). These upstream sequences are likewise derived from conserved nitrile hydratase regions and adapted to the codon usage of organisms having a different GC content.

A process as described at the outset, in which use is made of primers which are selected from the group consisting of

GCCAAGGTCGTCgcsmrsgcstgg (Seq. ID NO. 25)
GGCCGGTCCTGgggsctsccscc (Seq. ID NO. 26)
TCCTTGTACCAGGTCggsggsagscc (Seq. ID NO. 27)
GCCCGCCggncgcwbsgg (Seq. ID NO. 28)
GGCGCTAAAGTTGTTgcnmrrgcntgg (Seq. ID NO. 29)
TGGCCGGTTCTGggnytnccncc (Seq. ID NO. 30)
CAAATTCTTTATACCAAGTCggnggnarncc (Seq. ID NO. 31)
CCATATATCGCATTTCAGCTgwngwrtccca (Seq. ID NO. 32)
GGTCGTGGCCAAGgcntggrynga (Seq. ID NO. 33)
GGCCGGTCCTGggnytsccncc (Seq. ID NO. 34)
TCCTTGTACCAGGTCggnggsarncc (Seq. ID NO. 35)
GCGCATTTCGGCGswnswrtccca (Seq. ID NO. 36)

is therefore very particularly advantageous.

Using these primers, it was possible to detect nucleic acid sequences encoding nitrile hydratases, as well as other genes for putative "activators", in metagenome DNA libraries.

Accordingly, a next part of the subject matter of the present invention is constituted by the protein sequences which are required for constructing the activity of a nitrile hydratase, with these sequences possessing less than 100% homology, preferably less than 97%, more preferably less than 96%, even more preferably less than 95%, yet more preferably less than 90%, very preferably less than 85%, and extremely preferably less than 80%, homology, at the amino acid level, with such known protein sequences, and with the nucleic acid sequences encoding them being generated from part sequences which give a positive hybridization signal, under stringent conditions, with the primers according to the invention exhibiting the nucleic acid sequences having the sequences ID Nos. 1 to 12.

The positive hybridization is a prerequisite for it to be possible to find corresponding nucleic acid sequences using the PC reaction-based screening. Methods with which the skilled person is familar can then be used to obtain the corresponding recombinant protein sequences from these nucleic acid sequences.

These recombinant techniques are used to obtain organisms which are able to provide the protein sequence in question in a quantity which is adequate for an industrial process. The rec-protein sequences according to the invention are prepared using recombinant DNA methods which are known to the skilled person (Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York; Balbas, P. and Bolivar, F. (1990), Design and construction of expression plasmid vectors in E.coli, Methods Enzymol. 185, 14-37; Rodriguez, R.L. and Denhardt, D. T (eds) (1988), Vectors: a survey of molecular cloning vectors and their uses, 205-225, Butterworth, Stoneham). With regard to the general procedure (PCR, cloning, expression, etc.), the reader may be referred to the following literature and the references cited therein: Universal GenomeWalker™ Kit User Manual, Clontech, 3/2000, and the literature cited therein; Triglia T.; Peterson, M. G. and Kemp, D.J. (1988), A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences, Nucleic Acids Res. 16, 8186; Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York; Rodriguez, R.L. and Denhardt, D. T (eds) (1988), Vectors: a survey of molecular cloning vectors and their uses, Butterworth, Stoneham.

The corresponding nucleic acid sequences, which encode the protein sequences which have just been described, also constitute a next part of the present invention. These nucleic acid sequences are therefore all those which,

within the bounds of the degeneracy of the genetic code, encode the same protein sequence. This consequently likewise also encompasses those nucleic acid sequences which possess, at the nucleic acid level, a homology with the nucleic acid sequences found in accordance with the invention of at least 70 percent, or corresponding fragments of these nucleic acid sequences, which in turn encode protein sequences which are involved in constructing the activity of a nitrile hydratase. Preference is given to these nucleic acid sequences encoding protein sequences which are improved as compared with the protein sequences found in accordance with the invention.

Examples of nucleic acid sequences according to the invention are those of the odd-numbered Seq. ID Nos. 37 to 85. It is possible, as described above, to use the nucleic acid sequences which have been found to obtain the protein sequences according to the invention in high yields from rapidly growing host organisms, e.g. *E. coli*.

This takes place by incorporating (cloning) the nucleic acid sequences according to the invention into special expression systems which can be used to obtain the corresponding protein sequences recombinantly from preferred host organisms. A next aspect of the present invention is therefore constituted by a (articially prepared) expression system which exhibits one or more of the nucleic acid sequences according to the invention. In principle, all the systems which are familiar to the skilled person for this purpose are suitable for use as the expression system. These systems are preferably plasmids or vectors and microorganisms.

In principle, all the embodiments which are available to the skilled person for this purpose are suitable for use as plasmids or vectors. These plasmids and vectors can be found, for example, in Studier and coworkers (Studier, W. F.; Rosenberg A. H.; Dunn J. J.; Dubendroff J. W.; (1990),

Use of the T7 RNA polymerase to direct expression of cloned genes, Methods Enzymol. 185, 61-89) or in the brochures supplied by the companies Novagen, Promega, New England Biolabs, Clontech or Gibco BRL. Other preferred plasmids and vectors can be found in: Glover, D. M. (1985), DNA cloning: a practical approach, Vol. I-III, IRL Press Ltd., Oxford; Rodriguez, R.L. and Denhardt, D. T (eds) (1988), Vectors: a survey of molecular cloning vectors and their uses, 179-204, Butterworth, Stoneham; Goeddel, D. V. (1990), Systems for heterologous gene expression, Methods Enzymol. 185, 3-7; Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York.

Plasmids which can very preferably be used to clone the construct containing the nucleic acid sequences according to the invention into the host organism are: pUC18 (Roche Biochemicals), pKK-177-3H (Roche Biochemicals), pBTac2 (Roche Biochemicals), pKK223-3 (Amersham Pharmacia Biotech), pKK-233-3 (Stratagene) or pET (Novagen). Extreme preference is given to plasmids of the pET series.

As has been said, the recombinant microorganism into which the plasmids or vectors which contain the nucleic acid sequences according to the invention are cloned is used for replicating and isolating an adequate quantity of the recombinant protein sequence. The methods for doing this are well known to the skilled person (Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York). In principle, the microorganisms which are used can be any organisms which the skilled person takes into consideration for this purpose, for example yeasts, such as *Hansenula polymorpha*, *Pichia sp.*, *Saccharomyces cerevisiae*, prokaryotes, such as *E. coli* or *Bacillus subtilis*, or eukaryotes, such as mammalian cells, insect cells or plant cells. Preference is given to using *E. coli* strains for this purpose. The following are very

particularly preferred: *E. coli* XL1 Blue, NM 522, JM101, JM109, JM105, RR1, DH5α, TOP 10⁻ or HB101, BL21, BL21 (DE3), or BL21 (DE3) codon plus RIL.

Furthermore, the nucleic acid sequences which are in accordance with the invention (uneven-numbered Seq. ID Nos. 37 to 85) and which are additionally further improved, and which encode the protein sequences which are required for constructing the activity of a nitrile hydratase, are preferably suitable for preparing what are termed whole-cell catalysts. In principle, whole-cell catalysts of this type are recombinant microorganisms such as those which have just been mentioned. However, in addition to the cloned genes encoding a nitrile hydratase, these whole-cell catalysts also comprise additional enzymes which are required for decomposing nitriles into acids. As explained at the outset, these enzymes are enzymes which exhibit amidase activity.

For this reason, whole-cell catalysts which comprise at least one cloned gene for a protein sequence possessing (D- or L-)amidase activity and cloned genes encoding an active nitrile hydratase are also regarded as being recombinant microorganisms of the abovementioned type. Optionally, the whole-cell catalyst can comprise additional nucleic acid sequences which encode enzymes which are advantageous for decomposing a nitrile function into an acid function. These enzymes are, in particular, enzymes which are selected from the group of protein sequences possessing α-aminonitrile racemase activity, possessing cyanohydrin racemase activity, possessing α-hydroxycarboxylic acid racemase activity or possessing (α- or β-)-amino acid amide racemase activity.

In addition to producing the protein sequences according to the invention, which are required for constructing a nitrile hydratase activity, the whole-cell catalyst according to the invention preferably produces a protein sequence possessing L-amidase activity from rhizobium,

preferably *R. huautlense* DSM 14983 (WO2004/005517) or possessing D-amidase activity, e.g. that from *Variovorax* (EP 1318193).

Corresponding racemases are known, for example, from *Pseudomonas putida* and *Rhodococcus sp.* (Godtfredsen, S. E.; Clausen, K.; Ingvorsen, K.; Hermes, H. F.; Van Balken, J. A.; Meijer, E. M. (1989, EP 0 307 023; WO 8 901 525). Other amino acid amide racemases have been described in *Klebsiella oxytoca* by Hermes and coworkers (Hermes, H. F. M.; Peeters, W. P.; Peters, P. J. (1990), EP 0 383 403), as well as in *Agrobacterium rhizogenes* and *Ochrobacterium anthropi* (Boesten, W. H. J.; Raemakers-Franken, P. C.; Sonke, T.; Euverink, G. J. W.; WO 03106691). The advantage of using corresponding racemases is based on the fact that 100% of a racemic nitrile can be converted into the corresponding enantiomerically enriched acid.

An organism which is mentioned in DE10155928 as being a host organism is preferably used as whole-cell catalyst. The advantage of such an organism is that several enzyme systems are expressed simultaneously, which means that it is only necessary to grow one rec-organism for reacting a readily preparable nitrile or cyanohydrin or α-aminonitrile to give the corresponding enantiomerically enriched acid.

In order to adjust the expression of the nucleic acid sequences in question with regard to the turnover rates of the protein sequences (enzymes) which they encode, the corresponding nucleic acid sequences can be installed on different plasmids possessing different copy numbers and/or promoters of differing strength can be used so as to ensure that the nucleic acid sequences are expressed at different strengths. In enzyme systems which have been adjusted in this way, there is, advantageously, no accumulation of an intermediate compound, which might possibly have an inhibitory effect, and the reaction under consideration can take place at an optimal overall rate. However, this is sufficiently well known to the skilled person (Gellissen, G.; Piontek, M.; Dahlems, U.; Jenzelewski, V.; Gavagan, J.

W.; DiCosimo, R.; Anton, D. L.; Janowicz, Z. A. (1996), Recombinant Hansenula polymorpha as a biocatalyst. Coexpression of the spinach glycolate oxidase (GO) and the S. cerevisiae catalase T (CTT1) gene, Appl. Microbiol. Biotechnol. 46, 46-54; Farwick, M.; London, M.; Dohmen, J.; Dahlems, U.; Gellissen, G.; Strasser, A. W.; DE19920712).

It is consequently possible to use the present process to prepare the α and β subunits of nitrile hydratases using metagenome DNA libraries as the starting material. Accordingly, another part of the subject matter of the present invention is constituted by the nitrile hydratases which exhibit the protein sequences according to the invention for nitrile hydratase α subunits and β subunits and which can be prepared from the nucleic acid sequences which are made available by this process and which encode the α and β subunits according to the invention. As is demonstrated in the examples, active nitrile hydratases are also formed, in this connection, when any arbitrary α subunits are combined with any arbitrary β subunits. This thereby makes it possible to increase still further the diversity of possible nitrile hydratases.

Other parts of the subject matter of the present invention relate to the use of the nucleic acid sequences which are prepared by the process according to the invention for producing improved protein sequences which are required for constructing the activity of a nitrile hydratase. The way in which the skilled person proceeds when improving protein sequences on the basis of altering nucleic acid sequences is well known. In general, this is effected by using methods of mutagenesis. Any methods which are available to the skilled person for this purpose are suitable for being used as mutagenesis methods. In particular, these methods are saturation mutagenesis, random mutagenesis, in-vitro recombination methods and site-directed mutagenesis (Eigen, M. and Gardiner, W. (1984), Evolutionary molecular

engineering based on RNA replication, Pure Appl. Chem. 56, 967-978; Chen, K. and Arnold, F. (1991), Enzyme engineering for nonaqueous solvents: random mutagenesis to enhance activity of subtilisin E in polar organic media. Bio/Technology 9, 1073-1077; Horwitz, M. and Loeb, L. (1986), Promoters Selected From Random DNA-Sequences, Proc Natl Acad Sci USA 83, 7405-7409; Dube, D. and L. Loeb (1989), Mutants Generated By The Insertion Of Random Oligonucleotides Into The Active-Site Of The Beta-Lactamase Gene, Biochemistry 28, 5703-5707; Stemmer, P.C. (1994), Rapid evolution of a protein *in vitro* by DNA shuffling, Nature 370, 389-391 and Stemmer, P.C. (1994), DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. Proc Natl Acad Sci USA 91, 10747-10751).

The novel nucleic acid sequences which are obtained in this way are cloned into a host organism, and expressed, using the above-described methods, and the protein sequences which have been prepared in this way are detected by means of suitable screening methods and then isolated. In principle, any detection reactions which are possible for the molecules which are formed are suitable for the detection. Nitrile hydratase activities can be detected in a coupled enzymic test using amidases, with ammonium being formed as a by product. In principle, any possible reactions for detecting ammonia or ammonium ions, such as Nessler reagent (Vogel, A., I., (1989) Vogel´s textbook of quantitative chemical analysis, John Wiley & Sons, Inc., 5th ed., 679-698, New York), the indophenol reaction, also termed Berthelot´s reaction (Wagner, R., (1969) Neue Aspekte zur Stickstoffanalytik in der Wasserchemie, Vom Wasser, [Novel aspects of nitrogen analysis in water chemistry, about water] VCH-Verlag, vol. 36, 263-318, Weinheim), in particular the enzymic determination using glutamate dehydrogenase (Bergmeyer, H.,U., and Beutler, H.-O. (1985) Ammonia, in: Methods of Enzymatic Analysis, VCH-Verlag, 3rd Edition, Vol. 8: 454-461, Weinheim), or else

detection using ammonium-sensitive electrodes, are suitable for detecting it. Furthermore, HPLC methods are used for detecting amino acids, for example a derivative method based on o-phthaldialdehyde and N-isobutyryl-cysteine for enantiomerically separating amino acids (Brückner, H., Wittner R., and Godel H., (1991) Fully automated high-performance liquid chromatographic separation of DL-amino acids derivatized with o-Phthaldialdehyde together with N-isopropyl-cysteine. Application to food samples, Anal. Biochem. 144, 204-206). The amide which is directly formed by the nitrile hydratase reaction can likewise be detected using HPLC methods (e.g. reverse phase).

In a last embodiment of the present invention, the latter relates to the use of the nitrile hydratases according to the invention for preparing organic acid amides and acids, in particular enantiomerically enriched α-hydroxy acids or α-amino acids.

For this application, the nitrile hydratase in question can be used in free form, as a homogeneously purified compound or as a recombinantly prepared enzyme. Furthermore, the enzyme can also be employed as a constituent of an intact host organism or in combination with the disrupted, and, if desired, highly purified, cell mass of the host organism. It is likewise possible to use the enzyme in immobilized form (Sharma B. P.; Bailey L. F. and Messing R. A. (1982), Immobilisierte Biomaterialiern - Techniken und Anwendungen [Immobilized biomaterials - techniques and applications], Angew. Chem. 94, 836-852). The immobilization is advantageously effected by means of lyophilization (Paradkar, V. M.; Dordick, J. S. (1994), Aqueous-Like Activity of α-Chymotrypsin Dissolved in Nearly Anhydrous Organic Solvents, J. Am. Chem. Soc. 116, 5009-5010; Mori, T.; Okahata, Y. (1997), A variety of lipi-coated glycoside hydrolases as effective glycosyl transfer catalysts in homogeneous organic solvents, Tetrahedron Lett. 38, 1971-1974; Otamiri, M.; Adlercreutz, P.; Matthiasson, B. (1992),

Complex formation between chymotrypsin and ethyl cellulose as a means to solubilize the enzyme in active form in toluene, Biocatalysis 6, 291-305). Very particular preference is given to lyophilization in the presence of surface-active substances such as Aerosol OT or polyvinylpyrrolidone or polyethylene glycol (PEG) or Brij 52 (diethylene glycol monocetyl ether) (Kamiya, N.; Okazaki, S.-Y.; Goto, M. (1997), Surfactant-horseradish peroxidase complex catalytically active in anhydrous benzene, Biotechnol. Tech. 11, 375-378).

Greatest preference is given to immobilization on Eupergit®, in particular Eupergit C® and Eupergit 250L® (Röhm) (Eupergit.RTM. C, a carrier for immobilization of enzymes of industrial potential. Katchalski-Katzir, E.; Kraemer, D. M. Journal of Molecular Catalysis B: Enzymatic (2000), 10(1-3), 157-176).

Preference is also given to immobilization on Ni-NTA in combination with the polypeptide being provided with a His tag (hexahistidine) (Purification of proteins using polyhistidine affinity tags. Bornhorst, Joshua A.; Falke, Joseph J. Methods in Enzymology (2000), 326, 245-254).

It is likewise possible to conceive of using the polypeptides as CLECs (St. Clair, N.; Wang, Y.-F.; Margolin, A. L. (2000), Cofactor-bound cross-linked enzyme crystals (CLEC) of alcohol dehydrogenase, Angew. Chem. Int. Ed. 39, 380-383).

By using these procedures, it is possible to successfully generate polypeptides which are able to function in mixtures of aqueous and organic solvents, or in media which are entirely organic, from polypeptides which are rendered unstable by organic solvents.

The present invention describes the isolation of nitrile hydratases and their corresponding genes by using molecular genetic methods to tap the uncultured microbial diversity of different habitats. Using degenerate primers in connection with the PCR-based screening of metagenome DNA

libraries identifies nitrile hydratase genes and elucidates the part sequences of the PCR products which are obtained in this way. In a subsequent step, the complete DNA sequences of the genes are determined in order, after cloning and heterologous expression, to provide enzyme samples for activity profiling and investigating applications.

By means of rationally selecting soil samples, which might possibly contain nitriles, for preparing metagenome libraries (metagenome DNA libraries) and focusing on application-relevant substrates within the context of enrichment cultures, it is possible to enrich nitrile-converting microorganisms. In any case, the genetic screening in metagenome libraries provides a pool of corresponding nitrile hydratase genes for subsequent expression, which pool can, however, also serve as the basis for enzyme optimization by means of directional evolution.

Metagenome libraries which were prepared from four different habitats and sites (grassland, forest, sandy ecosystem and biofilm), and which contained more than 83 000 clones, were screened for identifying novel nitrile hydratases. The skilled person is familiar with the construction of such metagenome libraries (Schloss, PD; Handelsman, J (2003) Biotechnological prospects from metagenomics. Current Opinion in Biotechnology, 14(3): 303-310; Knietsch, AW, Tanja; BS; Henne, ADR (2003) Metagenomes of Complex Microbial Consortia Derived from Different Soils as Sources for Novel Genes Conferring Formation of Carbonyls from Short-Chain Polyols on Escherichia coli. Journal of Molecular Microbiology and Biotechnology 5(1): 46-56; Rondon, MR; August, PR; Bettermann AD; Brady, SF; Grossman, T H; Liles, MR; Loiacono, KA; Lynch, BA; MacNeil, IA; Mino,r C; Tiong, CL; Gilman, M; Osburne, MS; Clardy, J; Handelsman, J; Goodman, RM (2000) Cloning the soil

metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Applied and environmental microbiology 666): 2541-7). These clones together contain about 3000 MBp of DNA. The metagenome DNA libraries were screened for novel nitrile hydratases by means of a PCR screening which used degenerate oligonucleotides which were derived from conserved primary structural motifs of known nitrile hydratases. The degenerate primers employed were those depicted in sequence ID Nos. 25 to 36.

In this connection, the primers having the primer constituent A-01f can, in particular, in each case be combined with those having the primer constituents B-01r (size of the PCR product to be expected, approx. 210 bp) or C-01r (size of the PCR product to be expected, approx. 350 bp), while the primers having the primer constituent B-01f can be combined with those having the primer constituent C-01r (size of the PCR product to be expected, approx. 180 bp). In particular, the degenerate "core" part of the primers (underlined in sequence ID Nos. 25 to 36) is important for identifying novel nitrile hydratases while the nondegenerate segment can be varied.

The corresponding PC reaction can be carried out using known methods. When special polymerases are used, the PCR must be amended in accordance with the manufacturer's instructions.

It was possible to identify clones carrying genes for nitrile hydratases in all 5 of the metagenome libraries. The skilled person is familiar with methods for using the part sequences of the nitrile hydratase genes for identifying the gene-carrying clone and elucidating the full-length sequence (Schloss, PD; Handelsman, J (2003) Biotechnological prospects from metagenomics. Current Opinion in Biotechnology, 14(3): 303-310; Duran, R;

Nishiyama, M; Horinouchi, S; Beppy, T (1993) Characterization of nitrile hydratase genes cloned by DNA screening from Rhodococcus erythropolis. Biosci Biotech Biochem 57(8): 1323-1328).

It was thus possible to determine the full-length sequences (nucleic acid sequences) of a total of 12 genes encoding α subunits and a total of 10 genes encoding β subunits (figure 1). The sequences of three genes which could encode putative nitrile hydratase "activators" were also determined.

In order to prepare the nitrile hydratases, the genes encoding the α and β subunits were in each case placed separately under the control of the phage T7 gene 10 promoter. The vectors pET22b and, respectively, pET26b (Novagen, CN Bioscience, Inc.) were used for this purpose (figs. 5 and 6). Using a two-vector expression system made it possible to simply combine the nitrile hydratase subunits from different enzymes. Corresponding constructs were expressed in *E. coli* strain BL21 (DE3) CodonPlus RIL (Novagen, CN Bioscience, Inc.). The cells were incubated at 26°C using LB medium and induced with 0.5 mM IPTG on reaching a cell density of O.D.580 = 1.0. It was possible to deal with the problem, which occasionally arose, of protein aggregation following overexpression at 37°C by reducing the temperature down to 26°C. Coexpressing various chaperones (trigger factor, GroEL/GroES and Dnak/DnaJ/GrpE) did not, at 26°C, show any additional effect which went beyond that achieved by reducing the temperature.

The yields with regard to activity against benzonitrile varied considerably (fig. 2) with it not being possible to demonstrate any clear correlation between the protein quantity which was prepared and the activity which was produced. Thus, at 30 U/g of DBM, the yield in the case of the clone containing the nitrile hydratase M49bD9 (Seq. ID

No. 49/71) was relatively high even though it was not possible to detect any proteins for the nitrile hydratase in the SDS-PAGE analysis. This contrasts with the clone M12K24 (Seq. ID No. 39/63), which was determined to have an activity of about 2.5 U/g of DBM even though it was possible to overexpress both subunits relatively strongly and a substantial proportion could be found in the soluble fraction. The magnitude of the activity consequently depends to a high degree on the particular enzyme in question.

In the case of a variety of nitrile hydratases, an additional small open reading frame, which encodes a protein sequence which appears to be involved in activating the nitrile hydratase, was identified in the immediate vicinity of the genes for the subunits of the enzyme. These proteins were designated P12K (Seq. ID No. 81/83) or P14K (Seq. ID No. 85) since their molecular weight was about 12 or 14 kDa. While it was possible to prepare some nitrile hydratases in active form even without these P12K homologs, the presence of these proteins was essential for other nitrile hydratases to be expressed in active form. In order to investigate the influence of P12K homologs, whose genes were also found in three metagenome clones, on the expression of the corresponding nitrile hydratases, and also on that of nitrile hydratases from other clones, two of these genes were cloned into the vector pBBR1MCS5 (Kovach et al., 1995, Four new derivatives of the broad host range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassetes, Gene 166: 175-176) and, in this vector, placed under the control of the lac promoter. The corresponding constructs were designated pBBR5-P12K-M49bD9 and pBBR5-P12K-M3aG10 (figs. 6 and 7).

The nitrile hydratase genes M49bD9 (Seq. ID No. 49/71) and M3aG10 (Seq. ID No. 57/79) were expressed in the presence of the corresponding P12K homologs (Seq. ID No. 83 -

M49bD9; Seq. ID No. 85 - M3aG10) at 26°C in *E. coli* strain BL21 CodonPlus RIL. In both cases, it was possible to markedly overexpress the subunits of the nitrile hydratases.

The presence of the P12K homolog (Seq. ID No. 83) increased the activity of the clone M49bD9 nitrile hydratases (Seq. ID No. 49/71) by about a factor of 27, up to approx. 830 U/g of DBM (table 1). In the case of the clone M3aG10 enzyme (Seq. ID No. 57/79), it was possible to detect activity (approx. 23 U/g of DBM) for the first time under these conditions. These results verify that the presence of the P12K homologs (Seq. ID No. 85) can be crucial for increasing the activity yield.

Table 1: Activity of nitrile hydratases following coexpression with P12K homologs (Seq. ID Nos.83 and 85)

Clone	Without P12K	With P12K
M49bD9 (α,β)	30 U/g of DBM	826 U/g of DBM
M3aG10	0 U/g of DBM	23 U/g of DBM

Combining subunits from different metagenome clones opens up the possibility of generating nitrile hydratases which possess potentially novel substrate specificities. By combining α subunits with different β subunits, it would be possible to generate a great diversity of combinations of novel nitrile hydratases. Such a combination of subunits from unrelated nitrile hydratases has not thus far been disclosed in the literature.

In order to evaluate this possibility, the α subunits of clones M73dC9 and M15aA6 (Seq. ID Nos. 59 and 45), for which it was not possible to find any β subunits, were expressed with the β subunit of clone M12K24 (Seq. ID No. 63) in *E. coli* strain BL21 codon plus RIL.

Whereas both subunits could be expressed at approximately equal strengths when the a-M73dC9a/b-M12K24 (Seq. ID Nos. 59 and 63) pair was expressed, the α subunit of clone M15aA6 (Seq. ID No. 45) appears to have been expressed more strongly than does the β subunit of clone M12K24 (Seq. ID No. 63).

Combining the α subunit of clone M73dC9 (Seq. ID No. 59) and the β subunit of clone M12K24 (Seq. ID No. 63) surprisingly led to the formation of an active nitrile hydratase having an activity of approx. 0.07 U/g of DBM (fig. 3). This result verifies that it is in principle possible to prepare active enzymes by combining nitrile hydratase subunits from different clones. At the time of the invention, it was not possible to deduce this, as such, from the prior art.

Stringent conditions: in this present document, the expression "under stringent conditions" is understood as described in Sambrook et al. (Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York). A stringent hybridization in accordance with the present invention preferably exists when a positive hybridization signal is still observed after washing for 1 hour with 1 × SSC (150 mM sodium chloride, 15 mM sodium citrate, pH 7.0) and 0.1% SDS (sodium dodecyl sulfate) at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at 68°C, and more preferably for 1 hour with 0.2 × SSC and 0.1% SDS at 50°C, more preferably at 55°C, yet more preferably at 62°C and most preferably at 68°C.

Within the context of the invention, optically enriched (enantiomerically enriched, enantiomer enriched) compounds are understood as meaning the presence of one optical antipode at > 50 mol% in a mixture with the other antipode.

All types of single-stranded DNA, and DNA which is complementary thereto or double-stranded DNA (e.g. genomic DNA or cDNA), as well as RNA (e.g. mRNA), or mixtures thereof, are subsumed under the term nucleic acid sequences.

Improved nucleic acid sequences encode improved protein sequences. Improved protein sequences are those which exhibit an improvement, as compared with the original sequences, in regard to activity and/or selectivity and/or stability. According to the invention, this means that the proteins are more active and/or more selective or less selective, or are more stable under the reaction conditions employed. While the activity and the stability of the proteins should naturally be as high as possible for the industrial application, an improvement with regard to selectivity is said to have taken place when the substrate selectivity decreases but the enantioselectivity of the proteins is increased. This also applies to proteins as constituents of nitrile hydratases insofar as they help to confer the corresponding improved properties on the enzyme.

According to the invention, the claimed protein sequences and nucleic acid sequences also encompass those sequences which exhibit a homology (excluding the natural degeneracy) which is greater than 70% (with regard to the nucleic acid sequence) or 80% (with regard to the protein sequences), preferably greater than 90%, 91%, 92%, 93% or 94%, more preferably greater than 95% or 96% and particularly preferably greater than 97%, 98% or 99%, with one of these sequences, provided the mode of action or purpose of such a sequence is preserved. The expression "homology" (or identity), as used herein, can be defined by the equation H (%) = [1 - V/X] × 100, in which H denotes homology, X is the total number of nucleobases/amino acids in the comparison sequence and V is the number of different nucleobases/amino acids in the sequence to be considered, based on the

comparison sequence. In any case, the term nucleic acid sequences which encode polypeptides encompasses all the sequences which appear possible in accordance with the degeneracy of the genetic code.

Description of the figures:

Figure 1: Homology among the nitrile hydratase α and β subunits which were found by means of a genetic screening.

Figure 2: Activity of different metagenome nitrile hydratases toward benzonitrile following expression in *E. coli* BL21 (DE3) codon plus RIL, with □ and ■ without coexpression of the trigger factor (n.d.: not determined).

Figure 3: Activity of nitrile hydratases when subunits from different metagenome clones are combined.

Figure 4: The vector map shows the general arrangement of the α subunits in plasmid pET22, taking clone M49bD9 as an example.

Figure 5: The vector map shows the general arrangement of the β subunits in plasmid pET26 taking clone M49bD9 as an example.

Figure 6: The vector map shows the arrangement of the P12K protein from clone M49bD9 in plasmid pBBR5.

Figure 7: The vector map shows the arrangement of the P12K protein from clone M3aG10 in plasmid pBBR5.

Experimental section:

General PCR protocol:

Culturing microorganisms

The E. coli cells were cultured and stored as described in Sambrock et al. (Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York).

PCR protocol:

A typical PCR protocol is described below, with it being necessary to adapt the protocol in accordance with the manufacturer's instructions when using a different polymerase.

25 µl of HotStarTaq mastermix, Qiagen
(2.5 U of polymerase, 200 µM dNTPs, 1 × PCR buffer)

25 pmol of primer 01f

25 pmol of primer 01r

2 µl of template DNA (20 - 200 ng)

dist. water to 50 µl

Program:

Initial denaturation:	1 ×	15 min at 95°C
Cycle program:	35 ×	1 min at 95°C 1 min at 55 – 69°C 1 min at 72°C
Final elongation:	1 ×	7 min at 72°C

Last extension: 7 min at 72°C

Digestion with restriction enzymes

The DNA to be cut is provided with 5 U of restriction enzyme and the appurtenant buffer and, unless otherwise required, incubated at 37°C. Chromosomal DNA is digested with 10 U of enzyme. The incubation period is 1.5 - 2.5 hours.

Treatment with alkaline phosphatase

In order to prevent vectors which have only been cut with one restriction endonuclease from religating with themselves, the phosphate residue protruding at the 5´ end is removed with alkaline phosphatase. Circular DNA can only be formed once again by inserting a DNA fragment.

The vector which has been cut with a restriction endonuclease is incubated at 65°C for 15 minutes in order to stop the restriction endonuclease. The dephosphorylation buffer is then added and the vector is incubated with 1 U of shrimp alkaline phosphatase at 37°C for 10 min. The

enzyme is then separated off from the vector DNA by means of a subsequent gel electrophoresis.

Treatment with T4 DNA ligase

Vector and insert are used for the ligation in a ratio of 1:3. The volume is chosen to be as small as possible (7 - 20 µl). The mixture is incubated overnight at 16°C in ligation buffer and in the presence of 1 U of ligase.

Transformation

100 µl of competent cells are added by pipette to the ligation mixture, the constituents of which are mixed by being repeatedly drawn up into the pipette. After 30 min of incubation on ice, a heat shock step at 42°C is carried out for 45 sec and the mixture is incubated once again on ice for 2 min. 120-900 µl of SOC medium are added and the mixture is incubated at 37°C for 45 min while being agitated. The mixture is subsequently plated out and incubated overnight at 37°C.

Expression of metagenome nitrile hydratases

The constructs containing T7 promoters were expressed in accordance with the following protocol:
50 ml of LB_{amp100} medium containing 2 mM Fe citrate and in each case 50 µg of kanamycin and ampicillin/ml were inoculated in a ratio of 99:1 with an overnight culture. After an OD_{600} of approx. 0.5 had been reached, expression of the nitrile hydratases was induced with 1 mM IPTG (isopropylthiogalactoside). The cells were harvested approx. 24 hours after induction at 26°C.

Detection of activity using benzonitrile as substrate

The biotransformation was carried out on a 10 ml scale using approx. 100 mg of moist biomass (OD_{600} = 5) in potassium phosphate buffer (100 mM), pH 7.0. The incubation took place at 30°C and the substrate concentration was approx. 5 mM benzonitrile. Samples were taken every 5 - 10 min over a period of at most 1 hour. The sample volume was 100 µl and the reaction was stopped by adding 10 µl of 50% phosphoric acid.

The concentrations of benzonitrile and benzamide were then determined by means of HPLC:

Column: RP18 Phenomenex Hypersil ODS 5 µ column (with precolumn)
Mobile phase: 10 mM K2HPO4, (pH 2.3)
Flow rate: 1 ml/min
Wavelength: 202 nm
Injection volume: 20 µl
Duration of HPLC run: 12-15 min

The activity was calculated by calculating a µmol turnover after one minute, with one µmol corresponding to one U (unit). Specific activities are given in U per g of DBM or mg of protein.

Patent claims:

1. A degenerate primer constituent from the group consisting of

A-01f : gcsmrsgcstgg (Seq. ID NO. 1)
B-01f : ggsctsccscc (Seq. ID NO. 2)
B-01r : ggsggsagscc (Seq. ID NO. 3)
C-01r : ggncgcwbsgg (Seq. ID NO. 4)
A-01f : gcnmrrgcntgg (Seq. ID NO. 5)
B-01f : ggnytnccncc (Seq. ID NO. 6)
B-01r : ggnggnarncc (Seq. ID NO. 7)
C-01r : gwngwrtccca (Seq. ID NO. 8)
A-01f : gcntggrynga (Seq. ID NO. 9)
B-01f : ggnytsccncc (Seq. ID NO. 10)
B-01r : ggnggsarncc (Seq. ID NO. 11)
C-01r : swnswrtccca (Seq. ID NO. 12).

2. A process for preparing protein sequences which are required for constructing the activity of a nitrile hydratase, such that
a) a metagenome DNA library of a habitat is prepared,
b) this library is contacted with in each case at least one forward(f) primer and one reverse(r) primer exhibiting a degenerate nucleic acid sequence as claimed in claim 1,
c) a PCR is carried out using these primers,
d) the full-length sequences of the nucleic acids encoding protein sequences which are required for constructing the activity of a nitrile hydratase are generated from the part sequences which are obtained, and
e) these full-length sequences are cloned into a host organism and expressed.

3. The process as claimed in claim 2, characterized in that

in each case primer pairs composed of primers exhibiting the nucleic acid sequences A-01f and B-01r or C-01r and also B-01f and C-01r are used in the PCR.

4. The process as claimed in claim 2 and/or 3, characterized in that nucleic acid sequences selected from the group consisting of:

GCCAAGGTCGTC	(Seq. ID NO. 13)
GGCCGGTCCTG	(Seq. ID NO. 14)
TCCTTGTACCAGGTC	(Seq. ID NO. 15)
GCCCGCC	(Seq. ID NO. 16)
GGCGCTAATGTTGTT	(Seq. ID NO. 17)
TGGCCGGTTCTG	(Seq. ID NO. 18)
CAAATTCTTTATACCAAGTC	(Seq. ID NO. 19)
CCATATATCGCATTTCAGCT	(Seq. ID NO. 20)
GGTCGTGGCCAAG	(Seq. ID NO. 21)
GGCCGGTCCTG	(Seq. ID NO. 22)
TCCTTGTACCAGGTC	(Seq. ID NO. 23)
GCGCATTTCGGCG	(Seq. ID NO. 24)

are placed upstream of the degenerate nucleic acid sequences.

5. The process as claimed in one or more of the preceding claims 2 to 4, characterized in that use is made of primers which are selected from the group consisting of

GCCAAGGTCGTCgcsmrsgcstgg	(Seq. ID NO. 25)
GGCCGGTCCTGggsctsccscc	(Seq. ID NO. 26)
TCCTTGTACCAGGTCggsggsagscc	(Seq. ID NO. 27)
GCCCGCCggncgcwbsgg	(Seq. ID NO. 28)
GGCGCTAAAGTTGTTgcnmrrgcntgg	(Seq. ID NO. 29)
TGGCCGGTTCTGggnytnccncc	(Seq. ID NO. 30)
CAAATTCTTTATACCAAGTCggnggnarncc	(Seq. ID NO. 31)
CCATATATCGCATTTCAGCTgwngwrtccca	(Seq. ID NO. 32)
GGTCGTGGCCAAGgcntggrynga	(Seq. ID NO. 33)

GGCCGGTCCTGggnytsccncc (Seq. ID NO. 34)
TCCTTGTACCAGGTCggnggsarncc (Seq. ID NO. 35)
GCGCATTTCGGCGswnswrtccca (Seq. ID NO. 36).

6. A protein sequence which is required for constructing the activity of a nitrile hydratase and which has less than 100% homology, at the amino acid level, with such known protein sequences, with the nucleic acid sequences encoding it being generated from part sequences which give a positive hybridization signal, under stringent conditions, with the primers exhibiting the nucleic acid sequences of claim 1.

7. A nucleic acid sequence which encodes a protein sequence as claimed in claim 6.

8. An expression system which exhibits one or more nucleic acid sequences as claimed in claim 7.

9. A nitrile hydratase which exhibits protein sequences for α subunits and β subunits as claimed in claim 6.

10. The use of the nucleic acid sequences as claimed in claim 7 for producing improved protein sequences which are required for constructing the activity of a nitrile hydratase.

11. The use of the nitrile hydratases as claimed in claim 9 for preparing organic acid amides and acids.