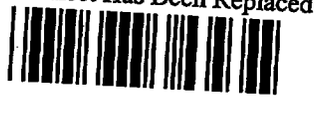
INDUSTRIAL PROCESS FOR PRODUCING HETEROLOGOUS PROTEINS IN *E. COLI* AND STRAINS USEFUL FOR SAID PROCESS

The present invention relates to a novel industrial process for producing heterologous proteins in *E. coli*. While for certain heterologous proteins with very high added value the cost price of the process for preparing them remains a factor which is negligible with compared to the purpose of the heterologous protein (in the pharmaceutical domain in particular), the development of the industrial production of heterologous proteins of lower added value in *E. coli* involves taking into account production factors such as the necessity of having an increased biomass and a very high content of heterologous proteins produced for the lowest possible cost, which cost should take account of the nature of the media, of the energetic and reagent yield and of the operating conditions. For industrial productions using reaction volumes which can reach several dozens of m^3, the simplest possible media and operating conditions will be sought. The present invention consists of the selection of an *E. coli* strain suitable for satisfying the conditions above, which are essential for economically satisfactory industrial production of heterologous proteins, independently of the value of the protein produced.

This Sheet Has Been Replaced

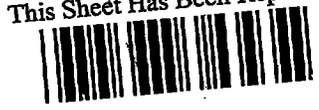


The strains of *E. coli* most commonly used for molecular biology studies derive from the strain K12 (Swartz. 1996, In Escherichia coli and Salmonella, Cellular and Molecular Biology, 2nd edition, ASM Press Washington, pp. 1693-1711). Derivatives of *E. coli* B, such as BL21, are also used for producing proteins, because of their physiological properties. A table of the strains most commonly used for producing recombinant proteins is given by Wingfield, 1997 (Current Protocols in Protein Science, Coligan *et al.* Ed. John Wiley & Sons, Inc. 5.0.1-5.0.3).

Many systems for expressing proteins in bacterial hosts have been described (Makrides, 1996, Microbiol. Rev. 60:512-538; Current Opinions in Biotechnology, 1996, 7). An expression system consists of a promoter, of its regulator, of a ribosome binding site followed by a restriction site which allows the insertion of the gene of interest, of a structure which can be used as a transcription terminator, optionally of genes the coexpression of which increases the quality of the protein of interest overexpressed, and of one or more vectors which make it possible to introduce these combinations into the host.

The promoter must have at least three properties in order to be used in a process for producing proteins (Makrides, 1996, mentioned above):

This Sheet Has Been Replaced



- it must be strong and cause the accumulation of the protein of interest, which can represent 10 to 50% of the total proteins of the host cell;
- it must be capable of being regulated so as to be able, as far as possible, to uncouple the biomass production phase from the protein production phase;
- it must be inducible (passage from a level of low transcriptional activity to a maximum level of transcriptional activity) using simple and inexpensive process conditions.

Many promoters have been described for expression in *E. coli* (Makrides, 1996, mentioned above; Weickert *et al.*, 1996, Current Opinions in Biotechnology 7 : 494-499). Among the homologous promoters used for producing proteins in *E. coli*, mention may be made of the *lac, trp, lpp, phoA, recA, araBAD, proU, cst-1, tetA, cadA, nar, tac, trc, lpp-lac, Psyn and cspA* promotors. Among the heterologous promoters used for producing proteins in *E. coli*, mention may be made of the *PL, PL-9G-50, PR-PL, T7,* λ*PL-PT7,* T3-*lac,* T5-*lac,* T4 gene 32, *nprM-lac,* VHb and Protein A promoters. A certain number of drawbacks are linked to these promoters. For some of them, mention may be made of the use of IPTG as the inducer molecule, the price of which can represent more than 14% of the cost of the medium. Others use regulation by temperature, which is difficult to implement on the scale of a 100 m^3 industrial fermenter.

This Sheet Has Been Replaced



The vectors most commonly used for expressing proteins in *E. coli* derive from the plasmid pBR322 (Swartz, 1996, mentioned above; Makrides, 1996, mentioned above). They are present in cells at a certain copy number, which is determined by the interaction of two RNAs encoded by the plasmid, RNAI and RAII (Polisky, 1988, Cell 55 : 929-932). The interaction of RNAI with RNAII inhibits the maturation of RNAII into a form required for the initiation of the replication of the plasmid. This interaction is modulated by the protein ROP, the gene of which is present on pBR322 but not on certain derivatives, such as the pUC-type plasmids (Lin-Chao and Cohen, 1991, Cell 65 : 1233-1242). With regard to regulation of the number of copies of the expression plasmid in *E. coli*, several strategies are mentioned (Swartz, 1996, mentioned above; Makrides, 1996, mentioned above). It will be appreciated in particular that a high number of copies of expression plasmid leads to a high level of messenger RNAs of the desired protein, but can be detrimental to the metabolism of the host cell (Bailey, 1993, Adv. Biochem. Eng. Biotechnol. 48 : 29-52).

The stability of the expression plasmids is an important criterion, all the more so given that industrial fermentations tend not to use antibiotics in the fermenters. Several strategies have been developed to stabilize expression plasmids, including the cloning

This Sheet Has Been Replaced



of the *cer* locus of the natural plasmid ColE1. This locus has been characterized (Leung *et al.*, 1985, DNA 4 : 351-355) and its insertion into multicopy plasmids has been described as having a beneficial effect on the stability of these plasmids (Summers and Sherratt, 1984, Cell 36 : 1097-1103).

While the strains and expression systems above make it possible to obtain good heterologous protein production yields, their use remains limited to the production of heterologous proteins with very high added value for which the cost price of the production system (bacterial strain, culture medium and conditions, raw materials) is minimal compared with the value of the protein produced. As examples of such proteins with very high added value, there are more particularly the heterologous proteins intended for pharmaceutical use, such as for example human growth factor, human alpha consensus interferon, human interleukins 1β, α1 and 2, human leukocyte interferon, human parathyroid hormone, human insulin, human serum albumin or human proapolipoprotein A-1 (Lee, 1996, Trends in Biotechnol. 14:98-105; Latta *et al.*, 1987, Bio/Technology 5 : 1309-1313).

However, for the mass production of chemical intermediates (Lee, 1997, Nature Biotech. 15 : 17-18) or for the production of enzymes for industrial use, in particular of the catalysts required for producing

This Sheet Has Been Replaced



chemical compounds, the cost price of the production system becomes a dominant factor to be taken into consideration in order to evaluate the technical advantage of said system.

For the production of heterologous proteins in bacteria, the productivity of the culture system employed can be significantly increased by using high cell density culturing strategies (S. Makrides, 1996, mentioned above; Wingfield, 1997, mentioned above). Among these is the fed-batch strategy (Jung *et al.*, 1988, Ann. Inst. Pasteur/Microbiol. 139 : 129-146; Kleman *et al.*, 1996, Appl. Environ. Microbiol. 62 : 3502-3507; Lee, 1996, mentioned above; Bauer and White, 1976, Biotechnol. Bioeng. 18 : 839-846). This strategy, combined with the use of a P_{trp} promoter, has made it possible to achieve significant productivities: 55 g of dry weight per liter, and 2.2 g of heterologous protein per liter (Jung *et al.*, 1988, mentioned above). Routine productions of 35 to 50 g of dry weight per liter are reported (Wingfield, 1997, mentioned above).

However, the strains and systems above do not make it possible to obtain culture densities which are sufficient for the industrial production of heterologous proteins for which the value (cost price) must be negligible compared to their purpose (in particular for the preparation of biological catalysts).

This Sheet Has Been Replaced



The present invention lies in the selection of a specific strain of *E. coli*, which is suitable for the industrial production of heterologous proteins. The strain which is useful for the process according to the invention is an *E. coli* strain W, more particularly the strain W referenced at the ATCC under the number 9637.

This strain W (ATCC 9637) is well known, and described in many publications (Davies & Mingioli, 1950, J. Bact., 60: 17-28; Doy and Brown, 1965, Biochim. Biophys. Acta, 104: 377-389; Brown and Doy, 1966, Biochim. Biophys. Acta, 118: 157-172; Wilson & Holden, 1969, J. Biol. Chem., 244: 2737-2742; Wilson & Holden, 1969, J. Biol. Chem., 244: 2743-2749; White, 1976, J. Gen. Microbiol., 96: 51-62; Shaw & Duncombe, 1963, Analyst 88: 694-701; Br. Pharmacopoeia, 1993, 2: A164-A169; Huang *et al.*, US 3,088,880; Hamsher *et al.*, US 3,905,868; Takahashi *et al.*, US 3,945,888; Huang *et al.*, US 3,239,427; Burkholder, 1951, Science, 114: 459-460; Prieto *et al.*, 1996, J. Bact., 178: 11-120; Lee 1996, mentioned above; Lee & Chang, 1995, Can. J. Microbiol, 41: 207-215; Lee *et al.*, 1994, Biotechnol. Bioeng., 44: 1337-1347; Lee & Chang, 1993, Biotechnology Letters. 15: 971-974; Bauer and White, 1976, mentioned above; Bauer and Shiloach, 1974, Biotechnol. Bioeng 16: 933-941; Gleiser and Bauer, 1981, Biotechnol. Bioeng., 23: 1015-1021; Lee and Chang, 1995, Advances in Biochem. Engine./Biotech.

This Sheet Has Been Replaced



52: 27-58). The strain W (ATCC9637) has thus been used for the production of 3-polyhydroxybutyric acid (PHA) after introduction of a plasmid carrying the operon of *Alcaligenes eutrophus* encoding enzymes involved in the PHA biosynthesis (Lee and Chang, 1993, mentioned above; Lee and Chang, 1995, mentioned above; Lee et al., 1994).

The strain W has also been used in high cell density cultures (Bauer and White, 1976, mentioned above; Bauer and Shiloach, 1974, mentioned above; Gleiser and Bauer, 1981, mentioned above; Lee and Chang, 1993, mentioned above; Lee et al., 1997, Biotechnology Techniques 11: 59-62). Biomasses of 125 g of dry weight per liter have thus been obtained (Lee and Chang, 1993, mentioned above) using sucrose as a carbon source.

However, this strain has never been described for the production of recombinant proteins. Furthermore, in combining a plasmid carrying the operon of *Alcaligenes eutrophus* encoding enzymes involved in PHA biosynthesis and a strategy of culturing the corresponding recombinant strain W at high cell density, Lee and Chang (1993, mentioned above) obtained worse PHA productivity than with a strain XL1-Blue derived from the strain K12 (Lee and Chang, 1995, mentioned above; Lee, 1996, mentioned above).

This Sheet Has Been Replaced



The present invention relates, therefore, to an industrial process for preparing heterologous proteins in *E. coli*, in which *E. coli* bacteria modified with a suitable system for expressing heterologous proteins are seeded and cultured in a suitable culture medium, characterized in that the strain of *E. coli* is an *E. coli* strain W. More preferably, the strain W is the strain W deposited at the ATCC under the number 9637.

According to one particular embodiment of the invention, the strain W is a derivative of the strain deposited at the ATCC under the number 9637, obtained by clonal selection or genetic manipulation.

According to the invention, the term "industrial process" is intended to mean any process in which the bacterial culture volume is greater than the usual culture volume employed in research laboratories. Generally, the term "industrial process" is intended to mean any process for which the culture volume is greater than 2 liters, preferably greater than or equal to 10 liters, more preferably greater than or equal to 20 liters, even more preferably greater than or equal to 50 liters. The process according to the invention is particularly suitable for culture volumes from several dozens of m^3 up to more than 100 m^3.

The suitable culture medium is a culture medium which is suitable for the production of a high

This Sheet Has Been Replaced



density of biomass and a high content of heterologous proteins produced. Several types of medium (defined, complex and semidefined) can be used for high cell density culturing (Lee, 1996, mentioned above). While the known media of the prior art, and in particular semidefined media, make it possible to accumulate good reproducibility of the composition of the medium and good productivity of the culture (Lee, 1996, mentioned above), the development of such a medium requires, however, empirical optimization for taking into account the economic constraints set out previously (Lee, 1996, mentioned above).

According to one preferential embodiment of the invention, the culture medium comprises sucrose as the main carbon source. According to the invention, the expression "main carbon source" is intended to mean that the sucrose represents at least 50% by weight of the total weight of the carbon sources of the culture medium, more preferably at least 75% by weight, even more preferably at least 85% by weight. According to a more preferential embodiment of the invention, the culture medium comprises substantially only sucrose as a carbon source. It is understood that, for the process according to the invention, the culture medium can comprise suitable additives so as to increase the overall yield of the invention. These additives can have the ancillary function of behaving as a carbon

This Sheet Has Been Replaced



source to the bacterial culture. However, these additives will not be considered as a carbon source for the purpose of the present invention if the *E. coli* W bacteria used in the process according to the invention cannot grow on said additives as the sole carbon source.

Advantageously, the amount of sucrose in the culture medium of the process according to the invention is between 0.1 and 300 g/l at the start of culturing, (before seeding), preferably between 0.5 and 200 g/l. It is understood that, since the sucrose constitutes the main carbon source of the medium according to the invention, the amount of sucrose will be decreasing during the process. In general, at the end of the reaction, the amount of sucrose in the culture medium at the end of the reaction is between 0 and 10 g/l.

According to one advantageous embodiment of the invention, the suitable culture medium also comprises a supplementary organic nitrogen source. This supplementary organic nitrogen source can consist of all organic nitrogen sources known to a person skilled in the art. Preferably, the supplementary organic nitrogen source consists of protein extracts. These protein extracts have more preferably the following composition: (in g amino acids per 100 g of product) alanine between 10 and 4, aspartic [lacuna] between 11

This Sheet Has Been Replaced



and 4, glycine between 22 and 2.5 and lysine between 7 and 4. Meat or potato peptones or proteins satisfy such a profile, is/are particularly preferred for the process according to the invention, more particularly the derivatives of potato proteins are preferred.

According to the invention, the expression "suitable system for expressing heterologous proteins" is intended to mean any expression system comprising regulation elements suitable for the expression of heterologous proteins in *E. coli* W. These regulation elements comprise in particular promoters, ribosome binding sites and transcription terminators.

Advantageously, the expression system comprises a P_{trp} promoter. The P_{trp} promoter has been used in several examples (EP Application 0 198 745; CIP Application No. 08/194,588; Application WO 97/04083; Latta *et al.*, 1987, Bio/Technology 5: 1309-1314; Denèfle *et al.*, 1987, Gene 56: 61-70). In particular, Latta *et al.* (1990, DNA Cell. Biol. 9: 129-137) have conducted a detailed study on the influence of regulatory sequences upstream of the promoter, and of tandem-duplicated promoter sequences, and on the influence of the coexpression of the TrpR repressor. Their reference construct, pXL534, was used as a basis for the construction of pXL642 (CIP Application No. 08/194,588), used in the examples which illustrate the present invention. Preferably, the P_{trp} promoter

This Sheet Has Been Replaced



comprises the nucleic acid sequence represented by sequence identifier No. 1 (SEQ ID NO 1).

According to one embodiment of the invention, in order to improve the level of expression of the heterologous protein, a coexpression of the molecular chaperones of *E. coli* GroESL (review by Makrides, 1996, mentioned above) is carried out. The increase in the intracellular concentration of the GroESL proteins makes it possible, in effect, to assist the folding of the recombinant protein and thus improve the level of active protein (Weicker et al., 1996, Curr. Opin. Biotechnol. 7: 494-499). The genes whose coexpression promotes the expression of the heterologous protein according to the invention, and its quality, are included in the expression system according to the invention.

According to the invention, the term "heterologous protein" is intended to mean any protein produced by the process according to the invention which is not naturally found in *E. coli* W, in the suitable expression system according to the invention. It can be a protein of nonbacterial origin, for example of animal, in particular human, or plant origin, or a protein of bacterial origin which is not naturally produced by *E. coli* W, or a protein of bacterial origin naturally produced by a bacterium other than *E. coli* W or a protein naturally produced by *E. coli* W, the

This Sheet Has Been Replaced

expression of which is controlled by regulation elements different from those of the expression system according to the invention, or finally, a protein which derives from the preceding ones after modification of certain elements of its primary structure.

Of course, the process according to the invention applies to any protein of interest the production of which requires a great accumulation of proteins before either extracting them and purifying them, totally or partially, or using them in a mixture with the biomass which will have made it possible to produce them. It is the case, for example, of enzymes which are useful for the biocatalysis of chemical reactions, and which can be used without a prior isolation and purification procedure, or also of enzymes which are used in the host bacterium in the process of growing, for the biotransformation of chemical compounds.

Advantageously, the heterologous protein is an enzyme produced in industrial amounts for a subsequent use as a chemical reaction catalyst. According to one particular embodiment of the invention, the enzyme is a nitrilase, advantageously a nitrilase of *Alcaligenes faecalis* (ATCC8750) described in patent application WO 98/18941 or a nitrilase of *Comamonas testosteroni sp.* described in CIP application No. 08/194,588, or an amidase such as those described

This Sheet Has Been Replaced



in applications WO 97/04083, EP 433 117 and EP 488 916, or a hydroxyphenylpyruvate dioxygenase described in application WO 96/38567.

The present invention also relates to an *E. coli* strain W as defined above, characterized in that it comprises a system for expressing heterologous proteins, in which the promoter is the P_{trp} promoter defined above.

The examples hereinbelow make it possible to illustrate the present invention without, however, seeking to limit the scope thereof.

The appended figures 1 to 3 represent maps of plasmids used in the various examples.

Figure 1 represents the map of the plasmid pRPA-BCAT41. The sites in brackets are sites which were eliminated during cloning. Ptrp: tryptophan promoter; nitB: nitrilase gene; TrrnB: transcription terminators; end ROP: end of the gene encoding the ROP protein (Chambers et al., 1988, Gene 68: 139-149); ORI: origin of replication; RNAI/II: RNAs involved in replication (Chambers et al., mentioned above); Tc: tetracyclin resistance gene.

Figure 2 reoresents 1 map of the plasmid pRPA-BCAT127. The sites between brackets have been eliminated during cloning. Ptrp: tryptophan promoter; nitB: nitrilase gene; TrrnB: transcription terminators; ORI: origin of replication; RNAI*/II: mutated RNAs

This Sheet Has Been Replaced



involved in replication; Cm: chloramphenicol resistance gene; cer: cer locus.

Figure 3 represents the map of the plasmid pRPA-BCAT103. The sites between brackets have been eliminated during cloning. Sm/Sp: streptomycin and spectinomycin resistance gene; parABCDE: par locus (Roberts and Helinski, 1992, J. Bacteriol. 174: 8119-8132); rep, mob, D20 and ori: regions involved in the replication and transfer of the plasmid (Scholtz et al., 1989, Gene 75: 271-288; Frey et al., 1992, Gene 113: 101-106).

The techniques used are conventional molecular biology and microbiology techniques known to a person skilled in the art and described, for example, by Ausubel *et al.*, 1987 (Current Protocols in Molecular Biology, John Wiley and Sons, New York), Maniatis *et al.*, 1982, (Molecular Cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), Coligan *et al.*, 1997 (Current Protocols in Protein Science, John Wiley & Sons, Inc).

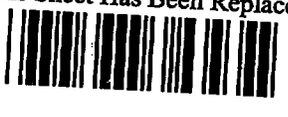
Example 1: Construction of the expression plasmids pBCAT29 and pBCAT41.

The 1.27 kb fragment conting the P_{trp} promoter, the ribosome binding site of the λ phage cII gene (RBScII) and the nitrilase gene of *Alcaligenes faecalis* ATCC8750 (*nitB*) was extracted from the plasmid pRPA6BCAT6 (application FR 96/13077) using the EcoRI

This Sheet Has Been Replaced

and XbaI restriction enzymes, so as to be cloned into the vector pXL642 (described in CIP application No. 08/194,588) opened with the same restriction enzymes. The resulting plasmid, pRPA-BCAT15, was opened with the StuI and BsmI enzymes, and the 4.3 kb fragment was ligated with the purified 136 bp StuI-BsmI fragment of pRPA-BCAT4 (application FR 96/13077) so as to produce the plasmid pRPA-BCAT19. The partial sequencing of pRPA-BCAT19 confirmed the replacement of the codon of the Asp279 residue of the nitrilase with the codon of an Asn279 residue. The 1.2 kb EcoRI-XbaI fragment of pRPA-BCAT19 containing the P_{trp}::RBScII::*nitB* fusion was then cloned into the vector pRPA-BCAT28 opened with the same enzymes, so as to produce the 6.2 kb plasmid pRPA-BCAT29. The vector pRPA-BCAT28 was obtained by ligating the 3.9 kb SspI-ScaI fragment of pXL642 (CIP application No. 08/194,588) with the 2.1 kb SmaI fragment of pHP45ΩTc (Fellay *et al.*, 1987, Gene 52: 147-154) in order to replace the ampicillin resistance marker with the tetracycline resistance marker. In destroying the NdeI site close to the origin of replication of the plasmid pRPA-BCAT29 by partial NdeI digestion and the action of *E. coli* Polymerase I (Klenow Fragment), a plasmid pRPA-BCAT41 was obtained, the map of which is represented in Figure 1. The sequence of the expression cassette is represented by sequence identifier No. 2 (SEQ ID NO 2).

This Sheet Has Been Replaced



Example 2: Expression of the nitrilase of A. faecalis ATCC8750 in "batch" *E. coli* K12, BL21 and W.

The plasmids pRPA-BCAT29 and pXL2035 (Levy-Schill *et al.*, 1995, Gene 161: 15-20) were introduced into the strains DH5α (CLONECH. Product reference C1021-1), BL21 (Novagen, product reference 69386-1) and W (ATCC9637) of *E. coli* by conventional electroporation. Expression cultures were prepared as described in Example 5 of application FR 96/13077, reducing the preculture time to 8 hours and fixing the expression time at 16 hours. The biomasses after expression were estimated according to the optical density of the cultures, read at 660 nm (OD660), using the following equation: biomass in gram of dry weight per liter of culture = OD660 × 0.35. The measurements of nitrilase activity of the cultures were carried out as described in application FR 96/13077. For each strain, two clones were analyzed and for each clone, the experiment was repeated. Table 1 contains, for each strain, the mean of the data obtained in the four experiments.

Table 1: *Biomass and activities of the strains harboring the plasmids pRPA-BCAT29 and pXL2035*

This Sheet Has Been Replaced



STRAINS	BIOMASS (g/l)	ACTIVITY (U)	PRODUCTIVITY (P)
DH5α	0.15	10.4	1.6
BL21	0.37	6.3	2.4
W	0.65	7.00	4.5

ABBREVIATIONS: g/l: gram of dry weight per liter of culture; U: kg of HMTBA formed per hour and per kg of dry weight; P: kg of HMTBA formed per hour and per liter of culture.

These data show that the strain W of *E. coli* (ATCC9637) is more effective at expressing the nitrilase NitB.

Example 3: Construction of pBCAT43.

The polyamide hydrolase gene of *Comamonas acidovorans* N12 described in application WO 97/04083 (*pamII*) was cloned into the vector pBCAT41. This polyamide hydrolase gene was amplified by PCR in the form of a 1.26 kb DNA fragment, while introducing, in the PCR primers, the EcoRI and NcoI restriction sites in the 5′ position of the gene and the XbaI restriction site in the 3′ position. This fragment was then treated successively with the EcoRI enzyme and Mung Bean nuclease. After extraction of the proteins with phenol-chloroform-isoamyl alcohol, the treatment was continued with an XbaI digestion. Similarly, the vector pRPA-BCAT41 was opened with the NdeI enzyme, and then treated with Mung Bean nuclease. After extraction of

This Sheet Has Been Replaced

the proteins with phenol-chloroform-isoamyl alcohol, the treatment was continued with an XbaI digestion. After ligation of these two samples, the plasmid pRPA-BCAT43 was obtained: it contains the P_{trp} promoter and the RBScII binding site separated from the translation start codon of the *pamII* gene by the sequence: AATACTTACACC.

Example 4: Expression of the polyamidase PamII in "batch" E. coli DH5α, BL21 and W.

The plasmid pRPA-BCAT43 was introduced into the strains DH5α, L21 and W of *E. coli* by conventional electroporation. Expression cultures were prepared as described in Example 2 above and varying the expression time from 14 to 24 hours. The biomasses after expression were estimated as in example 2 above. The measurements of polyamide hydrolase activity of the cultures were carried out as described in application WO 97/04083, with the following modifications:

- the cells were permeabilized with toluene by resuspending the cell pellets in a 100 mM trs-HCl, 5 mM EDTA, pH8, 1% toluene buffer so as to have a dry cell concentration of approximately 5 g/l; after vigorous shaking, the suspension is incubated for one hour at 4°C and then centrifuged, and finally, the pellets of permeabilized cells are taken up in a 100 mM, pH7, phosphate buffer.

This Sheet Has Been Replaced

- the hydrolysis activity was measured on the AB oligomer (one molecule of adipic acid condensed to one molecule of hexamethylenediamine) present at 2.5 g/l in the reaction medium containing 0.1 M potassium phosphate buffer at pH 7, and incubated at 30°C with stirring:
- 100 microliter samples are taken at regular intervals while adding to them the same volume of 0.2 N NaOH;
- the samples are analyzed by HPLC after ten-fold dilution in a solution of 50 mM H_3PO_4.

For each strain, from 1 to 24 clones were analyzed and for each clone, one to seven independent experiments were conducted. Table 2 contains, for each strain, the mean of the data obtained.

Table 2: *Biomass and activities of the strains harboring the plasmid pRPA-BCAT43*

STRAINS	NB CULTURES	ACTIVITY (U)	PRODUCTIVITY (P)
DH5a	11	0.77	0.3
BL21	3	1.4	1.8
W	24	2.1	2.6

ABBREVIATIONS: NB: number; U: g of AB hydrolyzed per hour and per g of dry weight; P: g of AB hydrolyzed per hour and per liter of culture.

This Sheet Has Been Replaced



These data show that the strain W (ATCC9637) of *E. coli* is more effective for expressing the PamII polyamidase.

Example 5: Construction and characterization of the plasmid pBCAT41-531.

The plasmid pRPA-BCAT41 underwent a mutagenesis step carried out with hydroxylamine as described in Miller 1992 (Mutagenesis. A short course in bacterial genetics. "A laboratory manual and handbook for *E. coli* and related bacteria", Cold Spring Harbor Laboratory Press, Unit 4, pp. 81-212) and Humphreys *et al.*, 1976 (Mol. Gen. Genet. 145: 101-108). Five micrograms of plasmid DNA purified on a cesium chloride gradient were incubated for 20 minutes at 80°C in a 50 mM sodium phosphate buffer, pH 6, containing 0.5 mM EDTA and 0.4 M NH_2OH. After the addition of an identical volume of 50 mM sodium phosphate buffer, pH 6, containing 0.5 mM EDTA, the reaction mixture was dialyzed against a large excess of 10 mM Tris-HCl buffer, pH 7.5, containing 1 mM EDTA and 100 mM NaCl. The plasmid DNA was then recovered by precipitation and approximately 20 ng of DNA was introduced by electroporation into the strain DH5α harboring the plasmid pXL2035. Among the transformants obtained, one clone was selected because the productivity of the culture was 3 times higher than that of a culture of the strain DH5α (pRPA-BCAT41, pXL2035). The plasmid

This Sheet Has Been Replaced



pRPA-BCAT41-531 that it was harboring was extracted and reintroduced into a new DH5α host harboring the plasmid pXL2035. Three clones were then analyzed under the conditions described in example 2, comparing them with 3 DH5α clones (pRPA-BCAT41, pXL2035), and the results are given in table in Table 3.

Table 3: *Biomass and activities of the strains harboring the plasmids pRPA-BCAT41, pRPA-BCAT41-531 and pXL2035*

Strains	**Biomass (g/l)**	**Activity (U)**	**Productivity (P)**
DH5α(pRPA-BCAT41, pXL2035)	0.21	12	2.5
DH5α(pRPA-BCAT41-531, pXL2035)	0.63	12	7.5

ABBREVIATIONS: g/l: gram of dry weight per liter of culture; U: kg of HMTBA formed per hour and per kg of dry weight; P: kg of HMTBA formed per hour and per liter of culture.

These results indicate that the improvement in the productivity of the cultures is correlated with the presence of the plasmid pRPA-BCAT41-531.

The 1.27 kb EcoRI-XbaI fragment containing the P_{trp}*::nitB* fusion was extracted from the plasmid pRPA-BCAT41 in order to be cloned in place of the one contained in pRPA-BCAT41-531. The resulting plasmid, pRPA-BCAT86, was introduced into the strain DH5a

This Sheet Has Been Replaced



(pXL2035) and 3 transformants were studied under conditions similar to those described above. The results are given in Table 4.

Table 4: *Biomass and activities of the strains harboring the plasmids pRPA-BCAT41, pRPA-BCAT41-531, pRPA-BCAT86 and pXL2035*

Strains	**Biomass (g/l)**	**Activity (U)**	**Productivity (P)**
DH5α(pRPA-BCAT41, pXL2035)	0.20	13.8	2.7
DH5α(pRPA-BCAT41-531, pXL2035)	0.68	11.0	7.4
DH5α(pRPA-BCAT86, pXL2035)	0.69	11.9	8.1

ABBREVIATIONS: g/l: gram of dry weight per liter of culture; U: kg of HMTBA formed per hour and per kg of dry weight; P: kg of HMTBA formed per hour and per liter of culture.

The results show that the improvement in the productivity of the cultures harboring pRPA-BCAT41-531 is not due to an improvement in the specific activity of the strain, and that this improvement is not caused by a mutation in the fragment carrying the P_{trp} promoter and the *nitB* gene.

This Sheet Has Been Replaced



Example 6: Characterization of a mutation carried by the plasmid pBCAT41-531 responsible for the improvement in productivity of the cultures of strains expressing nitrilase.

The analysis of the amount of protein produced by the strains of example 5, by polyacrylamide gel electrophoresis in the presence of SDS, showed that all these constructs led to levels of nitrilase polypeptide synthesis which were comparable among the strains described in this example. On the other hand, preparations of plasmid DNA of pRPA-bCAT41 and pRPA-BCAT41-531 prepared from equivalent amounts of biomass demonstrated that the plasmid pRPA-BCAT41-531 is present at lower number of copies than its parent pRPA-BCAT41. The sequencing of the 994 bp region of pRPA-BCAT41-531, which stretches from the Tth111I site and covers the origin of replication of the plasmid, revealed two differences with respect to the sequence of the corresponding region of pBR322 (GeneBank #J01749, name: SYNPBR322). By referring to the numbering given in the sequence J01749 (0 is the middle of the unique EcoRI site), we found that an insertion of an A had taken place after base 2319, and that the C of position 3039 is replaced with a T, in pRPA-BCAT41-531. The first difference can be attributed to an error during the action of the Klenow polymerase which was used to destroy one of the NdeI sites of pRPA-BCAT29,

This Sheet Has Been Replaced



and is located in a region which is not described as playing a role in the replication of pBR322 (Chambers *et al.*, 1988, Gene 68: 139-149). The second error corresponds to a transition, a characteristic effect of hydroxylamine on DNA (Drake and Baltz, 1976, Annu. Rev. Biochem. 45: 11-37), and is located at the second nucleotide of the region transcribed into RNA I involved in the replication of pBR322 (Chambers *et al.*, mentioned above). It is the latter mutation which is responsible for the lower number of copies of pRPA-BCAT41-531 in DH5α and which is responsible for the better nitrilase productivity of the cultures of the strain DH5α (pRPA-BCAT41-531, pXL2035).

Example 7: Expression of the nitrilase of A. faecalis ATCC 8750 in "fed-batch" (semicontinuous culture) E.coli BL21 and E. coli W.

The plasmids pRPa-BCAT41-531 and pXL2035 were introduced by electroporation into the strains BL21 (reference mentioned above) and W (ATCC9637) so as to give the RPA-BIOCAT594 [BL21 (pRPA-BCAT41-531, pXL2035)] and RPA-BIOCAT714 [W (pRPA-BCAT41-531, pXL2035)] strains, respectively. The recombinants E. coli BIOCAT 594 and E. coli BIOCAT 714 were cultured in 3.5 liter fermenters containing 2 liters of medium with the following composition:

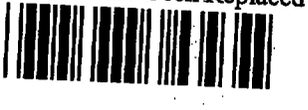
This Sheet Has Been Replaced



Compound	**Concentration in g/l**
KH_2PO_4	8
K_2HPO_4	6.3
$(NH_4)_2SO_4$	0.75
$MgSO_47H_2O$	2.5
Iron sulfate	0.04
$CaCl_2.2H_2O$	0.05
Manganese sulfate	0.01
Cobalt chloride	0.004
Zinc sulfate	0.002
Sodium molybdate	0.002
Copper chloride	0.002
Boric acid	0.0005
Citrate [lacuna].H2O	1.7
Glucose monohydrate	95
L-tryptophan	0.1
Meat peptone	5
Yeast extract	3

The pH is maintained at 7.0 by adding aqueous ammonia. The oxygen saturation is maintained at 20% by adding air in a proportion of 1 volume/volume of medium/minute and by stirring. The glucose is introduced at the start at a final concentration of 2 g/l. After having been totally consumed, it is introduced continuously from a stock solution with the following composition: 700 g/l glucose; 19.6 g/l

This Sheet Has Been Replaced



$MgSO_4.7H_2O$. The rate of addition is 2.2 g of glucose/h.l of medium.

After fermentation for 24 hours, the medium is recovered and centrifuged, and the dry weight is estimated in g/l. The enzymatic activity is measured following a protocol given in patent WO 96/09403. It is expressed in kilos of ammonium 3-hydroybutanoate formed per hour and per kilo of dry cells.

Strain	**Final biomass**	**Final activity**	**Yield on glucose**
BIOCAT 594 (BL21)	27 g/l	13	23%
BIOCAT 714 (W)	40 g/l	17	40%

In this example, it appears clearly that the nitrilase is expressed much better in E. coli W than in *E. coli* BL21, and that the recombinant *E. coli* W BIOCAT 714 grows much better than the recombinant E. coli BL21 BIOCAT 594.

Example 8: Influence of the organic nitrogen source of animal origin.

The E. coli strain W BIOCAT 714 is cultured in a 3.5 liter fermenter containing 2 liters of medium with the following composition:

This Sheet Has Been Replaced

Compound	**Concentration in the medium in g/l**
K_2HPO_4	8
$(NH_4)_2SO_4$	0.75
$MgSO_47H_2O$	2.5
Iron sulfate	0.04
$CaCl_2.2H_2O$	0.04
Manganese sulfate	0.026
Cobalt chloride	0.004
Zinc sulfate	0.013
Sodium molybdate	0.001
Copper chloride	0.001
Boric acid	0.00025
$AlCl_3$	0.00125
Citrate [lacuna].H_2O	1.7
Glucose monohydrate	95
L-tryptophan	0.1
Yeast extract	3

The pH is maintained at 7.0 by adding aqueous ammonia. The oxygen saturation is maintained at 20% by adding air in a proportion of 1 volume/volume of medium/minute and by stirring. The glucose is introduced at the start at a final concentration of 2 g/l. After having been totally consumed, it is introduced continuously from a stock solution with the following composition: 700 g/l glucose; 19.6 g/l

This Sheet Has Been Replaced



$MgSO_4.7H_2O$. The rate of addition is 2.2 g of glucose/h.l of medium.

An organic nitrogen source of animal origin is added to this medium.

Organic nitrogen source of animal origin	**Final biomass**	**Final activity**	**Yield on glucose**
None	30	2	40%
2.5 g/l of meat peptone	33	12	40%
5 g/l of meat peptone	40	25	45%
5 g/l of casein	35	20	43%

The use of an increasing concentration of organic nitrogen of animal origin significantly increases the specific activity of the cells.

Example 9: Influence of the organic nitrogen of plant origin.

The culture conditions are identical to those of example 8. In this example, organic nitrogen of plant origin is added.

This Sheet Has Been Replaced



Organic nitrogen source of animal origin	**Final biomass**	**Final activity**	**Yield on glucose**
None	30	2	40%
5 g/l of soybean peptone	31	4	40%
5 g/l of wheat peptone	32	5	40%
7.5 g/l of sodium hydrolysate of potato protein (Alburex SP; Roquette)	35	17	43%

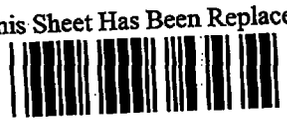
The addition of plant organic nitrogen does not give identical results depending on the origin.

Surprisingly, the addition of potato protein gives as good a result as the organic nitrogen of animal origin.

Example 10: Influence of the carbon source.

The E. coli strain W BIOCAT 714 is cultured in a 3.5 liter fermenter containing 2 liters of medium with the following composition:

This Sheet Has Been Replaced



Compound	**Concentration in the medium in g/l**
Corn-steep LAB2218 (Roquette)	40
Yeast extract	3
$MgSO_4 7H_2O$	2.5

The pH is maintained at 7.0 by adding aqueous ammonia. The oxygen saturation is maintained at 20% by adding air in a proportion of 1 volume/volume of medium/minute and by stirring. The carbon source is introduced at the start at a final concentration of 2 g/l. After having been totally consumed, it is introduced continuously from a stock solution with the following composition: 700 g/l carbon source; 19.6 g/l $MgSO_4.7H_2O$. The rate of addition is 2.2 g of glucose or of sucrose/h.l of medium.

The carbon source is varied in this example.

Carbon source	**Final biomass**	**Final activity**	**Yield on carbon**
Glucose monohydrate 90 g/l	38	11	45%
"Syrup zero" (EUROSUCRE) 90 g/l	38	17	45%

This Sheet Has Been Replaced



In this example, it is observed that the use of sucrose ("syrup zero") as a carbon source significantly increases the specific activity of the cells.

Example 11: Construction of a plasmid for coexpression of the TrpR regulator

A 434 bp DNA fragment which carries the *trpR* gene and its promoter was extracted from the plasmid pRPG9 (Gunsalus and Yanofsky, 1980, Proc. Natl. Aca. Sci. USA 77: 7117-7121) using the AatII and StuI restriction enzymes. This fragment was cloned into the plasmid pSL301 (Brosius, 1989, DNA 8: 759-777) by ligating it to the approximately 3.1 kb AatII-StuI fragment, so as to give the plasmid pRPA-BCAT30. The *trpR* gene and its promoter were then extracted from pRPA-BCAT30 in the form of a 475 bp EcoRI-NotI fragment in order to be cloned into the plasmid pXL2035 in place of a 240 bp EcoRI-NotI fragment. The resulting plasmid, pRPA-BCAT34, is therefore a derivative of pKT230 which allows the expression of the GroESL chaperones and of the TrpR regulator.

Example 12: Influence of the coexpression of GroESL and of TrpR.

The plasmid pRPA-BCAT34 was introduced by electroporation into the strains DH5α (pRPA-BCAT29),

This Sheet Has Been Replaced



BL21 (pRPA-BCAT29) and W (pRPA-BCAT29). Expression cultures of various strains were prepared as described in example 2, and the results are given in Table 5.

Table 5: *Biomass and activities of the strains harboring combinations the plasmids pRPA-BCAT29, pXL2035 and pRPA-BCAT34*

Combinations	**pRPA-BCAT29**		**pRPA-BCAT29 pXL2035**		**pRPA-BCAT29 pRPA-BCAT34**	
HOST	**U**	**P**	**U**	**P**	**U**	**P**
DH5-alpha	0.37	0.16	10.4	1.6	2.0	0.7
BL21	0	0.0	6.4	2.4	5.6	2.0
W	1.7	0.96	7.0	4.5	8.9	6.5

ABBREVIATIONS: U: activity, kg of HMTBA formed per hour and per kg of dry weight; P: productivity, kg of HMTBA formed per hour and per liter of culture.

The results show that the coexpression of GroESL makes it possible to increase the productivity of the cultures whatever the strain under consideration, by improving the specific activity of the cultures. This effect is correlated with an increase in the solubility of the nitrilase polypeptide, as shown by an analysis of the proteins by electrophoresis as described in application FR 96/13077. The effect of the coexpression of the TrpR

This Sheet Has Been Replaced



regulator is variable according to the strains, but makes it possible, in W, to improve the productivity of the cultures.

Example 13: Influence of the presence of a *cer* locus on pRPA-BCAT41

The 382 bp HpaII fragment containing the *cer* locus of the plasmid ColE1 (Leung *et al.*, 1985, DNA 4: 351-355) was cloned into the replicative form of the M13mp7 phage at one of the 2 AccI sites. The construct obtained then made it possible to extract, with the EcoRI enzyme, an approximately 430 bp fragment containing the *cer* locus, which was cloned into pRPA-BCAT41 at the EcoRI site, thereby producing the plasmid pRPA-BCAT66. This plasmid was introduced by electroporation into the strain W harboring the plasmid pRPA-BCAT34. Expression cultures of various strains were prepared as described in example 2, extending the duration of the expression cultures to 24 hours and studying three clones of each strain in a sole experiment. The mean results are given in Table 6.

Table 6: *Biomass and activities of the strains harboring the plasmids pRPA-BCAT41, pRPA-BCAT66 and pRPA-BCAT34*

This Sheet Has Been Replaced

Strains	**Biomass (g/l)**	**Activity (U)**	**Productivity (P)**
W (pRPA-BCAT41, pRPA-BCAT34)	2.1	6.9	14.5
DH5α (pRPA-BCAT66, pRPA-BCAT34)	1.8	10.0	18.0

ABBREVIATIONS: g/l: gram of dry weight per liter of culture; U: kg of HMTBA formed per hour and per kg of dry weight; P: kg of HMTBA formed per [lacuna]

These results show that adding the *cer* locus to the plasmid for expression of the nitrilase leads to an improvement in the productivity of the cultures.

Example 14: Construction of the plasmid pRPA-BCAT127

After elimination of the unique NdeI site of the plasmid pRPA-BCAT30 by digestion and formation of blunt ends with polymerase I (Klenow fragment), the *trpR* gene was extracted from this latter plasmid in the form of an approximately 300 bp fragment prepared by treatment with the AatII enzyme followed by the action of polymerase I (Klenow fragment), and then, after inactivation of the reaction mixture, by digestion with the SacII enzyme. This fragment was cloned into the pRPA-BCAT66 plasmid after opening this plasmid with Tth111 followed by treatment with polymerase I Klenow fragment) and, after inactivation, with SacII. The

This Sheet Has Been Replaced

plasmid pRPA-BCAT82 was thus obtained. Its origin of replication was replaced with that of the plasmid pRPA-BCAT41-531 by replacing the approximately 1.12 kb Bst1107I-Eam1105I fragment. The construct selected during this cloning, the plasmid pRPA-BCAT99, has an artefact which is in the form of a deletion of one nucleotide at the Eam1105I site, transforming this site into a unique PshAI site. The resistance marker of the plasmid pRPA-BCAT99 was then changed by cloning, between the AatII and PshAI sites, an approximately 1.07 kb AatII-PshAI fragment prepared after PCR amplification of the gene encoding chloramphenicol resistance from the matrix pACYC184 (New England Biolabs #401-M), using the primers Cm1 and Cm2, the sequence of which is:

Cm1 : 5'-CCCCCCGACAGCTGTCTTGCTTTCGAATTTCTGCC

Cm2 : 5'-TTGACGTCAGTAGCTGAACAGGAGGG

The plasmid thus obtained was called pRPA-BCAT123. It was then modified by eliminating the *trpR* gene in the form of an approximately 0.525 kb SacI-Bst1107I fragment, and reclosing the plasmid after forming blunt ends with the Pfu polymerase (15 minutes at 75°C in the buffer recommended by the manufacturer Stratagène, and in the presence of 0.2 mM of deoxynucleotides). The plasmid thus obtained is the plasmid pRPA-BCAT127, the map of which is represented schematically in Figure 2.

This Sheet Has Been Replaced



Example 15: Construction of the plasmids pRPA-BCAT98 and pRPA-BCAT103.

The plasmid pRPA-BCAT37, described in application FR 96/13077, was modified by replacing the approximately 3.2 kb SfiI-ScaI fragment with the approximately 2.42 kb SfiI-ScaI fragment of the plasmid RSF101D20 (Frey *et al.*, 1992, Gene 113: 101-106). This fragment contains a deletion in the 5' portion of the gene encoding the RepB primase, and reduces the frequency of transfer of the plasmid by 6 logs (Frey *et al.*, mentioned above). The plasmid thus obtained, pRPA-BCAT98, has several advantages: the loss of its mobilization functions makes it comply with the rules of industrial biosafety while at the same time retaining its properties of replication in Gram-negative bacteria.

The *par* locus (Gerlitz *et al.*, 1990, J. Bacteriol 172: 6194-6203) was then cloned on pRPA-BCAT98 as follows. The approximately 2.3 kb SphI-BamHI fragment of pGMA28 (Gerlitz *et al.*, mentioned above) was first cloned into the vector pUC18, thereby allowing its extraction in the form of a HindIII-EcoRI fragment so as to clone it into the vector pMTL22 (Chambers *et al.*, 1988, Gene 68: 139-49). The HindIII site was then destroyed by HindIII digestion and Klenow treatment. An approximately 2.38 kb fragment was then extracted with the PstI and BglII enzymes so as to be

This Sheet Has Been Replaced



cloned into the vector pXL2426 at the PstI and BamHI sites and to produce the vector pXL2572. The vector pXL2426 originates from the replacement of the 2.38 kb SfiI-EcoRV fragment of pXL2391 (application FR 96/13077) with the 1.47 kb SfiI-EcoRV fragment of RSF1010D20. The cloning on the plasmid pXL2572, at the NdeI and BamHI sites, of an approximately 0.960 bp NdeI-BamHI fragment of pRR71 (Weinstein *et al.*, 1992, J. Bacteriol. 174: 7486-7489) made it possible to reconstitute the *par* locus as a whole on the plasmid pXL2573. This locus was then extracted from pXL2573 in the form of a 2.6 kb EcoRI-blunt end (after treatment with PstI and Klenow) fragment in order to be cloned on the plasmid pRPA-BCAT98 opened with EcoRI and SacI, the latter end having been treated with the Pfu polymerase. The resulting plasmid was called pRPA-BCAT103 and its map is represented schematically in Figure 3.

Example 16: Use of the plasmids pRPA-BCAT98, pRPA-BCAT103 and pRPA-BCAT127 for expressing the nitrilase in W.

The plasmids pRPA-BCAT127, pRPA-BCAT98, pRPA-BCAT103, pXL2035 and pXL2231 (application FR 96/13077) were introduced into the strain W of *E. coli* by electroporation, and expression cultures were prepared under the conditions described in example 2, using the following antibiotics: 12 µg/ml tetracycline for pXL2231, 50 µg/ml kanamycin for pXL2035, 100 µg/ml

This Sheet Has Been Replaced



streptomycin for pRPA-BCAT98 and pRPA-BCAT103, and 20 µg/ml chloramphenicol for pRPA-BCAT127. For each combination of plasmids, two to three clones were analyzed, and the mean results are given in Table 7.

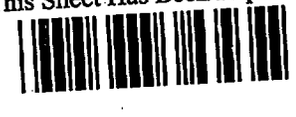
Table 7: *Biomass and activities of the strains harboring the plasmids pRPA-BCAT41-531, pRPA-BCAT127, pRPA-BCAT98, pRPA-BCAT103, pXL2035 and pXL2231*

Combination	**Biomass (g/l)**	**Activity (U)**	**Productivity (P)**
pBCAT127/pXL2231	1.43	4.9	7
pBCAT127/pBCAT103	1.75	7	12
pBCAT127/pBCAT98	1.72	11.2	19
pBCAT127/pXL2035	1.70	7.2	12
pBCAT41-531/pXL2035	1.36	5.9	8

ABBREVIATIONS: g/l: gram of dry weight per liter of culture; U: kg of HMTBA formed per hour and per kg of dry weight; P: kg of HMTBA formed per hour and per liter of culture

The combinations pRPA-BCAT127/pRPA-BCAT98 and pRPA-BCAT127/pRPA-BCAT103 allow an at least equivalent productivity to be obtained, using plasmids which are in conformity with the European criteria for biosafety.

This Sheet Has Been Replaced



LIST OF SEQUENCES

(iii) NUMBER OF SEQUENCES: 4

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 121 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAATTCCCTG TTGACAATTA ATCATCGAAC TAGTTAACTA GTACGCAGCT TGGCTGCAGG 60

TCGACCTGCA GCCAAGCTTG GGCATACATT CAATCAATTG TTATCTAAGG AAATACTTAC 120

A 121

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1793 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) CHARACTERISTIC:

(A) NAME/KEY: CDS

(B) POSITION: 123..1190

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

This Sheet Has Been Replaced



```
GAATTCCCTG TTGACAATTA ATCATCGAAC TAGTTAACTA GTACGCAGCT TGGCTGCAGG      60

TCGACCTGCA GCCAAGCTTG GGCATACATT CAATCAATTG TTATCTAAGG AAATACTTAC     120

AT ATG CAG ACA AGA AAA ATC GTC CGG GCA GCC GCC GTA CAG GCC GCC         167
   Met Gln Thr Arg Lys Ile Val Arg Ala Ala Ala Val Gln Ala Ala
     1               5                  10                  15

TCT CCC AAC TAC GAT CTG GCA ACG GGT GTT GAT AAA ACC ATT GAG CTG       215
Ser Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu
                 20                  25                  30

GCT CGT CAG GCC CGC GAT GAG GGC TGT GAC CTG ATC GTG TTT GGT GAA       263
Ala Arg Gln Ala Arg Asp Glu Gly Cys Asp Leu Ile Val Phe Gly Glu
             35                  40                  45

ACC TGG CTG CCC GGC TAT CCC TTC CAC GTC TGG CTG GGC GCA CCG GCC       311
Thr Trp Leu Pro Gly Tyr Pro Phe His Val Trp Leu Gly Ala Pro Ala
         50                  55                  60
```

This Sheet Has Been Replaced



```
TGG TCG CTG AAA TAC AGT GCC CGC TAC TAT GCC AAC TCG CTC TCG CTG       359
Trp Ser Leu Lys Tyr Ser Ala Arg Tyr Tyr Ala Asn Ser Leu Ser Leu
     65                  70                  75

GAC AGT GCA GAG TTT CAA CGC ATT GCC CAG GCC GCA CGG ACC TTG GGT       407
Asp Ser Ala Glu Phe Gln Arg Ile Ala Gln Ala Ala Arg Thr Leu Gly
 80                  85                  90                  95

ATT TTC ATC GCA CTG GGT TAT AGC GAG CGC AGC GGC GGC AGC CTT TAC       455
Ile Phe Ile Ala Leu Gly Tyr Ser Glu Arg Ser Gly Gly Ser Leu Tyr
                100                 105                 110

CTG GGC CAA TGC CTG ATC GAC GAC AAG GGC CAG ATG CTG TGG TCG CGT       503
Leu Gly Gln Cys Leu Ile Asp Asp Lys Gly Gln Met Leu Trp Ser Arg
            115                 120                 125

CGC AAA CTC AAA CCT ACA CAT GTT GAG CGC ACC GTG TTT GGT GAA GGT       551
Arg Lys Leu Lys Pro Thr His Val Glu Arg Thr Val Phe Gly Glu Gly
        130                 135                 140

TAT GCC CGA GAT CTG ATT GTG TCC GAC ACC GAG CTG GGC CGC GTC GGT       599
Tyr Ala Arg Asp Leu Ile Val Ser Asp Thr Glu Leu Gly Arg Val Gly
    145                 150                 155

GCC CTG TGC TGC TGG GAG CAC CTG TCC CCC TTG AGC AAG TAC GCG CTG       647
Ala Leu Cys Cys Trp Glu His Leu Ser Pro Leu Ser Lys Tyr Ala Leu
160                 165                 170                 175

TAC TCC CAG CAC GAA GCC ATT CAC ATT GCC GCC TGG CCG TCC TTT TCG       695
Tyr Ser Gln His Glu Ala Ile His Ile Ala Ala Trp Pro Ser Phe Ser
                180                 185                 190

CTG TAC AGC GAA CAG GCC CAT GCG CTC AGC GCC AAG GTG AAC ATG GCT       743
Leu Tyr Ser Glu Gln Ala His Ala Leu Ser Ala Lys Val Asn Met Ala
            195                 200                 205

GCC TCG CAA ATC TAT TCG GTT GAA GGC CAG TGC TTT ACC ATC GCC GCC       791
Ala Ser Gln Ile Tyr Ser Val Glu Gly Gln Cys Phe Thr Ile Ala Ala
        210                 215                 220

AGC AGT GTC GTC ACC CAG GAG ACA CTG GAC ATG CTG GAA GTA GGT GAA       839
Ser Ser Val Val Thr Gln Glu Thr Leu Asp Met Leu Glu Val Gly Glu
    225                 230                 235

CAC AAC GCC TCC CTG CTG AAA GTG GGC GGC GGC AGT TCC ATG ATT TTT       887
His Asn Ala Ser Leu Leu Lys Val Gly Gly Gly Ser Ser Met Ile Phe
240                 245                 250                 255

GCG CCG GAC GGA CGC ACA TTG GCT CCC TAC CTG CCA CAC GAT GCC GAA       935
Ala Pro Asp Gly Arg Thr Leu Ala Pro Tyr Leu Pro His Asp Ala Glu
                260                 265                 270

GGC CTG ATC ATT GCC GAT CTG AAC ATG GAA GAA ATT GCC TTC GCC AAG       983
Gly Leu Ile Ile Ala Asp Leu Asn Met Glu Glu Ile Ala Phe Ala Lys
            275                 280                 285

GCG ATC AAC GAC CCT GTG GGC CAC TAC TCC AAA CCC GAG GCC ACC CGT      1031
Ala Ile Asn Asp Pro Val Gly His Tyr Ser Lys Pro Glu Ala Thr Arg
        290                 295                 300
```

This Sheet Has Been Replaced



```
CTG GTA CTG GAC CTG GGG CAC CGT GAG CCC ATG ACT CGG GTG CAT TCC      1079
Leu Val Leu Asp Leu Gly His Arg Glu Pro Met Thr Arg Val His Ser
    305                 310                 315

AAA AGC GTG ATC CAG GAA GAA GCT CCC GAG CCG CAC GTG CAA AGT ACG      1127
Lys Ser Val Ile Gln Glu Glu Ala Pro Glu Pro His Val Gln Ser Thr
320                 325                 330                 335

GCT GCG CCC GTC GCC GTC AGC CAG ACT CAG GAC TCG GAT ACG CTA CTG      1175
Ala Ala Pro Val Ala Val Ser Gln Thr Gln Asp Ser Asp Thr Leu Leu
                340                 345                 350

GTG CAA GAA CCG TCC TGA CCCCAAAAGA TGACAAGGCC CGGGCAAACT             1223
Val Gln Glu Pro Ser
            355

GTCCGGGTCT TGATTCCTTC TGCGTCCCGG ATCCACTAGT TCTAGAGTCG ACCTGCAGGC    1283

ATGCAAGCTT GGGTCCCACC TGACCCCATG CCGAACTCAG AAGTGAAACG CCGTAGCGCC    1343

GATGGTAGTG TGGGGTCTCC CCATGCGAGA GTAGGGAACT GCCAGGCATC AAATAAAACG    1403

AAAGGCTCAG TCGAAAGACT GGGCCTTTCG TTTTATCTGT TGTTTGTCGG TGAACGCTCT    1463

CCTGAGTAGG ACAAATCCGC CGGGAGCGGA TTTGAACGTT GCGAAGCAAC GGCCCGGAGG    1523

GTGGCGGGCA GGACGCCCGC CATAAACTGC CAGGCATCAA ATTAAGCAGA AGGCCATCCT    1583

GACGGATGGC CTTTTTGCGT TTCTACAAAC TCTTCCTGTC GTCATATCTA CAAGCCATCC    1643

CCCCACAGAT ACGGTAAACT AGCCTCGTTT TTGCATCAGG AAAGCAGCTA TGAACCACTC    1703

CTTAAAACCC TGGAACACAT TTGGCATTGA TCATAATGCT CAGCACATTG TATGTGCCGA    1763

AGACGAACAA CAATTACTCA ATGCCCGCGG                                     1793
```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

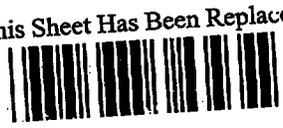
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
CCCCCCGACA GCTGTCTTGC TTTCGAATTT CTGCC                               35
```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

This Sheet Has Been Replaced



(A) LENGTH: 26 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TTGACGTCAG TAGCTGAACA GGAGGG 26

This Sheet Has Been Replaced



CLAIMS

1. Industrial process for preparing heterologous proteins in *E. coli*, in which *E. coli* bacteria modified with a suitable system for expressing heterologous proteins are seeded and cultured in a suitable culture medium, characterized in that the strain of *E. coli* is an *E. coli* strain W.

2. Process according to claim 1, characterized in that the strain W is the strain W deposited at the ATCC under the number 9637.

3. Process according to claim 1, characterized in that the strain W is a derivative of the strain deposited at the ATCC under the number 9637, obtained by clonal selection or genetic manipulation.

4. Process according to one of claims 1 to 3, characterized in that the suitable culture medium is a culture medium which is suitable for the production of a high density of biomass and a high content of heterologous proteins produced.

5. Process according to one of claims 1 to 4, characterized in that the culture medium comprises sucrose as the main carbon source.

6. Process according to claim 5, characterized in that the culture medium comprises substantially only sucrose as a carbon source.

This Sheet Has Been Replaced

* C L M S *

7. Process according to either of claims 5 and 6, characterized in that the amount of sucrose in the culture medium is between 0.1 and 300 g/l at the start of culturing, preferably between 0.5 and 200 g/l.

8. Process according to one of claims 1 to 7, characterized in that the suitable culture medium also comprises a supplementary organic nitrogen source.

9. Process according to claim 8, characterized in that the supplementary organic nitrogen source consists of protein extracts.

10. Process according to either of claims 7 and 8, characterized in that the protein extract has the following composition: (in g amino acids per 100 g of product) alanine between 10 and 4, aspartic [lacuna] between 11 and 4, glycine between 22 and 2.5 and lysine between 7 and 4.

11. Process according to one of claims 8 to 10, characterized in that the supplementary organic nitrogen source consists of meat or potato peptones or proteins, more particularly the derivatives of potato proteins.

12. Process according to one of claims 1 to 11, characterized in that the suitable system for expressing heterologous proteins comprises a P_{trp} promoter.

13. Process according to claim 12, characterized in that the P_{trp} promoter comprises the

This Sheet Has Been Replaced

nucleic acid sequence represented by sequence identifier no. 1 (SEQ ID NO 1).

14. Process according to one of claims 1 to 13, characterized in that the heterologous protein is an enzyme.

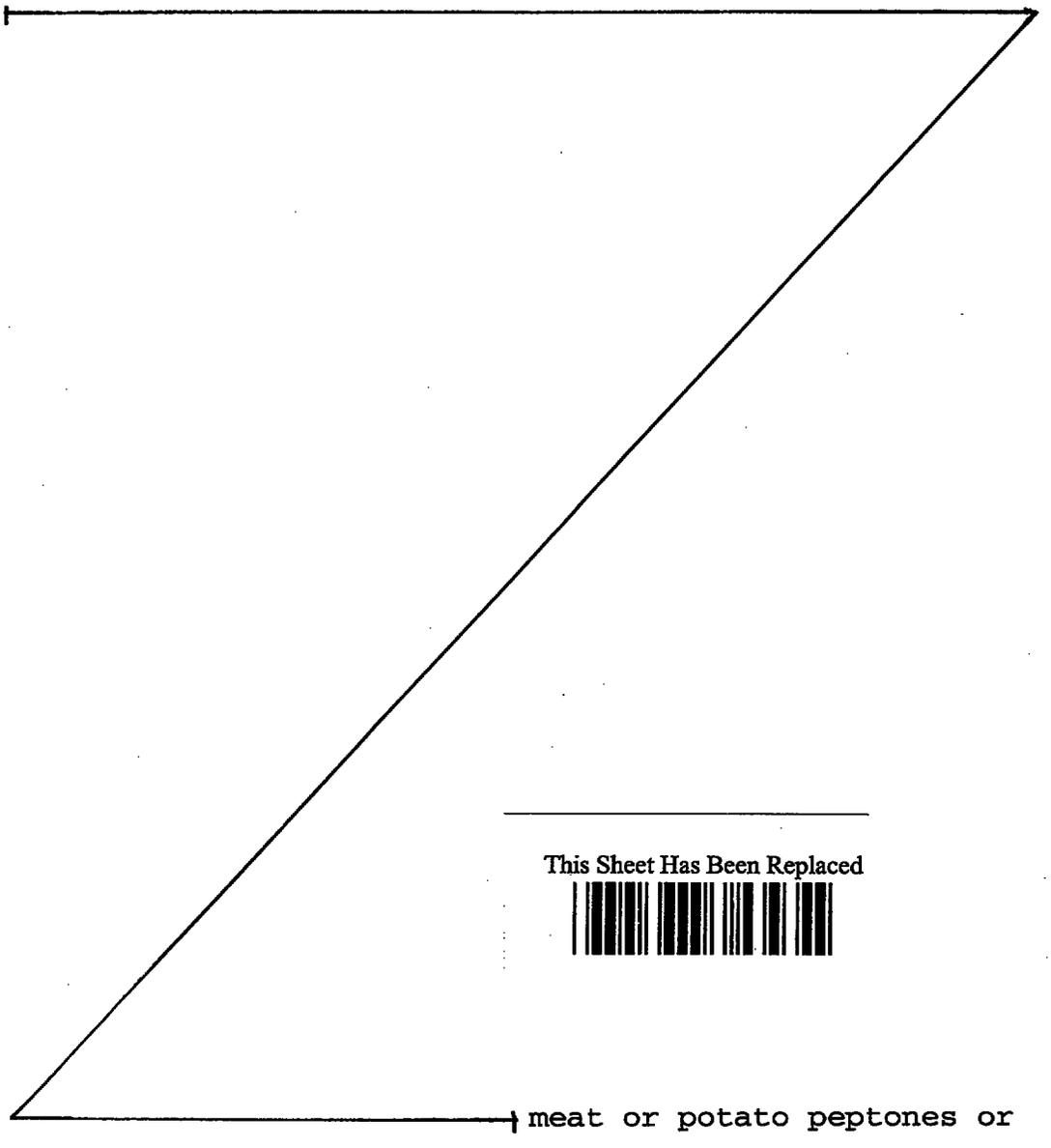
15. Process according to claim 14, characterized in that the enzyme is useful for the biocatalysis of chemical reactions.

16. Process according to claim 15, characterized in that the enzyme is a nitrilase.

17. *E. coli* strain W, characterized in that it comprises a system for expressing heterologous proteins, in which the promoter is the P_{trp} promoter.

18. Strain according to claim 16, characterized in that the P_{trp} promoter comprises the nucleic acid sequence represented by sequence identifier no. 1 (SEQ ID NO 1).

This Sheet Has Been Replaced



This Sheet Has Been Replaced

meat or potato peptones or proteins, more particularly the derivatives of potato proteins.

12. Process according to one of claims 1 to 11, characterized in that the suitable system for expressing heterologous proteins comprises a P_{trp} promoter.

13. Process according to claim 12, characterized in that the P_{trp} promoter comprises the

nucleic acid sequence represented by sequence identifier no. 1 (SEQ ID NO 1).

14. Process according to one of claims 1 to 13, characterized in that the heterologous protein is an enzyme.

15. Process according to claim 14, characterized in that the enzyme is useful for the biocatalysis of chemical reactions.

16. Process according to claim 15, characterized in that the enzyme is a nitrilase.

17. *E. coli* strain W, characterized in that it comprises a system for expressing heterologous proteins, in which the promoter is the P_{trp} promoter.

18. Strain according to claim 17, characterized in that the P_{trp} promoter comprises the nucleic acid sequence represented by sequence identifier no. 1 (SEQ ID NO 1).

This Sheet Has Been Replaced