(11) Application No. AU 2004293810 B2

(19) AUSTRALIAN PATENT OFFICE

(54) Title

Improved protein expression systems

(51)[6] International Patent Classification(s)

C12N 9/06 (2006.01) 8BMEP **C12N** 15/52
C12N 9/88 (2006.01) 20060101ALI2005100
C12N 15/52 (2006.01) 8BMEP **C12N** 15/78
C12N 15/78 (2006.01) 20060101ALI2005100
C12P 21/02 (2006.01) 8BMEP **C12P** 21/02
C12N 9/06 20060101ALI2005100
20060101AFI2005100 8BMEP
8BMEP **C12N** PCT/US2004/038884
9/88
20060101ALI2005100

(21) Application No: 2004293810 (22) Application Date: 2004.11.19

(87) WIPO No: WO05/052151

(30) Priority Data

(31) Number	(32) Date	(33) Country
60/523,420	2003.11.19	US
60/537,147	2004.01.16	US

(43) Publication Date : 2005.06.09

(71) Applicant(s)

Pfenex Inc.

(72) Inventor(s)

Chew, Lawrence C.; Ramseier, Thomas Martin; Badgley, Anne Kathryn; Schneider, Jane C.

(74) Agent/Attorney

Griffith Hack, Level 3 509 St Kilda Road, Melbourne, VIC, 3004

(56) Related Art

Barth, A.L. et al. CURRENT MICROBIOLOGY, 1998, vol. 36, pp. 190-195

Molina, L. et al. APPLIED AND ENVIRONMENTAL MICROBIOLOGY, 1998, vol. 64, no. 6, pp. 2072-2078.

WO 2003/062399 A2 (DIVERSA CORPORATION et al.)

Mergeay, M. et al. JOURNAL OF BACTERIOLOGY, 1978, vol. 135, no. 1, pp. 18-28.

Strych U. et al. CURRENT MICROBIOLOGY, 1994, vol. 29, pp. 353-359.

WO 2003/089455 (DOW GLOBAL TECHNOLOGIES INC.)

US 5,232,840 (OLINS)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 June 2005 (09.06.2005)

PCT

(10) International Publication Number
WO 2005/052151 A1

(51) International Patent Classification[7]: C12N 15/00, 15/74, 1/20, C12P 21/06, C07H 21/02, 21/04

(21) International Application Number:
PCT/US2004/038884

(22) International Filing Date:
19 November 2004 (19.11.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/523,420 19 November 2003 (19.11.2003) US
60/537,147 16 January 2004 (16.01.2004) US

(71) Applicant *(for all designated States except US)*: **DOW GLOBAL TECHNOLOGIES INC.** [US/US]; Washington Street, 1790 Building, Midland, MI 48674 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SCHNEIDER, Jane, C.** [US/US]; 2621 32nd Street, San Diego, CA 92104 (US). **CHEW, Lawrence, C.** [US/US]; 5590 Willowmere Lane, San Diego, CA 92130 (US). **BADGLEY, Anne, Kathryn** [US/US]; 12023 Old Pomerado Road, Poway, CA 92064 (US). **RAMSEIER, Thomas, Martin** [US/US]; 13033 Pomard Way, Poway, CA 92064 (US).

(74) Agent: **KNOWLES, Sherry, M.**: King & Spalding LLP, 191 Peachtree Street, 45th Floor, Atlanta, GA 30303-1763 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2005/052151 A1

(54) Title: IMPROVED PROTEIN EXPRESSION SYSTEMS

(57) Abstract: The present invention provides an improved expression system for the production of recombinant polypeptides utilizing auxotrophic selectable markers. In addition, the present invention provides improved recombinant protein production in host cells through the improved regulation of expression.

IMPROVED PROTEIN EXPRESSION SYSTEMS

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional patent application serial No. 60/523,420 filed November 19, 2003, entitled "Improved Pseudomonas Expression Systems with Auxotrophic Selection Markers," and U.S. Provisional patent application 60/537,147 filed January 16,2004, and entitled "Bacterial Expression Systems with Improved Repression."

FIELD OF THE INVENTION

The present invention provides an improved expression system for the production of recombinant polypeptides utilizing auxotrophic selectable markers. In addition, the present invention provides improved recombinant protein production in host cells through the improved regulation of expression.

BACKGROUND OF THE INVENTION

The use of bacterial cells to produce protein based therapeutics is increasing in commercial importance. One of the goals in developing a bacterial expression system is the production of high quality target polypeptides quickly, efficiently, and abundantly. An ideal host cell for such an expression system would be able to efficiently utilize a carbon source for the production of a target polypeptide, quickly grow to high cell densities in a fermentation reaction, express the target polypeptide only when induced, and grow on a medium that is devoid of regulatory and environmental concerns.

There are many hurdles to the creation of a superior host cell. First, in order to produce a recombinant polypeptide, an expression vector encoding the target protein must be inserted into the host cell. Many bacteria are capable of reverting back into an untransformed state, wherein the expression vector is eliminated from the host. Such revertants can decrease the fermentation efficiency of the production of the desired recombinant polypeptide.

Expression vectors encoding a target peptide typically include a selection marker in the vector. Often, the selection marker is a gene whose product is required for survival during the fermentation process. Host cells lacking the selection marker, such as revertants, are unable to survive. The use of selection markers during the fermentation process is intended to ensure that only bacteria containing the expression vector survive, eliminating

competition between the revertants and transformants and reducing the efficiency of fermentation.

The most commonly used selection markers are antibiotic resistance genes. Host cells are grown in a medium supplemented with an antibiotic capable of being degraded by the selected antibiotic resistance gene product. Cells that do not contain the expression vector with the antibiotic resistance gene are killed by the antibiotic. Typical antibiotic resistance genes include tetracycline, neomycin, kanamycin, and ampicillin. The presence of antibiotic resistance genes in a bacterial host cell, however, presents environmental, regulatory, and commercial problems. For example, antibiotic resistance gene-containing products (and products produced by the use of antibiotic resistance gene) have been identified as potential biosafety risks for environmental, human, and animal health. For example, see M. Droge et al., Horizontal Gene Transfer as a Biosafety issue: A natural phenomenon of public concern, J. Biotechnology. 64(1): 75-90 (17 Sept. 1998); Gallagher, D. M., and D. P. Sinn. 1983. Penicillin-induced anaphylaxis in a patient under hypotensive anaesthesia. *Oral Surg. Oral Med. Oral Pathol.* 56:361-364; Jorro, G., C. Morales, J. V. Braso, and A. Pelaez. 1996. Anaphylaxis to erythromycin. Ann. *Allergy Asthma Immunol.* 77:456-458; F. Gebhard & K. Smalla, Transformation of Acinetobacter sp. strain BD413 by transgenic sugar beet DNA, Appl. & Environ. Microbiol. 64(4):1550-54 (Apr. 1998); T. Hoffmann et al., Foreign DNA sequences are received by a wild type strain of Aspergillus niger after co-culture with transgenic higher plants, Curr. Genet. 27(1): 70-76 (Dec. 1994); DK Mercer et al., Fate of free DNA and transformation of the oral bacterium Streptococcus gordonoii DL1 by plasmid DNA in human saliva, Appl. & Environ. Microbiol. 65(1):6-10 (Jan 1999); R. Schubbert et al., Foreign (M13) DNA ingested by mice reaches peripheral leukocytes, spleen, and liver via the intestinal wall mucosa and can be covalently linked to mouse DNA, PNAS USA 94:961-66 (Feb. 4, 1997); and AA Salyers, Gene transfer in the mammalian intestinal tract, Curr. Opin. in Biotechnol. 4(3):294-98 (Jun 1993).

As a result of these concerns, many governmental food, drug, health, and environmental regulatory agencies, as well as many end users, require that antibiotic resistance gene nucleic acid be removed from products or be absent from organisms for use in commerce. In addition, evidence demonstrating clearance of the selection antibiotics from the final product must be provided in order to secure regulatory clearance. The United Kingdom, Canada, France, the European Community, and the United States have all addressed the use of antibiotic resistance genes in foods, animal feeds, drugs and drug production, including recombinant drug production. Clearance of these agents, and

especially demonstrating such clearance, is expensive, time consuming, and often only minimally effective.

Because of the concerns inherent in the use of antibiotic resistance genes for selection in the production of recombinant polypeptides, alternative selection methods have been examined.

Auxotrophic Selection Markers

Auxotrophic selection markers have been utilized as an alternative to antibiotic selection in some systems. For example, auxotrophic markers have been widely utilized in yeast, due largely to the inefficiency of antibiotic resistance selection markers in these host cells. See, for example, JT Pronk, (2002) "Auxotrophic yeast strains in fundamental and applied research," App. & Envirn. Micro. 68(5): 2095-2100; Boeke et al., (1984) "A positive selection for mutants lacking orotodine-5'-phosphate decarboxylase activity in yeast; 5-fluoro-orotic acid resistance," Mol. Gen. Genet. 197: 345-346; Botstein & Davis, (1982) "Principles and practice of recombinant DNA research with yeast," p.607-636, in JN Strathern, EW Jones. And JR Broach (ed.), The molecular biology of the yeast Saccharomyces cerevisiae, Metabolism and gene expression, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Cost & Boeke, (1996) "A useful colony color phenotype associated with the yeast selectable/counter selectable marker *MET15*," Yeast 12: 939-941. However, yeast expression systems due not provide the potential speed and efficiency for producing target proteins that bacterial systems do.

Auxotrophic marker selection in bacteria has also previously been described. See, for example, U.S. Pat. Nos. 4,920,048, 5,691,185, 6,291,245, 6,413,768, 6,752,994, Struhl et al. (1976) PNAS USA 73; 1471-1475;; MacCormick, C. A., et al., (1995) "Construction of a food-grade host/vector system for *Lactococcus lactis* based on the lactose operon," *FEMS Microbiol. Lett.* 127:105-109; Dickely et al. (1995), "Isolation of *Lactococcus lactis* nonsense suppressors and construction of a food-grade cloning vector," *Mol. Microbiol.* 15:839-847; Sørensen et al., (2000) "A food-grade cloning system for industrial strains of *Lactococcus lactis*," *Appl. Environ. Microbiol.* 66:1253-1258; Fiedler & Skerra, (2001) "proBA complementation of an auxotrophic E. coli strain improves plasmid stability and expression yield during fermenter production of a recombinant antibody fragment," Gene 274: 111-118.

The use of auxotrophic selection markers in the previously described commercial scale bacterial fermentation systems has drawbacks that limit their use. A major drawback, as noted in U.S. Patent. No. 6,413,768, is that nutritional auxotrophic selection marker

systems generally suffer from cross feeding. The term cross feeding refers to the ability of a first cell, auxotrophic for a particular metabolite, to survive in the absence of the metabolite by obtaining its supply of that metabolite from its environment, and typically, from the medium for which the cell is auxotrophic by utilizing excreted intermediates of the metabolite, the metabolite itself, or a prototrophic enabling molecule produced by a second cell, prototrophic for the metabolite absent from the medium. See also GR Barker et al., Biochem. J. 157(1):221-27 (1976) (cross feeding of thymine in E. coli); TJ Kerr & GJ Tritz, J. Bact. 115(3):982-86 (Sep. 1973) (cross feeding of NAD in E. coli auxotrophic for NAD synthesis); GA Sprenger et al., FEMS Microbiol. Lett. 37(3):299-304 (1986) (selection of nalidixic acid to avoid the cross feeding problem).

Because cross feeding allows revertant bacteria to survive, cross feeding decreases the overall capacity of the fermentation process to produce the desired product at efficient and maximized levels due to the presence of fewer target protein producing host cells.

Expression Vector Control

Another hurdle to the creation of the ideal host cell is the inefficient and low level production of target polypeptides in the fermentation process. Controlling expression of the target protein until optimal host cell densities and fermentation conditions are reached allows for a more efficient and larger yield of polypeptide. The reasons for this are several fold, including a more efficient utilization of a particular carbon source and the reduction of extended metabolic stresses on the host cell.

In many cases, however, repression of expression of the target protein during cell growth can be imperfect, resulting in a significant amount of expression prior to the particular induction phase. This "leaky" repression

results in host cell stress, inefficient utilization of carbon source due to metabolic energy being diverted from normal cell growth to transgene, and a delay in reaching optimal cell density induction points, resulting in a more lengthy and costly fermentation run, and often, a reduced yield of the target protein.

Disclosed herein is an improved expression system for the production of target proteins, wherein the production is efficient, regulatable, and performed in a medium that minimizes of regulatory and environmental concerns.

Also disclosed are organisms for use as host cells in an improved expression system for the production of target proteins.

Also disclosed are processes for the improved production of target proteins.

Also disclosed are novel constructs and nucleic acids for use in an improved expression system for the production of target proteins.

All references, including any patents or patent applications, cited in this specification are hereby incorporated by reference. No admission is made that any reference constitutes prior art. The discussion of the references states what their authors assert, and the applicants reserve the right to challenge the accuracy and pertinency of the cited documents. It will be clearly understood that, although a number of prior art publications are referred to herein, this reference does not constitute an admission that any of these documents forms part of the common general knowledge in the art, in Australia or in any other country.

SUMMARY OF THE INVENTION

A first aspect provides an auxotrophic *Pseudomonas fluorescens* host cell for use in a bacterial expression system, wherein said auxotrophic host cell comprises:

(a) a chromosomal *lacI* gene insert;

(b) a first nucleotide sequence encoding a recombinant polypeptide, wherein said first nucleotide sequence is operably linked to a promoter capable of directing expression of the first nucleotide sequence in said auxotrophic host cell; and

(c) a second nucleotide sequence encoding an auxotrophic selection marker, wherein the auxotrophic selection marker is at least one polypeptide that restores prototrophy to the auxotrophic host cell.

A second aspect provides a process for producing a recombinant polypeptide comprising:

(a) expressing a first nucleotide sequence encoding the recombinant polypeptide in a *Pseudomonad* cell, wherein said cell comprises a chromosomal *lacI* gene insert, and wherein said cell has been genetically modified to be auxotrophic for at least one metabolite;

(b) expressing a second nucleotide sequence encoding an auxotrophic selection marker, wherein the auxotrophic selection marker is a polypeptide that restores prototrophy to the auxotrophic cell; and

(c) growing the cell on a medium that lacks the metabolite for which the cell is auxotrophic, wherein the cell is grown to a density of about 20 g/L or more.

A third aspect provides a process for the production of a recombinant polypeptide in the absence of antibiotics comprising:

(a) selecting a *Pseudomonad* cell, wherein said cell comprises a chromosomal *lacI* gene insert, and wherein the cell has been genetically modified to induce an auxotrophy for at least one metabolite;

(b) introducing into the cell a nucleic acid construct comprising:

i. a first nucleotide sequence encoding the recombinant polypeptide, wherein said first nucleotide sequence is operably linked to a

promoter capable of directing expression of the first nucleotide sequence in said cell; and

ii. a second nucleotide sequence encoding an auxotrophic selection marker, wherein the auxotrophic selection marker is a polypeptide that restores prototrophy to the cell;

(c) expressing the recombinant polypeptide and the prototrophy restoring polypeptide in the cell; and

(d) growing the cell on a medium that lacks the metabolite for which the cell is auxotrophic, wherein the cell is grown to a density of about 20 g/L or more.

A fourth aspect provides a process for the production of a recombinant polypeptide in the absence of antibiotics comprising:

(a) selecting a *Pseudomonad* cell, wherein said cell comprises a chromosomal *lacI* gene insert, and wherein the cell has been genetically modified to induce an auxotrophy for at least one metabolite;

(b) introducing into the cell a nucleic acid construct comprising:

i. a first nucleotide sequence encoding the recombinant polypeptide, wherein said first nucleotide sequence is operably linked to a promoter capable of directing expression of the first nucleotide sequence in said cell; and

ii. a second nucleotide sequence encoding an auxotrophic selection marker, wherein the auxotrophic selection marker is a polypeptide that restores prototrophy to the cell;

(c) expressing the recombinant polypeptide and the prototrophy restoring polypeptide in the cell; and

(d) growing the cell on a medium that lacks the metabolite for which the cell is auxotrophic, wherein the cell is grown to a density of about 20 g/L or more, wherein said cell density is at least about 25 g/L, 30 g/L, 35 g/L, 40 g/L, 45 g/L, 50 g/L, 60 g/L,

70 g/L, 80 g/L, 90 g/L, 100 g/L, 110 g/L, 120 g/L, 130 g/L, 140 g/L, or 150 g/L, wherein the *lacI* gene is inserted in a levansucrase locus.

It has been discovered that bacterial protein production can be improved by selecting as a host cell a Pseudomonad organism that is capable of non-antibiotic resistant, auxotrophic selection, and/or contains a chromosomal insert of a *lacI* gene or derivative.

Specifically, it has been discovered that the Pseudomonad organism *Pseudomonas fluorescens* is particularly well suited for this purpose. To this end, it has been surprisingly discovered that *Pseudomonas fluorescens* does not exhibit adverse cross feeding inhibition under auxotrophic selection during the high-cell density fermentation of recombinant polypeptides. Such a discovery allows for the use of auxotrophic *Pseudomonas fluorescens* as host cells in the efficient production of high levels of recombinant polypeptides, overcoming the drawbacks inherent with the use of antibiotic resistance selection markers and the problems of auxotrophic cross feeding present in other bacterial expression systems.

It has also been surprisingly discovered that the use of a LacI-encoding gene other than as part of a whole or truncated *Plac-lacI-lacZYA* operon in Pseudomonads surprisingly resulted in substantially improved repression of pre-induction recombinant protein expression, higher cell densities in commercial-scale fermentation, and higher yields of the desired product in comparison with previously taught *lacI-lacZYA* Pseudomonad chromosomal insertion (U.S. Pat. No. 5,169,760). This *lacI* insertion is as effective in repressing Plac-Ptac family promoter-controlled transgenes as a multi-copy plasmid encoding a LacI repressor protein in *Pseudomonas fluorescens*, thereby eliminating the need to maintain a separate plasmid encoding a LacI repressor protein in the cell and reducing potential production inefficiencies caused by such

maintenance.

It has also been discovered that the use of dual *lac* operator sequences provides superior repression of recombinant protein expression prior to induction without a concomitant reduction in subsequent induction yields in *Pseudomonas fluorescens*.

Therefore, disclosed herein are Pseudomonad organisms for use as host cells in the improved production of proteins.

In one embodiment, the Pseudomonad organisms may be genetically modified to induce an auxotrophy. In a particular embodiment, the Pseudomonad organism may be *Pseudomonas fluorescens*. In one embodiment, the auxotrophy may be a result of genetic modifications to at least one nitrogenous base compound biosynthesis gene, or at least one amino acid biosynthesis gene. In a further embodiment, the genetic modification may be to a gene encoding an enzyme active in the uracil biosynthetic pathway, the thymidine biosynthetic pathway, or the proline biosynthetic pathway. In still a further embodiment, the genetic modification may be to the *pyrF* gene encoding orotidine-5'-phosphate decarboxylase, the *thyA* gene encoding thymidylate synthase, or the *proC* gene encoding Δ^1-pyrroline-5-carboxylate reductase.

In another embodiment, the Pseudomonad organisms may be genetically modified to provide at least one copy of a LacI-encoding gene inserted into the genome, other than as part of the whole or truncated Plac-lac*I*-*lacZYA* operon. In a particular embodiment, the Pseudomonad host cell may be *Pseudomonas fluorescens*. In one embodiment, the Pseudomonad may contain a native *E. coli lacI* gene encoding the LacI repressor protein. In another embodiment, the Pseudomonad cell may contain the *lacI*^{Q} gene. In still another embodiment, the Pseudomonad cell may contain the *lacI*^{Q1} gene.

In another embodiment, a Pseudomonad organism may comprise a nucleic acid construct containing a nucleic

acid comprising at least one *lacO* sequence involved in the repression of transgene expression. In a particular embodiment, the Pseudomonad host cell may be *Pseudomonas fluorescens*. In one embodiment, the nucleic acid construct may comprise more than one *lacO* sequence. In another embodiment, the nucleic acid construct may comprise at least one, and preferably more than one, *lacOid* sequence. In one embodiment, the nucleic acid construct may comprise a *lacO* sequence, or derivative thereof, located 3' of a Plac family promoter, and a *lacO* sequence, or derivative thereof, located 5' of a Plac family promoter. In a particular embodiment, the *lacO* derivative may be a *lacOid* sequence.

In a further embodiment, the Pseudomonad organisms may be genetically modified to induce an auxotrophy and further modified to contain a chromosomal insertion of a native *E.coli lacI* gene, $lacI^{Q}$ gene, or $lacI^{Q1}$ gene other than as part of a whole or truncated *Plac-lacI-lacZYA* operon. In another embodiment, the Pseudomonad organism may be further modified to contain a nucleic acid construct comprising at least one *lacO* sequence involved in the repression of transgene expression. In a particular embodiment, the Pseudomonad organism may be *Pseudomonas fluorescens*.

Also disclosed are nucleic acid sequences for use in the improved production of proteins.

In one embodiment, nucleic acid sequences may encode prototrophy-restoring enzymes for use in auxotrophic Pseudomonad host cells. In a particular embodiment, nucleic acid sequences may encode nitrogenous base compound biosynthesis enzymes purified from the organism *Pseudomonas fluorescens*. In one embodiment, nucleic acid sequences may encode the *pyrF* gene in *Pseudomonas fluorescens* (SEQ. ID No.s 1 and 3). In another embodiment, a nucleic acid sequence may encode the *thyA* gene in *Pseudomonas fluorescens* (SEQ. ID. No. 4). In still another embodiment, nucleic acid sequences may

encode an amino acid biosynthetic compound purified from the organism *Pseudomonas fluorescens*. In a particular embodiment, a nucleic acid sequence may encode the *proC* gene in *Pseudomonas fluorescens* (SEQ. ID No.s 6 and 8).

Also disclosed are amino acid sequences which are the products of the novel nucleic acid expression.

Also disclosed are nucleic acid constructs for use in the improved production of peptides.

In one embodiment, a nucleic acid construct may be used in transforming a Pseudomonad host cell comprising a) a nucleic acid sequence encoding a recombinant polypeptide, and b) a nucleic acid sequence encoding a prototrophy-enabling enzyme. In another embodiment, the nucleic acid construct may further comprise c) a Plac-Ptac family promoter. In still another embodiment, the nucleic acid construct may further comprise d) at least one *lacO* sequence, or derivative, 3' of a lac or tac family promoter. In yet another embodiment, the nucleic acid construct may further comprise e) at least one *lacO* sequence, or derivative, 5' of a lac or tac family promoter. In one embodiment, the derivative *lacO* sequence can be a *lacOid* sequence. In a particular embodiment, the Pseudomonad organism may be *Pseudomonas fluorescens*.

In one embodiment, nucleic acid constructs may be used as expression vectors in Pseudomonad organisms comprising a) a nucleic acid sequence encoding a recombinant polypeptide, b) a Plac-Ptac family promoter, c) at least one *lacO* sequence, or derivative, 3' of a lac or tac family promoter, d) at least one *lacO* sequence, or derivative, 5' of a lac or tac family promoter. In one embodiment, the derivative *lacO* sequence can be a *lacOid* sequence. In one embodiment, the nucleic acid construct may further comprise e) a prototrophy-enabling selection marker for use in an auxotrophic Pseudomonad cell. In a particular embodiment, the Pseudomonad organism may be *Pseudomonas fluorescens*.

Also disclosed are modified cells for use in the

improved production of proteins.

In one embodiment, an auxotrophic Pseudomonad cell may be provided that has a nucleic acid construct comprising i) a recombinant polypeptide, and ii) a prototrophy-enabling nucleic acid. In another embodiment, the nucleic acid construct may further comprise iii) a Plac-Ptac family promoter. In still another embodiment, the nucleic acid construct may further comprise iv) more than one *lacO* sequence. In one embodiment, the Pseudomonad may be an auxotrophic *Pseudomonas fluorescens* cell. In a further embodiment, the invention may further comprise auxotrophic Pseudomonad organisms, including *Pseudomonas fluorescens*, that have been further genetically modified to contain a chromosomal insertion of a native *E.coli lacI* gene, $lacI^{Q}$ gene, or $lacI^{Q1}$ gene other than as part of a whole or truncated *Plac-lacI-lacZYA* operon.

In another embodiment, a Pseudomonad cell may comprise a *lacI* transgene, or derivative thereof, other than as part of a whole or truncated *Plac-lacI-lacZYA* operon, inserted into the chromosome, and b) a nucleic acid construct comprising i) a recombinant polypeptide, and ii) a Plac-Ptac family promoter. In still another embodiment, the nucleic acid construct may further comprise iii) at least one *lacO* sequence, and preferably, more than one *lacO* sequence. In one embodiment, the *lacO* sequence may be a *lacOid* sequence. In one embodiment, the Pseudomonad may be further modified to induce auxotrophy. In one embodiment, the Pseudomonad cell may be a *Pseudomonas fluorescens*.

Also disclosed are processes of expressing recombinant polypeptides for use in improved protein protection.

In one embodiment of the process, a nucleic acid construct comprising nucleic acids encoding a) a recombinant polypeptide, and b) a prototrophy-restoring enzyme in a Pseudomonad that is auxotrophic for at least

one metabolite is expressed. In an alternative embodiment, the Pseudomonad may be auxotrophic for more than one metabolite. In one embodiment, the Pseudomonad may be a *Pseudomonas fluorescens* cell. In a particular embodiment, a recombinant polypeptide may be expressed in a Pseudomonad that is auxotrophic for a metabolite, or combination of metabolites, selected from the group consisting of a nitrogenous base compound and an amino acid. In a more particular embodiment, recombinant polypeptides may be expressed in a Pseudomonad that is auxotrophic for a metabolite selected from the group consisting of uracil, proline, and thymidine. In another embodiment, the auxotrophy can be generated by the knock-out of the host *pyrF*, *proC*, or *thyA* gene, respectively. An alternative embodiment, recombinant polypeptides may be expressed in an auxotrophic Pseudomonad cell that has been genetically modified through the insertion of a native *E.coli lacI* gene, $lacI^{Q}$ gene, or $lacI^{Q1}$ gene, other than as part of the PlacI-*lacI-lacZYA* operon, into the host cell's chromosome. In one particular embodiment, a vector encoding the recombinant polypeptide expressed in the auxotroph may comprise at least one *lacOid* operator sequences. In one particular embodiment, the vector encoding the recombinant polypeptide expressed in the auxotrophic host cell may comprise at least two *lac* operator sequences, or derivatives thereof. In still a further embodiment, the recombinant polypeptide may be driven by a Plac family promoter.

In another embodiment, the process may involve the use of Pseudomonad host cells that have been genetically modified to provide at least one copy of a LacI encoding gene inserted into the Pseudomonad host cell's genome, wherein the lacI encoding gene is other than as part of the PlacI-*lacI-lacZYA* operon. In one embodiment, the gene encoding the Lac repressor protein may be identical to that of native *E. coli lacI* gene. In another embodiment, the gene encoding the Lac repressor protein may be the

$lacI^Q$ gene. In still another embodiment, the gene encoding the Lac repressor protein may be the $lacI^{Q1}$ gene. In a particular embodiment, the Pseudomonad host cell may be *Pseudomonas fluorescens*. In another embodiment, the Pseudomonad may be further genetically modified to produce an auxotrophic cell. In another embodiment, the process may produce recombinant polypeptide levels of at least about 3 g/L, 4 g/L, 5 g/L, 6 g/L, 7 g/L, 8 g/L, 9 g/L or at least about 10 g/L. In another embodiment, the recombinant polypeptide may be expressed in levels of between 3 g/L and 100 g/L.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 represents a comparison of the performance of *P. fluorescens* dual-plasmid expression systems using a *pyrF* marker (Δ and □) against the performance of *P. fluorescens* dual-plasmid expression systems using only antibiotic resistance markers (♦). All data shown are averages of 9-multiple, representative 20-L fermentations, with IPTG being added to induce target enzyme expression during mid-exponential phase. The upper set of three curves presents relative cell density data, which is read with reference to the left vertical axis. The lower set of three curves presents relative enzyme activity data for the target enzyme produced in the corresponding fermentations, and is read with reference to the right vertical axis. ♦ - *P. fluorescens* containing pMYC plasmid having a tac promoter-controlled target enzyme expression cassette and a tetracycline resistance marker gene and

containing a pCN plasmid having a *lacI* repressor expression cassette and a kanamycin resistance marker gene. Variance bars shown are for these data points (n=4), and represent the normal variance typically observed for this expression system among different fermentation runs. △ - *P. fluorescens* strain with inactivated genomic *pyrF* containing pMYC plasmid having a tac promoter-controlled target enzyme expression cassette and a *pyrF* auxotrophic marker gene and containing pCN plasmid having a *lacI* repressor expression cassette and a kanamycin resistance marker gene. □ - *P. fluorescens* strain with inactivated genomic *pyrF* and *proC* containing pMYC plasmid having a tac promoter-controlled target enzyme expression cassette and a *pyrF* auxotrophic marker gene and containing pCN plasmid having a *lacI* repressor expression cassette and a *proC* auxotrophic marker gene.

Figure 2 represents a map of the plasmid pDOW1249-2.

Figure 3 represents a map of the plasmid pDOW1269-2.

Figure 4 represents a schematic of lac operator constructs. LacZ represents the positions of the native E.coli lacO sequences. tac DC239, DC240 represents the position of the native E.coli lac operator on a construct comprising a tac promoter and a nitrilase encoding nucleic acid. Opt lacO DC281 represents the position of the lacOid operator sequence on a construct comprising a tac promoter and a nitrilase encoding nucleic acid. Dual lacO DC262 represents the position of a lacOid operator sequence 5', and wild type lac operator sequence 3' of a tac promoter on a construct further comprising a nitrilase encoding nucleic acid.

Figure 5 represents a Western Blot analysis (UnBlot) of LacI protein accumulation in the *lacI* integrant strains grown in a shake flask gene expression medium. Broth samples were normalized to OD_{600}, combined with LDS NuPAGE sample buffer (Invitrogen), 50mM DTT and heated at 95°C for 40 min, then centrifuged briefly. Aliquots of 20 uL were loaded on a 10%, 1 mm NuPAGE Bis-Tris gel run in MOPS with antioxidant in the inner chamber. Detection of the LacI protein was accomplished with an in-gel hybridization method ("UnBlot", Pierce), using a polyclonal rabbit antibody to LacI (Stratagene cat. no. 217449-51) at 1:1000 and the secondary antibody, Stabilized Goat Anti-rabbit Horseradish Peroxidase Conjugated Antibody (Pierce) at 1:500. The

horseradish peroxidase was visualized with UnBlot Stable Peroxide and UnBlot Luminol Enhancer as according to the UnBlot kit.

Figure 6 represents the composite of nitrilase accumulation profiles of DC140, DC239 and DC240. Data were compiled from DC140 (n=5), DC239 (n=5) and DC240 (n=4) runs. Dc140 is represented by ■. DC239 is represented by □. DC240 is represented by □. Fermentation runs were performed over a 48 hour period.

DETAILED DESCRIPTION OF THE INVENTION

In one embodiment, the Pseudomonad organisms have been genetically modified to induce an auxotrophy. In a particular embodiment, the Pseudomonad organism is *Pseudomonas fluorescens*. In one embodiment, the auxotrophy is a result of genetic modifications to at least one nitrogenous base compound biosynthesis gene, or at least one amino acid biosynthesis gene. In a further embodiment, the genetic modification is to a gene encoding an enzyme active in the uracil biosynthetic pathway, the thymidine biosynthetic pathway, or the proline biosynthetic pathway. In still a further embodiment, the genetic modification is to the *pyrF* gene encoding orotidine-5'-phosphate decarboxylase, the *thyA* gene encoding thymidilate synthase, or the *proC* gene encoding Δ^1-pyrroline-5-carboxylate reductase.

In another embodiment, the present invention provides Pseudomonad organisms that have been genetically modified to provide at least one copy of a LacI-encoding gene inserted into the genome, other than as part of the PlacI-*lacI-lacZYA* operon. In a particular embodiment, the Pseudomonad host cell is *Pseudomonas fluorescens*. In one embodiment, the Pseudomonad contains a native *E. coli lacI* gene encoding the LacI repressor protein.. In another embodiment, the Pseudomonad cell contains the $lacI^Q$ gene. In still another embodiment, the Pseudomonad cell contains the $lacI^{Q1}$ gene.

In another embodiment, a Pseudomonad organism is provided comprising a nucleic acid construct containing a nucleic acid comprising at least one *lacO* sequence involved in the repression of transgene expression. In a particular embodiment, the Pseudomonad host cell is *Pseudomonad fluorescens*. In one embodiment, the nucleic acid construct comprises more than one *lacO* sequence. In another embodiment, the nucleic acid construct comprises at least one, and preferably more than one, *lacOid* sequence. In one embodiment, the nucleic

acid construct comprises a *lacO* sequence, or derivative thereof, located 3' of a Plac family promoter, and a *lacO* sequence, or derivative thereof, located 5' of a Plac family promoter. In a particular embodiment, the *lacO* derivative is a *lacOid* sequence.

In a further embodiment, the present invention provides Pseudomonad organisms that have been genetically modified to induce an auxotrophy and further modified to contain a chromosomal insertion of a native *E.coli lacI* gene, *lacI Q* gene, or *lacIQ1* gene other than as part of a whole or truncated *Plac-lacI-lacZYA* operon. In another embodiment, the Pseudomonad organism is further modified to contain a nucleic acid construct comprising at least one *lacO* sequence involved in the repression of transgene expression. In a particular embodiment, the Pseudomonad organism is a *Pseudomonas fluorescens*.

The host cell provided by the present invention for use in an expression system producing recombinant polypeptides can be selected from the "Pseudomonads and closely related bacteria" or from a Subgroup thereof, as defined below. In one embodiment, the host cell is selected from the genus *Pseudomonas*. In a particular embodiment, the particular species of *Pseudomonas* is *P. fluorescens*. In a particular embodiment, the host cell is *Pseudomonas fluorescens* biotype A or biovar I.

Definitions

In the claims which follow and in the description of the invention, except where the context requires otherwise due to express language or necessary implication, the word "comprise" or variations such as "comprises" or "comprising" is used in an inclusive sense, i.e. to specify the presence of the stated features but not to preclude the presence or addition of further features in various embodiments of the invention.

The term "isolated" refers to nucleic acid, protein, or peptide that is substantially or essentially free from

other material components, for example, which can be cellular components.

The term "fragment" means a portion or partial sequence of a nucleotide, protein, or peptide sequence.

As used herein, the term "percent total cell protein" means the amount of protein or peptide in the host cell as a percentage of aggregate cellular protein.

The term "operably attached", as used herein, refers to any configuration in which the transcriptional and any translational regulatory elements are covalently attached to the encoding sequence in such disposition(s), relative to the coding sequence, that in and by action of the host cell, the regulatory elements can direct the expression of the coding sequence.

The term "auxotrophic", as used herein, refers to a cell which has been modified to eliminate or reduce its ability to produce a specific substance required for growth and metabolism.

As used herein, the term "percent total cell protein" means a measure of the fraction of total cell protein that represents the relative amount of a given protein expressed by the cell.

The term "prototrophy," as used herein, refers to a cell that is capable of producing a specific substance required for growth and metabolism.

As used herein, the term "homologous" or means either i) a protein or peptide that has an amino acid sequence that is substantially similar (i.e., at least 70, 75, 80, 85, 90, 95, or 98%) to the sequence of a given original protein or peptide and that retains a desired function of the original protein or peptide or ii) a nucleic acid that has a sequence that is substantially similar (i.e., at least 70, 75, 80, 5, 90, 95, or 98%) to the sequence of a given nucleic acid and that retains a desired function of the original nucleic acid sequence. In all of the embodiments of this invention and disclosure, any disclosed protein, peptide or nucleic acid can be substituted with a homologous or substantially homologous protein, peptide or nucleic acid that retains a desired function. In all of the embodiments of this invention and disclosure, when any nucleic acid is disclosed, it should be assumed that the invention also includes all nucleic acids that hybridize to the disclosed nucleic acid.

In one non-limiting embodiment, the non-identical amino acid sequence of the homologous polypeptide can be amino acids that are members of any one of the 15 conservative or semi-conservative groups shown in Table 1.

TABLE 1. SIMILAR AMINO ACID SUBSTITUTION GROUPS

Conservative Groups (8)	Semi-Conservative Groups (7)
Arg, Lys	Arg, Lys, His
Asp, Glu	Asn, Asp, Glu, Gln
Asn, Gln	
Ile, Leu, Val	Ile, Leu, Val, Met, Phe
Ala, Gly	Ala, Gly, Pro, Ser, Thr
Ser, Thr	Ser, Thr, Tyr
Phe, Tyr	Phe, Trp, Tyr
Cys (non-cystine), Ser	Cys (non-cystine), Ser, Thr

Amino acid sequences provided herein are represented by the following abbreviations:

A	Ala	alanine
P	Pro	proline
B		aspartate or asparagine
Q	Gln	glutamine
C	Cys	cysteine
R	Arg	arginine

D	Asp	aspartate
S	Ser	serine
E	Glu	glutamate
T	Thr	threonine
F	Phe	phenylalanine
G	Gly	glycine
V	Val	valine
H	His	histidine
W	Trp	tryptophan
I	Ile	isoleucine
Y	Tyr	tyrosine
Z		glutamate or glutamine
K	Lys	lysine
L	Leu	leucine
M	Met	methionine
N	Asn	asparagine

I. SELECTION OF PSEUDOMONADS AND RELATED BACTERIA AS HOST CELLS

The present invention provides the use of Pseudomonads and related bacteria as host cells in the improved production of proteins.

Auxotrophic Selection Efficiency

It has been discovered that Pseudomonads have the ability to utilize auxotrophic selection markers for the maintenance of protein expressing plasmids without the drawbacks typically associated with other systems, such as plasmid instability and cross-feeding.

Auxotrophic markers, in other host cell systems, are not always sufficient to maintain plasmids in every cell, especially during fermentations where loss of the plasmid may give plasmid-less cells a selective advantage, resulting in the accumulation of a large fraction of nonproductive cells, reducing product formation. Such revertant strains are especially troublesome if they have the ability to cross-feed the auxotrophic metabolite from prototrophic enabled bacteria. For example, use of the *trp* operon on a plasmid in an *E. coli* tryptophan auxotroph was not sufficient to prevent a large proportion of plasmid-less cells from accumulating, until combined with the *valS* gene (encoding valyl t-RNA synthetase) in a *valS*^{ts} host (Skogman, S. G.; Nilsson, J., Temperature-dependent retention of a tryptophan-operon-bearing plasmid in *Escherichia coli*. *Gene* 1984, 31, (1-3), 117-22.) Presumably, the cells containing the *trp* operon on a plasmid secreted enough tryptophan or related molecules to allow growth of plasmid-less cells. Likewise, using the *LEU2* gene on a xylitol-reductase

production plasmid in *leu2* mutant yeast resulted in plasmid loss; up to 80% of a fed-batch culture was made up of cells without a production plasmid, because leucine was secreted by plasmid-containing cells into the broth (Meinander, N. Q.; Hahn-Haegerdal, B., Fed-batch xylitol production with two recombinant *Saccharomyces cerevisiae* strains expressing XYL1 at different levels, using glucose as a cosubstrate: a comparison of production parameters and strain stability. *Biotechnology and Bioengineering* 1997, 54, (4), 391-399).

It has been discovered that *Pseudomonas fluorescens* (*Pf*) does not exhibit the inherent problems associated with cross-feeding observed in other host cell systems, for example, *E.coli* and yeast. While not wanting to be bound by any particular theory, it is thought that auxotrophic *Pseudomonas fluorescens* is a particularly suitable organism for use as a host cell because of the observed inability of a *Pf* auxotrophic cell to out compete a auxotrophic cell containing a prototrophic-enabling plasmid on a supplemented medium that contains the auxotrophic metabolite, indicating an innate difficulty of an *Pf* auxotroph to import the required metabolite. Because of this, *Pf* auxotrophic cells that lose the selection marker plasmid do not gain a selective advantage over *Pf* auxotrophic cells containing the selection marker, even in the presence of a supplemental metabolite, greatly reducing any potential effects of cross-feeding. Because of the reduced effects of cross-feeding, production yields of the recombinant polypeptide in a fermentation run are not reduced due to the presence of non-recombinant polypeptide producing cells.

LacI Insert

It has been discovered that Pseudomonads are able to use a single-copy *lacI* transgene, other than as part of a whole or truncated Plac-*lacI*-*lacZYA* operon, chromosomal insert to effectively repress protein expression until induction.

Transcription initiation from regulated promoters by RNA polymerase is activated or deactivated by the binding or releasing of a regulatory protein. Thus, regulated promoters include those that participate in negative control (i.e. repressible promoters), wherein the gene encoding the target polypeptide of interest is expressed only when the promoter is free of the regulator protein (i.e. a "repressor" protein), and those that participate in positive control, wherein the gene is expressed only when the promoter is bound by the regulator protein (i.e. an "activator" protein).

One of the most common classes of repressible promoters used in bacterial expression systems is the family of Plac-based promoters. The family of Plac-based promoters originates with the native *E. coli* lactose operon, referred to as the "lac" operon, also symbolized as "*lacZYA*," the expression of which is regulated by the expression product of the *lacI* gene. The native *E. coli* structure of the operon is "PlacI-*lacI*-PlacZ-*lacZYA*," wherein the native *E. coli* Plac promoter is represented by "PlacZ" (also called "PlacZYA"). "PlacI" represents the native promoter for the *lacI* gene, and "*lacI*" represents the gene encoding the lac repressor, i.e. the LacI protein. "*lacZYA*" represents the operon encoding the lactose utilization pathway.

The LacI-regulated promoters include, among others, the native *E. coli* lactose operon promoter ("Plac"). In addition, improved mutants have also been discovered, as have intra promoter hybrids of Plac, such as the "Ptac" promoter, "Ptrc" promoter, and the "PtacII" promoters. The Ptac promoter in *E. coli*, for example, is 3-fold stronger than the Plac promoter when fully derepressed. Therefore, it is frequently used for promoting high level regulated gene expression in *E. coli*. However, while the Plac promoter is 1,000-fold repressed by LacI, while the Ptac promoter is only 50-fold repressed under similar conditions (Lanzer, M. & H. Bujard. 1988. Proc. Natl. Acad. Sci. USA. 85:8973). Repression of the *E. coli* Ptac promoter or other lac related promoters, depends upon the concentration of the repressor, LacI. (De Boer, et al., 1983. Proc. Natl. Acad. Sci. USA. 78:21-25). As set forth above, release from repression can occur through the addition of an inducer which reduces the affinity of the repressor for its specific DNA binding site, in this case, the *lac* operator (lacO). Alternatively, a reduction in the concentration of the repressor relative to the molar concentration of specific DNA binding sites on the plasmid can also derepress the promoter.

If the *lacI* gene is located on a high copy number cloning plasmid, then a large amount of inducer is required to initiate expression because of the large amount of repressor produced in such a system.

In commercial production systems, the lac repressor is typically encoded by a gene whose expression is constitutive, i.e. non-regulated, thus providing an intracellular environment in which the desired transgene, encoding the desired target protein, is repressed until a desired host cell biomass or cell density is achieved. At that time, a quantity of a small molecule known as an inducer whose presence is effective to dissociate the repressor from the transgene, is added to the cell culture and taken up by the host cell, thereby permitting transcription of the transgene. In the case of lac repressor proteins, the inducer can be lactose or a non-metabolized, gratuitous inducer such as isopropyl-beta-D-thio-galactoside ("IPTG"). The selected point in time at which the inducer is to be added is referred to as the "induction phase."

A variety of lac repressor genes have been identified as useful for the repression of Plac family promoters present on recombinant polypeptide expression vectors. These include the native *E. coli lacI* gene and/or by variants thereof, including the $lacI^{Q}$ and $lacI^{Q1}$ genes that encode the same LacI protein, but at a higher expression level. For example, the $lacI^{Q}$ mutation is a single CG to TA change at -35 of the promoter region of *lacI* (Calos, M. 1978. Nature 274:762) which causes a 10-fold increase in LacI expression in *E. coli* (Mueller-Hill, B., et al. 1968. Proc. Natl. Acad. Sci. USA. 59:1259). Wild-type *E. coli* cells have a concentration of LacI of 10^{-8} M or about 10 molecules per cell, with 99% of the protein present as a tetramer (Fickert, R. & B. Mueller-Hill 1992. J. Mol. Biol. 226:59). Cells containing the $lacI^{Q}$ mutation contain about 100 molecules per cell or 10^{-7} M LacI. As a result, a number of bacterial expression systems have been developed in which Plac family promoter controlled transgenes, resident in plasmids, are maintained in host cells expressing LacI proteins at different levels, thereby repressing the desired transgene until a chosen "induction phase" of cell growth.

In many cases, however, repression of expression of the target protein during cell growth can be imperfect, resulting in a significant amount of expression prior to the particular induction phase. This "leaky" repression results in host cell stress, inefficient utilization of carbon source due to metabolic energy being diverted from normal cell growth to transgene, and a delay in reaching optimal cell density induction points, resulting in a more lengthy and costly fermentation run, and often, a reduced yield of the target protein.

One common strategy for improving repression of Plac-family promoter-driven transgenes has been to place a *lacI* or a *lacI*^{Q} gene on the plasmid bearing the Plac-family promoter-driven target gene (e.g. see MJR Stark in Gene 51:255-67 (1987) and E Amann et al. in Gene:301-15 (1988)). However, this often results in overproduction of the Lac repressor protein, which then requires use of an even higher inducer concentration to restore induction levels of the transgene to overcome the decrease in recombinant protein production. Moreover, the use of a second plasmid containing the *lacI* gene, separate from the plasmid containing the Plac-family promoter-driven target gene, requires the use of two different selection marker genes in order to maintain both plasmids in the expression host cell: one selection marker gene for each of the two different plasmids. The presence of the second selection marker gene, i.e. the selection marker gene for the second plasmid, in turn requires the use of either: 1) a separate antibiotic in the case of an antibiotic-resistance selection marker gene, which is costly and disadvantageous from a health/safety regulatory perspective; or 2) a separate metabolic deficiency in the host cell genome, in the case of an auxotrophic selection marker gene, which requires the additional work of mutating the host cell.

It has surprisingly been discovered that a *lacI* insertion, other than as part of a whole or truncated Plac-*lacI-lacZYA* operon, is as effective in repressing Plac-Ptac family promoter-controlled transgenes as a multi-copy plasmid encoding a LacI repressor protein in *Pseudomonas fluorescens*. This surprising discovery eliminates the need to maintain a separate plasmid encoding a LacI repressor protein in the cell, or eliminates the need to define an additional auxotrophic selection marker, and further reduces the potential production inefficiencies caused by such maintenance of a *lacI* containing plasmid.

In a previous attempt to regulate transgene expression in *Pseudomonas*, an *E.coli* PlacI-*lacI-lacZYA* operon that has been deleted of the lacZ promoter region, but that retains the constitutive PlacI promoter, was chromosomally inserted (See U.S. Pat. No. 5,169,760). The deletion allows for constitutive expression of the gene products of the lac operon. However, the inserted operon contains the *E.coli lacY* gene, which encodes for the lactose transporter protein lactose permease. Lactose permease is capable of transporting lactose, or similar derivatives, into the host cell from the medium. The presence of lactose permease may lead to increased importation of lactose-like contaminants from the medium, ultimately resulting in derepression of the Plac family promoter prior to induction. Furthermore, expression of the *lac* operon *lacZ*, *lacY*, and *lacA* gene products may result in the inefficient dedication of carbon utilization resources to these products, resulting in increased metabolic

stress on the cells, and delaying the establishment of a high cell density for induction. In addition, the larger *lacI-lacZYA* fusion operon may produce increased message instability compared to a *lacI* insert alone in a host cell.

It has been surprisingly discovered that the use of a LacI-encoding gene other than as part of a whole or truncated PlacI-*lacI-lacZYA* operon in Pseudomonads surprisingly resulted in substantially improved repression of pre-induction recombinant protein expression, higher cell densities in commercial-scale fermentation, and higher yields of the desired product in comparison with previously taught *lacI-lacZYA* Pseudomonad chromosomal insertion (U.S. Pat. No. 5,169,760).

Additional attempts to utilize derivative lacI genes, such as $lacI^{Q}$ and $lacI^{Q1}$, which are expressed at greater levels than lacI due to promoter modifications, have also been described. CG Glascock & MJ Weickert describe *E. coli* strains in which a separate LacI protein-encoding gene was present in the chromosome of the host cell in an attempt to assess the level of control of a plasmid-borne Ptac-driven target gene. See CG Glascock & MJ Weickert, "Using chromosomal $lacI^{Q1}$ to control expression of genes on high-copy number plasmids in *Escherichia coli*," Gene 223(1-2):221-31(1998); See also WO 97/04110. Among the LacI protein-encoding genes tested were *lacI*, $lacI^{Q}$, and $lacI^{Q1}$. The results obtained for the *lacI* gene and the $lacI^{Q}$ gene demonstrated inferior levels of repression of the Ptac-driven target gene when present on a high-copy number plasmid, resulting in substantial levels of pre-induction target gene expression. Only the high expressing $lacI^{Q1}$gene provided sufficient repression in that system.

Such a strategy, however, has the potential to increase costs by increasing the amount of inducer required to sufficiently derepress the promoter at induction, and decreasing yields due to the inability of the inducer to sufficiently bind all of the constitutively expressed repressor protein.

Comparatively, it has surprisingly been discovered that a single-copy *lacI* chromosomal insert was sufficient to repress Plac-Ptac family promoter driven transgene expression. Such a discovery allows potential cost saving measures on the amount of inducer used, and provides additional flexibility in the development of *Pseudomonas fluorescens* as a host cell in the improved production of proteins.

Pseudomonas Organisms

Pseudomonads and closely related bacteria, as used herein, is co-extensive with the group defined herein as "Gram(-) Proteobacteria Subgroup 1." " Gram(-) Proteobacteria

Subgroup 1" is more specifically defined as the group of Proteobacteria belonging to the families and/or genera described as falling within that taxonomic "Part" named "Gram-Negative Aerobic Rods and Cocci" by R.E. Buchanan and N.E. Gibbons (eds.), Bergey's Manual of Determinative Bacteriology, pp. 217-289 (8th ed., 1974) (The Williams & Wilkins Co., Baltimore, MD, USA) (hereinafter "Bergey (1974)"). Table 4 presents the families and genera of organisms listed in this taxonomic "Part."

TABLE 1. FAMILIES AND GENERA LISTED IN THE PART, "GRAM-NEGATIVE AEROBIC RODS AND COCCI" (IN BERGEY (1974))

Family I. Pseudomonadaceae	*Gluconobacter* *Pseudomonas* *Xanthomonas* *Zoogloea*
Family II. Azotobacteraceae	*Azomonas* *Azotobacter* *Beijerinckia* *Derxia*
Family III. Rhizobiaceae	*Agrobacterium* *Rhizobium*
Family IV. Methylomonadaceae	*Methylococcus* *Methylomonas*
Family V. Halobacteriaceae	*Halobacterium* *Halococcus*
Other Genera	*Acetobacter* *Alcaligenes* *Bordetella* *Brucella* *Francisella* *Thermus*

"Gram(-) Proteobacteria Subgroup 1" contains all Proteobacteria classified there under, as well as all Proteobacteria that would be classified according to the criteria used in forming that taxonomic "Part." As a result, "Gram(-) Proteobacteria Subgroup 1" excludes, e.g.: all Gram-positive bacteria; those Gram-negative bacteria, such as the

Enterobacteriaceae, which fall under others of the 19 "Parts" of this Bergey (1974) taxonomy; the entire "Family V. Halobacteriaceae" of this Bergey (1974) "Part," which family has since been recognized as being a non-bacterial family of Archaea; and the genus, *Thermus*, listed within this Bergey (1974) "Part," which genus which has since been recognized as being a non-Proteobacterial genus of bacteria.

"Gram(-) Proteobacteria Subgroup 1" further includes those Proteobacteria belonging to (and previously called species of) the genera and families defined in this Bergey (1974) "Part," and which have since been given other Proteobacterial taxonomic names. In some cases, these re-namings resulted in the creation of entirely new Proteobacterial genera. For example, the genera *Acidovorax, Brevundimonas, Burkholderia, Hydrogenophaga, Oceanimonas, Ralstonia, and Stenotrophomonas*, were created by regrouping organisms belonging to (and previously called species of) the genus *Pseudomonas* as defined in Bergey (1974). Likewise, e.g., the genus *Sphingomonas* (and the genus *Blastomonas*, derived therefrom) was created by regrouping organisms belonging to (and previously called species of) the genus *Xanthomonas* as defined in Bergey (1974). Similarly, e.g., the genus *Acidomonas* was created by regrouping organisms belonging to (and previously called species of) the genus *Acetobacter* as defined in Bergey (1974). Such subsequently reassigned species are also included within "Gram(-) Proteobacteria Subgroup 1" as defined herein.

In other cases, Proteobacterial species falling within the genera and families defined in this Bergey (1974) "Part" were simply reclassified under other, existing genera of Proteobacteria. For example, in the case of the genus *Pseudomonas, Pseudomonas enalia* (ATCC 14393), *Pseudomonas nigrifaciens* (ATCC 19375), and *Pseudomonas putrefaciens* (ATCC 8071) have since been reclassified respectively as *Alteromonas haloplanktis*, *Alteromonas nigrifaciens*, and *Alteromonas putrefaciens*. Similarly, e.g., *Pseudomonas acidovorans* (ATCC 15668) and *Pseudomonas testosteroni* (ATCC 11996) have since been reclassified as *Comamonas acidovorans* and *Comamonas testosteroni*, respectively; and *Pseudomonas nigrifaciens* (ATCC 19375) and *Pseudomonas piscicida* (ATCC 15057) have since been reclassified respectively as *Pseudoalteromonas nigrifaciens* and *Pseudoalteromonas piscicida*. Such subsequently reassigned Proteobacterial species are also included within "Gram(-) Proteobacteria Subgroup 1" as defined herein.

"Gram(-) Proteobacteria Subgroup 1" also includes Proteobacterial species that have since been discovered, or that have since been reclassified as belonging, within the Proteobacterial families and/or genera of this Bergey (1974) "Part." In regard to Proteobacterial families, "Gram(-) Proteobacteria Subgroup 1" also includes Proteobacteria

classified as belonging to any of the families: Pseudomonadaceae, Azotobacteraceae (now often called by the synonym, the "Azotobacter group" of Pseudomonadaceae), Rhizobiaceae, and Methylomonadaceae (now often called by the synonym, "Methylococcaceae"). Consequently, in addition to those genera otherwise described herein, further Proteobacterial genera falling within "Gram(-) Proteobacteria Subgroup 1" include: 1) Azotobacter group bacteria of the genus *Azorhizophilus;* 2) Pseudomonadaceae family bacteria of the genera *Cellvibrio, Oligella, and Teredinibacter*; 3) Rhizobiaceae family bacteria of the genera *Chelatobacter, Ensifer, Liberibacter* (also called "*Candidatus Liberibacter*"), and *Sinorhizobium*; and 4) Methylococcaceae family bacteria of the genera *Methylobacter, Methylocaldum, Methylomicrobium, Methylosarcina, and Methylosphaera.*

In one embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 1," as defined above.

In another embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 2." "Gram(-) Proteobacteria Subgroup 2" is defined as the group of Proteobacteria of the following genera (with the total numbers of catalog-listed, publicly-available, deposited strains thereof indicated in parenthesis, all deposited at ATCC, except as otherwise indicated): *Acidomonas* (2); *Acetobacter* (93); *Gluconobacter* (37); *Brevundimonas* (23); *Beijerinckia* (13); *Derxia* (2); *Brucella* (4); *Agrobacterium* (79); *Chelatobacter* (2); *Ensifer* (3); *Rhizobium* (144); *Sinorhizobium* (24); *Blastomonas* (1); *Sphingomonas* (27); *Alcaligenes* (88); *Bordetella* (43); *Burkholderia* (73); *Ralstonia* (33); *Acidovorax* (20); *Hydrogenophaga* (9); *Zoogloea* (9); *Methylobacter* (2); *Methylocaldum* (1 at NCIMB); *Methylococcus* (2); *Methylomicrobium* (2); *Methylomonas* (9); *Methylosarcina* (1); *Methylosphaera*; *Azomonas* (9); *Azorhizophilus* (5); *Azotobacter* (64); *Cellvibrio* (3); *Oligella* (5); *Pseudomonas* (1139); *Francisella* (4); *Xanthomonas* (229); *Stenotrophomonas* (50); and *Oceanimonas* (4).

Exemplary host cell species of "Gram(-) Proteobacteria Subgroup 2" include, but are not limited to the following bacteria (with the ATCC or other deposit numbers of exemplary strain(s) thereof shown in parenthesis): *Acidomonas methanolica* (ATCC 43581); *Acetobacter aceti* (ATCC 15973); *Gluconobacter oxydans* (ATCC 19357); *Brevundimonas diminuta* (ATCC 11568); *Beijerinckia indica* (ATCC 9039 and ATCC 19361); *Derxia gummosa* (ATCC 15994); *Brucella melitensis* (ATCC 23456), *Brucella abortus* (ATCC 23448); *Agrobacterium tumefaciens* (ATCC 23308), *Agrobacterium radiobacter* (ATCC 19358), *Agrobacterium rhizogenes* (ATCC 11325); *Chelatobacter heintzii* (ATCC 29600); *Ensifer adhaerens* (ATCC 33212); *Rhizobium leguminosarum* (ATCC 10004); *Sinorhizobium fredii* (ATCC 35423); *Blastomonas natatoria* (ATCC 35951); *Sphingomonas paucimobilis*

(ATCC 29837); *Alcaligenes faecalis* (ATCC 8750); *Bordetella pertussis* (ATCC 9797); *Burkholderia cepacia* (ATCC 25416); *Ralstonia pickettii* (ATCC 27511); *Acidovorax facilis* (ATCC 11228); *Hydrogenophaga flava* (ATCC 33667); *Zoogloea ramigera* (ATCC 19544); *Methylobacter luteus* (ATCC 49878); *Methylocaldum gracile* (NCIMB 11912); *Methylococcus capsulatus* (ATCC 19069); *Methylomicrobium agile* (ATCC 35068); *Methylomonas methanica* (ATCC 35067); *Methylosarcina fibrata* (ATCC 700909); *Methylosphaera hansonii* (ACAM 549); *Azomonas agilis* (ATCC 7494); *Azorhizophilus paspali* (ATCC 23833); *Azotobacter chroococcum* (ATCC 9043); *Cellvibrio mixtus* (UQM 2601); *Oligella urethralis* (ATCC 17960); *Pseudomonas aeruginosa* (ATCC 10145), *Pseudomonas fluorescens* (ATCC 35858); *Francisella tularensis* (ATCC 6223); *Stenotrophomonas maltophilia* (ATCC 13637); *Xanthomonas campestris* (ATCC 33913); and *Oceanimonas doudoroffii* (ATCC 27123).

In another embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 3." "Gram(-) Proteobacteria Subgroup 3" is defined as the group of Proteobacteria of the following genera: *Brevundimonas; Agrobacterium; Rhizobium; Sinorhizobium; Blastomonas; Sphingomonas; Alcaligenes; Burkholderia; Ralstonia; Acidovorax; Hydrogenophaga; Methylobacter; Methylocaldum; Methylococcus; Methylomicrobium; Methylomonas; Methylosarcina; Methylosphaera; Azomonas; Azorhizophilus; Azotobacter; Cellvibrio; Oligella; Pseudomonas; Teredinibacter; Francisella; Stenotrophomonas; Xanthomonas*; and *Oceanimonas*.

In another embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 4." "Gram(-) Proteobacteria Subgroup 4" is defined as the group of Proteobacteria of the following genera: *Brevundimonas; Blastomonas; Sphingomonas; Burkholderia; Ralstonia; Acidovorax; Hydrogenophaga; Methylobacter; Methylocaldum; Methylococcus; Methylomicrobium; Methylomonas; Methylosarcina; Methylosphaera; Azomonas; Azorhizophilus; Azotobacter; Cellvibrio; Oligella; Pseudomonas; Teredinibacter; Francisella; Stenotrophomonas; Xanthomonas*; and *Oceanimonas*.

In an embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 5." "Gram(-) Proteobacteria Subgroup 5" is defined as the group of Proteobacteria of the following genera: *Methylobacter; Methylocaldum; Methylococcus; Methylomicrobium; Methylomonas; Methylosarcina; Methylosphaera; Azomonas; Azorhizophilus; Azotobacter; Cellvibrio; Oligella; Pseudomonas; Teredinibacter; Francisella; Stenotrophomonas; Xanthomonas*; and *Oceanimonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 6." "Gram(-) Proteobacteria Subgroup 6" is defined as the group of Proteobacteria of the following genera: *Brevundimonas; Blastomonas; Sphingomonas; Burkholderia; Ralstonia; Acidovorax; Hydrogenophaga; Azomonas; Azorhizophilus; Azotobacter; Cellvibrio; Oligella; Pseudomonas; Teredinibacter; Stenotrophomonas; Xanthomonas*; and *Oceanimonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 7." "Gram(-) Proteobacteria Subgroup 7" is defined as the group of Proteobacteria of the following genera: Azomonas; *Azorhizophilus; Azotobacter; Cellvibrio; Oligella; Pseudomonas; Teredinibacter; Stenotrophomonas; Xanthomonas*; and *Oceanimonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 8." "Gram(-) Proteobacteria Subgroup 8" is defined as the group of Proteobacteria of the following genera: *Brevundimonas; Blastomonas; Sphingomonas; Burkholderia; Ralstonia; Acidovorax; Hydrogenophaga; Pseudomonas; Stenotrophomonas; Xanthomonas;* and *Oceanimonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 9." "Gram(-) Proteobacteria Subgroup 9" is defined as the group of Proteobacteria of the following genera: *Brevundimonas; Burkholderia; Ralstonia; Acidovorax; Hydrogenophaga; Pseudomonas; Stenotrophomonas*; and *Oceanimonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 10." "Gram(-) Proteobacteria Subgroup 10" is defined as the group of Proteobacteria of the following genera: *Burkholderia; Ralstonia; Pseudomonas; Stenotrophomonas*; and *Xanthomonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 11." "Gram(-) Proteobacteria Subgroup 11" is defined as the group of Proteobacteria of the genera: *Pseudomonas; Stenotrophomonas*; and *Xanthomonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 12." "Gram(-) Proteobacteria Subgroup 12" is defined as the group of Proteobacteria of the following genera: *Burkholderia; Ralstonia; Pseudomonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 13." "Gram(-) Proteobacteria Subgroup 13" is defined as the group of Proteobacteria of the following genera: *Burkholderia; Ralstonia; Pseudomonas*; and *Xanthomonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 14." "Gram(-) Proteobacteria Subgroup 14" is defined as the group of Proteobacteria of the following genera: *Pseudomonas* and *Xanthomonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 15." "Gram(-) Proteobacteria Subgroup 15" is defined as the group of Proteobacteria of the genus *Pseudomonas*.

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 16." "Gram(-) Proteobacteria Subgroup 16" is defined as the group of Proteobacteria of the following *Pseudomonas* species (with the ATCC or other deposit numbers of exemplary strain(s) shown in parenthesis): *Pseudomonas abietaniphila* (ATCC 700689); *Pseudomonas aeruginosa* (ATCC 10145); *Pseudomonas alcaligenes* (ATCC 14909); *Pseudomonas anguilliseptica* (ATCC 33660); *Pseudomonas citronellolis* (ATCC 13674); *Pseudomonas flavescens* (ATCC 51555); *Pseudomonas mendocina* (ATCC 25411); *Pseudomonas nitroreducens* (ATCC 33634); *Pseudomonas oleovorans* (ATCC 8062); *Pseudomonas pseudoalcaligenes* (ATCC 17440); *Pseudomonas resinovorans* (ATCC 14235); *Pseudomonas straminea* (ATCC 33636); *Pseudomonas agarici* (ATCC 25941); *Pseudomonas alcaliphila; Pseudomonas alginovora; Pseudomonas andersonii; Pseudomonas asplenii* (ATCC 23835); *Pseudomonas azelaica* (ATCC 27162); *Pseudomonas beijerinckii* (ATCC 19372); *Pseudomonas borealis; Pseudomonas boreopolis* (ATCC 33662); *Pseudomonas brassicacearum; Pseudomonas butanovora* (ATCC 43655); *Pseudomonas cellulosa* (ATCC 55703); *Pseudomonas aurantiaca* (ATCC 33663); *Pseudomonas chlororaphis* (ATCC 9446, ATCC 13985, ATCC 17418, ATCC 17461); *Pseudomonas fragi* (ATCC 4973); *Pseudomonas lundensis* (ATCC 49968); *Pseudomonas taetrolens* (ATCC 4683); *Pseudomonas cissicola* (ATCC 33616); *Pseudomonas coronafaciens; Pseudomonas diterpeniphila; Pseudomonas elongata* (ATCC 10144); *Pseudomonas flectens* (ATCC 12775); *Pseudomonas azotoformans; Pseudomonas brenneri; Pseudomonas cedrella; Pseudomonas corrugata* (ATCC 29736); *Pseudomonas extremorientalis; Pseudomonas fluorescens* (ATCC 35858); *Pseudomonas gessardii; Pseudomonas libanensis; Pseudomonas mandelii* (ATCC 700871); *Pseudomonas marginalis* (ATCC 10844); *Pseudomonas migulae; Pseudomonas mucidolens* (ATCC 4685); *Pseudomonas orientalis; Pseudomonas rhodesiae; Pseudomonas synxantha* (ATCC 9890); *Pseudomonas tolaasii* (ATCC 33618); *Pseudomonas veronii* (ATCC 700474); *Pseudomonas frederiksbergensis; Pseudomonas geniculata* (ATCC 19374); *Pseudomonas gingeri; Pseudomonas graminis; Pseudomonas grimontii; Pseudomonas halodenitrificans; Pseudomonas halophila; Pseudomonas hibiscicola* (ATCC 19867); *Pseudomonas huttiensis* (ATCC 14670); *Pseudomonas hydrogenovora; Pseudomonas jessenii* (ATCC 700870); *Pseudomonas kilonensis; Pseudomonas lanceolata* (ATCC 14669); *Pseudomonas lini; Pseudomonas marginata* (ATCC 25417); *Pseudomonas mephitica* (ATCC 33665);

Pseudomonas denitrificans (ATCC 19244); *Pseudomonas pertucinogena* (ATCC 190); *Pseudomonas pictorum* (ATCC 23328); *Pseudomonas* psychrophila; *Pseudomonas* fulva (ATCC 31418); *Pseudomonas monteilii* (ATCC 700476); *Pseudomonas mosselii; Pseudomonas oryzihabitans* (ATCC 43272); *Pseudomonas plecoglossicida* (ATCC 700383); *Pseudomonas putida* (ATCC 12633); *Pseudomonas reactans; Pseudomonas spinosa* (ATCC 14606); *Pseudomonas balearica; Pseudomonas luteola* (ATCC 43273); *Pseudomonas stutzeri* (ATCC 17588); *Pseudomonas amygdali* (ATCC 33614); *Pseudomonas avellanae* (ATCC 700331); *Pseudomonas caricapapayae* (ATCC 33615); *Pseudomonas cichorii* (ATCC 10857); *Pseudomonas ficuserectae* (ATCC 35104); *Pseudomonas fuscovaginae; Pseudomonas meliae* (ATCC 33050); *Pseudomonas syringae* (ATCC 19310); *Pseudomonas viridiflava* (ATCC 13223); *Pseudomonas thermocarboxydovorans* (ATCC 35961); *Pseudomonas thermotolerans; Pseudomonas thivervalensis; Pseudomonas vancouverensis* (ATCC 700688); *Pseudomonas wisconsinensis*; and *Pseudomonas xiamenensis.*

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 17." "Gram(-) Proteobacteria Subgroup 17" is defined as the group of Proteobacteria known in the art as the "fluorescent Pseudomonads" including those belonging, e.g., to the following *Pseudomonas* species: *Pseudomonas azotoformans; Pseudomonas brenneri; Pseudomonas cedrella; Pseudomonas corrugata; Pseudomonas extremorientalis; Pseudomonas fluorescens; Pseudomonas gessardii; Pseudomonas libanensis; Pseudomonas mandelii; Pseudomonas marginalis; Pseudomonas migulae; Pseudomonas mucidolens; Pseudomonas orientalis; Pseudomonas rhodesiae; Pseudomonas synxantha; Pseudomonas tolaasii*; and *Pseudomonas veronii.*

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 18." "Gram(-) Proteobacteria Subgroup 18" is defined as the group of all subspecies, varieties, strains, and other sub-special units of the species *Pseudomonas fluorescens*, including those belonging, e.g., to the following (with the ATCC or other deposit numbers of exemplary strain(s) shown in parenthesis): *Pseudomonas fluorescens* biotype A, also called biovar 1 or biovar I (ATCC 13525); *Pseudomonas fluorescens* biotype B, also called biovar 2 or biovar II (ATCC 17816); *Pseudomonas fluorescens* biotype C, also called biovar 3 or biovar III (ATCC 17400); *Pseudomonas fluorescens* biotype F, also called biovar 4 or biovar IV (ATCC 12983); *Pseudomonas fluorescens* biotype G, also called biovar 5 or biovar V (ATCC 17518); *Pseudomonas fluorescens* biovar VI; *Pseudomonas fluorescens* Pf0-1; *Pseudomonas fluorescens* Pf-5 (ATCC BAA-477); *Pseudomonas fluorescens* SBW25; and *Pseudomonas fluorescens subsp. cellulosa* (NCIMB 10462).

The host cell can be selected from "Gram(-) Proteobacteria Subgroup 19." "Gram(-) Proteobacteria Subgroup 19" is defined as the group of all strains of *Pseudomonas fluorescens* biotype A. A particularly particular strain of this biotype is *P. fluorescens* strain MB101 (see US Patent No. 5,169,760 to Wilcox), and derivatives thereof.

In one embodiment, the host cell is any of the Proteobacteria of the order Pseudomonadales. In a particular embodiment, the host cell is any of the Proteobacteria of the family Pseudomonadaceae.

In a particular embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 1." In a particular embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 2." In a particular embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 3." In a particular embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 5." In a particular embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 7." In a particular embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 12." In a particular embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 15." In a particular embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 17." In a particular embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 18." In a particular embodiment, the host cell is selected from "Gram(-) Proteobacteria Subgroup 19."

Additional *P. fluorescens* strains that can be used in the present invention include *Pseudomonas fluorescens* Migula and *Pseudomonas fluorescens* Loitokitok, having the following ATCC designations: [NCIB 8286]; NRRL B-1244; NCIB 8865 strain CO1; NCIB 8866 strain CO2; 1291 [ATCC 17458; IFO 15837; NCIB 8917; LA; NRRL B-1864; pyrrolidine; PW2 [ICMP 3966; NCPPB 967; NRRL B-899]; 13475; NCTC 10038; NRRL B-1603 [6; IFO 15840]; 52-1C; CCEB 488-A [BU 140]; CCEB 553 [IEM 15/47]; IAM 1008 [AHH-27]; IAM 1055 [AHH-23]; 1 [IFO 15842]; 12 [ATCC 25323; NIH 11; den Dooren de Jong 216]; 18 [IFO 15833; WRRL P-7]; 93 [TR-10]; 108 [52-22; IFO 15832]; 143 [IFO 15836; PL]; 149 [2-40-40; IFO 15838]; 182 [IFO 3081; PJ 73]; 184 [IFO 15830]; 185 [W2 L-1]; 186 [IFO 15829; PJ 79]; 187 [NCPPB 263]; 188 [NCPPB 316]; 189 [PJ227; 1208]; 191 [IFO 15834; PJ 236; 22/1]; 194 [Klinge R-60; PJ 253]; 196 [PJ 288]; 197 [PJ 290]; 198 [PJ 302]; 201 [PJ 368]; 202 [PJ 372]; 203 [PJ 376]; 204 [IFO 15835; PJ 682]; 205 [PJ 686]; 206 [PJ 692]; 207 [PJ 693]; 208 [PJ 722]; 212 [PJ 832]; 215 [PJ 849]; 216 [PJ 885]; 267 [B-9]; 271 [B-1612]; 401 [C71A; IFO 15831; PJ 187]; NRRL B-3178 [4; IFO 15841]; KY 8521; 3081; 30-21; [IFO 3081]; N; PYR; PW; D946-B83 [BU 2183; FERM-P 3328]; P-2563 [FERM-P 2894; IFO 13658]; IAM-1126 [43F]; M-1; A506 [A5-06]; A505 [A5-05-1]; A526

[A5-26]; B69; 72; NRRL B-4290; PMW6 [NCIB 11615]; SC 12936; A1 [IFO 15839]; 1847 [CDC-EB]; F 1848 [CDC 93]; NCIB 10586; P17; F-12; AmMS 257; PRA25; 6133D02; 6519E01; N1; SC15208; BNL-WVC; NCTC 2583 [NCIB 8194]; H13; 1013 [ATCC 11251; CCEB 295]; IFO 3903; 1062; or Pf-5.

II. AUXOTROPHIC SELECTION MARKERS

Disclosed herein are Pseudomonads and related cells that have been genetically modified to induce auxotrophy for at least one metabolite. The genetic modification can be to a gene or genes encoding an enzyme that is operative in a metabolic pathway, such as an anabolic biosynthetic pathway or catabolic utilization pathway. Preferably, the host cell has all operative genes encoding a given biocatalytic activity deleted or inactivated in order to ensure removal of the biocatalytic activity. In a particular embodiment, the Pseudomonad is a *Pseudomonas fluorescens* cell.

One or more than one metabolic activity may be selected for knock-out or replacement. In the case of native auxotrophy(ies), additional metabolic knockouts or replacements can be provided. Where multiple activities are selected, the auxotrophy-restoring selection markers can be of a biosynthetic-type (anabolic), of a utilization-type (catabolic), or may be chosen from both types. For example, one or more than one activity in a given biosynthetic pathway for the selected compound may be knocked-out; or more than one activity, each from different biosynthetic pathways, may be knocked-out. The corresponding activity or activities are then provided by at least one recombinant vector which, upon transformation into the cell, restores prototrophy to the cell.

Compounds and molecules whose biosynthesis or utilization can be targeted to produce auxotrophic host cells include: lipids, including, for example, fatty acids; mono- and disaccharides and substituted derivatives thereof, including, for example, glucose, fructose,

sucrose, glucose-6-phosphate, and glucuronic acid, as well as Entner-Doudoroff and Pentose Phosphate pathway intermediates and products; nucleosides, nucleotides, dinucleotides, including, for example, ATP, dCTP, FMN, FAD, NAD, NADP, nitrogenous bases, including, for example, pyridines, purines, pyrimidines, pterins, and hydro-, dehydro-, and/or substituted nitrogenous base derivatives, such as cofactors, for example, biotin, cobamamide, riboflavine, thiamine; organic acids and glycolysis and citric acid cycle intermediates and products, including, for example, hydroxyacids and amino acids; storage carbohydrates and storage poly(hydroxyalkanoate) polymers, including, for example, cellulose, starch, amylose, amylopectin, glycogen, poly-hydroxybutyrate, and polylactate.

In one embodiment, the biocatalytic activity(ies) knocked out to produce the auxotrophic host cell is selected from the group consisting of: the lipids; the nucleosides, nucleotides, dinucleotides, nitrogenous bases, and nitrogenous base derivatives; and the organic acids and glycolysis and citric acid cycle intermediates and products. Preferably, the biocatalytic activity(ies) knocked out is selected from the group consisting of: the nucleosides, nucleotides, dinucleotides, nitrogenous bases, and nitrogenous base derivatives; and the organic acids and glycolysis and citric acid cycle intermediates and products. More preferably, the biocatalytic activity(ies) knocked out is selected from the group consisting of: the pyrimidine nucleosides, nucleotides, dinucleotides, nitrogenous bases, and nitrogenous base derivatives; and the amino acids.

A given transgenic host cell may use one or more than one selection marker or selection marker system. For example, one or more biosynthesis selection marker(s) or selection marker system(s) according to the present invention may be used together with each other, and/or may be used in combination with a utilization-type selection marker or selection marker system according to the present invention. In any one of these prototrophy-enabling embodiments, the host cell may also contain one or more non-auxotrophic selection marker(s) or selection marker system(s). Examples of non-auxotrophic selection marker(s) and system(s) include, for example: toxin-resistance marker genes such as antibiotic-resistance genes that encode an enzymatic activity that degrades an antibiotic; toxin-resistant marker genes, such as, for example, imidazolinone-resistant mutants of acetolactate synthase ("ALS;" EC 2.2.1.6) in which mutation(s) are expressed that make the enzyme insensitive to toxin-inhibition exhibited by versions of the enzyme that do not contain such mutation(s). The compound(s) may exert this effect directly; or the compound(s) may exert this effect indirectly, for example, as a result of metabolic action of the cell that converts the compound(s) into toxin form or as a result of combination of the compound(s) with at least one further compound(s).

Bacterial-host-operative genes encoding such marker enzymes can be obtained from the bacterial host cell strain chosen for construction of the knock-out cell, from other bacteria, or from other organisms, and may be used in native form or modified (e.g., mutated or sequence recombined) form. For example, a DNA coding sequence for an enzyme exhibiting the knocked out biocatalytic activity may be obtained from one or more organisms and then operatively attached to DNA regulatory elements operative within the host cell. In specific, all of the chosen host's intracellular genes that encode a selected enzymatic activity are knocked-out; the bacterial knock-out host is then transformed with a vector containing at

least one operative copy of a native or non-native gene encoding an enzyme exhibiting the activity lost by the bacterial knockout.

Bacterial and other genes encoding such enzymes can be selected and obtained through various resources available to one of ordinary skill in the art. These include the nucleotide sequences of enzyme coding sequences and species-operative DNA regulatory elements. Useful on-line InterNet resources include, e.g.,: (1) the ExPASy proteomics facility (see the ENZYME and BIOCHEMICAL PATHWAYS MAPS features) of the Swiss Institute of Bioinformatics (Bâtiment Ecole de Pharmacie, Room 3041; Université de Lausanne; 1015 Lausanne-Dorigny; Switzerland) available at, e.g., http://us.expasy.org/; and (2) the GenBank facility and other Entrez resources (see the PUBMED, PROTEIN, NUCLEOTIDE, STRUCTURE, GENOME, et al. features) offered by the National Center for Biotechnology Information (NCBI, National Library of Medicine, National Institutes of Health, U.S. Dept. of Health & Human Services; Building 38A; Bethesda, Maryland, USA) and available at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi.

The selected coding sequence may be modified by altering the genetic code thereof to match that employed by the bacterial host cell, and the codon sequence thereof may be enhanced to better approximate that employed by the host. Genetic code selection and codon frequency enhancement may be performed according to any of the various methods known to one of ordinary skill in the art, e.g., oligonucleotide-directed mutagenesis. Useful on-line InterNet resources to assist in this process include, e.g.: (1) the Codon Usage Database of the Kazusa DNA Research Institute (2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818 Japan) and available at http://www.kazusa.or.jp/codon/; and (2) the Genetic Codes tables available from the NCBI Taxonomy database at http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi?mode=c. For example, *Pseudomonas* species are reported as utilizing Genetic Code Translation Table 11 of the NCBI Taxonomy site, and at the Kazusa site as exhibiting the codon usage frequency of the table shown at http://www.kazusa.or.jp/codon/cgibin/.

In a particular embodiment, *Pseudomonas fluorescens* can be used as the host cell. In one embodiment, *Pseudomonas fluorescens* provides at least one auxotrophic selection marker gene. In an alternative embodiment, *Pseudomonas fluorescens* provides all auxotrophic selection marker genes. In a particular embodiment, *Pseudomonas fluorescens* can both be the host cell and provide at least one, and preferably all, auxotrophic selection marker genes.

30

Biosynthetic Nucleoside and Nitrogenous Base Selection Markers

In one embodiment, a biosynthetic enzyme involved in anabolic metabolism can be chosen as the auxotrophic selection marker. In particular, the biosynthetic enzyme can be selected from those involved in biosynthesis of the nucleosides, nucleotides, dinucleotides, nitrogenous bases, and nitrogenous base derivatives.

In a particular embodiment at least one purine-type biosynthetic enzyme can be chosen as an auxotrophic selection marker. Such purine biosynthetic enzymes include, for example, adenine phosphoribosyltransferases, adenylosuccinate lyases, adenylosuccinate synthases, GMP synthases, IMP cyclohydrolases, IMP dehydrogenases, phosphoribosylamine-glycine ligases, phosphoribosyl-aminoimidazolecarboxamide formyltransferases, phosphoribosylaminoimidazole carboxylases, phosphoribosyl aminoimidazolesuccinocarboxamide synthases, phosphoribosyl-formylglycinamidine cyclo ligases, phosphoribosyl-formylglycinamidine synthases, phosphoribosyl-glycinamide formyltransferases, ribose-phosphate diphosphokinases, and ribose-5-phosphate-ammonia ligases.

In another particular embodiment, a pyrimidine-type biosynthetic enzyme can be chosen as an auxotrophic selection marker. Such pyrimidine-type biosynthetic include enzymes involved in biosynthesis of UMP, such as carbamate kinase (EC 2.7.2.2), carbamoyl-phosphate synthase (EC 6.3.5.5), aspartate carbamoyltransferase (EC 2.1.3.2), dihydroorotase (EC 3.5.2.3), dihydroorotate dehydrogenase (EC 1.3.3.1), orotate phosphoribosyltransferase ("OPRT;" EC 2.4.2.10), and orotidine-5'-phosphate decarboxylase ("ODCase;" EC 4.1.1.23).

Examples of genes encoding pyrimidine-type biosynthetic enzymes are well known. In the case of bacterial synthesis of UMP, examples of useful genes include: *arcC* genes, encoding carbamate kinases; *carA* and *carB* genes, collectively encoding carbamoyl-phosphate synthases; *pyrB* genes, encoding aspartate carbamoytransferases; *pyrC* genes, encoding dihydroorotases; *pyrD* genes, singly or collectively encoding dihydroorotate dehydrogenases; *pyrE* genes encoding orotate phosphoribosyltransferases; and *pyrF* genes, encoding orotidine-5'-phosphate decarboxylases.

In a particular embodiment, an expression system according to the present invention will utilize a *pyrF* auxotrophic selection marker gene. *pyrF* genes encode ODCase, an enzyme required for the bacterial pyrimidine nucleotide biosynthesis pathway, by which the cell performs *de novo* synthesis of pyrimidine nucleotides proper (UTP, CTP), as well as

pyrimidine deoxynucleotides (dTTP, dCTP). The pathway's initial reactants are ATP, an amino group source (i.e. ammonium ion or L-glutamine), and a carboxyl group source (i.e. carbon dioxide or bicarbonate ion); the pathway's ultimate product is dTTP, with dCTP, UTP, and CTP also being formed in the process. Specifically, the bacterial *de novo* pyrimidine nucleotide biosynthesis pathway begins with the formation of carbamoyl phosphate. Carbamoyl phosphate is synthesized either: (a) by action of carbamate kinase (EC 2.7.2.2), encoded by the *arcC* gene; or, more commonly, (b) by action of the glutamine-hydrolyzing, carbamoyl-phosphate synthase (EC 6.3.5.5), whose small and large subunits are encoded by the *carA* and *carB* genes, respectively. Carbamoyl phosphate is then converted to UDP by the following six-step route: 1) conversion of carbamoyl phosphate to N-carbamoyl-L-aspartate, by aspartate carbamoyltransferase (EC 2.1.3.2), encoded by *pyrB*; then 2) conversion thereof to (S)-dihydroorotate, by dihydroorotase (EC 3.5.2.3), encoded by *pyrC*; then 3) conversion thereof to orotate, by dihydroorotate dehydrogenase (EC 1.3.3.1), encoded by *pyrD* gene(s); then 4) conversion thereof to orotidine-5'-monophosphate ("OMP"), by orotate phosphoribosyltransferase ("OPRT;" EC 2.4.2.10), encoded by *pyrE*; and then 5) conversion thereof to uridine-5'-monophosphate ("UMP"), by orotidine-5'-phosphate decarboxylase ("ODCase;" EC 4.1.1.23), encoded by *pyrF*. The UMP is then utilized by a variety of pathways for synthesis of pyrimidine nucleotides (UTP, CTP, dTTP, dCTP), nucleic acids, nucleoproteins, and other cellular metabolites.

In bacteria in which one or more of the *carA*, *carB*, or *pyrB-pyrF* genes has become inactivated or lost, or mutated to encode a non-functional enzyme, the cell can still thrive if uracil is added to the medium, provided that the cell contains a functioning uracil salvage pathway. Most bacteria contain a native uracil salvage pathway, including the Pseudomonads and related species. In a uracil salvage pathway, the cell imports and converts exogenous uracil into UMP, to synthesize the required pyrimidine nucleotides. In this, uracil is reacted with 5-phosphoribosyl-1-pyrophosphate to form UMP, by the action of either uracil phosphoribosyltransferase (EC 2.4.2.9), encoded by the *upp* gene, or by the bifunctional, pyrimidine operon regulatory protein ("pyrR bifunctional protein"), encoded by *pyrR*. The resulting UMP is then converted to UDP, and then the subsequent pyrimidine nucleotides, as described above.

Consequently, a *pyrF*(-) Pseudomonad or related cell can be maintained on uracil-containing medium. After a *pyrF* gene-containing DNA construct is transfected into the *pyrF*(-) cell and expressed to form a functioning ODCase enzyme, the resulting combined *pyrF*(+) plasmid-host cell system can be maintained in a medium lacking uracil.

The coding sequence of the *pyrF* gene for use in a Pseudomonad or related host cell can be provided by any gene encoding an orotidine-5'-phosphate decarboxylase enzyme ("ODCase"), provided that the coding sequence can be transcribed, translated, and otherwise processed by the selected Pseudomonad or related host cell to form a functioning ODCase. The *pyrF* coding sequence may be a native sequence, or it may be an engineered sequence resulting from, for example, application of one or more sequence-altering, sequence-combining, and/or sequence-generating techniques known in the art. Before use as part of a *pyrF* selection marker gene, the selected coding sequence may first be improved or optimized in accordance with the genetic code and/or the codon usage frequency of a selected Pseudomonad or related host cell. Expressible coding sequences will be operatively attached to a transcription promoter capable of functioning in the chosen host cell, as well as all other required transcription and translation regulatory elements. A native coding sequence for a *pyrF* gene as described above may be obtained from a bacterium or from any other organism, provided that it meets the above-described requirements.

In one embodiment, the *pyrF* coding sequence is isolated from the Pseudomonad or related host cell in which it is intended to be used as a selection marker. The entire *pyrF* gene (including the coding sequence and surrounding regulatory regions) can be isolated there from. In a particular embodiment, a bacterium providing the *pyrF* gene or coding sequence will be selected from the group consisting of a member of the order *Pseudomonadales*, a member of the suborder *Pseudomonadineae*, a member of the family *Pseudomonadaceae*, a member of the tribe *Pseudomonadeae*, a member of the genus *Pseudomonas*, and a member of the *Pseudomonas fluorescens* species group (i.e. the "fluorescent pseudomonads"). In a particular embodiment, the bacterium will belong to the species, *Pseudomonas fluorescens*.

In a particular embodiment, the *pyrF* gene contains the nucleic acid sequence of SEQ ID NO. 1 (Table 2), or a variant thereof. Alternatively, the ODCase encoded by the pyrF gene contains the amino acid sequence of SEQ ID NO. 2 (Table 3), a variant thereof, or a variant having a codon sequence redundant therewith, in accordance with the genetic code used by a given host cell according to the present invention.

Alternatively, the *pyrF* gene contains a nucleic acid sequence encoding an ODCase enzyme selected from the group consisting of a nucleic acid sequence at least 70%, 75%, 80%, 85%, 88%, 90%, and 95% homologous to SEQ ID No. 1. Likewise, the *pyrF* gene encodes an ODCase selected from the group consisting of an amino acid sequence at least 70%, 75%, 80%, 85%, 88%, 90%, and 95% homologous to SEQ ID No. 2.

In another embodiment, the *pyrF* gene can contain a coding sequence having a nucleotide sequence at least 90%, 93%, 95%, 96%, 97%, 98% or 99% homologous to the nucleotide sequence of nucleotides 974-1669 of SEQ ID NO:1.

In a particular embodiment, the *pyrF* gene can contain a coding sequence having a codon sequence that hybridizes to the anti-codon sequence of SEQ ID NO:3 (Table 4), when hybridization has been performed under highly stringent hybridization conditions, or can have a codon sequence redundant therewith. In a particularly particular embodiment, the *pyrF* gene will contain the nucleotide sequence of SEQ ID No. 3

TABLE 2 – PSEUDOMONAS FLUORESCENS *PYRF* NUCLEIC ACID SEQUENCE

gatcagttgcggagccttggggtcatcccccagtttctgacgcaggcgcgacaccagcaagtcgatgctgcggtcga aagcctcgatggaacgcccacgggccgcgtccagcagctgttcgcggctcagcacacgccgcgggcgttcgataa acacccacaacaaacgaaactcggcgttggacagcggcaccaccaggccgtcatcggccaccagctggcgcagta cgctgttcaggcgccaagtgtcgaaacggatattggcccgctgttcggtgcggtcatcacgcacccggcgcaggatg gtctggatacgcgcgaccagttcccggggttcgaacggcttggacatatagtcgtctgcccccagttccaggccgatg atgcggtcggtgggttcgcagcgggcggtgagcatcaggatcggaatgtccgattcggcgcgcagccagcggcac aatgtcagcccgtcttcgcccggcagcatcaggtcgagcaccaccacatcgaaggtctccgcttgcatggcctggcg catggcgatgccgtcggtgacgcctgaggcgagaatattgaagcgtgccaggtagtcgatcagcagttcgcggatcg gcacgtcgtcgtcgacaatcagcgcgcgggtgttccagcgcttgtcttcggcgatcaccgcgtcttttggcgcttcgttt acagggtcgcaaggggtatgcatagcgaggtcatctgcctggttgtggctgtcagcataggcgcccagttccagggc tggaagtgctgggcgggcggtcatgtgcgcgaggctagccgggcggcgtattgggggcgtgtcgtgaatgtatcgg gcttgaaacaattgccttgaatcgccggtattgggcgcttgatcggtatttaccgatcatcggatcccgcaacggcgctg cttgcgctacaatccgcgccgatttcgacttgcctgagagcccattccaatgtccgtctgccagactcctatcatcgtcg ccctggattaccccaccgtgacgccgcactgaagctggctgaccagttggaccccaagctttgccgggtcaaggtc ggcaaggaattgttcaccagttgcgcggcggaaatcgtcggcaccctgcgggacaaaggcttcgaagtgttcctcga cctcaaattccatgacatccccaacaccacggcgatggccgtcaaagccgcggccgagatgggcgtgtggatggtc aatgtgcactgctccggtggcctgcgcatgatgagcgcctgccgcgaagtgctggaacagcgcagcggccccaaa ccgttgttgatcggcgtgaccgtgctcaccagcatggagcgcgaagacctggcgggcattggcctggatatcgagcc gcaggtgcaagtgttgcgcctggcagccctggcgcagaaagccggcctcgacggcctggtgtgctcagccctgga agcccaggccctgaaaaacgcacatccgtcgctgcaactggtgacaccgggtatccgtcctaccggcagcgcccag gatgaccagcgccgtatcctgacccgcgccaggccctggatgcgggctctgactacctggtgatcggccggccga tcagccaggcggcggatcctgcaaaagcgttggcagcggtcgtcgccgagatcgcctgatttttagagtgagcaaaa aatgtgggagctggcttgcctgcgatagtatcaactcggtatcacttagaaaccgagttgcttgcatcgcaggcaagcc agctcccacatttgtttttgtggtgtgtcagctgactttgagcaccaacttcccgaagttctcgccgttgaacagcttcatc agcgtttccgggaatgtctccagcccttcgacaatatcttccttgctcttgagcttgccctgggccatccagccggccatt tcctgacccgccgccgcgaagttcgccgcgtggtccatcaccacaaagccttccatacgcgcacggttgaccagcaa tgacaggtagttcgccgggcctttgaccgcttccttgttgttgtactggctgattgcaccgcaaatcaccacgcgggcttt gagcgccaggcggctgagcaccgcgtcgagaatatcgccgccgacgttatcgaaatacacgtccacgccttgggg cactcgcgcttgagggcggcgggcacgtcttcgcttttgtagtcgatggcggcgtcgaagcccagctcatcgaccag gaacttgcacttctcggcgccaccggcgatccccactacgcgacagcctttgagcttagcgatctgcccggcgatgct gcccacggcaccggcggcgccggagatcaccacggtgtcaccggctttcggtgcgccggtctccagcagagcaaa gtaggccgtcatgccggtcatgcccagggcggacaggtagcggggcaggggcgccagcttggggtccaccttata gaaaccacgggggctcgccaaggaagtaatcctgcacgcccagtgcaccgttcacgtagtcccccaccgcgaagttc ggatggttcgaggcaagcaccttgcctacgcccagggcgcgcatcacttcgccgatgcctaccggtgggatgtagg acttgccttcattcatccagccacgca	SEQ ID NO. 1

TABLE 3 – PSEUDOMONAS FLUORESCENS ODCASE AMINO ACID SEQUENCE

Met Ser Val Cys Gln Thr Pro Ile Ile Val Ala Leu Asp Tyr Pro Thr Arg Asp Ala Ala Leu Lys Leu Ala Asp Gln Leu Asp Pro Lys Leu Cys Arg Val Lys Val Gly Lys Glu Leu Phe Thr Ser Cys Ala Ala Glu Ile Val Gly Thr Leu Arg Asp Lys Gly Phe Glu Val Phe Leu Asp Leu Lys Phe His Asp Ile Pro Asn Thr Thr Ala Met Ala Val Lys Ala Ala Ala Glu Met Gly Val Trp Met Val Asn Val His Cys Ser Gly Gly Leu Arg Met Met Ser Ala Cys Arg Glu Val Leu Glu Gln Arg Ser Gly Pro Lys Pro Leu Leu Ile Gly Val Thr Val Leu Thr Ser Met Glu Arg Glu Asp Leu Ala Gly Ile Gly Leu Asp Ile Glu Pro Gln Val Gln Val Leu Arg Leu Ala Ala Leu Ala Gln Lys Ala Gly Leu Asp Gly Leu Val Cys Ser Ala Leu Glu Ala Gln Ala Leu Lys Asn Ala His Pro Ser Leu Gln Leu Val Thr Pro Gly Ile Arg Pro Thr Gly Ser Ala Gln Asp Asp Gln Arg Arg Ile Leu Thr Pro Arg Gln Ala Leu Asp Ala Gly Ser Asp Tyr Leu Val Ile Gly Arg Pro Ile Ser Gln Ala Ala Asp Pro Ala Lys Ala Leu Ala Ala Val Val Ala Glu Ile Ala	SEQ ID NO. 2

TABLE 4 - PSEUDOMONAS FLUORESCENS *PYRF* NUCLEIC ACID SEQUENCE

atgtccgtctgccagactcctatcatcgtcgccctggattaccccacccgtgacgccgcactgaag ctggctgaccagttggaccccaagctttgccgggtcaaggtcggcaaggaattgttcaccagttgc gcggcggaaatcgtcggcaccctgcgggacaaaggcttcgaagtgttcctcgacctcaaattccat gacatccccaacaccacgggcgatggccgtcaaagccgcggccgagatgggcgtgtggatggtcaat gtgcactgctccggtggcctgcgcatgatgagcgcctgccgcgaagtgctggaacagcgcagcggc cccaaaccgttgttgatcggcgtgaccgtgctcaccagcatggagcgcgaagacctggcgggcatt ggcctggatatcgagccgcaggtgcaagtgttgcgcctggcagccctggcgcagaaagccggcctc gacggcctggtgtgctcagccctggaagcccaggccctgaaaaacgcacatccgtcgctgcaactg gtgacaccgggtatccgtcctaccggcagcgcccaggatgaccagcgccgtatcctgaccccgcgc caggccctggatgcgggctctgactacctggtgatcggccggccgatcagccaggcggcggatcct gcaaaagcgttggcagcggtcgtcgccgagatcgcc	SEQ. ID No. 3

In an alternate embodiment, an expression system according to the present invention will utilize a *thyA* auxotrophic selection marker gene. *thyA* genes encode thymidylate synthase (EC 2.1.1.45), an enzyme required for the bacterial pyrimidine nucleotide biosynthesis pathway. Since DNA contains thymine (5-methyluracil) as a major base instead of uracil, the synthesis of thymidine monophospate (dTMP or thymidylate) is essential to provide dTTP (thymidine triphosphate) needed for DNA replication together with dATP, dGTP, and dCTP. Methylation of dUMP by thymidylate synthase utilizing 5,10-methylenetetrahydrofolate as the source of the methyl group generates thymidylate. Thymidylate synthesis can be interrupted, and consequently the synthesis of DNA arrested, by the removal, inhibition, or disruption of thymidylate synthase.

In bacteria in which the *thyA* gene has become inactivated or lost, or mutated to encode a non-functional enzyme, the cell can still thrive if exogenous thymidine is added to the medium.

35

In *Pseudomonas fluorescens*, the addition of an *E.coli tdk* gene, encoding thymidine kinase, is required for survival on exogenous thymidine. Therefore, prior to selection, a plasmid comprising a *tdk* gene can be used to transform *thyA*(-) *P. fluorescens* host cells, generating a *thyA*(-)/p*tdk* cell, allowing survival on a thymidine containing medium. Alternatively, a *tdk* gene producing a functional thymidylate synthase enzyme capable of utilizing exogenous thymidine in *Pseudomonas fluorescens* can be inserted into the genome, producing a *thyA*(-)/*tdk*(+) host cell. After a *thyA* gene-containing DNA construct is transfected into the *thyA*(-)/*ptdk* cell and expressed to form a functioning thymidylate synthase enzyme, the resulting combined *thyA*(+) plasmid-host cell system can be maintained in a medium lacking thymidine.

The coding sequence of the *thyA* gene for use in a Pseudomonad or related host cell can be provided by any gene encoding a thymidylate synthase enzyme ("TS"), provided that the coding sequence can be transcribed, translated, and otherwise processed by the selected Pseudomonad or related host cell to form a functioning TS. The *thyA* coding sequence may be a native sequence, or it may be an engineered sequence resulting from, for example, application of one or more sequence-altering, sequence-combining, and/or sequence-generating techniques known in the art. Before use as part of a *thyA* selection marker gene, the selected coding sequence may first be improved or optimized in accordance with the genetic code and/or the codon usage frequency of a selected Pseudomonad or related host cell. Expressible coding sequences will be operatively attached to a transcription promoter capable of functioning in the chosen host cell, as well as all other required transcription and translation regulatory elements. A native coding sequence for a *thyA* gene as described above may be obtained from a bacterium or from any other organism, provided that it meets the above-described requirements.

In one embodiment, the *thyA* coding sequence is isolated from the Pseudomonad or related host cell in which it is intended to be used as a selection marker. The entire *thyA* gene (including the coding sequence and surrounding regulatory regions) can be isolated there from. In a particular embodiment, a bacterium providing the *thyA* gene or coding sequence will be selected from the group consisting of a member of the order *Pseudomonadales*, a member of the suborder *Pseudomonadineae*, a member of the family *Pseudomonadaceae,* a member of the tribe *Pseudomonadeae*, a member of the genus *Pseudomonas*, and a member of the *Pseudomonas fluorescens* species group (i.e. the "fluorescent pseudomonads"). In a particular embodiment, the bacterium will belong to the species, *Pseudomonas fluorescens*.

In a particular embodiment, the *thyA* gene contains the nucleic acid sequence of SEQ ID NO. 4 (Table 5). Alternatively, the TS encoded by the *thyA* gene contains the amino acid sequence of SEQ ID NO. 5 (Table 6), a variant thereof, or a variant having a codon sequence redundant therewith, in accordance with the genetic code used by a given host cell according to the present invention.

TABLE 5 - PSEUDOMONAS FLUORESCENS *THYA* NUCLEIC ACID SEQUENCE

atgaagcaatatctcgaactactgaacgacgtcgtgaccaatggattgaccaagggcgatcgcac cggcaccggcaccaaagccgtatttgcccgtcagtatcggcataacttggccgacggcttcccgc tgctgaccaccaagaagcttcatttcaaaagtatcgccaacgagttgatctggatgttgagcggc aacaccaacatcaagtggctcaacgaaaatggcgtgaaaatctggcacgagtgggccaccgaaga cggcgacctgggcccggtgtacggcgagcaatggaccgcctggccgaccaaggacggcggcaaga tcaaccagatcgactacatggtccacaccctcaaaaccaaccccaacagccgccgcatcctgttt catggctggaacgtcgagtacctgccggacgaaaccaagagcccgcaggagaacgcgcgcaacgg caagcaagccttgccgccgtgccatctgttgtaccaggcgttcgtgcatgacgggcatctgtcga tgcagttgtatatccgcagctccgacgtcttcccggcctgccgtacaacaccgccgcgttggcc ttgctgactcacatgctggctcagcaatgcgacctgatccctcacgagatcatcgtcaccaccgg cgacacccatgcttacagcaaccacatggaacagatccgcacccagctggcgcgtacgccgaaaa agctgccggaactggtgatcaagcgtaaacctgcgtcgatctacgattacaagtttgaagacttt gaaatcgttggctacgacgccgacccgagcatcaaggctgacgtggctatctga	SEQ ID NO. 4

TABLE 6 - PSEUDOMONAS FLUORESCENS TS AMINO ACID SEQUENCE

MKQYLELLNDVVTNGLTKGDRTGTGTKAVFARQYRHNLADGFPLLTTKKLHFKSIANELIWMLSG NTNIKWLNENGVKIWDEWATEDGDLGPVYGEQWTAWPTKDGGKINQIDYMVHTLKTNPNSRRILF HGWNVEYLPDETKSPQENARNGKQALPPCHLLYQAFVHDGHLSMQLYIRSSDVFLGLPYNTAALA LLTHMLAQQCDLIPHEIIVTTGDTHAYSNHMEQIRTQLARTPKKLPELVIKRKPASIYDYKFEDF EIVGYDADPSIKADVAI	SEQ ID NO. 5

Biosynthetic Amino Acid Selection Markers

In an alternative embodiment, the biosynthetic enzyme involved in anabolic metabolism chosen as the auxotrophic selection marker can be selected from those involved in the biosynthesis of amino acids. In particular embodiments, the biosynthetic amino acid enzymes are selected from the group consisting of enzymes active in the biosynthesis of: the Glutamate Family (Glu; Gln, Pro, and Arg); the Aspartate Family (Asp; Asn, Met, Thr, Lys, and Ile); the Serine Family (Ser; Gly and Cys); the Pyruvate Family (Ala, Val, and Leu); the Aromatic Family (Trp, Phe, and Tyr); and the Histidine Family (His). Examples of genes and enzymes involved in these biosynthetic pathways include: the Glutamate Family member *arg*, *gdh*, *gln*, and, *pro* genes, including, for example, *argA-argH*, *gdhA*, *glnA*, *proA*, *proC*; the Aspartate Family member *asd*, *asn*, *asp*, *dap*, *lys*, *met*, and *thr* genes, including, for example, *asnA*, *asnB*, *aspC*, *dapA*, *dapB*, *dapD-dapF*, *lysA*, *lysC*, *metA-metC*, *metE*, *metH*, *metL*, *thrA-thrC*; the Serine Family member *cys*, *gly*, and *ser* genes, including, for example, *cysE*, *cysK*,

glyA, *serA-serC*; the Aromatic Family member *aro*, *phe*, *trp*, and *tyr* genes, including, for example, *aroA-aroH*, *aroK*, *aroL*, *trpAtrpE*, *tyrA*, and *tyrB*; and the Histidine Family member *his* genes, including *hisA-hisD*, *hisF-hisH*.

In a further particular embodiment, the auxotrophic selection marker can be selected from enzymes involved in the biosynthesis of members of the Glutamate Family. Examples of useful Glutamate Family auxotrophic selection markers include the following, listed with representative examples of their encoding genes: *argA*, encoding N-acetylglutamate synthases, amino acid acetyltransferases; *argB*, encoding acetylglutamate kinases; *argC*, encoding N-acetyl-gammaglutamylphosphate reductases; *argD*, encoding acetylornithine delta-aminotransferases; *argE*, encoding acetylornithine deacetylases; *argF* and *argI*, encoding ornithine carbamoyltransferases; *arg*G, encoding argininosuccinate synthetases; *argH*, encoding argininosuccinate lyases; *gdhA*, encoding glutamate dehydrogenases; *glnA*, encoding glutamine synthetases; *proA*, encoding gamma-glutamylphosphate reductases; *proB*, encoding gamma-glutamate kinases; and *proC*, encoding pyrroline-5-carboxylate reductases.

In one embodiment, an amino acid biosynthesis selection marker gene can be at least one member of the proline biosynthesis family, in particular *proA*, *proB*, or *proC*. In a particular embodiment, the proline biosynthesis selection marker gene can comprise a *proC* gene. *proC* genes encode an enzyme catalyzing the final step of the proline biosynthesis pathway. In bacteria, the proline (i.e. L-proline) biosynthesis pathway comprises a three-enzyme process, beginning with L-glutamic acid. The steps of this process are: 1) conversion of L-glutamic acid to L-glutamyl-5-phosphate, by glutamate-5-kinase ("GK;" EC 2.7.2.11), encoded by *proB*; then 2a) conversion thereof to L-glutamate-5-semialdehyde, by glutamate-5-semialdehyde dehydrogenase (EC 1.2.1.41), also known as glutamyl-5-phosphate reductase ("GPR"), encoded by *proA*, followed by 2b) spontaneous cyclization thereof to form Δ^1-pyrroline-5-carboxylate; and then 3) conversion thereof to L-proline, by Δ^1-pyrroline-5-carboxylate reductase ("P5CR;" EC 1.5.1.2), encoded by *proC*. In most bacteria, *proC* encodes the P5CR subunit, with the active P5CR enzyme being a homo-multimer thereof.

In bacteria in which one or more of the *proA, proB, or proC* genes has become inactivated or lost, or mutated to encode a non-functional enzyme, the cell can still thrive if proline is added to the medium. Consequently, a *proC*(-) Pseudomonad or related cell can be maintained on a proline-containing medium. After a *proC* gene-containing DNA construct is transfected into the *proC*(-) cell and expressed to form a functioning P5CR enzyme, the

resulting combined *proC*(+) plasmid-host cell system can be maintained in a medium lacking proline.

The coding sequence of the *proC* gene for use in a Pseudomonad or related host cell can be provided by any gene encoding an Δ^1-pyrroline-5-carboxylate reductase enzyme (P5CR), provided that the coding sequence can be transcribed, translated, and otherwise processed by the selected Pseudomonad or related host cell to form a functioning P5CR. The *proC* coding sequence may be a native sequence, or it may be an engineered sequence resulting from, for example, application of one or more sequence-altering, sequence-combining, and/or sequence-generating techniques known in the art. Before use as part of a *proC* selection marker gene, the selected coding sequence may first be improved or optimized in accordance with the genetic code and/or the codon usage frequency of a selected Pseudomonad or related host cell. Expressible coding sequences will be operatively attached to a transcription promoter capable of functioning in the chosen host cell, as well as all other required transcription and translation regulatory elements. A native coding sequence for a *proC* gene as described above may be obtained from a bacterium or from any other organism, provided that it meets the above-described requirements.

In one embodiment, the *proC* coding sequence is isolated from the Pseudomonad or related host cell in which it is intended to be used as a selection marker. The entire *proC* gene (including the coding sequence and surrounding regulatory regions) can be isolated therefrom. In a particular embodiment, a bacterium providing the *proC* gene or coding sequence will be selected from the group consisting of a member of the order *Pseudomonadales*, a member of the suborder *Pseudomonadineae*, a member of the family *Pseudomonadaceae*, a member of the tribe *Pseudomonadeae*, a member of the genus *Pseudomonas*, and a member of the *Pseudomonas fluorescens* species group (i.e. the "fluorescent pseudomonads"). In a particular embodiment, the bacterium will belong to the species, *Pseudomonas fluorescens*.

In a particular embodiment, the *proC* gene contains the nucleic acid sequence of SEQ ID NO. 6 (Table 7), or a variant thereof. Alternatively, the P5CR encoded by the *proC* gene contains the amino acid sequence of SEQ ID NO. 7 (Table 8), a variant thereof, or a variant having a codon sequence redundant therewith, in accordance with the genetic code used by a given host cell according to the present invention.

Alternatively, the *proC* gene contains a nucleic acid sequence encoding an P5CR enzyme that is at least 70%, 75%, 80%, 85%, 88%, 90%, and 95% homologous to SEQ ID

No. 6. Likewise, the *proC* gene encodes an ODCase that is at least 70%, 75%, 80%, 85%, 88%, 90%, and 95% homologous to SEQ ID No. 7.

In another embodiment, the *proC* gene can contain a coding sequence at least 90%, 93%, 95%, 96%, 97%, 98% or 99% homologous to the nucleotide sequence of SEQ. ID NO. 8 (Table 9).

In a particular embodiment, the *proC* gene can contain a coding sequence having a codon sequence that hybridizes to the anti-codon sequence of SEQ ID NO. 8, when hybridization has been performed under stringent hybridization conditions, or can have a codon sequence redundant therewith. In a particularly particular embodiment, the *proC* gene will contain the nucleotide sequence of SEQ ID NO. 8.

TABLE 7 – *PSEUDOMONAS FLUORESCENS PROC* NUCLEIC ACID SEQUENCE

gccctgagttggcacttcatcggccccattcaatcgaacaagactcgtgccatcgccgagcacttcgcttgggt gcactccgtggaccgcctgaaaatcgcacaacgcctgtccgaacaacgcccggccgacctgccgccgctca atatctgcatccaggtcaatgtcagtggcgaagccagcaagtccggctgcacgcccgctgacctgccggccct ggccacagcgatcagcgccctgccgcgcttgaagctgcgggcttgatggcgattcccgagccgacgcaag accgggcggagcaggatgcggcgttcgccacggtgcgcgacttgcaagccagcttgaacctggcgctggac acactttccatgggcatgagccacgaccttgagtcggccattgcccaaggcgccacctgggtgcggatcggta ccgccctgtttggcgcccgcgactacggccagccgtgaaatggctgacatccctcgaaataaggacctgtcat gagcaacacgcgtattgcctttatcggcgccggtaacatggcggccagcctgatcggtggcctgcgggccaa gggcctggacgccgagcagatccgcgccagcgacccccggtgccgaaacccgcgagcgcgtcagagccga acacggtatccagaccttcgccgataacgccgaggccatccacggcgtcgatgtgatcgtgctggcggtcaag ccccaggccatgaaggccgtgtgcgagagcctgagcccgagcctgcaaccccatcaactggtggtgtcgatt gccgctggcatcacctgcgccagcatgaccaactggctcggtgcccagcccattgtgcgctgcatgcccaaca cccggcgctgctgcgccagggcgtcagcggtttgtatgccactggcgaagtcaccgcgcagcaacgtgacc aggcccaggaactgctgtctgcggtgggcatcgccgtgtggctggagcaggaacagcaactggatgcggtc accgccgtctccggcagcggcccggcttacttcttcctgttgatcgaggccatgacggccgcaggcgtcaagc tgggcctgccccacgacgtggccgagcaactggcggaacaaaccgccctgggcgccgccaagatggcggt cggcagcgaggtggatgccgccgaactgcgccgtcgcgtcacctcgccaggtggtaccacacaagcggcta ttgagtcgttccaggccgggggctttgaagccctggtggaaacagcactgggtgccgccgcacatcgttcagc cgagatggctgagcaactgggcaaatagtcgtcccttaccaaggtaatcaaacatgctcggaatcaatgacgct gccattttcatcatccagaccctgggcagcctgtacctgctgatcgtactgatgcgctttatcctgcaactggtgc gtgcgaacttctacaacccgctgtgccagttcgtggtgaaggccacccaaccgctgctcaagccgctgcgccg ggtgatcccgagcctgttcggcctggacatgtcgtcgctggtgctggcgctgttgctgcagattttgctgttcgtg gtgatcctgatgctcaatggataccaggccttcaccgtgctgctgttgccatggggcctgatcgggatttttctgc tgttcctgaagatcatttttctggtcgatgatcatcagcgtgatcctgtcctgggtcgcaccgggtagccgtagccc gggtgccgaattggtggctcagatcaccgagccggtgctggcacccttccgtcgcctgattccgaacctgggt ggcctggatatctcgccgatcttcgcgtttatc	SEQ ID NO. 6

TABLE 8 – *PSEUDOMONAS FLUORESCENS P5CR* AMINO ACID SEQUENCE

Met Ser Asn Thr Arg Ile Ala Phe Ile Gly Ala Gly Asn Met Ala Ala Ser Leu Ile Gly Gly Leu Arg Ala Lys Gly Leu Asp Ala Glu Gln Ile Arg Ala Ser Asp Pro Gly Ala Glu Thr Arg Glu Arg Val Arg Ala Glu His Gly Ile Gln Thr Phe	SEQ ID NO. 7

40

Ala Asp Asn Ala Glu Ala Ile His Gly Val Asp Val Ile Val Leu Ala Val Lys Pro Gln Ala Met Lys Ala Val Cys Glu Ser Leu Ser Pro Ser Leu Gln Pro His Gln Leu Val Val Ser Ile Ala Ala Gly Ile Thr Cys Ala Ser Met Thr Asn Trp Leu Gly Ala Gln Pro Ile Val Arg Cys Met Pro Asn Thr Pro Ala Leu Leu Arg Gln Gly Val Ser Gly Leu Tyr Ala Thr Gly Glu Val Thr Ala Gln Gln Arg Asp Gln Ala Gln Glu Leu Leu Ser Ala Val Gly Ile Ala Val Trp Leu Glu Gln Glu Gln Gln Leu Asp Ala Val Thr Ala Val Ser Gly Ser Gly Pro Ala Tyr Phe Phe Leu Leu Ile Glu Ala Met Thr Ala Ala Gly Val Lys Leu Gly Leu Pro His Asp Val Ala Glu Gln Leu Ala Glu Gln Thr Ala Leu Gly Ala Ala Lys Met Ala Val Gly Ser Glu Val Asp Ala Ala Glu Leu Arg Arg Arg Val Thr Ser Pro Gly Gly Thr Thr Gln Ala Ala Ile Glu Ser Phe Gln Ala Gly Gly Phe Glu Ala Leu Val Glu Thr Ala Leu Gly Ala Ala Ala His Arg Ser Ala Glu Met Ala Glu Gln Leu Gly Lys	

TABLE 9 – *PSEUDOMONAS FLUORESCENS PROC* NUCLEIC ACID SEQUENCE

atgagcaacacgcgtattgcctttatcggcgccggtaacatggcggccagcctgatcggtggc ctgcgggccaagggcctggacgccgagcagatccgcgccagcgaccccggtgccgaaacccgc gagcgcgtcagagccgaacacggtatccagaccttcgccgataacgccgaggccatccacggc gtcgatgtgatcgtgctggcggtcaagccccaggccatgaaggccgtgtgcgagagcctgagc ccgagcctgcaaccccatcaactggtggtgtcgattgccgctggcatcacctgcgccagcatg accaactggctcggtgcccagcccattgtgcgctgcatgcccaacaccccggcgctgctgcgc cagggcgtcagcggtttgtatgccactggcgaagtcaccgcgcagcaacgtgaccaggcccag gaactgctgtctgcggtgggcatcgccgtgtggctggagcaggaacagcaactggatgcggtc accgccgtctccggcagcggccccggcttacttcttcctgttgatcgaggccatgacggccgca ggcgtcaagctgggcctgccccacgacgtggccgagcaactggcggaacaaaccgccctgggc gccgccaagatggcggtcggcagcgaggtggatgccgccgaactgcgccgtcgcgtcacctcg ccaggtggtaccacacaagcggctattgagtcgttccacgccgggggctttgaagccctggtg gaaacagcactgggtgccgccgcacatcgttcagccgagatggctgagcaactgggcaaa	SEQ ID NO. 8

Utilization Selection Markers

In one embodiment, an enzyme involved in the catabolic utilization of metabolites can be chosen as the auxotrophic selection marker. In particular, the enzymes can be selected from those involved in the utilization of a carbon source. Examples of such enzymes include, for example, sucrases, lactases, maltases, starch catabolic enzymes, glycogen catabolic enzymes, cellulases, and poly(hydroxyalkanoate) depolymerases. If the bacterial host cell exhibits native catabolic activity of the selected type, it can be knocked-out before transformation with the prototrophy-restoring vector. Bacteria exhibiting native auxotrophy for these compounds can also be used in their native state for such transformation. In those embodiments in which a compound not importable or diffusible into the cell can be selected and supplied to the medium, the prototrophy restoring or prototrophy-enabling enzyme(s) can be secreted for use. In that case, the secreted enzyme(s) can degrade the compound extracellularly to produce smaller compounds, for example glucose, that are diffusible or importable into the cell, by selecting or designing the coding sequence of the enzyme(s) to include a coding sequence for a secretion signal peptide operative within the chosen host cell.

In these embodiments, the prototrophy-restorative gene can be selected or be engineered to include a coding sequence for a secretion signal peptide operative within the chosen host cell to obtaining transport of the enzyme across the cytoplasmic membrane. In either of these embodiments, or those in which the selected compound is importable or diffusible into the cell, the cell will be grown in medium supplying no other carbon source apart from the selected compound.

In a carbon-source-utilization-based marker system, every prototrophy-restorative or prototrophy-enabling carbon-source utilization enzyme can be involved in utilization of only one carbon source. For example, two genes from the same catabolic pathway may be expressed together on one vector or may be co-expressed separately on different vectors in order to provide the prototrophy. Specific examples of such multi-gene carbon-source-utilization-based marker systems include, for example, the use of glycogen as the sole carbon source with transgenic expression of both a glycogen phosphorylase and an (alpha-1,4) glucantransferase; and the use of starch as the sole carbon source with transgenic expression of both an alpha-amylase, and an alpha(1 ->6) glucosidase. However, the selected single- or multi-gene carbon-source marker system can be used simultaneously with other types of marker system(s) in the same host cell, provided that the only carbon source provided to the cell is the compound selected for use in the carbon-source catabolic selection marker system.

Other examples of useful enzymes for biochemical-utilization-type activities are well known in the art, and can include racemases and epimerases that are capable of converting a non-utilizable D-carbon source, supplied to the cell, to a nutritive L-carbon source. Examples of these systems include, for example: a D-acid or a D-acyl compound used with trangenic expression of the corresponding racemase; and lactate used with transgenically expressed lactate racemase.

Similarly, where an amino acid biosynthetic activity has been selected for use in the marker system, the auxotrophy may also be overcome by supplying the cell with both a non-utilizable R-amino acid and an R-amino acid racemase or epimerase (EC 5.1.1) that converts the R-amino acid into the corresponding L-amino acid for which the cell is auxotrophic.

Trait Stacking

A plurality of phenotypic changes can also be made to a host cell, before or after insertion of an auxotrophic selection marker gene, for target gene expression, according to the present invention. For example, the cell can be genetically engineered, either simultaneously or sequentially, to exhibit a variety of enhancing phenotypic traits. This

process is referred to as "trait stacking". A *pryF* deletion may be present as one such phenotypic trait. In such a strain, a *pyrF* gene can be used on a suicide vector as both a selectable marker and a counterselectable marker (in the presence of 5'-fluoroorotic acid) in order to effect a cross-in/cross-out allele exchange of other desirable traits. Thus, a *pyrF* gene may be used in a process for "trait stacking" a host cell. In such a process, a suicide vector containing such a *pyrF* gene can be transformed into the host cell strain in a plurality of separate transformations; in each such procedure the re-establishment of the *pyrF* phenotype can be used to create, ad infinitum, subsequent genetically-enhancing phenotypic change. Thus, not only can the *pyrF* gene itself provide a trait, it can be used to obtain additional phenotypic traits in a process of trait-stacking.

In one embodiment, the present invention provides auxotrophic Pseudomonads and related bacteria that have been further genetically modified to induce additional auxotrophies. For example, a *pyrF*(-) auxotroph can be further modified to inactivate another biosynthetic enzyme present in an anabolic or catabolic pathway, such as through the inactivation of a *proC* gene or a *thyA* gene. In this way, multiple auxotrophies in the host cell can be produced.

In another embodiment, genetic alterations can be made to the host cell in order to improve the expression of recombinant polypeptides in the host cell. Further modifications can include genetic alterations that allow for a more efficient utilization of a particular carbon source, thereby optimizing the overall efficiency of the entire fermentation.

In one particular embodiment, auxotrophic host cells are further modified by the insertion of a *lacI* containing transgene into the host chromosome. Preferably, the *lacI* transgene, or derivate thereof, is other than part of a

whole or truncated structural gene containing *PlacI-lacI-lacZYA* construct.

Modifications to induce Auxotrophism

A Pseudomonad or related host cell selected for use in an expression system according to the present disclosure can be deficient in its ability to express any functional biocatalyst exhibiting the selected auxotrophic activity. For example, where an orotidine-5'-phosphate decarboxylase activity is selected, the host cell can be deficient in its ability to express a) any *pyrF* gene product (i.e. any functional ODCase enzyme), and b) any effective replacement therefore (i.e. any other biocatalyst having ODCase activity). In a one embodiment, the host cell will be made biocatalytically-deficient for the selected activity by altering its genomic gene(s) so that the cell cannot express, from its genome, a functional

enzyme involved in the targeted auxotrophy (i.e. ODCase). In other words, the prototrophic cell (activity(+) cell) will become auxotrophic through the "knock-out" of a functional enzymatic encoding gene involved in the targeted prototrophic pathway (i.e. an activity(-) cell). This alteration can be done by altering the cell's genomic coding sequence(s) of the gene(s) encoding the selected activty(ies). In one embodiment, the coding sequence alteration(s) will be accomplished by introducing: insertion or deletion mutation(s) that change the coding sequence reading frame(s); substitution or inversion mutations that alter a sufficient number of codons; and/or deletion mutations that delete a sufficiently large group of contiguous codons there from capable of producing a non-functional enzyme.

In a one embodiment in which the host cell strain has also provided the auxotrophic gene(s) for use as selection marker(s) therein, preferably each of the selected gene's transcription promoter and/or transcription terminator element(s) can also be inactivated by introduction of mutation(s), including deletion mutations. For example, the transcription element inactivation can be optionally performed in addition to the coding sequence alteration(s) described above. In a one embodiment in which the host cell strain has also provided the auxotrophic selection marker gene(s), all of the selected gene(s)'s DNA can be deleted from the host cell genome.

Such knock-out strains can be prepared according to any of the various methods known in the art as effective. For example, homologous recombination vectors containing homologous targeted gene sequences 5' and 3' of the desired nucleic acid deletion sequence can be transformed into the host cell. Ideally, upon homologous recombination, a desired targeted enzymatic gene knock-out can be produced.

Specific examples of gene knock-out methodologies include, for example: Gene inactivation by insertion of a polynucleotide has been previously described. See, e.g., DL Roeder & A Collmer, *Marker-exchange mutagenesis of a pectate lyase isozyme gene in Erwinia chrysanthemi*, J Bacteriol. 164(1):51-56 (1985). Alternatively, transposon mutagenesis and selection for desired phenotype (such as the inability to metabolize benzoate or anthranilate) can be used to isolate bacterial strains in which target genes have been insertionally inactivated. See, e.g., K Nida & PP Cleary, *Insertional inactivation of streptolysin S expression in Streptococcus pyogenes*, J Bacteriol. 155(3):1156-61 (1983). Specific mutations or deletions in a particular gene can be constructed using cassette mutagenesis, for example, as described in JA Wells et al., *Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites*, Gene 34(2-3):315-23 (1985);

whereby direct or random mutations are made in a selected portion of a gene, and then incorporated into the chromosomal copy of the gene by homologous recombination.

In one embodiment, both the organism from which the selection marker gene(s) is obtained and the host cell in which the selection marker gene(s) is utilized can be selected from a prokaryote. In a particular embodiment, both the organism from which the selection marker gene(s) is obtained and the host cell in which a selection marker gene(s) is utilized can be selected from a bacteria. In another embodiment, both the bacteria from which the selection marker gene(s) is obtained and the bacterial host cell in which a selection marker gene(s) is utilized, will be selected from the Proteobacteria. In still another embodiment, both the bacteria from which the selection marker gene(s) is obtained and the bacterial host cells in which a selection marker gene(s) is utilized, can be selected from the Pseudomonads and closely related bacteria or from a Subgroup thereof, as defined below.

In a particular embodiment, both the selection marker gene(s) source organism and the host cell can be selected from the same species. Preferably, the species will be a prokaryote; more preferably a bacterium, still more preferably a Proteobacterium. In another particular embodiment, both the selection marker gene(s) source organism and the host cell can be selected from the same species in a genus selected from the Pseudomonads and closely related bacteria or from a Subgroup thereof, as defined below. In one embodiment, both the selection marker gene(s) source organism and the host cell can be selected from a species of the genus *Pseudomonas*, particularly the species *Pseudomonas fluorescens*, and preferably the species *Pseudomonas fluorescens* biotype A.

III. LACI INSERTION

Disclosed herein are Pseudomonads and related cells

that have been genetically modified to contain a chromosomally insert *lacI* transgene or derivative, other than as part of a whole or truncated *PlacI-lacI-lacZYA* operon. In one embodiment, the *lacI* insert provides stringent expression vector control through the expression of the LacI repressor protein which binds to the *lacO* sequence or derivative on the vector, and inhibits a Plac-Ptac family promoter on the vector. The result is reduced basal levels of recombinant polypeptide expression prior to induction.

In one embodiment, Pseudomonad host cells containing a chromosomal insertion of a native *E.coli lacI* gene, or *lacI* gene derivative such as $lacI^{Q}$ or $lacI^{Q1}$, are provided wherein the *lacI* insert is other than part of a whole or truncated, structural gene-containing *PlacI-lacI-*

lacZYA construct. Other derivative *lacI* transgenes useful in the present invention include: *lacI* derivatives that have altered codon sequences different from a native *lacI* gene (for example, the native *E. coli lacI* gene contains a 'gtg' initiation codon, and this may be replaced by an alternative initiation codon effective for translation initiation in the selected expression host cell, e.g., 'atg'); *lacI* derivatives that encode LacI proteins having mutated amino acid sequences, including temperature-sensitive *lacI* mutants, such as that encoded by $lacI^{ts}$ (or "*lacI*(Ts)"), which respond to a shift in temperature in order to achieve target gene induction, e.g., a shift up to 42° C (see, e.g., Bukrinsky *et al.*, *Gene* 70:415-17 (1989); N Hasan & W Szybalski, *Gene* 163(1):35-40 (1995); H Adari *et al.*, *DNA Cell Biol.* 14:945-50 (1995)); LacI mutants that respond to the presence of alternative sugars other than lactose in order to achieve induction, e.g., arabinose, ribose, or galactose (see, e.g., WO 99/27108 for Lac Repressor Proteins with Altered Responsivity); and LacI mutants that exhibit at least wild-type binding to lac operators, but enhanced sensitivity to an inducer (e.g., IPTG), or that exhibit enhanced binding to lac operators, but at least wild-type de-repressibility (see, e.g., L Swint-Kruse *et al.*, *Biochemistry* 42(47):14004-16 (2003)).

In a particular embodiment, the gene encoding the Lac repressor protein inserted into the chromosome is identical to that of native *E. coli lacI* gene, and has the nucleic acid sequence of SEQ ID NO. 9 (Table 10). In another embodiment, the gene inserted into the host chromosome encodes the Lac repressor protein having the amino acid sequence of SEQ ID NO. 10 (Table 11).

TABLE 10 – NUCLEIC ACID SEQUENCE OF NATIVE E.COLI *LACI* GENE

Gacaccatcgaatggcgcaaaacctttcgcggtatggcatgatagcgcccggaagagagtca attcagggtggtgaatgtgaaaccagtaacgttatacgatgtcgcagagtatgccggtgtct cttatcagaccgtttcccgcgtggtgaaccaggccagccacgtttctgcgaaaacgcgggaa aaagtggaagcggcgatggcggagctgaattacattcccaaccgcgtggcacaacaactggc gggcaaacagtcgttgctgattggcgttgccacctccagtctggccctgcacgcgccgtcgc aaattgtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtggtggtgtcgatg gtagaacgaagcggcgtcgaagcctgtaaagcggcggtgcacaatcttctcgcgcaacgcgt cagtgggctgatcattaactatccgctggatgaccaggatgccattgctgtggaagctgcct gcactaatgttccggcgttatttcttgatgtctctgaccagacacccatcaacagtattatt ttctcccatgaagacggtacgcgactgggcgtggagcatctggtcgcattgggtcaccagca aatcgcgctgttagcgggcccattaagttctgtctcggcgcgtctgcgtctggctggctggc ataaatatctcactcgcaatcaaattcagccgatagcggaacgggaaggcgactggagtgcc atgtccggttttcaacaaaccatgcaaatgctgaatgagggcatcgttcccactgcgatgct ggttgccaacgatcagatggcgctgggcgcaatgcgcgccattaccgagtccgggctgcgcg ttggtgcggatatctcggtagtgggatacgacgataccgaagacagctcatgttatatcccg ccgtcaaccaccatcaaacaggattttcgcctgctggggcaaaccagcgtggaccgcttgct gcaactctctcagggccaggcggtgaagggcaatcagctgttgcccgtctcactggtgaaaa gaaaaaccaccctggcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcatta atgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatg tgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttg tgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacggatt cactggccgtcgttttacaacgtcgtga	SEQ ID NO 9

TABLE 11 – AMINO ACID SEQUENCE OF LACI REPRESSOR

Met Lys Pro Val Thr Leu Tyr Asp Val Ala Glu Tyr Ala Gly Val Ser Tyr Gln Thr Val Ser Arg Val Val Asn Gln Ala Ser His Val Ser Ala Lys Thr Arg Glu Lys Val Glu Ala Ala Met Ala Glu Leu Asn Tyr Ile Pro Asn Arg Val Ala Gln Gln Leu Ala Gly Lys Gln Ser Leu Leu Ile Gly Val Ala Thr Ser Ser Leu Ala Leu His Ala Pro Ser Gln Ile Val Ala Ala Ile Lys Ser Arg Ala Asp Gln Leu Gly Ala Ser Val Val Val Ser Met Val Glu Arg Ser Gly Val Glu Ala Cys Lys Ala Ala Val His Asn Leu Leu Ala Gln Arg Val Ser Gly Leu Ile Ile Asn Tyr Pro Leu Asp Asp Gln Asp Ala Ile Ala Val Glu Ala Ala Cys Thr Asn Val Pro Ala Leu Phe Leu Asp Val Ser Asp Gln Thr Pro Ile Asn Ser Ile Phe Ser His Glu Asp Gly Thr Arg Leu Gly Val Glu His Leu Val Ala Leu Gly His Gln Gln Ile Ala Leu Leu Ala Gly Pro Leu Ser Ser Val Ser Ala Arg Leu Arg Leu Ala Gly Trp His Lys Tyr Leu Thr Arg Asn Gln Ile Gln Pro Ile Ala Glu Arg Glu Gly Asp Trp Ser Ala Met Ser Gly Phe Gln Gln Thr Met Gln Met Leu Asn Glu Gly Ile Val Pro Thr Ala Met Leu Val Ala Asn Asp Gln Met Ala Leu Gly Ala Met Arg Ala Ile Thr Glu Ser Gly Leu Arg Val Gly Ala Asp Ile Ser Val Val Gly Tyr Asp Asp Thr Glu Asp Ser Ser Cys Tyr Ile Pro Pro Ser Thr Thr Ile Lys Gln Asp Phe Arg Leu Leu Gly Gln Thr Ser Val Asp Arg Leu Leu Gln Leu Ser Gln Gly Gln Ala Val Lys Gly Asn Gln Leu Leu Pro Val Ser Leu Val Lys Arg Lys Thr Thr Leu Ala Pro Asn Thr Gln Thr Ala Ser Pro Arg Ala Leu Ala Asp Ser Leu Met Gln Leu Ala Arg Gln Val Ser Arg Leu Glu Ser Gly Gln	SEQ ID NO. 10

In an alternative embodiment, the inserted *lacI* transgene is a derivative of the native *E.coli lacI* gene. In one particular embodiment, the *lacI* derivative gene is the $lacI^Q$ gene having the nucleic acid sequence of SEQ ID NO. 11 (Table 12). The $lacI^Q$ variant is identical to the native *E.coli lacI* gene except that it has a single point mutation in the -35 region of the promoter which increases the level of lacI repressor by 10-fold in *E.coli*. See, for example, MP Calos, Nature 274 (5673): 762-65 (1978).

TABLE 12 - NUCLEIC ACID SEQUENCE OF $LACI^Q$ GENE

gacaccatcgaatggtgcaaaacctttcgcggtatggcatgatagcgcccggaagagagtca attcagggtggtgaatgtgaaaccagtaacgttatacgatgtcgcagagtatgccggtgtct cttatcagaccgtttcccgcgtggtgaaccaggccagccacgtttctgcgaaaacgcgggaa aaagtggaagcggcgatggcggagctgaattacattcccaaccgcgtggcacaacaactggc gggcaaacagtcgttgctgattggcgttgccacctccagtctggccctgcacgcgccgtcgc aaattgtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtggtggtgtcgatg gtagaacgaagcggcgtcgaagcctgtaaagcggcggtgcacaatcttctcgcgcaacgcgt cagtgggctgatcattaactatccgctggatgaccaggatgccattgctgtggaagctgcct gcactaatgttccggcgttatttcttgatgtctctgaccagacacccatcaacagtattatt ttctcccatgaagacggtacgcgactgggcgtggagcatctggtcgcattgggtcaccagca aatcgcgctgttagcgggcccattaagttctgtctcggcgcgtctgcgtctggctggctggc ataaatatctcactcgcaatcaaattcagccgatagcggaacgggaaggcgactggagtgcc atgtccggttttcaacaaaccatgcaaatgctgaatgagggcatcgttcccactgcgatgct ggttgccaacgatcagatggcgctgggcgcaatgcgcgccattaccgagtccgggctgcgcg ttggtgcggatatctcggtagtgggatacgacgataccgaagacagctcatgttatatcccg ccgtcaaccaccatcaaacaggattttcgcctgctggggcaaaccagcgtggaccgcttgct gcaactctctcagggccaggcggtgaagggcaatcagctgttgcccgtctcactggtgaaaa gaaaaaccaccctggcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcatta atgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatg tgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttg tgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacggatt cactggccgtcgttttac	SEQ ID NO. 11

In still another embodiment, the *lacI* derivate gene is the *lacI*^{Q1} gene having the nucleic acid sequence of SEQ ID NO. 12 (Table 13). The *lacI*^{Q1} variant has a rearrangement which substitutes a -35 region whose nucleotide sequence exactly matches that of the *E.coli* -35 region consensus sequence, resulting in expression that is 100-fold higher than the native promoter in *E. coli*. See, for example, MP Colas & JH Miller, Mol. & Gen. Genet. 183(3): 559-60(1980).

TABLE 13 – NUCLEIC ACID SEQUENCE OF *LACI*^{Q1} GENE

agcggcatgcatttacgttgacaccacctttcgcggtatggcatgatagcgcccggaagaga gtcaattcagggtggtgaatgtgaaaccagtaacgttatacgatgtcgcagagtatgccggt gtctcttatcagaccgtttcccgcgtggtgaaccaggccagccacgtttctgcgaaaacgcg ggaaaaagtggaagcggcgatggcggagctgaattacattcccaaccgcgtggcacaacaac tggcgggcaaacagtcgttgctgattggcgttgccacctccagtctggccctgcacgcgccg tcgcaaattgtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtggtggtgtc gatggtagaacgaagcggcgtcgaagcctgtaaagcggcggtgcacaatcttctcgcgcaac gcgtcagtgggctgatcattaactatccgctggatgaccaggatgccattgctgtggaagct gcctgcactaatgttccggcgttatttcttgatgtctctgaccagacacccatcaacagtat tattttctcccatgaagacggtacgcgactgggcgtggagcatctggtcgcattgggtcacc agcaaatcgcgctgttagcgggcccattaagttctgtctcggcgcgtctgcgtctggctggc tggcataaatatctcactcgcaatcaaattcagccgatagcggaacgggaaggcgactggag tgccatgtccggttttcaacaaaccatgcaaatgctgaatgagggcatcgttcccactgcga tgctggttgccaacgatcagatggcgctgggcgcaatgcgcgccattaccgagtccgggctg cgcgttggtgcggatatctcggtagtgggatacgacgataccgaagacagctcatgttatat cccgccgtcaaccaccatcaaacaggattttcgcctgctggggcaaaccagcgtggaccgct tgctgcaactctctcagggccaggcggtgaagggcaatcagctgttgcccgtctcactggtg aaaagaaaaaccaccctggcgcccaatacgcaaaccgcctctccccgcgcgttggccgattc attaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaatt aatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtat gttgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgaccatgattacg gattcactggccgtcgttttac	SEQ ID NO. 12

In the present invention, the host cell chromosome can be modified by insertion of at least one nucleic acid sequence containing at least one copy of a gene encoding a LacI protein, the gene being capable of use by the cell to, preferably, constitutively express the encoded LacI protein, and the polynucleotide containing the gene being other than a PlacI-*lacI-lacZYA* nucleic acid sequence (i.e. a Plac(-) version of the PlacI-*lacI-lacZYA* operon) or a PlacI-*lacI-lacZ* polynucleotide (i.e. a structural lac utilization operon gene-containing portion of such a Plac(-) operon, such as an at least partially truncated version of a PlacI-*lacI-lacZYA* nucleic acid sequence).

The gene encoding the chosen LacI protein is preferably constitutively expressed. This may be accomplished by use of any promoter that is constitutively expressed in the selected expression host cell. For example, a native *E. coli* PlacI may be operably attached to

the selected LacI coding sequence, or a different constitutively expressed promoter may be operably attached thereto. In some cases, a regulated promoter may be used, provided that the regulated promoter is maintained throughout fermentation in a state wherein the LacI protein is continually expressed there from. In a particular embodiment, a lac or tac family promoter is utilized in the present invention, including Plac, Ptac, Ptrc, PtacII, PlacUV5, lpp-PlacUV5, lpp-lac, nprM-lac, T7lac, T5lac, T3lac, and Pmac.

Genomic Insertion Sites

Chromosomal insertion may be performed according to any technique known in the art. For example, see: DS Toder, "Gene replacement in *Pseudomonas aeruginosa*," *Methods in Enzymology* 235:466-74 (1994); and J Quandt & MF Hynes, "Versatile suicide vectors which allow direct selection for gene replacement in Gram negative bacteria," *Gene* 127(1):15-21 (1993). Transposon-type insertion techniques such as are known in the art, followed by selection, may also be used; see, e.g., IY Goryshin & WS Reznikoff, "Tn5 in vitro transposition," *Journal of Biological Chemistry* 273(13):7367-74 (1998). Alternatively, gene transfection by (non-lytic) phage transduction may also be used for chromosomal insertion; see, e.g., JH Miller, *Experiments in Molecular Genetics* (1972) (Cold Spring Harbor Lab., NY).

Sites within the bacterial expression host cell chromosome that are useful places in which to insert the *lacI* gene(s), or derivative thereof, include any location that is are not required for cell function under the fermentation conditions used, for example within any gene whose presence, transcription, or expression is important for the healthy functioning of the cell under the fermentation conditions used. Illustrative examples of such insertion sites include, but are not limited to: sucrose import and metabolism genes (e.g., *sacB*), fructose import and catabolism genes (e.g., fructokinase genes, 1-phosphofructokinase genes), aromatic carbon source import and utilization genes (e.g., anthranilate operon genes, such as *antABC* genes, benzoate operon genes, as *benABCD* genes), beta-lactamase genes (e.g., *ampC*, *bll1*, *blc* genes, *blo* genes, *blp* genes), alkaline phosphatase genes (e.g., *phoA*), nucleobase or nucleotide biosynthetic genes (e.g., *pyrBCDEF* genes), amino acid biosynthetic genes (e.g., *proABC* genes), aspartate semi-aldehyde dehydrogenase genes (e.g., *asd*), 3-isopropylmalate dehydrogenase genes (e.g., *leuB*), and anthranilate synthase genes (e.g., *trpE*).

In any embodiment in which the genomic insertion has resulted in or is concomitant with an auxotrophy, then either the host cell will be grown in media supplying an effective replacement metabolite to the cell to overcome (and avoid) the lethal effect, or a replacement gene will be provided in the host cell that expresses a biocatalyst effective to restore the corresponding prototrophy, e.g., as a selection marker gene. The gene or genes selected for deletion or inactivation (i.e. "knock-out") in constructing a metabolic auxotroph can be any gene encoding an enzyme that is operative in a metabolic pathway. The enzyme can be one that is involved in the anabolic biosynthesis of molecules that are necessary for cell survival. Alternatively, the enzyme can be one that is involved in the catabolic utilization of molecules that are necessary for cell survival. Preferably, all operative genes encoding a given biocatalytic activity are deleted or inactivated in order to ensure removal of the targeted enzymatic activity from the host cell in constructing the auxotrophic host cell. Alternatively, the host cell can exhibit a pre-existing auxotrophy (i.e. native auxotrophy), wherein no further genetic modification via deletion or inactivation (knock-out) need be performed.

For example, an amino acid biosynthetic gene (e.g., a *proA*, *proB*, or *proC* gene) or a nucleobase or nucleotide biosynthetic gene (e.g., *pyrB*, *pyrC*, *pyrD*, *pyrE*, or *pyrF*) may be used as the insertion site, in which case a necessary biosynthetic activity is normally disrupted, thus producing an auxotrophy. In such a case, either: 1) the medium is supplemented to avoid metabolic reliance on the biosynthetic pathway, as with a proline or uracil supplement; or 2) the auxotrophic host cell is transformed with a further gene that is expressed and thus replaces the biocatalyst(s) missing from the biosynthetic pathway, thereby restoring prototrophy to the cell, as with a metabolic selection marker gene such as *proC, pyrF, or thyA*. In a particular embodiment, the *lacI* transgene, or variant thereof, is inserted into a cell that is concomitantly or subsequently auxotrophically induced through the knock-out of a gene, or combination of genes, selected from the group consisting of *pyrF*, *thyA*, and *proC*. In a specific embodiment, a native *E.coli lacI*, $lacI^Q$, or $lacI^{Q1}$ transgene is inserted into a cell that is concomitantly or subsequently rendered auxotrophic through the knock-out of *pyrF*. In another specific embodiment, a native *E.coli lacI*, $lacI^Q$, or $lacI^{Q1}$ transgene is inserted into a cell that is concomitantly of subsequently rendered auxotrophic through the knock-out of *proC*. In still a further embodiment, a native *E.coli lacI*, $lacI^Q$, or $lacI^{Q1}$ transgene is inserted into a cell that is concomitantly or subsequently rendered auxotrophic through the knock-out of *pyrF* and *proC*.

In another embodiment, a native *E.coli lacI*, $lacI^Q$, or $lacI^{Q1}$ transgene, or derivative thereof, can be inserted into the Levansucrase locus of the host cell. For example, in one

particular embodiment, a native *E.coli lacI*, *lacI*^{Q}, or *lacI*^{Q1} transgene, or derivative thereof, can be inserted in the Levansucrase gene locus of *Pseudomonas fluorescens*. In particular, a native *E.coli lacI*, *lacI*^{Q}, or *lacI*^{Q1} transgene, or derivative thereof, can be inserted into the Levansucrase gene locus of *Pseudomonas fluorescens* having the nucleic acid sequence of SEQ ID. NO. 13 (Table 14).

TABLE 14 –OPEN READING FRAME OF *PF* LEVAN SUCRASE GENE LOCUS

ctacccagaacgaagatcagcgcctcaatggcctcaaggttctactggtcgatgattcagcc gaagtcgttgaggtgctgaacatgctgctggaaatggaaggcgcccaagtgagcgccttcag cgacccttgagcgcgcttgaaacagcccgggatgcccattacgacgtgattatttcggaca tcggcatgccgaaaatgaatggccatgagctgatgcagaagctgcgtaaagtaggccacctt cgacaggctcccgccatcgccttaacgggctatggcgctggcaatgaccagaaaaaggcgac tgaatcgggctttaatgcgcatgtcagcaaacccgttggccatgattcgctcatcaccttga tcgaaaaactgtgccgctcccgcccctaggcgtggggcaggcgttcaagggtagatgaactg agaaaagcgcacggacgcgcccgtttctggtcgcgacacctgggtatccacgctgcccaccg tgtcgctgcgcaaggtcaggtacaacacggcctggccggcgctgtcactcagcatccagacg ctcacacccctcccggccgccctggcctggagcgggctgaggctgcagcatctcgatattgaa accgcgcagcagctcaccgctcaactcgacctccagggggttcctgggccttaccttgcacat gaatcaccagcccatcggaggcgccattgcgcaaaaagcgttggtactccacgcgcaactgc ccatcggcactgcgcacctcgcggctgctcagcggcccgctggaaaacagccctgccaagct caagccgatcagcaccagcagcgcgtaccaacccacccgctcaaagcgccagaccttgcgct gcaaggccatgttttcctgcaccggataattgcggctgtgtaagtcgtcagggtctgggttg ttcatagcggggcccggactcaacccttgctgtgctcgggagaagacggccccttggtgaca ccccgtgggccggcaatcgcccatatcgcagcgcccagaaacggcagcaccacgactaccgc actccagcctgccttgctggccgaggcgttatcgctgcgccagatgctgttgatgatccacg catcgagcagtacgaggatcactgccaggcctatccagaagtaagtggtttgcatgatgcac ctccaggttatgtaacttttggtgcgcgggtgcgggcagggttcatttattttaggttctct gcctggcgcttggttttgccgccatcatgcgggcaacttcgccgatctacttaatgatcgaac ctcttcaaacaagacaagctgaaacgtctcagctcctataaaaagccaaatcatgcacaaat gcatttttgccttgaccacgggaatcgagtcttctaaagtcaaatcactgtatatgaatac agtaatttgattccctcatggacgagacttactatgaaaagcacccctcgaaatttggca aaacaccccatcaacccagcctgtggacccgcgccgatgcgcttaaagtgcatgcggacgac cccaccaccacccagccgctggtcagcgcgaacttcccggtattgagtgacgaggtgtttat ctgggacaccatgccgctgcgtgatatcgacggcaacatcacctccgtcgatggctggtcgg tgatcttcaccctcaccgcggatcgccacccgaacgacccgcaatacctcgatcagaatggc aactacgacgtcatccgcgactggaacgatcgccatggccgggcaaagatgtactactggtt ctcccgcaccggcaaagactggaagctcggcggccgagtgatggctgaaggggtttcgccca ccgtgcgcgaatgggccggcacgccgatcctgttgaacgagcaaggcgaagtagacctgtac tacaccgccgtcacgcccggcgcgaccatcgtcaaggtgcgtggccgcgtggtgaccaccga gcatggcgtcagcctggtgggctttgagaaggtcaagccgctgttcgaggcggacggcaaga tgtaccagaccgaagcgcaaaatgcgttctggggctttcgcgatccatggccgttccgcgac ccgaaagacggcaagctgtacatgctgttcgaaggtaacgtggccggcgaaccgcggctcgca caaggtcggtaaagccgaaatcggcgacgtgccgccaggttatgaagacgtcggtaactcgc gcttccagactgcctgcgtcggtatcgccgtggcccgcgacgaagacggcgacgactgggaa atgctgccaccgctgctgaccgcggtgggcgtcaacgaccagaccgaacgcccgcacttcgt gttccaggacggcaagtactacctgttcaccatcagccacaccttcacctacgccgacggcg tgaccggcccggacggcgtgtacggcttcgtcgccgattcgctgttcggtccgtatgtgccg ttgaacgcgtctggtctggtactgggcaacccgtcctcccaaccgttccagacctactcgca ctgcgtcatgcccaacggcctggtgacctccttcatcgacagcgtaccgaccgacgacaccg gcacgcagatccgtatcggcggcaccgaagcaccgacggtgggcatcaagatcaaagggcag caaacgtttgtggtcgctgagtatgactacggttacatcccgccgatgctcgacgttacgct caagtaaccggaggctatgaggtagcggctttgagctcgatgacaaacccgcggtgaatatt cgctgcacctgtggcgagggagcttgctcccggttgggccggacagccgccatcgcaggcaa gccagctccacattttggttcctggggcgtcagggaggtatgtgtcggctgaggggccgtc acgggagcaagctcctcgccacaggttcaacagcccattgggtggatattcaggaaataga aatgcctgcaccattgagttgagtc	SEQ ID NO. 13

IV. LACO SEQUENCES

Attempts to repress the leakiness of a promoter must be balanced by the potential concomitant reduction in target recombinant polypeptide expression. One approach to further repress a promoter and reduce the leakiness of the promoters is to modify regulatory elements known as operator sequences, to increase the capacity of the associated repressor protein to bind to the operator sequence without reducing the potential expression of the target recombinant polypeptide upon induction.

It has been discovered that the use of a dual *lac* operator in *Pseudomonas fluorescens* offers superior repression of pre-induction recombinant protein expression without concomitant reductions in induced protein yields.

In one embodiment, a Pseudomonad organism is provided comprising a nucleic acid construct containing a nucleic acid comprising at least one *lacO* sequence involved in the repression of transgene expression. In a particular embodiment, the Pseudomonad host cell is *Pseudomonad fluorescens*. In one embodiment, the nucleic acid construct comprises more than one *lacO* sequence. In another embodiment, the nucleic acid construct comprises at least one, and preferably more than one, *lacOid* sequence. In one embodiment, the nucleic acid construct comprises a *lacO* sequence, or derivative thereof, located 3' of a promoter, and a *lacO* sequence, or derivative thereof, located 5' of a promoter. In a particular embodiment, the *lacO* derivative is a *lacOid* sequence.

In another embodiment of the present invention, nucleic acid constructs comprising more than one *lac* operator sequence, or derivative thereof for use in a Pseudomonad host cell is provided. In one embodiment, at least one *lac* operator sequence may be a $lacO_{id}$ sequence.

The native *E.coli lac* operator acts to down regulate expression of the *lac* operon in the absence of an inducer. To this end, the *lac* operator is bound by the LacI repressor protein, inhibiting transcription of the operon. It has been determined that the LacI protein can bind simultaneously to two *lac* operators on the same DNA molecule. See, for example, Muller et al., (1996) "Repression of lac promoter as a function of distance, phase, and quality of an auxiliary lac operator," J.Mol.Biol. 257: 21-29. The repression is mediated by the promoter-proximal operator O_1 and the two auxiliary operators O_2 and O_3, located 401 base pairs downstream of O_1 within the coding region of the *lacZ* gene and 92 bp upstream of O_1, respectively (See Figure 4). Replacement of the native *E.coli lac* operator sequences with an ideal *lac* operator (O_{id}) results in increased repression of the native *lac* operon in *E.coli*. See

Muller et al., (1996) "Repression of *lac* promoter as a function of distance, phase, and quality of an auxiliary *lac* operator", J.Mol.Biol. 257: 21-29.

The *lacO* sequence or derivative can be positioned in the *E. coli* native O_1 position with respect to a promoter. Alternatively, the *lacO* sequence or derivative can be positioned in the *E.coli* O_3 position with respect to a promoter. In another embodiment, the *lacO* sequence or derivative can be located in the E.coli native O_1 position, the native O_3 position, or both with respect to a promoter. In one embodiment, the nucleic acid construct contains at least one *lacOid* sequence either 5' to the promoter sequence or 3' to the promoter sequence. In a particular embodiment, the nucleic acid construct contains a *lacOid* sequence 3' of a promoter, and at least one *lacO* sequence, or derivative, 5' of a promoter. In an alternative embodiment, the nucleic acid construct contains a *lacOid* sequence 5' of a promoter, and at least one *lacO* sequence, or derivative, 3' of a promoter. In still another embodiment, the nucleic acid construct contains a *lacOid* sequence both 5' and 3' of a promoter.

In a particular embodiment, the *lacO* sequence is *lacOid* represented by SEQ ID NO. 14, or a sequence substantially homologous. In another embodiment, a *lacOid* sequence of SEQ. ID. NO. 59, or a sequence substantially homologous to SEQ ID NO. 59 is employed.

TABLE 15 - *LACOID* SEQUENCE

5'-AATTGTGAGCGCTCACAATT-3'	SEQ ID NO. 14
5' – tgtgtggAATTGTGAGCGCTCACAATTccacaca-3'	SEQ ID NO. 59

V. ISOLATED NUCLEIC ACIDS AND AMINO ACIDS

Disclosed herein are nucleic acid sequences are provided for use in the improved production of proteins.

In one embodiment, nucleic acid sequences encoding prototrophy-restoring enzymes for use in an auxotrophic

Pseudomonad host cells are provided. In a particular embodiment, nucleic acid sequences encoding nitrogenous base compound biosynthesis enzymes purified from the organism *Pseudomonas fluorescens* are provided. In one embodiment, nucleic acid sequences encoding the *pyrF* gene in *Pseudomonas fluorescens* is provided (SEQ. ID No.s 1 and 3). In another embodiment, a nucleic acid sequence encoding the *thyA* gene in *Pseudomonas fluorescens* is provided (SEQ. ID. No. 4). In still another embodiment, nucleic acid sequences encoding an amino acid biosynthetic compound purified from the organism *Pseudomonas fluorescens* are provided. In a particular embodiment, a nucleic acid sequence encoding the *proC* gene in *Pseudomonas fluorescens* is provided (SEQ. ID No.s 6 and 8).

Also disclosed are amino acid sequences for use in the improved production of proteins.

In one embodiment, amino acid sequences of nitrogenous base compound biosynthesis enzymes purified from the organism *Pseudomonas fluorescens* are provided. In one embodiment, the amino acid sequence containing SEQ. ID No. 2 is provided. In another embodiment, an amino acid sequence containing SEQ. ID No. 5 is provided. In still another embodiment, amino acid sequences of an amino acid biosynthetic compound purified from the organism *Pseudomonas fluorescens* is provided. In a particular embodiment, an amino acid sequence containing SEQ. ID No. 7 is provided.

One embodiment of the present invention is novel isolated nucleic acid sequences of the *Pseudomonas fluorescens pyrF* gene (Table 2, Seq. ID No. 1; Table 4, Seq. ID No. 3). Another aspect of the present invention provides isolated peptide sequences of the *Pseudomonas fluorescens pyrF* gene (Table 3, Seq. ID No. 2). Nucleic and amino acid sequences containing at least 90, 95, 98 or 99% homologous to Seq. ID Nos. 1, 2, or 3 are provided. In addition, nucleotide and peptide sequences that contain

at least 10, 15, 17, 20 or 25, 30, 40, 50, 75, 100, 150, 250, 350, 500, or 1000 contiguous nucleic or amino acids of Seq ID Nos 1, 2, or 3 are also provided. Further provided are fragments, derivatives and analogs of Seq. ID Nos. 1, 2, or 3. Fragments of Seq. ID Nos. 1, 2, or 3 can include any contiguous nucleic acid or peptide sequence that includes at least about 10 bp, 15 bp, 17 bp, 20 bp, 50 bp, 100 bp, 500 bp, 1 kbp, 5 kbp or 10 kpb.

Another embodiment of the present invention is novel isolated nucleic acid sequences of the *Pseudomonas fluorescens thyA* gene (Table 5, Seq. ID No. 4). Another aspect of the present invention provides isolated peptide sequences of the *Pseudomonas fluorescens thyA* gene (Table 6, Seq. ID No. 5). Nucleic and amino acid sequences containing at least 90, 95, 98 or 99% homologous to Seq. ID Nos. 4 or 5 are provided. In addition, nucleotide and peptide sequences that contain at least 10, 15, 17, 20 or 25, 30, 40, 50, 75, 100, 150, 250, 350, 500, or 1000 contiguous nucleic or amino acids of Seq ID Nos 4 or 5 are also provided. Further provided are fragments, derivatives and analogs of Seq. ID Nos. 4 or 5. Fragments of Seq. ID Nos. 4 or 5 can include any contiguous nucleic acid or peptide sequence that includes at least about 10 bp, 15 bp, 17 bp, 20 bp, 50 bp, 100 bp, 500 bp, 1 kbp, 5 kbp or 10 kpb.

Another embodiment of the present invention is novel isolated nucleic acid sequences of the *Pseudomonas fluorescens proC* gene (Table 7, Seq. ID No. 6; Table 9, Seq. ID No. 8).

Also disclosed are isolated peptide sequences of the *Pseudomonas fluorescens proC* gene (Table 8, Seq. ID No. 7). Nucleic and amino acid sequences containing at least 90, 95, 98 or 99% homologous to Seq. ID Nos. 6, 7, or 8 are provided. In addition, nucleotide and peptide sequences that contain at least 10, 15, 17, 20 or 25, 30, 40, 50, 75, 100, 150, 250, 350, 500, or 1000 contiguous nucleic or amino acids of Seq ID Nos 6, 7, or 8 are also

provided. Further provided are fragments, derivatives and analogs of Seq. ID Nos. 6, 7, or 8. Fragments of Seq. ID Nos. 6, 7, or 8 can include any contiguous nucleic acid or peptide sequence that includes at least about 10 bp, 15 bp, 17 bp, 20 bp, 50 bp, 100 bp, 500 bp, 1 kbp, 5 kbp or 10 kpb.

Sequence Homology

Sequence homology is determined according to any of various methods well known in the art. Examples of useful sequence alignment and homology determination methodologies include those described below.

Alignments and searches for homologous sequences can be performed using the U.S. National Center for Biotechnology Information (NCBI) program, MegaBLAST (currently available at http://www.ncbi.nlm.nih.gov/BLAST/). Use of this program with options for percent identity set at 70% for amino acid sequences, or set at 90% for nucleotide sequences, will identify those sequences with 70%, or 90%, or greater homology to the query sequence. Other software known in the art is also available for aligning and/or searching for homologous sequences, *e.g.*, sequences at least 70% or 90% homologous to an information string containing a promoter base sequence or activator-protein-encoding base sequence according to the present invention. For example, sequence alignments for comparison to identify sequences at least 70% or 90% homologous to a query sequence can be performed by use of, *e.g.*, the GAP, BESTFIT, BLAST, FASTA, and TFASTA programs available in the GCG Sequence Analysis Software Package (available from the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705), with the default parameters as specified therein, plus a parameter for the extent of homology set at 70% or 90%. Also, for example, the CLUSTAL program (available in the PC/Gene software package from Intelligenetics, Mountain View, Cal.) may be

used.

These and other sequence alignment methods are well known in the art and may be conducted by manual alignment, by visual inspection, or by manual or automatic application of a sequence alignment algorithm, such as any of those embodied by the above-described programs. Various useful algorithms include, *e.g.*: the similarity search method described in W.R. Pearson & D.J. Lipman, *Proc. Nat'l Acad. Sci. USA* 85:2444-48 (Apr 1988); the local homology method described in T.F. Smith & M.S. Waterman, in *Adv. Appl. Math.* 2:482-89 (1981) and in *J. Molec. Biol.* 147:195-97 (1981); the homology alignment method described in S.B. Needleman & C.D. Wunsch, *J. Molec. Biol.* 48(3):443-53 (Mar 1970); and the various methods described, *e.g.*, by W.R. Pearson, in *Genomics* 11(3):635-50 (Nov 1991); by W.R. Pearson, in *Methods Molec. Biol.* 24:307-31 and 25:365-89 (1994); and by D.G. Higgins & P.M. Sharp, in *Comp. Appl'ns in Biosci.* 5:151-53 (1989) and in *Gene* 73(1):237-44 (15 Dec 1988).

Nucleic acid hybridization performed under highly stringent hybridization conditions is also a useful technique for obtaining sufficiently homologous sequences for use herein.

VI. NUCLEIC ACID CONSTRUCTS

Disclosed herein are nucleic acid constructs for use in the improved production of peptides.

In one embodiment, a nucleic acid construct for use in transforming a Pseudomonad host cell comprising a) a nucleic acid sequence encoding a recombinant polypeptide, and b) a nucleic acid sequence encoding a prototrophy-enabling enzyme is provided. In another embodiment, the nucleic acid construct further comprises c) a Plac-Ptac family promoter. In still another embodiment, the nucleic acid construct further comprises d) at least one *lacO* sequence, or derivative, 3' of a lac or tac family promoter. In yet another embodiment, the nucleic acid

construct further comprises e) at least one *lacO* sequence, or derivative, 5' of a lac or tac family promoter. In one embodiment, the derivative *lacO* sequence can be a *lacOid* sequence. In a particular embodiment, the Pseudomonad organism is *Pseudomonas fluorescens*.

In one embodiment of the present invention, nucleic acid constructs are provided for use as expression vectors in Pseudomonad organisms comprising a) a nucleic acid sequence encoding a recombinant polypeptide, b) a Plac-Ptac family promoter, c) at least one *lacO* sequence, or derivative, 3' of a lac or tac family promoter, d) at least one *lacO* sequence, or derivative, 5' of a lac or tac family promoter. In one embodiment, the derivative *lacO* sequence can be a *lacOid* sequence. In one embodiment, the nucleic acid construct further comprises e) a prototrophy-enabling selection marker for use in an auxotrophic Pseudomonad cell. In a particular embodiment, the Pseudomonad organism is *Pseudomonas fluorescens*.

In one embodiment of the present invention, a nucleic acid construct is provided comprising nucleic acids that encode at least one biosynthetic enzyme capable of transforming an auxotrophic host cell to prototrophy. The biosynthetic enzyme can be any enzyme capable of allowing an auxotrophic host cell to survive on a selection medium that, without the expression of the biosynthetic enzyme, the host cell would be incapable of survival due to the auxotrophic metabolic deficiency. As such, the biosynthetic enzyme can be an enzyme that complements the metabolic deficiency of the auxotrophic host by restoring prototrophic ability to grow on non-auxotrophic metabolite supplemented media.

In one particular embodiment, the present invention provides a nucleic acid construct that encodes a functional orotodine-5'-phosphate decarboxylase enzyme that complements an *pyrF*(-) auxotrophic host. In a particular embodiment, the nucleic acid construct contains the nucleic acid sequence of SEQ ID NO. 1 or 3. In an alternative embodiment, the nucleic acid construct contains a nucleic acid sequence that encodes the amino acid sequence of SEQ ID NO. 2.

In another particular embodiment, the present invention provides a nucleic acid construct that encodes a functional thymidylate synthase enzyme that complements a *thyA* (-) auxotrophic host. In a particular embodiment, the nucleic acid construct contains the nucleic acid sequence of SEQ ID NO. 4. In an alternative embodiment, the nucleic acid construct contains a nucleic acid sequence that encodes the amino acid sequence of SEQ ID NO. 5.

In a further particular embodiment, the present invention provides a nucleic acid construct that encodes a functional Δ^1-pyrroline-5-carboxylate reductase enzyme that complements a *proC* (-) auxotrophic host. In a particular embodiment, the nucleic acid construct contains the nucleic acid sequence of SEQ ID NO. 6 or 8. In an alternative embodiment, the nucleic acid construct contains the nucleic acid sequence that encodes the amino acid sequence of SEQ ID NO. 7.

In an alternative embodiment, the present invention provides a nucleic acid construct that encodes at least one biosynthetic enzyme capable of transforming an auxotrophic host cell to prototrophy and an additional non-auxotrophic selection marker. Examples of non-auxotrophic selection markers are well known in the art, and can include markers that give rise to colorimetric/chromogenic or a luminescent reaction such as *lacZ* gene, the *GUS* gene, the *CAT* gene, the *luxAB* gene, antibiotic resistance selection markers such as amphotericin B, bacitracin, carbapenem, cephalosporin, ethambutol, fluoroquinolones, isonizid, cephalosporin, methicillin, oxacillin, vanomycin, streptomycin, quinolines, rifampin, rifampicin, sulfonamides, ampicillin, tetracycline, neomycin, cephalothin, erythromycin, streptomycin,

kanamycin, gentamycin, penicillin, and chloramphenicol resistance genes, or other commonly used non-auxotrophic selection markers.

In another embodiment, the expression vector can comprise more than one biosynthetic enzyme capable of transforming an auxotrophic host cell to prototrophy. The biosynthetic enzymes can be any enzymes capable of allowing an auxotrophic host cell to survive on a selection medium that, without the expression of the biosynthetic enzyme, the host cell would be incapable of survival due to the auxotrophic metabolic deficiency. As such, the biosynthetic enzymes can be enzymes that complement the metabolic deficiencies of the auxotrophic host by restoring prototrophic ability to grow on non-auxotrophic metabolite supplemented media. For example, an expression vector comprise a first and second prototrophy-enabling selection marker gene, allowing the host cell containing the construct to be maintained under either or both of the conditions in which host cell survival requires the presence of the selection marker gene(s). When only one of the marker-gene dependent survival conditions is present, the corresponding marker gene must be expressed, and the other marker gene(s) may then be either active or inactive, though all necessary nutrients for which the cell remains auxotrophic will still be supplied by the medium. This permits the same target gene, or the same set of covalently linked target genes, encoding the desired transgenic product(s) and/or desired transgenic activity(ies), to be maintained in the host cell continuously as the host cell is transitioned between or among different conditions. The coding sequence of each of the chosen selection marker genes independently can be operatively attached to either a constitutive or a regulated promoter.

In a particular embodiment, the nucleic acid vector comprises a nucleic acid construct that encodes a functional orotodine-5'-phosphate decarboxylase enzyme and a functional Δ^1-pyrroline-5-carboxylate reductase enzyme that can complement a *pyrF*(-) auxotrophic host cell, a *proC*(-) auxotrophic host cell, or a *pyrF*(-)/*proC*(-) dual-auxotrophic host cell. In a particular embodiment, the nucleic acid construct comprises the nucleic acid sequences of SEQ ID NO. 1 or 3, and SEQ ID. NO. 6 or 8. In an alternative embodiment, the nucleic acid construct contains a nucleic acid sequence that encodes the amino acid sequences of SEQ ID NO. 2 and 7.

In an alternative particular embodiment, the nucleic acid vector comprises a nucleic acid construct that encodes a functional orotodine-5'-phosphate decarboxylase enzyme and a functional thymidylate synthase enzyme that can complement a *pyrF*(-) auxotrophic host cell, a *thyA*(-) auxotrophic host cell, or a *pyrF*(-)/*thyA*(-) dual-auxotrophic host cell. In a particular embodiment, the nucleic acid construct comprises the nucleic acid sequences of SEQ ID NO.

1 or 3, and SEQ ID. NO. 4. In an alternative embodiment, the nucleic acid construct contains a nucleic acid sequence that encodes the amino acid sequences of SEQ ID NO. 2 and 5.

In a particular embodiment, the nucleic acid vector comprises a nucleic acid construct that encodes a functional Δ^1-pyrroline-5-carboxylate reductase enzyme and a thymidylate synthase enzyme that can complement a *proC*(-) auxotrophic host cell, a *thyA*(-) auxotrophic host cell, or a *proC*(-)/*thyA*(-) dual-auxotrophic host cell. In a particular embodiment, the nucleic acid construct comprises the nucleic acid sequences of SEQ ID NO. 4, and SEQ ID. NO. 6 or 8. In an alternative embodiment, the nucleic acid construct contains a nucleic acid sequence that encodes the amino acid sequences of SEQ ID NO. 5 and 7.

Promoters

In a fermentation process, once expression of the target recombinant polypeptide is induced, it is ideal to have a high level of production in order to maximize efficiency of the expression system. The promoter initiates transcription and is generally positioned 10-100 nucleotides upstream of the ribosome binding site. Ideally, a promoter will be strong enough to allow for recombinant polypeptide accumulation of around 50% of the total cellular protein of the host cell, subject to tight regulation, and easily (and inexpensively) induced.

The promoters used in accordance with the present invention may be constitutive promoters or regulated promoters. Examples of commonly used inducible promoters and their subsequent inducers include lac (IPTG), lacUV5 (IPTG), tac (IPTG), trc (IPTG), P_{syn} (IPTG), trp (tryptophan starvation), araBAD (l-arabinose), lpp[a] (IPTG), lpp-lac (IPTG), phoA (phosphate starvation), recA (nalidixic acid), proU (osmolarity), cst-1 (glucose starvation), tetA (tretracylin), cadA (pH), nar (anaerobic conditions), PL (thermal shift to 42° C), cspA (thermal shift to 20° C), T7 (thermal induction), T7-lac operator (IPTG), T3-lac operator (IPTG), T5-lac operator (IPTG), T4 gene32 (T4 infection), nprM-lac operator (IPTG), Pm (alkyl- or halo-benzoates), Pu (alkyl- or halo-toluenes), Psal (salicylates), and VHb (oxygen). See, for example, Makrides, S.C. (1996) Microbiol. Rev. 60, 512-538; Hannig G. & Makrides, S.C. (1998) TIBTECH 16, 54-60; Stevens, R.C. (2000) Structures 8, R177-R185. See, e.g.: J. Sanchez-Romero & V. De Lorenzo, Genetic Engineering of Nonpathogenic *Pseudomonas* strains as Biocatalysts for Industrial and Environmental Processes, in Manual of Industrial Microbiology and Biotechnology (A. Demain & J. Davies, eds.) pp.460-74 (1999) (ASM Press, Washington, D.C.); H. Schweizer, Vectors to express foreign genes and techniques to monitor gene expression for Pseudomonads, Current Opinion in Biotechnology, 12:439-445 (2001); and R. Slater & R. Williams, The Expression of Foreign DNA in Bacteria, in

Molecular Biology and Biotechnology (J. Walker & R. Rapley, eds.) pp.125-54 (2000) (The Royal Society of Chemistry, Cambridge, UK).

A promoter having the nucleotide sequence of a promoter native to the selected bacterial host cell can also be used to control expression of the transgene encoding the target polypeptide, e.g, a *Pseudomonas* anthranilate or benzoate operon promoter (Pant, Pben). Tandem promoters may also be used in which more than one promoter is covalently attached to another, whether the same or different in sequence, e.g., a Pant-Pben tandem promoter (interpromoter hybrid) or a Plac-Plac tandem promoter.

Regulated promoters utilize promoter regulatory proteins in order to control transcription of the gene of which the promoter is a part. Where a regulated promoter is used herein, a corresponding promoter regulatory protein will also be part of an expression system according to the present invention. Examples of promoter regulatory proteins include: activator proteins, e.g., *E. coli* catabolite activator protein, MalT protein; AraC family transcriptional activators ; repressor proteins, e.g., *E. coli* LacI proteins; and dual-faction regulatory proteins, e.g., *E. coli* NagC protein. Many regulated-promoter/promoter-regulatory-protein pairs are known in the art.

Promoter regulatory proteins interact with an effector compound, i.e. a compound that reversibly or irreversibly associates with the regulatory protein so as to enable the protein to either release or bind to at least one DNA transcription regulatory region of the gene that is under the control of the promoter, thereby permitting or blocking the action of a transcriptase enzyme in initiating transcription of the gene. Effector compounds are classified as either inducers or co-repressors, and these compounds include native effector compounds and gratuitous inducer compounds. Many regulated-promoter/promoter-regulatory-protein/effector-compound trios are known in the art. Although an effector compound can be used throughout the cell culture or fermentation, in a particular embodiment in which a regulated promoter is used, after growth of a desired quantity or density of host cell biomass, an appropriate effector compound is added to the culture in order to directly or indirectly result in expression of the desired target gene(s).

By way of example, where a lac family promoter is utilized, a *lacI* gene, or derivative thereof such as a $lacI^{Q}$ or $lacI^{Q1}$ gene, can also be present in the system. The *lacI* gene, which is (normally) a constitutively expressed gene, encodes the Lac repressor protein (LacI protein) which binds to the lac operator of these promoters. Thus, where a lac family promoter is utilized, the *lacI* gene can also be included and expressed in the expression system. In the case of the lac promoter family members, e.g., the tac promoter, the effector

60

compound is an inducer, preferably a gratuitous inducer such as IPTG (isopropyl-β-D-1-thiogalactopyranoside, also called "isopropylthiogalactoside").

In a particular embodiment, a lac or tac family promoter is utilized in the present invention, including Plac, Ptac, Ptrc, PtacII, PlacUV5, lpp-PlacUV5, lpp-lac, nprM-lac, T7lac, T5lac, T3lac, and Pmac.

Other Elements

Other regulatory elements can be included in an expression construct, including *lacO* sequences and derivatives, as discussed above. Such elements include, but are not limited to, for example, transcriptional enhancer sequences, translational enhancer sequences, other promoters, activators, translational start and stop signals, transcription terminators, cistronic regulators, polycistronic regulators, tag sequences, such as nucleotide sequence "tags" and "tag" peptide coding sequences, which facilitates identification, separation, purification, or isolation of an expressed polypeptide, including His-tag, Flag-tag, T7-tag, S-tag, HSV-tag, B-tag, Strep-tag, polyarginine, polycysteine, polyphenylalanine, polyaspartic acid, (Ala-Trp-Trp-Pro)n, thioredoxin, beta-galactosidase, chloramphenicol acetyltransferase, cyclomaltodextrin gluconotransferase, CTP:CMP-3-deoxy-D-manno-octulosonate cytidyltransferase, trpE or trpLE, avidin, streptavidin, T7 gene 10, T4 gp55, Staphylococcal protein A, streptococcal protein G, GST, DHFR, CBP, MBP, galactose binding domain, Calmodulin binding domain, GFP, KSI, c-myc, ompT, ompA, pelB, , NusA, ubiquitin, and hemosylin A.

At a minimum, a protein-encoding gene according to the present invention can include, in addition to the protein coding sequence, the following regulatory elements operably linked thereto: a promoter, a ribosome binding site (RBS), a transcription terminator, translational start and stop signals. Useful RBSs can be obtained from any of the species useful as host cells in expression systems according to the present invention, preferably from the selected host cell. Many specific and a variety of consensus RBSs are known, e.g., those described in and referenced by D. Frishman et al., Starts of bacterial genes: estimating the reliability of computer predictions, Gene 234(2):257-65 (8 Jul 1999); and B.E. Suzek et al., A probabilistic method for identifying start codons in bacterial genomes, Bioinformatics 17(12):1123-30 (Dec 2001). In addition, either native or synthetic RBSs may be used, e.g., those described in: EP 0207459 (synthetic RBSs); O. Ikehata et al., Primary structure of nitrile hydratase deduced from the nucleotide sequence of a *Rhodococcus* species

and its expression in *Escherichia coli*, Eur. J. Biochem. 181(3):563-70 (1989) (native RBS sequence of AAGGAAG). Further examples of methods, vectors, and translation and transcription elements, and other elements useful in the present invention are described in, e.g.: US Patent No. 5,055,294 to Gilroy and US Patent No. 5,128,130 to Gilroy et al.; US Patent No. 5,281,532 to Rammler et al.; US Patent Nos. 4,695,455 and 4,861,595 to Barnes et al.; US Patent No. 4,755,465 to Gray et al.; and US Patent No. 5,169,760 to Wilcox.

Vectors

Transcription of the DNA encoding the enzymes of the present invention by a Pseudomonad host can further be increased by inserting an enhancer sequence into the vector or plasmid. Typical enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp in size that act on the promoter to increase its transcription.

Generally, the recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the Pseudomonad host cell, e.g., the prototrophy restoring genes of the present invention, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters have been described above. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and in certain embodiments, a leader sequence capable of directing secretion of the translated polypeptide. Optionally, and in accordance with the present invention, the heterologous sequence can encode a fusion polypeptide including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for use with *P. fluorescens* in expressing enzymes are constructed by inserting a structural DNA sequence encoding a desired target polypeptide together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable hosts for transformation in accordance with the present disclosure include various species within the genera *Pseudomonas*, and particularly particular is the host cell strain of *Pseudomonas fluorescens*.

Vectors are known in the art as useful for expressing recombinant proteins in host cells, and any of these may be modified and used for expressing the genes according to the present invention. Such vectors include, e.g., plasmids, cosmids, and phage expression vectors. Examples of useful plasmid vectors that can be modified for use on the present

invention include, but are not limited to, the expression plasmids pBBR1MCS, pDSK519, pKT240, pML122, pPS10, RK2, RK6, pRO1600, and RSF1010. Further examples can include pALTER-Ex1, pALTER-Ex2, pBAD/IIis, pBAD/Myc-His, pBAD/gIII, pCal-n, pCal-n-EK, pCal-c, pCal-Kc, pcDNA 2.1, pDUAL, pET-3a-c, pET 9a-d, pET-11a-d, pET-12a-c, pET-14b, pET15b, pET-16b, pET-17b, pET-19b, pET-20b(+), pET-21a-d(+), pET-22b(+), pET-23a-d(+), pET24a-d(+), pET-25b(+), pET-26b(+), pET-27b(+), pET28a-c(+), pET-29a-c(+), pET-30a-c(+), pET31b(+), pET-32a-c(+), pET-33b(+), pET-34b(+), pET35b(+), pET-36b(+), pET-37b(+), pET-38b(+), pET-39b(+), pET-40b(+), pET-41a-c(+), pET-42a-c(+), pET-43a-c(+), pETBlue-1, pETBlue-2, pETBlue-3, pGEMEX-1, pGEMEX-2, pGEX1λT, pGEX-2T, pGEX-2TK, pGEX-3X, pGEX-4T, pGEX-5X, pGEX-6P, pHAT10/11/12, pHAT20, pHAT-GFPuv, pKK223-3, pLEX, pMAL-c2X, pMAL-c2E, pMAL-c2g, pMAL-p2X, pMAL-p2E, pMAL-p2G, pProEX HT, pPROLar.A, pPROTet.E, pQE-9, pQE-16, pQE-30/31/32, pQE-40, pQE-50, pQE-70, pQE-80/81/82L, pQE-100, pRSET, and pSE280, pSF380, pSE420, pThioHis, pTrc99A, pTrcHis, pTrcHis2, pTriEx-1, pTriEx-2, pTrxFus. Other examples of such useful vectors include those described by, e.g.: N. Hayase, in Appl. Envir. Microbiol. 60(9):3336-42 (Sep 1994); A.A. Lushnikov et al., in Basic Life Sci. 30:657-62 (1985); S. Graupner & W. Wackernagel, in Biomolec. Eng. 17(1):11-16. (Oct 2000); H.P. Schweizer, in Curr. Opin. Biotech. 12(5):439-45 (Oct 2001); M. Bagdasarian & K.N. Timmis, in Curr. Topics Microbiol. Immunol. 96:47-67 (1982); T. Ishii et al., in FEMS Microbiol. Lett. 116(3):307-13 (Mar 1, 1994); I.N. Olekhnovich & Y.K. Fomichev, in Gene 140(1):63-65 (Mar 11, 1994); M. Tsuda & T. Nakazawa, in Gene 136(1-2):257-62 (Dec 22, 1993); C. Nieto et al., in Gene 87(1):145-49 (Mar 1, 1990); J.D. Jones & N. Gutterson, in Gene 61(3):299-306 (1987); M. Bagdasarian et al., in Gene 16(1-3):237-47 (Dec 1981); H.P. Schweizer et al., in Genet. Eng. (NY) 23:69-81 (2001); P. Mukhopadhyay et al., in J. Bact. 172(1):477-80 (Jan 1990); D.O. Wood et al., in J. Bact. 145(3):1448-51 (Mar 1981); and R. Holtwick et al., in Microbiology 147(Pt 2):337-44 (Feb 2001).

Further examples of expression vectors that can be useful in *Pseudomonas* host cells include those listed in Table 16 as derived from the indicated replicons.

TABLE 16. SOME EXAMPLES OF USEFUL EXPRESSION VECTORS

Replicon	Vector(s)
pPS10	pCN39, pCN51
RSF1010	pKT261-3

	pMMB66EH
	pEB8
	pPLGN1
	pMYC1050
RK2/RP1	pRK415
	pJB653
pRO1600	pUCP
	pBSP

The expression plasmid, RSF1010, is described, e.g., by F. Heffron et al., in Proc. Nat'l Acad. Sci. USA 72(9):3623-27 (Sep 1975), and by K. Nagahari & K. Sakaguchi, in J. Bact. 133(3):1527-29 (Mar 1978). Plasmid RSF1010 and derivatives thereof are particularly useful vectors in the present invention. Exemplary, useful derivatives of RSF1010, which are known in the art, include, e.g., pKT212, pKT214, pKT231 and related plasmids, and pMYC1050 and related plasmids (see, e.g., US Patent Nos. 5,527,883 and 5,840,554 to Thompson et al.), such as, e.g., pMYC1803. Plasmid pMYC1803 is derived from the RSF1010-based plasmid pTJS260 (see US Patent No. 5,169,760 to Wilcox), which carries a regulated tetracycline resistance marker and the replication and mobilization loci from the RSF1010 plasmid. Other exemplary useful vectors include those described in US Patent No. 4,680,264 to Puhler et al.

In a one embodiment, an expression plasmid is used as the expression vector. In another embodiment, RSF1010 or a derivative thereof is used as the expression vector. In still another embodiment, pMYC1050 or a derivative thereof, or pMYC1803 or a derivative thereof, is used as the expression vector.

VII. EXPRESSION OF RECOMBINANT POLYPEPTIDES IN PSEUDOMONAD HOST CELLS

Disclosed herein are processes of expressing recombinant polypeptides for use in improved protein production are provided.

In one embodiment, the process provides expression of a nucleic acid construct comprising nucleic acids encoding a) a recombinant polypeptide, and b) a prototrophy-restoring enzyme in a Pseudomonad that is auxotrophic for at least one metabolite. In an alternative embodiment, the Pseudomonad is auxotrophic for more than one metabolite. In one embodiment, the Pseudomonad is a *Pseudomonas fluorescens* cell. In a particular

embodiment, a recombinant polypeptide is expressed in a Pseudomonad that is auxotrophic for a metabolite, or combination of metabolites, selected from the group consisting of a nitrogenous base compound and an amino acid. In a more particular embodiment, recombinant polypeptides are expressed in a Pseudomonad that is auxotrophic for a metabolite selected from the group consisting of uracil, proline, and thymidine. In another embodiment, the auxotrophy can be generated by the knock-out of the host *pyrF*, *proC,* or *thyA* gene, respectively. An alternative embodiment, recombinant polypeptides are expressed in an auxotrophic Pseudomonad cell that has been genetically modified through the insertion of a native *E.coli lacI* gene, $lacI^Q$ gene, or $lacI^{Q1}$ gene, other than as part of the PlacI-*lacI*-*lacZYA* operon, into the host cell's chromosome. In one particular embodiment, the vector containing the recombinant polypeptide expressed in the auxotrophic host cell comprises at least two *lac* operator sequences, or derivatives thereof. In still a further embodiment, the recombinant polypeptide is driven by a Plac family promoter.

In another embodiment, the process involves the use of Pseudomonad host cells that have been genetically modified to provide at least one copy of a LacI encoding gene inserted into the Pseudomonad host cell's genome, wherein the lacI encoding gene is other than as part of the PlacI-*lacI*-*lacZYA* operon. In one embodiment, the gene encoding the Lac repressor protein is identical to that of native *E. coli lacI* gene. In another embodiment, the gene encoding the Lac repressor protein is the $lacI^Q$ gene. In still another embodiment, the gene encoding the Lac repressor protein is the $lacI^{Q1}$ gene. In a particular embodiment, the Pseudomonad host cell is *Pseudomonas fluorescens*. In another embodiment, the Pseudomonad is further genetically modified to produce an auxotrophic cell. In another embodiment, the process produces recombinant polypeptide levels of at least about 3 g/L, 4 g/L, 5 g/L 6 g/L, 7 g/L, 8 g/L, 9 g/L or at least about 10 g/L. In another embodiment, the recombinant polypeptide is expressed in levels of between 3 g/L and 100 g/L.

The method generally includes:

a) providing a Pseudomonad host cell, preferably a *Pseudomonas fluorescens*, as described in the present invention,

b) transfecting the host cell with at least one nucleic acid expression vector comprising i) a target recombinant polypeptide of interest, and, in the case of the utilization of an auxotrophic host, ii) a gene encoding a prototrophy enabling enzyme that, when expressed, overcomes the auxotrophy of the host cell;

c) growing the host cell in a growth medium that provides a selection pressure effective for maintaining the nucleic acid expression vector containing the recombinant polypeptide of interest in the host cell; and

d) expressing the target recombinant polypeptide of interest.

The method can further comprise transfecting the host cell with at least once nucleic acid expression vector further comprising iii) a Plac family promoter, and optionally iv) more than one *lac* operator sequences. In one embodiment, at least one *lac* operator sequence may be a $lacO_{id}$ sequence. Preferably, the expression system is capable of expressing the target polypeptide at a total productivity of polypeptide of at least 1 g/L to at least 80 g/L. In a particular embodiment, the recombinant polypeptide is expressed at a level of at least 3 g/L, 4g/L, 5g/L, 6 g/L, 7 g/L, 8 g/L, 9 g/L, 10 g/L, 12 g/L, 15 g/L, 20 g/L, 25 g/L, 30 g/L, 35 g/L, 40 g/L, 45 g/L, 50 g/L, 60 g/L, 70 g/L, or at least 80 g/L. In a particular embodiment, a lac or tac family promoter is utilized in the present invention, including Plac, Ptac, Ptrc, PtacII, PlacUV5, lpp-PlacUV5, lpp-lac, nprM-lac, T7lac, T5lac, T3lac, and Pmac.

In one embodiment, at least one recombinant polypeptide can be expressed in a Pseudomonad cell that is auxotrophic for one metabolite, wherein the auxotrophy serves as a selection marker for the maintenance of the nucleic acid expression vector encoding the polypeptide of interest and the prototrophy-enabling enzyme. Alternatively, more than one recombinant polypeptide can be expressed in a Pseudomonad cell that is auxotrophic for one metabolite, wherein the nucleic acids encoding the recombinant polypeptides can be contained on the same vector, or alternatively, on multiple vectors.

In yet another embodiment, more than one expression vector encoding different target polypeptides can be maintained in a Pseudomonad host cell auxotrophic for at least one metabolite, wherein one expression vector contains a nucleic acid encoding a prototrophic-enabling enzyme and a first target polypeptide of interest, and a second expression vector contains a nucleic acid encoding an alternative, non-auxotrophic selection marker and a second polypeptide of interest.

In another embodiment, at least one recombinant polypeptide can be expressed in a Pseudomonad cell that is auxotrophic for more than one metabolite, wherein the multiple auxotrophies serve as selection markers for the maintenance of nucleic acid expression vectors. For example, an expression vector may be utilized in which a first and second prototrophy-enabling selection marker gene are present. If both marker genes are located on the same DNA construct, the host cell containing the construct may be maintained under either or both of the conditions in which host cell survival requires the presence of the

selection marker gene(s). When only one of the marker-gene dependent survival conditions is present, the corresponding marker gene must be expressed, and the other marker gene(s) can then be either active or inactive, though all necessary nutrients for which the cell remains auxotrophic will still be supplied by the medium. This permits the same target gene, or the same set of covalently linked target genes, encoding the desired transgenic product(s) and/or desired transgenic activity(ies), to be maintained in the host cell continuously as the host cell is transitioned between or among different conditions. If each of the two selection marker genes is located on a different DNA construct, then, in order to maintain both of the DNA constructs in the host cell, both of the marker-gene dependent survival conditions are present, and both of the corresponding marker gene must be expressed. This permits more than one non-covalently linked target gene or set of target gene(s) to be separately maintained in the host cell. The coding sequence of each of the chosen selection marker genes independently can be operatively attached to either a constitutive or a regulated promoter.

Dual-target-gene examples of such a multi-target-gene system include, but are not limited to: (1) systems in which the expression product of one of the target genes interacts with the other target gene itself; (2) systems in which the expression product of one of the target genes interacts with the other target gene's expression product, e.g., a protein and its binding protein or the α- and β- polypeptides of an αn-βn protein; (3) systems in which the two expression products of the two genes both interact with a third component, e.g., a third component present in the host cell; (4) systems in which the two expression products of the two genes both participate in a common biocatalytic pathway; and (5) systems in which the two expression products of the two genes function independently of one another, e.g., a bi-clonal antibody expression system.

In one example of a dual-target-gene system of the above-listed type (1), a first target gene can encode a desired target protein, wherein the first target gene is under the control of a regulated promoter; the second target gene may then encode a protein involved in regulating the promoter of the first target gene, e.g., the second target gene may encode the first target gene's promoter activator or repressor protein. In an example in which the second gene encodes a promoter regulatory protein for the first gene, the coding sequence of the second gene can be under the control of a constitutive promoter. In one embodiment, the second gene will be part of a separate DNA construct that is a maintained in the cell as a high-copy-number construct with a copy number of at least 10, 20, 30, 40, 50, or more than 50 copies being maintained in the host cell.

In another embodiment, the present invention provides the use of more than one *lacO* sequence on an expression vector in the production of recombinant polypeptides in Pseudomonads, particularly in *Pseudomonas fluorescens*.

Also disclosed is a method of producing a recombinant polypeptide comprising transforming a bacterial host cell that is a member of the Pseudomonads and closely related bacteria having at least one chromosomally inserted copy of a Lac repressor protein encoding a *lacI* transgene, or derivative thereof such as $lacI^{Q}$ or $lacI^{Q1}$, which transgene is other than part of a whole or truncated structural gene containing *PlacI-lacI-lacZYA* construct with a nucleic acid construct encoding at least one target recombinant polypeptide. The nucleic acid encoding at least one target recombinant polypeptide can be operably linked to a Plac family promoter, in which all of the Plac family promoters present in the host cell are regulated by Lac repressor proteins expressed solely from the *lacI* transgene inserted in the chromosome. Optionally, the expression system is capable of expressing the target polypeptide at a total productivity of at least 3 g/L to at least 10 g/L. Preferably, the expression system is capable of expressing the target polypeptide at a total productivity of polypeptide of at least 3 g/L, 4g/L, 5g/L, 6 g/L, 7 g/L, 8 g/L, 9 g/L, or at least 10 g/L.

In one embodiment, the present invention provides a method of expressing recombinant polypeptides in an expression system utilizing auxotrophic Pseudomonads or related bacteria that have been further genetically modified to provide at least one copy of a LacI encoding gene inserted into the cell's genome, other than as part of the PlacI-*lacI-lacZYA* operon. In a particular embodiment, a recombinant polypeptide is expressed in an auxotrophic *Pseudomonas fluorescens* host cell containing a *lacI* transgene insert. In another particular embodiment, a recombinant polypeptide is expressed in an auxotrophic *Pseudomonas fluorescens* host cell containing a $lacI^{Q}$

transgene insert. In still another particular embodiment, a recombinant polypeptide is expressed in an auxotrophic *Pseudomonas fluorescens* host cell containing a *lacI*^{Q1} transgene insert. The *Pseudomonas fluorescens* host can be auxotrophic for a biochemical required by the cell for survival. In a particular embodiment, the *Pseudomonas fluorescens* cell is auxotrophic for a nitrogenous base. In a particular embodiment, the *Pseudomonas fluorescens* is auxotrophic for a nitrogenous base selected from the group consisting of thymine and uracil. In a particularly particular embodiment, the *Pseudomonas fluorescens* host cell's auxotrophy is induced by a genetic modification to a *pyrF* or *thyA* gene rendering the associated encoded product non-functional. In an alternative embodiment, the Pseudomonas fluorescens cell is auxotrophic

for an amino acid. In a particular embodiment, the Pseudomonas fluorescens is auxotrophic for the amino acid proline. In a particularly particular embodiment, the Pseudomonas fluorescens host cell's auxotrophy is induced by a genetic modification to a *proC* gene rendering the associated encoded product non-functional.

Transformation

Transformation of the Pseudomonad host cells with the vector(s) may be performed using any transformation methodology known in the art, and the bacterial host cells may be transformed as intact cells or as protoplasts (i.e. including cytoplasts). Exemplary transformation methodologies include poration methodologies, e.g., electroporation, protoplast fusion, bacterial conjugation, and divalent cation treatment, e.g., calcium chloride treatment or $CaCl/Mg^{2+}$ treatment, or other well known methods in the art. See, e.g., Morrison, J. Bact., 132:349-351 (1977); Clark-Curtiss & Curtiss, Methods in Enzymology, 101:347-362 (Wu et al., eds, 1983), Sambrook et al., Molecular Cloning, A Laboratory Manual (2nd ed. 1989); Kriegler, Gene Transfer and Expression: A Laboratory Manual (1990); and Current Protocols in Molecular Biology (Ausubel et al., eds., 1994)).

Selection

Preferably, cells that are not successfully transformed are selected against following transformation, and continuously during the fermentation. The selection marker can be an auxotrophic selection marker or a traditional antibiotic selection marker. When the cell is auxotrophic for multiple nutrient compounds, the auxotrophic cell can be grown on medium supplemented with all of those nutrient compounds until transformed with the prototrophy-restoring vector. Where the host cell is or has been made defective for multiple biosynthetic activities, the prototrophy-restorative marker system(s) can be selected to restore one or more or all of the biosynthetic activities, with the remainder being compensated for by continuing to provide, in the medium, the still-lacking nutrients. In selection marker systems in which more than one biosynthetic activity, and/or more than one prototrophy, is restored, the plurality of selection marker genes may be expressed together on one vector or may be co-expressed separately on different vectors. Even where a single metabolite is the target of the selection marker system, multiple biosynthetic activities may be involved in the selection marker system. For example, two or more genes encoding activities from the same anabolic pathway may be expressed together on one vector or may be co-expressed separately on

different vectors, in order to restore prototrophy in regard to biosynthesis of the compound that is the product of the pathway.

Where the selection marker is an antibiotic resistance gene, the associated antibiotic can be added to the medium to select against non transformed and revertant cells, as well known in the art.

Fermentation

As used herein, the term "fermentation" includes both embodiments in which literal fermentation is employed and embodiments in which other, non-fermentative culture modes are employed. Fermentation may be performed at any scale. In one embodiment, the fermentation medium may be selected from among rich media, minimal media, a mineral salts media; a rich medium may be used, but is preferably avoided. In another embodiment either a minimal medium or a mineral salts medium is selected.

In still another embodiment, a minimal medium is selected. In yet another embodiment, a mineral salts medium is selected. Mineral salts media are particularly particular.

Prior to transformation of the host cell with a nucleic acid construct encoding a prototrophic enabling enzyme, the host cell can be maintained in a media comprising a supplemental metabolite, or analogue thereof, that complements the auxotrophy. Following transformation, the host cell can be grown in a media that is lacking the complementary metabolite that the host cell is auxotrophic for. In this way, host cells that do not contain the selection marker enabling prototrophy are selected against. Likewise cells expressing recombinant proteins from expression vectors containing an antibiotic resistance selection marker gene can be maintained prior to transformation on a medium lacking the associated antibiotic used for selection. After transformation and during the fermentation, an antibiotic can be added to the medium, at concentrations known in the art, to select against non-transformed and revertant cells.

Mineral salts media consists of mineral salts and a carbon source such as, e.g., glucose, sucrose, or glycerol. Examples of mineral salts media include, e.g., M9 medium, *Pseudomonas* medium (ATCC 179), Davis and Mingioli medium (see, BD Davis & ES Mingioli, in J. Bact. 60:17-28 (1950)). The mineral salts used to make mineral salts media include those selected from among, e.g., potassium phosphates, ammonium sulfate or chloride, magnesium sulfate or chloride, and trace minerals such as calcium chloride, borate, and sulfates of iron, copper, manganese, and zinc. No organic nitrogen source, such as peptone, tryptone, amino acids, or a yeast extract, is included in a mineral salts medium.

Instead, an inorganic nitrogen source is used and this may be selected from among, e.g., ammonium salts, aqueous ammonia, and gaseous ammonia. A particular mineral salts medium will contain glucose as the carbon source. In comparison to mineral salts media, minimal media can also contain mineral salts and a carbon source, but can be supplemented with, e.g., low levels of amino acids, vitamins, peptones, or other ingredients, though these are added at very minimal levels.

In one embodiment, media can be prepared using the components listed in Table 16 below. The components can be added in the following order: first $(NH_4)HPO_4$, KH_2PO_4 and citric acid can be dissolved in approximately 30 liters of distilled water; then a solution of trace elements can be added, followed by the addition of an antifoam agent, such as Ucolub N 115. Then, after heat sterilization (such as at approximately 121°C), sterile solutions of glucose $MgSO_4$ and thiamine-HCL can be added. Control of pH at approximately 6.8 can be achieved using aqueous ammonia. Sterile distilled water can then be added to adjust the initial volume to 371 minus the glycerol stock (123 mL). The chemicals are commercially available from various suppliers, such as Merck. This media can allow for high cell density cultivation (HCDC) for growth of *Pseudomonas* species and related bacteria. The HCDC can start as a batch process which is followed by two-phase fed-batch cultivation. After unlimited growth in the batch part, growth can be controlled at a reduced specific growth rate over a period of 3 doubling times in which the biomass concentration can increased several fold. Further details of such cultivation procedures is described by Riesenberg, D.; Schulz, V.; Knorre, W. A.; Pohl, H. D.; Korz, D.; Sanders, E. A.; Ross, A.; Deckwer, W. D. (1991) "High cell density cultivation of Escherichia coli at controlled specific growth rate" J Biotechnol: 20(1) 17-27.

The expression system according to the present disclosure can be cultured in any fermentation format.

For example, batch, fed-batch, semi-continuous, and continuous fermentation modes may be employed herein.

The expression systems according to the present disclosure are useful for transgene expression at any scale (i.e. volume) of fermentation. Thus, e.g., microliter-scale, centiliter scale, and deciliter scale fermentation volumes may be used; and 1 Liter scale and larger fermentation volumes can be used. In one embodiment, the fermentation volume will be at or above 1 Liter. In another embodiment, the fermentation volume will be at or above 5 Liters, 10 Liters, 15 Liters, 20 Liters, 25 Liters, 50 Liters, 75 Liters, 100 Liters, 200 Liters, 500 Liters, 1,000 Liters, 2,000 Liters, 5,000 Liters, 10,000 Liters or 50,000 Liters.

In the present invention, growth, culturing, and/or fermentation of the transformed host cells is performed within a temperature range permitting survival of the host cells, preferably a temperature within the range of about 4°C to about 55°C, inclusive.

Cell Density

An additional advantage in using *Pseudomonas fluorescens* in expressing recombinant proteins includes the ability of *Pseudomonas fluorescens* to be grown in high cell densities compared to *E. coli* or other bacterial expression systems. To this end, *Pseudomonas fluorescens* expressions systems according to the present invention can provide a cell density of about 20 g/L or more. The *Pseudomonas fluorescens* expressions systems according to the present invention can likewise provide a cell density of at least about 70 g/L, as stated in terms of biomass per volume, the biomass being measured as dry cell weight.

In one embodiment, the cell density will be at least 20 g/L. In another embodiment, the cell density will be at least 25 g/L, 30 g/L, 35 g/L, 40 g/L, 45 g/L, 50 g/L, 60 g/L, 70 g/L, 80 g/L, 90 g/L., 100 g/L, 110 g/L, 120 g/L, 130 g/L, 140 g/L, or at least 150 g/L.

In another embodiments, the cell density at induction will be between 20 g/L and 150 g/L;, 20 g/L and 120 g/L; 20 g/L and 80 g/L; 25 g/L and 80 g/L; 30 g/L and 80 g/L; 35 g/L and 80 g/L; 40 g/L and 80 g/L; 45 g/L and 80 g/L; 50 g/L and 80 g/L; 50 g/L and 75 g/L; 50 g/L and 70 g/L; 40 g/L and 80 g/L.

Expression Levels of Recombinant Protein

The expression systems according to the present invention can express transgenic polypeptides at a level at between 5% and 80% total cell protein (%tcp). In one embodiment, the expression level will be at or above 5%, 8%, 10%, 12%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, or 80% tcp.

Isolation and Purification

The recombinant proteins produced according to this invention may be isolated and purified to substantial purity by standard techniques well known in the art, including, but not limited to, ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, nickel chromatography, hydroxylapatite chromatography, reverse phase chromatography, lectin chromatography, preparative

electrophoresis, detergent solubilization, selective precipitation with such substances as column chromatography, immunopurification methods, and others. For example, proteins having established molecular adhesion properties can be reversibly fused a ligand. With the appropriate ligand, the protein can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form. The fused protein is then removed by enzymatic activity. In addition, protein can be purified using immunoaffinity columns or Ni-NTA columns. General techniques are further described in, for example, R. Scopes, Protein Purification: Principles and Practice, Springer-Verlag: N.Y. (1982); Deutscher, Guide to Protein Purification, Academic Press (1990); U.S. Pat. No. 4,511,503; S. Roe, Protein Purification Techniques: A Practical Approach (Practical Approach Series), Oxford Press (2001); D. Bollag, et al., Protein Methods, Wiley-Lisa, Inc. (1996); AK Patra et al., Protein Expr Purif, 18(2): p/ 182-92 (2000); and R. Mukhija, et al., Gene 165(2): p. 303-6 (1995). See also, for example, Ausubel, et al. (1987 and periodic supplements); Deutscher (1990) "Guide to Protein Purification," Methods in Enzymology vol. 182, and other volumes in this series; Coligan, et al. (1996 and periodic Supplements) Current Protocols in Protein Science Wiley/Greene, NY; and manufacturer's literature on use of protein purification products, e.g., Pharmacia, Piscataway, N.J., or Bio-Rad, Richmond, Calif. Combination with recombinant techniques allow fusion to appropriate segments, e.g., to a FLAG sequence or an equivalent which can be fused via a protease-removable sequence. See also, for example., Hochuli (1989) Chemische Industrie 12:69-70; Hochuli (1990) "Purification of Recombinant Proteins with Metal Chelate Absorbent" in Setlow (ed.) Genetic Engineering, Principle and Methods 12:87-98, Plenum Press, NY; and Crowe, et al. (1992) QIAexpress: The High Level Expression & Protein Purification System QUIAGEN, Inc., Chatsworth, Calif.

Detection of the expressed protein is achieved by methods known in the art and includes, for example, radioimmunoassays, Western blotting techniques or immunoprecipitation.

The recombinantly produced and expressed enzyme can be recovered and purified from the recombinant cell cultures by numerous methods, for example, high performance liquid chromatography (HPLC) can be employed for final purification steps, as necessary.

Certain proteins expressed in this invention may form insoluble aggregates ("inclusion bodies"). Several protocols are suitable for purification of proteins from inclusion bodies. For example, purification of inclusion bodies typically involves the extraction, separation and/or purification of inclusion bodies by disruption of the host cells, e.g., by incubation in a buffer of 50 mM TRIS/HCL pH 7.5, 50 mM NaCl, 5 mM MgCl.sub.2, 1 mM

DTT, 0.1 mM ATP, and 1 mM PMSF. The cell suspension is typically lysed using 2-3 passages through a French Press. The cell suspension can also be homogenized using a Polytron (Brinknan Instruments) or sonicated on ice. Alternate methods of lysing bacteria are apparent to those of skill in the art (see, e.g., Sambrook et al., supra; Ausubel et al., supra).

If necessary, the inclusion bodies can be solubilized, and the lysed cell suspension typically can be centrifuged to remove unwanted insoluble matter. Proteins that formed the inclusion bodies may be renatured by dilution or dialysis with a compatible buffer. Suitable solvents include, but are not limited to urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to about 8 M). Although guanidine hydrochloride and similar agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of immunologically and/or biologically active protein. Other suitable buffers are known to those skilled in the art.

Alternatively, it is possible to purify the recombinant proteins or peptides from the host periplasm. After lysis of the host cell, when the recombinant protein is exported into the periplasm of the host cell, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to those skilled in the art. To isolate recombinant proteins from the periplasm, for example, the bacterial cells can be centrifuged to form a pellet. The pellet can be resuspended in a buffer containing 20% sucrose. To lyse the cells, the bacteria can be centrifuged and the pellet can be resuspended in ice-cold 5 mM MgSO.sub.4 and kept in an ice bath for approximately 10 minutes. The cell suspension can be centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of skill in the art.

An initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest. One such example can be ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol includes adding saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This concentration will precipitate the most hydrophobic of proteins. The precipitate is then discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the

protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, either through dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

The molecular weight of a recombinant protein can be used to isolated it from proteins of greater and lesser size using ultrafiltration through membranes of different pore size (for example, Amicon or Millipore membranes). As a first step, the protein mixture can be ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of the protein of interest. The retentate of the ultrafiltration can then be ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

Recombinant proteins can also be separated from other proteins on the basis of its size, net surface charge, hydrophobicity, and affinity for ligands. In addition, antibodies raised against proteins can be conjugated to column matrices and the proteins immunopurified. All of these methods are well known in the art. It will be apparent to one of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (e.g., Pharmacia Biotech).

Renaturation and Refolding

Insoluble protein can be renatured or refolded to generate secondary and tertiary protein structure conformation. Protein refolding steps can be used, as necessary, in completing configuration of the recombinant product. Refolding and renaturation can be accomplished using an agent that is known in the art to promote dissociation/association of proteins. For example, the protein can be incubated with dithiothreitol followed by incubation with oxidized glutathione disodium salt followed by incubation with a buffer containing a refolding agent such as urea.

Recombinant protein can also be renatured, for example, by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be refolded while immobilized on a column, such as the Ni NTA column by using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation can be performed over a period of 1.5 hours or more. After renaturation the proteins can be eluted by the addition of 250 mM immidazole. Immidazole can be removed by a final dialyzing step against PBS or 50

mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein can be stored at 4.degree. C. or frozen at -80.degree. C.

Other methods include, for example, those that may be described in MH Lee et al., Protein Expr. Purif., 25(1): p. 166-73 (2002), W.K. Cho et al., J. Biotechnology, 77(2-3): p. 169-78 (2000), Ausubel, et al. (1987 and periodic supplements), Deutscher (1990) "Guide to Protein Purification," Methods in Enzymology vol. 182, and other volumes in this series, Coligan, et al. (1996 and periodic Supplements) Current Protocols in Protein Science Wiley/Greene, NY, S. Roe, Protein Purification Techniques: A Practical Approach (Practical Approach Series), Oxford Press (2001); D. Bollag, et al., Protein Methods, Wiley-Lisa, Inc. (1996).

VI. Recombinant Polypeptides

The present invention provides improved protein production in bacterial expression systems. Examples of recombinant polypeptides that can be used in the present invention include polypeptides derived from prokaryotic and eukaryotic organisms. Such organisms include organisms from the domain Archea, Bacteria, Eukarya, including organisms from the Kingdom Protista, Fungi, Plantae, and Animalia.

Types of proteins that can be utilized in the present invention include non-limiting examples such as enzymes, which are responsible for catalyzing the thousands of chemical reactions of the living cell; keratin, elastin, and collagen, which are important types of structural, or support, proteins; hemoglobin and other gas transport proteins; ovalbumin, casein, and other nutrient molecules; antibodies, which are molecules of the immune system; protein hormones, which regulate metabolism; and proteins that perform mechanical work, such as actin and myosin, the contractile muscle proteins.

Other specific non-limiting polypeptides include molecules such as, e.g., renin, a growth hormone, including human growth hormone; bovine growth hormone; growth hormone releasing factor; parathyroid hormone; thyroid stimulating hormone; lipoproteins; .alpha.1-antitrypsin; insulin A-chain; insulin B-chain; proinsulin; thrombopoietin; follicle stimulating hormone; calcitonin; luteinizing hormone; glucagon; clotting factors such as factor VIIIC, factor IX, tissue factor, and von Willebrands factor; anti-clotting factors such as Protein C; atrial naturietic factor; lung surfactant; a plasminogen activator, such as urokinase or human urine or tissue-type plasminogen activator (t-PA); bombesin; thrombin; hemopoietic growth factor; tumor necrosis factor-alpha and -beta; enkephalinase; a serum albumin such as human serum albumin; mullerian-inhibiting

substance; relaxin A-chain; relaxin B-chain; prorelaxin; mouse gonadotropin-associated peptide; a microbial protein, such as beta-lactamase; Dnase; inhibin; activin; vascular endothelial growth factor (VEGF); receptors for hormones or growth factors; integrin; protein A or D; rheumatoid factors; a neurotrophic factor such as brain-derived neurotrophic factor (BDNF), neurotrophin-3, -4, -5, or -6 (NT-3, NT-4, NT-5, or NT-6), or a nerve growth factor such as NGF-.beta.; cardiotrophins (cardiac hypertrophy factor) such as cardiotrophin-1 (CT-1); platelet-derived growth factor (PDGF); fibroblast growth factor such as aFGF and bFGF; epidermal growth factor (EGF); transforming growth factor (TGF) such as TGF-alpha and TGF-beta, including TGF-.beta.1, TGF-.beta.2, TGF-.beta.3, TGF-.beta.4, or TGF-.beta.5; insulin-like growth factor-I and -II (IGF-I and IGF-II); des(1-3)-IGF-I (brain IGF-I), insulin-like growth factor binding proteins; CD proteins such as CD-3, CD-4, CD-8, and CD-19; erythropoietin; osteoinductive factors; immunotoxins; a bone morphogenetic protein (BMP); an interferon such as interferon-alpha, -beta, and -gamma; colony stimulating factors (CSFs), e.g., M-CSF, GM-CSF, and G-CSF; interleukins (ILs), e.g., IL-1 to IL-10; anti-HER-2 antibody; superoxide dismutase; T-cell receptors; surface membrane proteins; decay accelerating factor; viral antigen such as, for example, a portion of the AIDS envelope; transport proteins; homing receptors; addressins; regulatory proteins; antibodies; and fragments of any of the above-listed polypeptides.

The recombinant peptides to be expressed by according to the present invention can be expressed from polynucleotides in which the target polypeptide coding sequence is operably attached to transcription and translation regulatory elements to form a functional gene from which the host cell can express the protein or peptide. The coding sequence can be a native coding sequence for the target polypeptide, if available, but will more preferably be a coding sequence that has been selected, improved, or optimized for use in the selected expression host cell: for example, by synthesizing the gene to reflect the codon use bias of a *Pseudomonas* species such as *Pseudomonas fluorescens*. The gene(s) that result will have been constructed within or will be inserted into one or more vector, which will then be transformed into the expression host cell. Nucleic acid or a polynucleotide said to be provided in an "expressible form" means nucleic acid or a polynucleotide that contains at least one gene that can be expressed by the selected bacterial expression host cell.

Extensive sequence information required for molecular genetics and genetic engineering techniques is widely publicly available. Access to complete nucleotide sequences of mammalian, as well as human, genes, cDNA sequences, amino acid sequences and genomes can be obtained from GenBank at the URL address

http://www.ncbi.nlm.nih.gov/Entrez. Additional information can also be obtained from GeneCards, an electronic encyclopedia integrating information about genes and their products and biomedical applications from the Weizmann Institute of Science Genome and Bioinformatics (http://bioinformatics.weizmann.ac.il/cards/), nucleotide sequence information can be also obtained from the EMBL Nucleotide Sequence Database (http://www.ebi.ac.uk/embl/) or the DNA Databank or Japan (DDBJ, http://www.ddbj.nig.ac.jp/; additional sites for information on amino acid sequences include Georgetown's protein information resource website (http://www-nbrf.georgetown.edu/pir/) and Swiss-Prot (http://au.expasy.org/sprot/sprot-top.html).

EXAMPLES

Example 1

Construction of a *pyrF* Selection Marker System in a *P. fluorescens* Host Cell Expression System

Reagents were acquired from Sigma-Aldrich (St. Louis MO) unless otherwise noted. LB is 10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl in a gelatin capsule (BIO 101). When required, uracil (from BIO101, Carlsbad CA) or L-proline was added to a final concentration of 250 ug/mL, and tetracycline was added to 15 ug/mL. LB/5-FOA plates contain LB with 250 mM uracil and 0.5 mg/mL 5-fluoroorotic acid (5-FOA). M9 media consists of 6 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 1 g/L NH_4Cl, 0.5 g/L NaCl, 10 mM $MgSO_4$, 1x HoLe Trace Element Solution, pH7. Glucose was added to a final concentration of 1%. The 1000x HoLe Trace Element Solution is 2.85 g/L H_3BO_3, 1.8 g/L $MnCl_2 \cdot 4H2O$, 1.77 g/L sodium tartrate, 1.36 g/L $FeSO_4 \cdot 7H2O$, 0.04 g/L $CoCl_2 \cdot 6H_2O$, 0.027 g/L $CuCl_2 \cdot 2H_2O$, 0.025 g/L $Na_2MoO_4 \cdot 2H_2O$, 0.02 g/L $ZnCl_2$.

Oligonucleotides Used Herein

MB214pyrF1 (NotI site in bold)
5'-**GCGGCCGC**TTTGGCGCTTCGTTTACAGG-3' (SEQ ID NO:14)

MB214pyrR1 (PvuI site in bold; KpnI site in underlined bold)
5'-**CGATCGGGTACC**TGTCGAAGGGCTGGAGACAT-3' (SEQ ID NO: 15)

pyrFPstF (PstI site in bold)
5'-AA**CTGCAG**GATCAGTTGCGGAGCCTTGG-3' (SEQ ID NO: 16)

pyrFoverlap
5'-TGCTCACTCTAAAAATCTGGAATGGGCTCTCAGGC-3' (SEQ ID NO: 17)

pyrFXbaR2 (XbaI site in bold)
5'-GC**TCTAGA**TGCGTGGCTGGATGAATGAA-3' (SEQ ID NO: 18)

pyranalF
5'-GGCGTCGAACAGGTAGCCTT-3' (SEQ ID NO:19)

pyranalR
5'-CTCGCCTCCTGCCACATCAA-3' (SEQ ID NO:20)

M13F(-40)
5'-CAGGGTTTTCCCAGTCACGA-3' (SEQ ID NO:21)

Cloning of a *pyrF* gene from *P. fluorescens*

The *pyrF* gene was cloned from *P. fluorescens* by polymerase chain reaction (PCR) amplification, using primers MB214pyrF1 and MB214pyrR1 that bind 297 bp upstream from the *pyrF* gene start codon and 212 bp downstream of its stop codon, respectively. Restriction sites were included at the 5' ends of the primers to facilitate further cloning reactions The amplified region upstream of the *pyrF* open reading frame (ORF) was estimated as long enough to include the native promoter upstream of *pyrF*. A strong stem-loop structure at 14-117 bp downstream of the *pyrF* ORF, which may be a transcription terminator, was also included in the downstream flanking region.

To PCR-amplify the *pyrF* gene, the high-fidelity PROOFSTART DNA polymerase was mixed in a 50 uL reaction volume containing buffer provided by the manufacturer (Qiagen, Valencia CA) 0.3 mM dNTPs (Promega, Madison, WI), 1 uM each of MB214pyrF1 and MB214pyrR1 primers, and about 0.3 μg of genomic DNA from *P. fluorescens* MB214. The amplification conditions were 5 min at 95°C, followed by 35 cycles of a 30 sec denaturation at 94°C, 30 sec annealing at 57°C, and a 2 min extension at 72°C, followed by a final step at 72°C for 10 min. The reaction was separated on a 1% gel of SEAKEM GTG agarose (from BioWhittaker Molecular Applications , Rockland ME). The expected 1.2 kb band was excised from the gel and purified by extraction on a ULTRAFREE-DA centrifugal gel

nebulizer from Millipore (Bedford MA) column and de-salted into Tris-HCl buffer with a MICROBIOSPIN 6 P-6 polyacrylamide spin column (from Bio-Rad, Hercules CA).

The cloned gene contained a single ORF, encoding orotidine 5'phosphate decarboxylase. The identity of the gene was further confirmed as *pyrF* by its high similarity (P-value of 3.3×10^{-78}) along the entire length of the gene (209 out of 232 residues) to the *pyrF* gene from *P. aeruginosa*, which had been previously reported (Strych *et al.*, 1994). The *P. fluorescens* strain used was found to contain no other copies of any *pyrF* genes.

Sequencing was performed by The Dow Chemical Company. The *pyrF* sequence is presented within SEQ ID NO:1.

Construction of a *pyrF*(-) *P. fluorescens*

To construct a *pyrF*(-) *P. fluorescens*, the cell's genomic *pyrF* gene was altered by deleting of the ORF between and including the gene's start and stop codons. The deletion was made by fusing *in vitro* the upstream and downstream regions flanking the *pyrF* region on a nonreplicating plasmid, then using allele exchange, i.e. homologous recombination, to replace the endogenous *pyrF* gene in MB101 with the deletion allele.

To construct the fusion of the flanking regions, the "Megaprimer" method (Barik 1997) was used, whereby the region upstream and then downstream of the desired deletion were subsequently amplified by PCR using an overlapping primer with homology on both sides of the desired deletion, so that the flanking regions become linked, leaving out the *pyrF* ORF. The upstream region was amplified from MB214 genomic DNA using the Proofstart polymerase (Qiagen) as described above, with the primers pyrFPstF and pyrFoverlap, and an extension time of 1 minute. After gel purification using binding to glass milk (GENECLEAN Spin Kit from Bio101, Carlsbad, Calif., USA), the 1 kB product was used as the "Megaprimer" for the second amplification.

Because there was difficulty amplifying the desired product in this second step, a template containing the genomic *pyrF* region was made by PCR amplification in order to increase the template quantity. HOTSTARTAQ DNA polymerase (from Qiagen, Valencia CA) was used with *P. fluorescens* genomic DNA and the pyrFPstF and pyrFXbaR2 primers. The Megaprimer and the pyrFXbaR2 primer were then used with this template and HOTSTARTAQ polymerase, to amplify the deletion product by PCR, using amplification conditions of 15

min at 95°C, followed by 30 cycles of a 30 sec denaturation at 94°C, 30 sec annealing at 59°C, and a 2 min extension at 72°C, followed by a final step at 72°C for 3 min. The expected 2 kB band was separated from a number of other products by gel electrophoresis, and then gel purified as above and cloned into plasmid pCR2.1Topo (from Invitrogen, Carlsbad CA) according to instructions from the manufacturer, to form pDOW1215-7 Sequencing the PCR-amplified region of pDOW1215-7 showed that there were 3 mutations introduced by the amplification process; all three changes were within 112 bp downstream of the stop codon for *pyrF*. Sequencing through this area was difficult, because the process of the reaction stopped in this area. Analysis by M-FOLD (GCG) of the secondary structure of RNA that would be encoded by this area showed the presence of a very stable stem-loop structure and a run of uridine residues that is characteristic of a rho-independent transcription terminator. None of the mutations occurred in the open reading frame.

pDOW1215-7 was used to delete the chromosomal *pyrF* gene in MB101. To do this, first, electrocompetent *P. fluorescens* cells made according to the procedure of Artiguenave et al. (1997), were transformed with 0.5 μg of the purified plasmid. Transformants were selected by plating on LB medium with kanamycin at 50 μg/mL. This plasmid cannot replicate in *P. fluorescens*, therefore kanamycin resistant colonies result from the plasmid integrating into the chromosome. The site of integration of the plasmid was analyzed by PCR using the HOTSTARTAQ polymerase and primers pryanalF and M13F(-40), annealing at 57°C and with an extension time of 4 min. One out of the 10 isolates (MB101::pDOW1215-7#2) contained an insertion of pDOW1215-7 into the downstream region (2.8 kB analytical product) and in the other nine were in the upstream region (2.1 kb analytical product).

Second, to identify strains that had lost the integrated plasmid by recombination between the homologous regions the following analytical PCR procedure was used: MB101::pDOW1215-7#2 was inoculated from a single colony into LB supplemented with 250 mM uracil, grown overnight, and then plated onto LB-uracil and 500 μg/mL 5-fluoroorotic acid (5-FOA – Zymo Research, Orange CA). Eight colonies were analyzed by PCR with HOTSTARTAQ and primers pyranalF and pyranalR, annealing at 57°C and extending for 4 min. The expected size of the amplified product from the parent MB101 was 3.2 kB, or if the *pyrF* gene was deleted, then 2.5 kB. Each of the colonies gave rise to the 2.5 kB band expected from a deletion of *pyrF*. The first three isolates were purified and named PFG116, PFG117, and PFG118 (also known as DC36). The three isolates exhibit the phenotype expected from a *pyrF* deletion, i.e. they are sensitive to kanamycin, uracil is

required for growth, and they are resistant to 5-FOA. The DNA sequence of PFG118 was identical to that of the amplified regions in pDOW1215-7; i.e. the three mutations in the stem-loop structure immediately downstream from *pyrF* were incorporated into the PFG118 genome, along with the *pyrF* deletion.

Use of the *pyrF* Gene as a Selection Marker in *P. fluorescens* Expression System

The ability of the *pyrF* gene to act as a selectable marker was tested by cloning it into a pMYC expression plasmid containing both an existing tetracycline resistance marker and the target enzyme coding sequence under the control of the tac promoter. For this, the plasmid pMYC5088 was digested at 37°C for 2 hr with *SnaBI* in a 50 uL reaction using NEB Buffer 4 and 0.1 mg/mL of bovine serum albumen (BSA) (from New England Biolabs, Beverly MA).The reaction mixture was then treated at 70°C for 20 min to inactivate the enzyme, then gel-purified as described above. 60 ng of the *SnaB*I-digested pMYC5088 was ligated to 50 ng of the MB214pyrF1 - MB214pyrR1 PCR product using the FAST-LINK DNA Ligation Kit (Epicentre Technologies, Madison WI). After 1 hr at 25°C, the reaction was stopped by treating the mixture at 70°C for 20 min. The result was then transformed into chemically-competent JM109 *E. coli* cells (Promega Corp., Madison WI) using conditions recommended by the manufacturer.

Transformants were selected on LB medium containing tetracycline at 15 μg/mL. Plasmid DNA was prepared from 12 isolates using the QiaPrep Spin Miniprep Kit (Qiagen, Valencia CA) and screened with *NotI* and *EcoR*I, which indicated that one isolate contained the desired clone, pDOW1249-2 (Figure 2). The plasmid pDOW1249-2 was transformed into *pyrF*(-) *P. fluorescens* containing a pCN plasmid containing a *lacI* repressor expression cassette and a kanamycin resistance marker gene. Isolates were tested in shake flasks and in 20-L fermentors.

Isolates were grown in minimal salts medium and kanamycin, but no tetracycline, so that the only selective pressure for the pDOW1249-2 plasmid was provided by the ability of the *pyrF* gene on the plasmid to complement the *pyrF* deletion in the chromosome. As determined by SDS-PAGE analysis, the amount of target protein produced by the new strain in the shake flask test was similar to that of the control strain, a genomically *pyrF*(+) *P. fluorescens* control system containing the same two plasmids, but for the absence of the *pyrF* gene in pDOW1249-2, and grown on the same medium but further supplemented with tetracycline in order to maintain the plasmid (data not shown). Two strains were chosen for

further analysis at the 20-L scale, based on the amount of target protein seen on the SDS-PAGE gel and OD_{575} values in shake flasks. Both strains showed a level of accumulation of target protein within the normal range observed for the control strain (Figure 1).

Example 2

Construction of a *pyrF* – *proC* Dual Auxotrophic Selection Marker System in a *P. fluorescens* Host Cell Expression System

Oligonucleotides Used Herein

proC1
5'-ATAT**GAGCTC**CGACCTTGAGTCGGCCATTG-3' (SEQ ID NO:22)

proC2
5'-ATAT**GAGCTC**GGATCCAGTACGATCAGCAGGTACAG-3' (SEQ ID NO:23)

proC3
5'-AGCAACACGCGTATTGCCTT-3' (SEQ ID NO:24)

proC5
5'-GCCCTTGAGTTGGCACTTCATCG-3' (SEQ ID NO:25)

proC6
5'-GATAAACGCGAAGATCGGCGAGATA-3' (SEQ ID NO:26)

proC7
5'-CCGAGCATGTTTGATTAGACAGGTCCTTATTTCGA-3' (SEQ ID NO:27)

proC8
5'-TGCAACGTGACGCAAGCAGCATCCA-3' (SEQ ID NO:28)

proC9
5'-GGAACGATCAGCACAAGCCATGCTA-3' (SEQ ID NO: 29)

genF2
5'-ATAT**GAGCTC**TGCCGTGATCGAAATCCAGA-3' (SEQ ID NO:30)

genR2
5'-ATAT**GGATCC**CGGCGTTGTGACAATTTACC-3' (SEQ ID NO:31)

XbaNotDraU2 linker
5'-TCTAGAGCGGCCGCGTT-3' (SEQ ID NO: 32)

XbaNotDraL linker
5'-GCGGCCGCTCTAGAAAC-3' (SEQ ID NO: 33)

Cloning of *proC* from *P. fluorescens* and Formation of a pCN Expression Plasmid Containing *proC*

Replacing antibiotic resistant gene in pCN51lacI with proC

The *proC* ORF and about 100 bp of adjacent upstream and downstream sequence was amplified from MB101 genomic DNA using proC1 and proC2, an annealing temperature of 56°C and a 1 min extension. After gel purification of the 1 kB product and digestion with *Sac*I, the fragment was cloned into *Sac*I-digested pDOW1243 (a plasmid derived from pCN51*lacI* by addition of a polylinker and replacement of *kanR* with the gentamycin resistance gene), to create pDOW1264-2. This plasmid was tested in the *proC*(-) mutant strain PFG932 for its ability to regulate amylase synthesis from pDOW1249-2. Expressed target enzyme production levels at the 20-L scale was similar to that of the dual-antibiotic-resistance marker control strain DC88 (data not shown).

The *genR* antibiotic marker gene was then removed from the pDOW1264-2 (Figure 3) to create an antibiotic-marker-free plasmid with *proC* and *lac*I. Removing the *genR* gene was accomplished by restriction digestion of pDOW1264-2 with *BamH*I, purification of the 6.1 kB fragment, ligation to itself, and electroporation into the *proC*(-) *P. fluorescens* host PFG1016. Isolates were checked by restriction digestion using *EcoR*I. The resulting plasmid was named pDOW1306-6. Analytical restriction digests with *EcoRI* and sequencing across the *BamHI* junction verified the identity of the plasmid and the proper orientations of the genes therein.

Sequencing was performed by The Dow Chemical Company. The *proC* sequence is presented within SEQ ID NO:4.

Construction of Target Enzyme Expression Plasmid Containing a *pyrF* Marker in Place of an Antibiotic Resistance Marker

The antibiotic-marker-free production plasmid, pDOW1269-2, containing a target enzyme-encoding gene under control of a *tac* promoter, was constructed by restriction digestion of pDOW1249-2 with *PvuI* to remove the *tetR/tetA* genes. Derived from pMYC5088 by insertion of the *pyrF* gene from MB214, pDOW1249-2 was prepared as described in Example 1. The 10.6 kB *PvuI* fragment was gel-purified, ligated to itself, transformed into PFG118/pCN51*lacI* by electroporation and spread on M9 glucose medium containing kanamycin (to retain the pCN51*lac*I). Plasmid DNA was isolated and analytical restriction digests with *NcoI* were carried out; two isolates showed a restriction digest that

was consistent with the expected bands. Both isolates were sequenced across the *PvuI* junction, which verified the identity of the plasmids and the proper orientations of the genes therein.

Construction of a Pseudomonas fluorescens Strain with Genomic Deletions of pyrF and proC

PFG118, a *P. fluorescens* MB101 strain with a deletion of *pyrF*, was described in Example 1.

Construction of pDOW1261-2, a Vector for Gene Replacement and Deletion

The vector pDOW1261-2 was designed to create clean deletions of genomic DNA, using marker exchange by the cross-in/cross-out method (Toder 1994; Davison 2002), by combining the following properties:

- a ColEI replication origin that functions only in *E. coli* and not in *P. fluorescens*;
- a selectable marker *(tetR/tet*A) for integration of the plasmid into the chromosome;
- a counterselectable marker *(pyrF)* that allows for selection for loss of the inserted plasmid (as long as the host strain is *pyrF*-); cells that lose the *pyrF* gene are resistant to 5-FOA; and
- a blunt-end cloning site, *SrfI,* which has an uncommon 8 bp recognition site - if the desired insert lacks the site, the efficiency of insertion can be increased by adding *SrfI* (Stratagene, La Jolla CA) to the ligation reaction to re-cleave vectors that ligate without an insert.

To construct this vector, a 5 kB *PstI* to *EcoRI* fragment containing the *tet*R, *tet*A, and *pyrF* genes was cloned into pCRScriptCAM (Stratagene, La Jolla CA) that had been digested with *PstI* and *Eco*RI, creating pDOW1261-2.

Construction of a Vector to Delete *proC* from the Chromosome

To construct a deletion of *proC*, the copies of the flanking regions upstream and downstream of the *proC* gene were joined together by PCR, and then cloned into the pDOW1261-2 gene replacement vector. The proC7 primer, which bridges the *proC* ORF,

was designed to delete the entire coding sequence from the ATG start codon to the TAG stop codon. An additional 16 bp downstream of the stop codon was also included in the deletion.

To make the PCR fusion of regions upstream and downstream from *proC*, the Megaprimer method of PCR amplification was used (Barik 1997). To make the megaprimer, the 0.5 kB region directly upstream of the *proC* open reading frame was amplified by PCR from MB214 genomic DNA, using primers proC5 and proC7. Primer proC7 overlaps the regions upstream and downstream of the *proC* ORF. The polymerase chain reaction was carried out with 1 uM of primers, 200 uM each of the four dNTPs, and Herculase high-fidelity polymerase (Stratagene, La Jolla CA) in the buffer recommended by the vendor. Herculase is a high-fidelity enzyme that consists mostly of *Pfu* polymerase, which leaves blunt ends. The amplification program was 95°C for 2 min, 30 cycles of 95°C for 30 sec, 50°C for 30 sec, and 72°C for 1 min per kB, followed by 10 min at 72°C. The amplified products were separated by 1% agarose gel electrophoresis in TBE and visualized using ethidium bromide. A gel slice containing the DNA was cut from the gel and purified as above

The 1.3 kB region downstream from the *proC* gene was amplified using primers proC3 and proC6, to serve as a template for subsequent reactions. The same amplification protocol was used, except for an annealing temperature of 60°C. The reaction was checked on an agarose gel, and then purified using the StrataPrep PCR Purification Kit (Stratagene, La Jolla CA).

In the second step to make the deletion fusion, the megaprimer was used as one of the primers in a PCR reaction along with primer proC6, and with the proC3-proC6 PCR reaction as the template. An annealing temperature of 61°C and extension time of 2 min was used. The 1 kB PCR product was purified and blunt-end ligated into the suicide vector pDOW1261-2 that had been digested with *SrfI*. *SrfI* was included in the ligation in order to decrease background caused by re-ligation of the vector, as according to instructions from the manufacturer (pCRScriptCam Cloning Kit – Stratagene, La Jolla CA). The ligation was transformed into DH10 β (Gibco BRL Life Technologies, now Invitrogen, Carlsbad CA) by electroporation (2 mM gap cuvette, 25 μF, 2.25 kV, 200 Ohms) (Artiguenave et al. 1997), and isolates were screened using the *DraIII* restriction enzyme. The PCR amplified region of each isolate was sequenced by The Dow Chemical Company; isolate pDOW1305-6 was verified as containing the correct genomic DNA sequence.

Formation of the *P. fluorescens pyrF-proC* Double Deletion

To make a doubly deleted strain, PFG118 was transformed with pDOW1305-6 by electroporation as described above. Analytical PCR on the colonies with primers proC8 and the M13/pUC Reverse Sequencing Primer (-48) (which hybridizes to the plasmid only) (New England Biolabs, Beverly MA), using HotStarTaq (Qiagen, Valencia CA), an annealing temperature of 59°C and an extension time of 4 min, showed that 9 out of 22 isolates had the plasmid integrated into the region upstream from *proC*, and 7 out of 22 had the plasmid integrated downstream (data not shown). Three of each orientation were purified to single colonies. The three isolates PFG118::1305-6.1, -6.8, -6.10 have an insertion in the region upstream, and the three isolates PFG118::1305-6.2, -6.3, -6.9 have an insertion in the region downstream.

To select for cells that have carried out a homologous recombination between the plasmid and the chromosome genes thereby leaving a deletion, PFG118::1305-6.1 and -6.2 were grown to stationary phase in 50 mL of LB with uracil and proline supplementation and then plated on LB-5-FOA with uracil and proline supplementation. Cells that lose the integrated plasmid by recombination also lose the *pyrF* gene, and are therefore expected to be resistant to 5-FOA which would otherwise be converted into a toxic compound. PCR analysis with proC8 and proC9 was carried out to distinguish between cells that had lost the plasmid and regenerated the original sequence, and those that had left the deletion. Two isolates with the expected 1.3 kB band were chosen from each of the two selections and named PFG1013, PFG1014, PFG1015 and PFG1016 (also known as DC164). All four isolates were unable to grow on M9 glucose unless both proline and uracil were added, and were tetracycline-sensitive. The genomic region of PFG118 (wild type *proC*) and PFG1016 (*proC* deletion) was amplified by PCR (primers proC8 and proC9, HotStarTaq polymerase, 63°C annealing and 3 min extension) and sequenced. The region between proC5 and proC6 of strain PFG1016 was identical to the parent, except for the expected 835 bp deletion.

Construction of a Dual Auxotrophic Selection Marker Expression System PFG1016/pDOW1306-6 pDOW1269-2

Plasmids were isolated from strain PFG118 pCN51*lacI* pDOW1269-2 by HISPEED Plasmid Midi Kit (Qiagen, Valencia CA). The pDOW1269-2 was partially purified from the pCN51*lacI* by agarose gel electrophoresis and then electroporated into PFG1016 pDOW1306-6 . Transformants were selected on M9/glucose without supplementation.

Because there was a possibility that some of the pCN51*lacI* contaminating the pDOW1269-2 preparation would also be cotransformed into the cells, three isolates from each transformation were tested for sensitivity to kanamycin, the antibiotic marker carried on pCN51*lacI*; all six were found sensitive. All six strains were found to express the target enzyme, in a test of target enzyme activity. PCR analysis showed that all six also contained the chromosomal *proC* deletion.

Restriction digestion of plasmids isolated from the transformants was consistent with the expected pattern.

Performance Testing of the Dual Auxotrophic Marker Expression System in Shake Flasks

The six strains were then tested in shake flasks as described above in Example 1. Induction of target enzyme expression was initiated at 26 hours by addition of IPTG. The OD_{575} for all six strains was comparable to that of the dual-antibiotic-resistance marker expression system control, DC88. Target enzyme production levels in all six were also comparable to that of the control, as assessed by SDS-PAGE. The two strains that achieved the highest OD_{575}, strains 1046 and 1048, were selected for further characterization.

Performance Testing of the Dual Auxotrophic Marker Expression System in 20-L Bioreactors

Strains 1046 and 1048 were subsequently tested in 20-L bioreactors. Induction of target enzyme expression was initiated at 26 hours by addition of IPTG. Both strains achieved performance levels within the normal range for the DC88 control strain, for both OD_{575} and target enzyme activity. The performance averages of these two strains are shown in Figure 1. Restriction digests of plasmids prepared from samples taken at the seed stage and at a time just before the 26-hour start of induction showed a pattern consistent with that expected. Analytical PCR of genomic DNA carried out on the same samples demonstrated the retention of the *proC* deletion and the *pyrF* deletion. Aliquots of the 25 hr samples were plated on tetracycline-, gentamycin-, or kanamycin-supplemented media; no cell growth was observed, thus confirming the absence of antibiotic resistance gene activity.

Analysis of strain 1046 (also known as DC167) in 20-L bioreactors was repeated twice with similar results. Plasmid stability at the seed stage and after 25 hours of fermentation (immediately before induction) was tested by replica plating from samples that

had been diluted and plated on complete media. Both plasmids were present in more than 97% of the colonies examined, indicating the lack of cross feeding revertants able to survive without the plasmid and the stable maintenance of the expression vector in Pseudomonas fluorescens.

Results

Both of the *pyrF* expression systems performed as well as the control system in which only antibiotic resistance markers were used (Figure 1). For the control strain, there is no negative effect of cross-feeding, since any importation of exogenous metabolites from lysed cells does not decrease or remove the selection pressures provided by the antibiotics in the medium. The expected decreases in performance expected as a result of cross-feeding on the two *pyrF* expression systems were surprisingly not observed.

Example 3

Chromosomal Integration of *lacI*, $lacI^Q$ and $lacI^{Q1}$ in *P. fluorescens*

Three *P. fluorescens* strains have been constructed, each with one of three different *Escherichia coli lacI* alleles, *lacI* (SEQ ID NO:9), $lacI^Q$ (SEQ ID NO:11), and $lacI^{Q1}$ (SEQ ID NO:12), integrated into the chromosome. The three strains exhibit differing amounts of LacI repressor accumulation. Each strain carries a single copy of its *lacI* gene at the levansucrase locus (SEQ ID NO:13) of *P. fluorescens* DC36, which is an MB101 derivative (see TD Landry *et al.*, "Safety evaluation of an α–amylase enzyme preparation derived from the archaeal order Thermococcales as expressed in *Pseudomonas fluorescens* biovar I," *Regulatory Toxicology and Pharmacology* 37(1): 149-168(2003)) formed by deleting the *pyrF* gene thereof, as described above.

No vector or other foreign DNA sequences remain in the strains. The strains are antibiotic-resistance-gene free and also contain a *pyrF* deletion, permitting maintenance, during growth in uracil un-supplemented media, of an expression plasmid carrying a *pyrF*+ gene. Protein production is completely free of antibiotic resistance genes and antibiotics.

MB214 contains the *lacI-lacZYA* chromosomal insert described in U.S. Pat. 5,169,760. MB214 also contains a duplication in the C-terminus of the LacI protein, adding about 10 kDa to the molecular weight of the LacI repressor.

Construction of Vector pDOW1266-1 for Insertion of Genes into the Levansucrase Locus

Plasmid pDOW1266-1 was constructed by PCR amplification of the region upstream of and within the *P. fluorescens* gene for levansucrase (SEQ ID NO:13), replacing the start codon with an *Xba*I site, using the Megaprimer method, see A Barik, "Mutagenesis and Gene Fusion by Megaprimer PCR," in BA White, *PCR Cloning Protocols* 173-182 (1997) (Humana). PCR was performed using HERCULASE polymerase (Stratagene, Madison WI, USA) using primers *LEV1* (SEQ ID NO:34) and *LEV2* (SEQ ID NO:35), and *P. fluorescens* MB214 genomic DNA as a template (see below for oligonucleotide sequences). Primer *LEV2* (SEQ ID NO:35) contains the sequence that inserts an *Xba*I site. The reaction was carried out at 95°C for 2 min, 35 cycles of [95°C for 30 sec, 58°C for 30 sec, 72°C for 1 min],followed by 10 min at 72°C. The 1 kB product was gel purified and used as one of the primers in the next reaction, along with *LEV3* (SEQ ID NO:36), using MB214 genomic DNA as a template and the same conditions except that extension time was 2 min. The 2 kB product was gel purified and re-amplified with *LEV2* (SEQ ID NO: 35) and *LEV3* (SEQ ID NO. 36) in order to increase the quantity.

Oligonucleotides used

LEV1 (SEQ ID NO:34)

5'-TTCGAAGGGGTGCTTTTTCTAGAAGTAAGTCTCGTCCATGA

LEV2 (SEQ ID NO:35)

5'-CGCAAGGTCAGGTACAACAC

LEV3 (SEQ ID NO:36)

5'-TACCAGACCAGAGCCGTTCA

LEV7 (SEQ ID NO:37)

5'-CTACCCAGAACGAAGATCAG

LEV8 (SEQ ID NO:38)

5'-GACTCAACTCAATGGTGCAGG

BglXbaLacF (SEQ ID NO:39)

5'-AGATCTCTAGAGAAGGCGAAGCGGCATGCATTTACG

lacIR4 (SEQ ID NO:40)

5'-ATATTCTAGAGACAACTCGCGCTAACTTACATTAATTGC

Lacpro9 (SEQ ID NO:41)

5'-ATATTCTAGAATGGTGCAAAACCTTTCGCGGTATGGCATGA

LacIQF (SEQ ID NO:42)

5'-GCTCTAGAAGCGGCATGCATTTACGTTGACACC

LacINXR (SEQ ID NO:43)

5'-AGCTAGCTCTAGAAAGTTGGGTAACGCCAGGGT

lacIQ1 (SEQ ID NO:44)

5'-AGTAAGCGGCCGCAGCGGCATGCATTTACGTTGACACCACCTT
TCGCGGTATGGCATG

The Oligos Below were Used for Analytical Sequencing Only

lacIF1 (SEQ ID NO:45)

5'-ACAATCTTCTCGCGCAACGC

lacIF2 (SEQ ID NO:46)

5'-ATGTTATATCCCGCCGTTAA

lacIR1 (SEQ ID NO:47)

5'-CCGCTATCGGCTGAATTTGA

lacIR2 (SEQ ID NO:48)
5'-TGTAATTCAGCTCCGCCATC

SeqLev5AS (SEQ ID NO:49)
5'-TATCGAGATGCTGCAGCCTC

SeqLev3S (SEQ ID NO:50)
5'-ACACCTTCACCTACGCCGAC

LEV10 (SEQ ID NO:51)
5'-TCTACTTCGCCTTGCTCGTT

The *LEV2 - LEV3* amplification product was cloned into the *SrfI* site of pDOW1261-2, a suicide vector that contains *P. fluorescens pyrF*+ gene as a selection marker to facilitate selection for cross-outs. The new plasmid was named pDOW1266-1. The amplified region was sequenced.

Cloning the *lacI* Genes into Insertion Vector pDOW1266-1

The *E.coli lacI* gene was amplified from pCN51lacI with primers *BglXbaLacF* (SEQ ID 10 NO:39) and *lacIR4* (SEQ ID NO. 40), using HERCULASE polymerase (annealing at 62°C and extension time of 2 min). After gel purification and digestion with *Xba*I, the *lacI* gene was cloned into the *Xba*I site of pDOW1266-1, to make pDOW1310. The $lacI^Q$ gene was created by PCR amplification using pCN51lacI as a template with 15 primers *lacpro9* (SEQ ID NO. 41) and *lacIR4* (SEQ ID NO. 40), using HERCULASE polymerase (annealing at 62°C and extension time of 2 min). After gel purification and digestion with *Xba*I, it was cloned into the *Xba*I site of pDOW1266-1, to make pDOW1311.

The $lacI^{Q1}$ gene was created by amplifying the *lacI* gene from *E. coli* K12 (ATCC47076) using primers *lacIQ1* (SEQ ID NO. 44) and *lacINXR* (SEQ ID NO. 43) and cloned into pCR2.1Topo (Invitrogen, Carlsbad, CA, USA), to make pCR2-lacIQ1. The $lacI^{Q1}$ gene was reamplified from pCR2-lacIQ1 using primers *lacIQF* (SEQ ID NO. 42) and *lacINXR* (SEQ ID NO. 43) with Herculase polymerase (61°C annealing, 3 min extension time,

35 cycles). After gel purification and digestion with *Xba*I, the PCR product was cloned into the *Xba*I site of pDOW1266-1, to make pDOW1309.

The PCR amplified inserts in pCR2-lacIQ1, pDOW1310, pDOW1311, and pDOW1309 were sequenced (using primers *lacIF1* (SEQ ID NO:45), *lacIF2* (SEQ ID NO. 46), *lacIR1* (SEQ ID NO. 47), *lacIR2* (SEQ ID NO. 48), *SeqLev5AS* (SEQ ID NO. 49), *SeqLev3S* (SEQ ID NO. 50), and *LEV10* (SEQ ID NO. 51)) to insure that no mutations had been introduced by the PCR reaction. In each case, an orientation was chosen in which the *lacI* was transcribed in the same direction as the levansucrase gene. Although the levansucrase promoter is potentially able to control transcription of *lacI*, the promoter would only be active in the presence of sucrose, which is absent in the fermentation conditions used.

Construction of P. fluorescens Strains with Integrated lacI Genes at the Levansucrase Locus

The vectors pDOW1309, pDOW1310, and pDOW1311 were introduced into DC36 by electroporation, first selecting for integration of the vector into the genome with tetracycline resistance. Colonies were screened to determine that the vector had integrated at the levansucrase locus by PCR with primers *LEV7* (SEQ ID NO. 37) and *M13R* (from New England Biolabs, Gloucester MA, USA). To select for the second cross-over which would leave the *lacI* gene in the genome, the isolates were grown in the presence of 5'-fluoroorotic acid and in the absence of tetracycline. Recombination between the duplicated regions in the genome either restores the parental genotype, or leaves the *lacI* gene. The resulting isolates were screened for sensitivity to tetracycline, growth in the absence of uracil, and by PCR with primers *LEV7* (SEQ ID NO. 37) and *LEV8* (SEQ ID NO. 38). The names of the new strains are shown in Table 17. To obtain sequence information for genomic regions, PCR products were sequenced directly, see E Werle, "Direct sequencing of polymerase chain reaction products," *Laboratory Methods for the Detection of Mutations and Polymorphisms in DNA* 163-174 (1997). For each strain, the sequencing confirmed the identity of the promoter, the orientation of the *lacI* variant relative to the flanking regions, and whether there were any point mutations relative to the parental sequence. The sequences of DC202 and DC206 were as expected. The sequence of DC204 showed a point mutation within the levansucrase open reading frame, downstream of $lacI^Q$, which did not change any coding sequence and therefore is inconsequential.

TABLE 17. *P. FLUORESCENS* STRAINS WITH *LACI* ALLELES INTEGRATED INTO THE GENOME

Strain Designation	Plasmid used to make the *lacI* insertion	Genotype
DC202	pDOW1310-1	*pyrF lev::lacI*
DC204	pDOW1311-4	*pyrF lev::lacI*^{Q}
DC206	pDOW1309oriA	*pyrF lev::lacI*^{Q1}

Analysis of Relative Concentration of LacI in the lacI Integrants, Compared to pCN51lacI

UnBlot is a method analogous to Western analysis, in which proteins are detected in the gel without the need for transfer to a filter. The technique was carried out following the directions from Pierce Biotechnology (Rockford, IL, USA), the manufacturer. Analysis using UnBlot showed that the amount of LacI in each of the new integrant strains was higher than in MB214. MB214 contains the *lacI-lacZYA* insert described in U.S. Pat. 5,169,760. The relative concentration of LacI in the *lacI*^{Q} and *lacI*^{Q1} integrants was about the same as in strains carrying pCN51lacI, the multi-copy plasmid containing *lacI*. See Figure 5.

A dilution series was carried out in order to asscss more precisely the relative difference in LacI concentration in MB214, DC202 (*lacI* integrated) and DC206 (*lacI*^{Q1} integrated). MB101pCN51lacI, DC204 and DC206 have about 100 times more LacI than MB214, whereas DC202 has about 5 times more.

Example 4

Nitrilase Expression and Transcription

Strain DC140 was constructed by introducing into *P. fluorescens* MB214 a tetracycline- resistant broad-host-range plasmid, pMYC1803 (WO 2003/068926), into which a nitrilase gene (G DeSanthis *et al., J. Amer. Chem. Soc.* 125:11476-77 (2003)), under the control of the Ptac promoter, had been inserted. In order to compare regulation of un-induced expression of the target gene in DC202 and DC206 with MB214, the same nitrilase gene was cloned onto a pMYC1803 derivative where the tetracycline-resistance gene has been replaced by a *pyrF* selection marker. The new plasmid, pDOW2415, was then electroporated into DC202 and DC204, resulting in DC239 and DC240, respectively. DC140, DC239 and

DC240 were cultured in 20 L fermentors by growth in a mineral salts medium fed with glucose or glycerol, ultimately to cell densities providing biomasses within the range of about 20 g/L to more than 70 g/L dry cell weight (See WO 2003/068926). The gratuitous inducer of the Ptac promoter, IPTG, was added to induce expression.

The ratio of pre-induction nitrilase activities of DC140 to DC239 toDC240 was 6:2:1. RNA analysis by Northern blots of the same samples revealed the same ranking of de-repression. Based on densitometric measurements, the ratio of un-induced transcript levels of DC140:DC239:DC240 was 2.4:1.4:1.0. Shortly after induction (30 min) with 0.3mM IPTG, the levels of transcript of all the strains were the same. Post-induction nitrilase productivity rates were also comparable. This indicated that the concentration of inducer used was sufficient to fully induce the Ptac promoter in these three strains despite their different LacI protein levels. However, fermentations of the most derepressed strain, DC140, suffered significant cell lysis accompanied with loss of nitrilase activity after approximately 24 hours post-induction. Induction of the improved, more tightly regulated strains, DC239 and DC240, could be extended to more than 48 hours post induction, while maintaining high nitrilase productivity, with the ultimate result of a doubling of nitrilase yields. See Figure 6.

Results

The examples indicate It that the use of a LacI-encoding gene other than as part of a whole or truncated Plac-*lacI-lacZYA* operon in Pseudomonads resulted in substantially improved repression of pre-induction recombinant protein expression, higher cell densities in commercial-scale fermentation, and higher yields of the desired product in comparison with previously taught *lacI-lacZYA* Pseudomonad chromosomal insertion (U.S. Pat. No. 5,169,760). The results also indicated that the *lacI* insertion is as effective in producing LacI repressor protein in *Pseudomonas fluorescens*, thereby eliminating the need to maintain a separate plasmid encoding a LacI repressor protein in the cell and reducing potential production inefficiencies caused by such maintenance.

In addition to being antibiotic free, derepression during the growth stage in DC239 and DC240 was up to 10 fold less than the MB214 strain. Pre-induction nitrilase activity levels of DC239 and DC240 averaged 0.4 U/ml in more than 13 separate fermentations, and cell density and nitrilase expression in DC239 and DC240 did not decay during extended induction phase, as it did in DC140. Given the higher derepression, DC239 and DC240

fermentation runs decreased the time of the growth phase by more than 30%, reducing fermentation costs.

Example 5

Construction of *tac* promoter with a single optimal *lac* operator and with two *lac* operators

The native *tac* promoter only has a single native *lac* operator, AATTGTGAGCGGATAACAATT, at the O1 position (Figure 4). In the first construct, pDOW1418, the native operator was replaced by the more symmetrical *lacOid* operator sequence 5'-AATTGTGAGC GCTCACAATT – 3' (SEQ. ID. NO. 14) (JR. Sadler, H. Sasmor and JL. Betz. PNAS. 1983 Nov; 80 (22): 6785-9). A 289 bp *HindIII/ SpeI* fragment containing the *tac* promoter and the native *lacO* sequence was removed from a pMYC1803 derivative, pDOW2118, and replaced by a *HindIII/SpeI* fragment isolated from an SOE PCR amplification product containing the symmetrical *lacOid* sequence. The SOE PCR primers (RC-3 and RC-9) incorporated 4 nucleotide changes that produced the optimized/symmetrical *lacO* sequence (three base pair substitutions and one base pair deletion). The *HindIII/SpeI* promoter fragment of the resulting plasmid, pDOW2201, was cloned into the nitrilase expression plasmid based on pMYC1803, to replace the native *tac* promoter, resulting in pDOW1414. This expression cassette was then transferred onto the *pyrF(+)* plasmid pDOW1269, resulting in pDOW1418 by exchanging *DraI/XhoI* fragments. Plasmid pDOW1418 was then transformed into host strain DC206, resulting in strain DC281 (See Figure 4).

Oligonucleotides used

RC-3 (SEQ ID NO:52)

5'- GTGAGCGCTCACAATTCCACACAGGAAAACAG

RC-4 (SEQ ID NO:53)

5'- TTCGGGTGGAAGTCCAGGTAGTTGGCGGTGTA

RC-9 (SEQ ID NO:54)

5'- GAATTGTGAGCGCTCACAATTCCACACATTATACGAGC

RC-10 (SEQ ID NO:55)

5'- ATTCAGCGCATGTTCAACGG

In the second construct, pDOW1416, the *lacOid* operator, 5'-AATTGTGAGC GCTCACAATT-3' (SEQ ID. No. 14), was inserted 52 nucleotides up-stream (5') of the existing native lacO1 by PCR. PCR amplification of the promoter region using the Megaprimer method was performed using a pMYC1803 derivative, pMYC5088, and the following primers AKB-1 and AKB-2 as a first step. The resulting PCR product was combined with primer AKB-3 in a second round of PCR amplification using the same template. After purification and digestion with *HindIII* and *SpeI*, the promoter fragment containing the dual operators was cloned into the *HindIII* and *SpeI* sites of plasmid pMYC5088 resulting in pDOW1411. Introduction of the second operator introduced a unique *MfeI* site immediately upstream of the optimal operator. The *XhoI/SpeI* vector fragment with promoter regions of pDOW1411 was then ligated with the compatible fragment of the pMYC1803 derivative bearing the nitrilase gene, forming pDOW1413. Subsequent ligation of the *MfeI/XhoI* expression cassette fragment of pDOW1413 to the compatible vector fragment of pDOW1269 resulted in pDOW1416; which when transformed into DC206, formed the strain DC262.

Oligonucleotides used

AKB-1 (SEQ ID NO:56)

5'-

ACGGTTCTGGCAAACAATTGTGAGCGCTCACAATTTATTCTGAAATGAGC

AKB-2 (SEQ ID NO:57)

5'- GCGTGGGCGGTGTTTATCATGTTC

AKB-3 (SEQ ID NO:58)

5'- TACTGCACGCACAAGCCTGAACA

Nitrilase Derepression

97

Northern blot analysis was performed pre and post induction on MB214, DC202, and DC206. MB214, DC202, and DC206 were transformed with a nitrilase expression vector containing the wild type *lacO* sequence in the O_1 position 3' of the tac promoter, creating MB214 wtO_1, DC202wtO_1 (DC239), and DC206wtO_1 (DC240), as described above. DC206 was transformed with a nitrilase expression vector containing a *lacOid* sequence in place of the wild type *lacO* sequence at the O_1 position 3' of the tac promoter as described above, creating DC206Oid (DC281). DC206 was also transformed with a nitrilase expression vector containing a wild type *lacO* sequence at the O_1 position 3' of the tac promoter and a *lacOid* sequence at the O_3 position 5' of the tac promoter, creating the dual *lacO* containing DC206wtO_1 O_3id (DC263).

Northern blot analysis indicated a greater repression by the strain containing the Dual *lacO* sequence (DC206wtO_1 O_3id (DC263)) cassette prior to induction. The greater repression of pre-induction expression is especially useful when producing toxic proteins, since basal levels of pre-induction toxic proteins result in the delayed entry of the cell into the growth phase, and, potentially, lower overall yields of the product.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. An auxotrophic *Pseudomonas fluorescens* host cell for use in a bacterial expression system, wherein said auxotrophic host cell comprises:
(a) a chromosomal *lacI* gene insert;
(b) a first nucleotide sequence encoding a recombinant polypeptide, wherein said first nucleotide sequence is operably linked to a promoter capable of directing expression of the first nucleotide sequence in said auxotrophic host cell; and
(c) a second nucleotide sequence encoding an auxotrophic selection marker, wherein the auxotrophic selection marker is at least one polypeptide that restores prototrophy to the auxotrophic host cell.

2. A process for producing a recombinant polypeptide comprising:
(a) expressing a first nucleotide sequence encoding the recombinant polypeptide in a *Pseudomonad* cell, wherein said cell comprises a chromosomal *lacI* gene insert, and wherein said cell has been genetically modified to be auxotrophic for at least one metabolite;
(b) expressing a second nucleotide sequence encoding an auxotrophic selection marker, wherein the auxotrophic selection marker is a polypeptide that restores prototrophy to the auxotrophic cell; and
(c) growing the cell on a medium that lacks the metabolite for which the cell is auxotrophic, wherein the cell is grown to a density of about 20 g/L or more.

3. A process for the production of a recombinant polypeptide in the absence of antibiotics comprising:
(a) selecting a *Pseudomonad* cell, wherein said cell comprises a chromosomal *lacI* gene insert, and wherein

the cell has been genetically modified to induce an auxotrophy for at least one metabolite;

(b) introducing into the cell a nucleic acid construct comprising:

i. a first nucleotide sequence encoding the recombinant polypeptide, wherein said first nucleotide sequence is operably linked to a promoter capable of directing expression of the first nucleotide sequence in said cell; and

ii. a second nucleotide sequence encoding an auxotrophic selection marker, wherein the auxotrophic selection marker is a polypeptide that restores prototrophy to the cell;

(c) expressing the recombinant polypeptide and the prototrophy restoring polypeptide in the cell; and

(d) growing the cell on a medium that lacks the metabolite for which the cell is auxotrophic, wherein the cell is grown to a density of about 20 g/L or more.

4. The process of claim 2 or claim 3, wherein the Pseudomonad is *Pseudomonas fluorescens*.

5. The cell or process of any one of claims 1 to 4, wherein the cell is auxotrophic for more than one metabolite.

6. The cell or process of any one of claims 1 to 5, wherein the cell is auxotrophic for uracil or proline or both uracil and proline.

7. The cell or process of any one of claims 1 to 6, wherein the auxotrophic selection marker is an enzyme active in the biosynthesis of a metabolite required for cell survival.

8. The cell or process of claim 7, wherein the enzyme is orotodine-5'-phosphate decarboxylase or Δ^1-pyrroline-5-carboxylate reductase.

9. The cell or process of claim 8, wherein orotodine-5'-phosphate decarboxylase is encoded by the nucleic acid sequence of SEQ ID NO: 1 or SEQ ID NO: 3, or comprises the amino acid sequence of SEQ ID NO: 2.

10. The cell or process of claim 8, wherein Δ^1-pyrroline-5-carboxylate reductase is encoded by the nucleic acid sequence of SEQ ID NO: 6 or SEQ ID NO: 8, or comprises the amino acid sequence of SEQ ID NO: 7.

11. The cell or process of any one of claims 1 to 10, wherein the cell comprises a *pyrF* gene knockout and/or a *proC* gene knockout.

12. The cell or process of claim 11, wherein the *pyrF* gene comprises the nucleic acid of SEQ ID NO: 1 or SEQ ID NO: 3, and/or the *proC* gene comprises the nucleic acid of SEQ ID NO: 6 or SEQ ID NO: 8.

13. The cell or process of any one of claims 1 to 12, wherein said cell lacks a nucleic acid sequence encoding an antibiotic selection marker and wherein said cell is grown in the absence of an antibiotic.

14. The cell or process of any one of claims 1 to 13, wherein the *lacI* gene is other than as part of *PlacI-lacI-lacZYA* operon.

15. The cell or process of any one of claims 1 to 14, wherein the *lacI* gene is selected from the group consisting of *lacI*, $lacI^Q$ and $lacI^{Q1}$.

16. The cell or process of any one of claims 1 to 15, wherein the first nucleotide sequence further comprises at least one *lac* operator sequence.

17. The cell or process of claim 16, wherein the first nucleotide sequence comprises a first lac operator sequence located 5' of a promoter and a second lac operator sequence located 5' of said promoter.

18. The cell or process of claim 16, wherein the at least one lac operator sequence is a *lacOid* sequence.

19. The cell or process of claim 18, wherein the *lacOid* sequence is selected from the group consisting of SEQ ID NO: 14 and SEQ ID NO: 59.

20. The cell or process of any one of claims 1 to 19, wherein the *lacI* gene is inserted in a levansucrase locus.

21. A process for the production of a recombinant polypeptide in the absence of antibiotics comprising:
(a) selecting a *Pseudomonad* cell, wherein said cell comprises a chromosomal *lacI* gene insert, and wherein the cell has been genetically modified to induce an auxotrophy for at least one metabolite;
(b) introducing into the cell a nucleic acid construct comprising:
i. a first nucleotide sequence encoding the recombinant polypeptide, wherein said first nucleotide sequence is operably linked to a promoter capable of directing expression of the first nucleotide sequence in said cell; and
ii. a second nucleotide sequence encoding an auxotrophic selection marker, wherein the auxotrophic selection marker is a polypeptide that restores prototrophy to the cell;

(c) expressing the recombinant polypeptide and the prototrophy restoring polypeptide in the cell; and

(d) growing the cell on a medium that lacks the metabolite for which the cell is auxotrophic, wherein the cell is grown to a density of about 20 g/L or more, wherein said cell density is at least about 25 g/L, 30 g/L, 35 g/L, 40 g/L, 45 g/L, 50 g/L, 60 g/L, 70 g/L, 80 g/L, 90 g/L, 100 g/L, 110 g/L, 120 g/L, 130 g/L, 140 g/L, or 150 g/L.

22. A cell according to claim 1 or a process according to any one of claims 2, 3 or 21, substantially as hereinbefore described, with reference to any one of the examples, and, or figures.

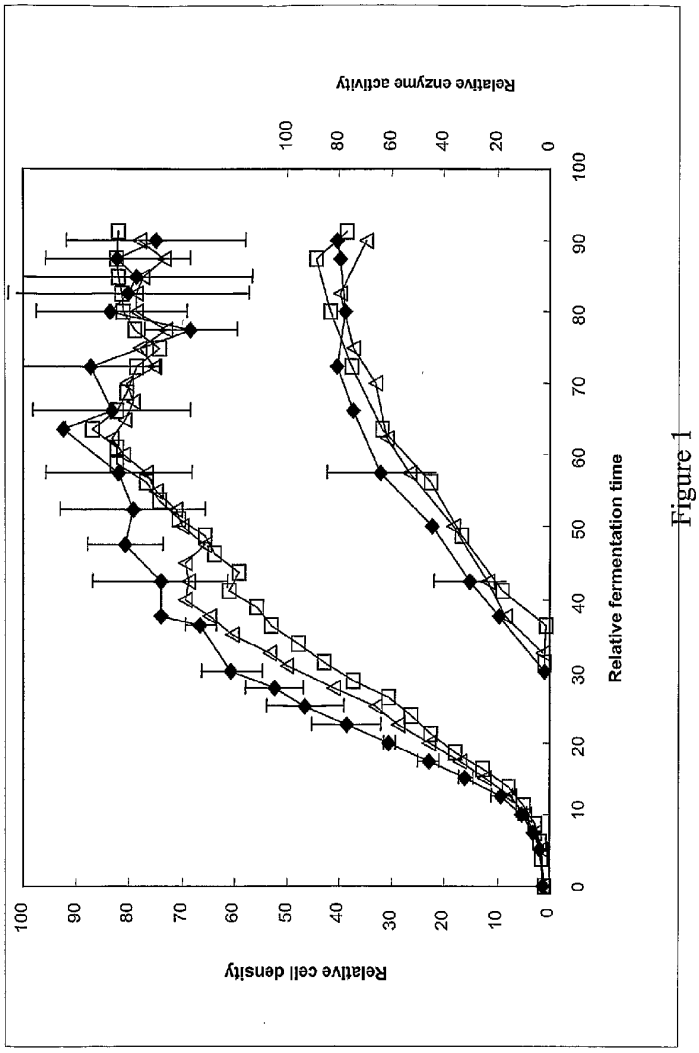
Relative cell density
Relative enzyme activity
Relative fermentation time
0
10
20
30
40
50
60
70
80
90
100

Figure 1

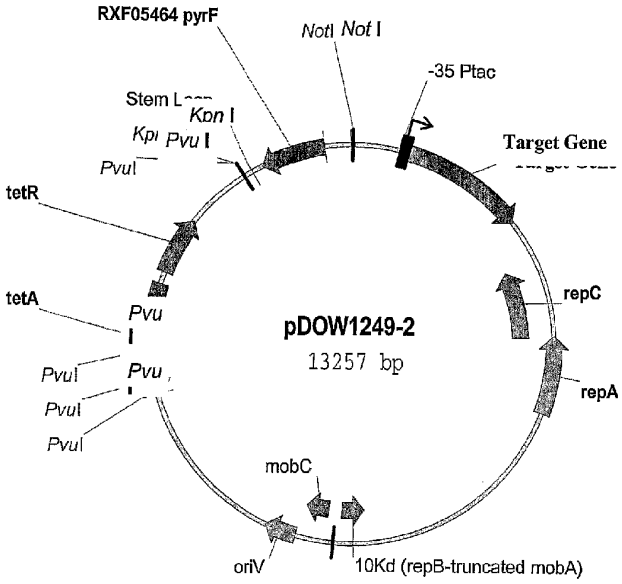
RXF05464 pyrF
NotI Not I
-35 Ptac
Stem L
Kpn I
Kpi Pvu I
PvuI
Target Gene
tetR
tetA
Pvu I
pDOW1249-2
13257 bp
repC
repA
PvuI
Pvu
PvuI
PvuI
mobC
oriV
10Kd (repB-truncated mobA)

Figure 2

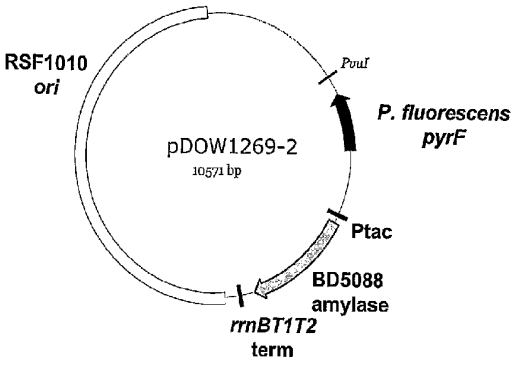
RSF1010
ori
PoxtF
P. fluorescens
pyrF
pDOW1269-2
10571 bp
Ptac
BD5088
amylase
rrnBT1T2
term

Figure 3

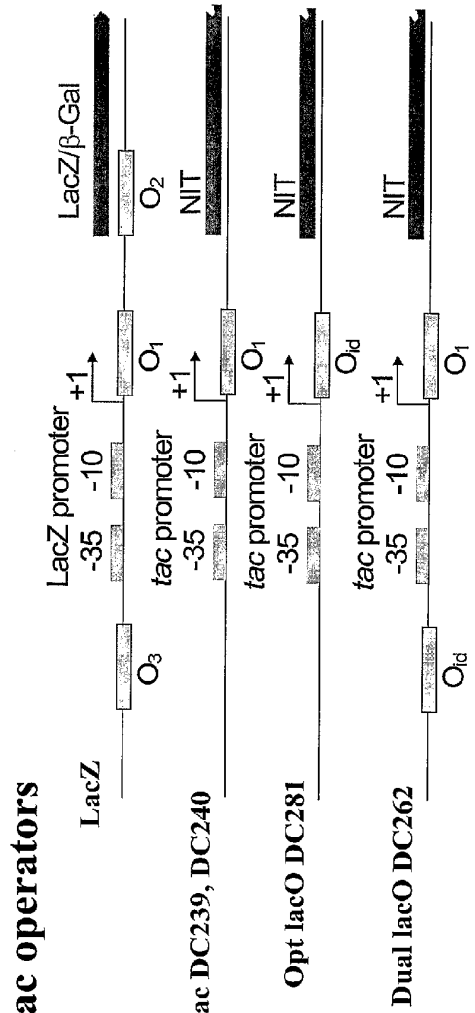
Lac operators
LacZ
LacZ promoter
-35 -10
+1
LacZ/β-Gal
O3
O1
O2
tac DC239, DC240
tac promoter
-35 -10
+1
O1
NIT
Opt lacO DC281
tac promoter
-35 -10
+1
Oid
NIT
Dual lacO DC262
tac promoter
-35 -10
+1
O1
NIT
Oid
wtO1 5′ AATTGTGAGCGGATAACAATT 3′
Oid 5′ AATTGTGAGC GCTCACAATT 3′

Figure 4

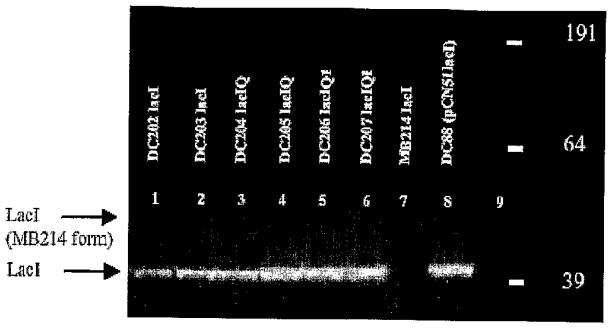
DC202 lacI
DC203 lacI
DC204 lacIQ
DC205 lacIQ
DC206 lacIQ1
DC207 lacIQ1
MB214 lacI
DC88 (pCN51lacI)
1
2
3
4
5
6
7
8
9
191
64
39
LacI
(MB214 form)
LacI

Figure 5

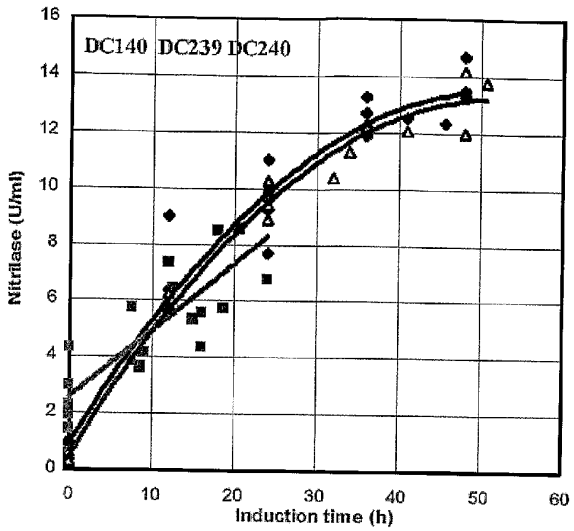
DC140 DC239 DC240
Nitrilase (U/ml)
Induction time (h)
0
2
4
6
8
10
12
14
16
0
10
20
30
40
50
60

Figure 6