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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 January 2004 (15.01.2004)

PCT

(10) International Publication Number
WO 2004/005221 A2

(51) International Patent Classification[7]: C07C

(21) International Application Number:
PCT/US2003/020840

(22) International Filing Date: 3 July 2003 (03.07.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/393,422 3 July 2002 (03.07.2002) 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)*: **RETALLACK, Diane, M.** [US/US]; 13420 Carriage Road, Poway, CA 92064 (US). **SUBRAMANIAN, Venkiteswaran** [US/US]; 3980 Corte Mar De Hierba, San Diego, CA 92130 (US).

(74) Agents: **SCOTT, Mark, S.** et al.; The Dow Chemical Company, Intellectual Property, P.O. Box 1967, Midland, MI 48641-1967 (US).

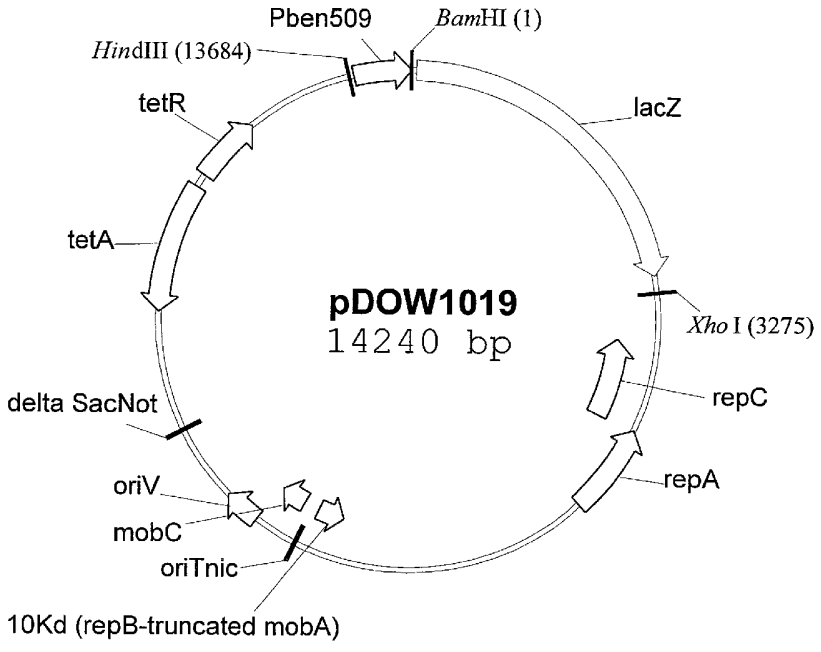
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU,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

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

Published:
— *without international search report and to be republished upon receipt of that report*

[Continued on next page]

(54) Title: BENZOATE- AND ANTHRANILATE-INDUCIBLE PROMOTERS



(57) Abstract: Novel benzoate- or anthranilate-inducible promoters, and novel tandem promoters, and variants and improved mutants thereof, useful for commercial prokaryotic fermentation systems, nucleic acid constructs containing the promoters, expression systems using them, methods for expressing proteins by use thereof, and proteins expressed thereby.

WO 2004/005221 A2

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.

BENZOATE- AND ANTHRANILATE-INDUCIBLE PROMOTERS

BACKGROUND

Gene expression in bacteria, as in any organism, requires that a promoter be present in the regulatory region located 5' (*i.e.* upstream) from the coding sequence in order to direct the gene's transcription. Promoters are classified as constitutive promoters and regulated promoters. In commercially useful bacterial expression systems, regulated promoters have proven particularly useful because they permit increase in the organismal biomass while a desired gene(s) is inactive. This allows the host organism to devote maximal energy to cell division and growth. When the regulated promoter is then activated/induced, more cells will be available to express the desired gene(s), thereby increasing the yield of the desired gene product(s).

Regulated promoters include: (1) activatable promoters, *i.e.* promoters that are inactive until an activator protein binds to the 5' regulatory region; and (2) repressible promoters, *i.e.* promoters that are inactive while the 5' regulatory region is bound by a repressor protein. Some genes or operons are regulated by more than one mechanism. For example, some bacterial genes and operons are subject to both a first, activation or derepression regulatory mechanism, and a second regulatory mechanism, called "catabolite repression." Catabolite repression, also called "glucose catabolite repression" or "carbon catabolite repression," is a phenomenon in which gene(s) under the control of a regulated promoter are also maintained in an unexpressed state until the concentration of glucose (the primary carbon source) falls below a threshold level, *e.g.*, until conditions of glucose starvation. In other words, such a gene(s) cannot be expressed until two conditions are met: 1) glucose reduction/starvation and 2) activation or derepression of the regulated promoter. The occurrence of only one or the other condition is not sufficient to achieve expression of such gene(s). Among the genes and operons that have been found subject to catabolite repression are many that encode enzymes and/or pathways needed to utilize non-glucose carbon sources, *i.e.* alternative carbon sources.

The mechanism by which catabolite repression is effected is still undergoing intense scrutiny. In the case of some catabolite-repressed operons in *E. coli*, a transcriptional level

of control has been assigned, in which catabolite repression is overcome by an "activatable promoter" mechanism. For example, the *E. coli* lactose operon (*lacZYA*) is maintained in an untranscribable state until glucose starvation permits a "catabolite activator protein" to bind to the operon's 5' regulatory region; then, when lactose is present, Lac repressor protein is removed from a separate site(s) (the lac operator(s)) in the 5' regulatory region, causing derepression, and transcription is initiated. Both conditions, *i.e.* both glucose starvation and the presence of lactose, are required for formation of lac operon-encoded mRNA in *E. coli*.

In some cases, post-transcriptional controls are suspected. For example, there is evidence that, in Pseudomonads and closely related species, catabolite repression involving the *crc* gene is mediated post-transcriptionally. This is seen from studies of the regulation of *bdkR* [Ref. 7]. The *bdkR* protein, a transcriptional activator, is involved in the regulation of expression of branched-chain keto acid dehydrogenase in *Pseudomonas putida*. The data presented show that, in rich media, there is no *bdkR* protein detectable in wild type *P. putida*, despite the presence of *bdkR* transcripts. However, in a mutant *P. putida* in which *crc* is impaired or inactivated, *bkdR* protein is detected, *bdkR* transcript levels are slightly lower than those found in the wild type strain, and the transcript of the *bdkR*-regulated gene, *bdkA,* is induced about four-fold. Moreover, mutations identified in mutants in which the catabolite repression of *bdkR* is overcome, have been mapped to the *crc* gene, or to its cognate gene, *vacB*. In *Shigella flexneri*, the *vacB* protein regulates virulence genes post-transcriptionally; this presents additional, although circumstantial, evidence that *crc* acts post-transcriptionally [Ref. 13].

In commercial, prokaryotic systems, one of the key technological challenges associated with the production of proteins and chemicals by fermentation is total control of the transgene expression. The promoter selected for use in expressing the transgene of interest should have the following qualities. It should:

1. Separate growth from reaction;
2. Control gene expression for efficient/maximum product yield;
3. Induce the gene of interest at low/no cost; and
4. Allow no significant level of transcription in the repressed or non-induced state.

For these reasons, regulated promoters are relied upon extensively. In particular, the lac promoter (*i.e.* the lacZ promoter) and its derivatives, especially the tac and trc promoters described in U.S. Patent No. 4,551,433 to DeBoer, and the related promoters listed in Table 1, are relied upon almost exclusively.

Table 1. Commercial lac Promoters, Derivatives & Relatives

Promoter	Commercial Inducer	Bacterial Host Cell(s)	Reference(s)
Ptac16	IPTG	*E.coli*, Pseudomonads	3, 4
Ptac17	IPTG	*E.coli*, Pseudomonads	3
PlacUV5	IPTG	*E.coli*, Pseudomonads	3
Plac	IPTG	*E. coli*	3, 4
Plac(down)	IPTG	*E. coli*	3
T7	IPTG	*E.coli*, Pseudomonads	3, 4

In a typical commercial, bacterial fermentation system, the host cell contains a construct in which a tac promoter is operably attached to a gene or operon whose expression is desired. The *lacI* gene, which is a constitutively expressed gene that encodes the Lac repressor protein which binds to the lac operator, is also included in the bacterial host cell (multiple copies of the lacI gene are usually included therein). After growth of a desired quantity or density of biomass, an inducer is added to the culture in order to derepress the tac promoter and permit expression of the desired gene or operon.

In commercial fermentation systems using a lac-type promoter, such as the tac promoter, the gratuitous inducer, IPTG (isopropyl-β-D-1-thiogalactopyranoside, also called "isopropylthiogalactoside"), is almost universally employed. However, IPTG is expensive and must be carefully controlled since it is significantly toxic to biological systems. Standard IPTG preparations are currently available at about US$18 per gram or about US$125 per 10 grams; these IPTG preparations also contain dioxane, which is likewise toxic to biological systems. Dioxane-free IPTG is also available on the market, but costs about twice the price of standard IPTG (*i.e.* currently about US$36 per gram or about US$250 per 10 grams). In addition to the problems of expense and high toxicity to the fermentation system itself, in situations in which the expression product or fermentation product is to be marketed, environmental and health regulatory issues arise in regard to the

presence of IPTG therein, since IPTG also poses toxicity risks to humans, animals, and other biological organisms.

As a result, there is a need for promoters that are both useful for commercial fermentation production systems and activated by non- or low-toxicity inducers.

A further drawback of the use of lac promoters and their derivatives is that these promoters are "leaky" in that, even in a native, repressed state, the promoter permits a relatively high background level of expression. Therefore, multiple copies of the LacI repressor protein gene are usually included within the expression host cell in order to increase the degree of repression of the lac-type promoter. As a result, there is a need for promoters that are both useful for commercial fermentation systems and readily susceptible of being tightly controlled in an inactive state until induced.

In light of these concerns, several other, non-lac-type promoters have been proposed for controlling gene expression in commercial, prokaryotic fermentation systems (see Table 2).

Table 2. Proposed Inducible Commercial non-lac Promoters

Promoter	Inducer	Bacterial Host Cell(s)	Reference(s)
λP_R	High temperature	*E.coli*, Pseudomonads	3, 4
λP_L	High temperature	*E.coli*, Pseudomonads	3, 4
Pm	Alkyl- or halo-benzoates	Pseudomonads	4, 5
Pu	Alkyl- or halo-toluenes	Pseudomonads	5
Psal	Salicylates	Pseudomonads	5
Para	Arabinose in the absence of glucose	*E. coli*	5

In regard to the first two promoters listed in Table 2, promoters induced by high temperatures are problematic: since high temperatures can be harmful to the host cell culture; since it is often impractical to generate an even temperature spike throughout the large-scale, commercial fermentation volume; and since it is preferred to operate commercial fermentation equipment at lower temperatures than required for such induction. The other four suggested promoters listed in Table 2 have, to the inventors' knowledge, not

been demonstrated to function well in large-scale fermentation conditions; also, the alkyl- and halo-toluene inducers of the Pu promoter are significantly toxic to biological systems.

Thus, there remains a need for promoters that are useful for commercial fermentation production systems, activated by low-cost, low-toxicity chemical inducers, and tightly controlled.

In addition, in order to facilitate control of gene expression for production of proteins (and other expression products) and chemicals (processed by action of the expression products and/or the host cell) in a common fermentation platform using one prokaryotic organism, it is desirable to have a library of expression cassettes. These cassettes would each contain one or more of a variety of promoters that are of differing strengths, and/or induced under different growth conditions or by different chemicals. These expression cassettes would then be linked to various genes of interest to achieve total control of those genes under fermentation friendly conditions. The identification and optimization of a wide variety of growth-phase-dependent or chemically-inducible promoters is thus essential for control of (trans)gene expression during fermentation in such a fermentation platform.

Moreover, the construction of genetic circuits in which activation or induction of a first gene or operon leads to repression or activation of one or more subsequent genes or operons has been suggested as a means for very fine control of gene expression. Both linear (*e.g.*, serial and cascade) and circular (*e.g.*, daisy-chain) genetic circuits have been created. See, *e.g.*, U.S. Patent Pub. No. 20010016354 A1 of Cebolla Ramirez *et al.* These genetic circuits require a number of different promoters in order to function, and, in commercial fermentation, genetic circuits would need to rely upon promoters that are effective in commercial fermentation conditions. Thus, there is a need in the field of genetic circuits for a greater variety of promoters useful in commercial fermentation.

As noted above, promoters for use in commercial fermentation systems should be tightly regulated so that expression occurs only upon induction, preferably effected late in the fermentation run. The chemicals used to induce the promoters must be low cost, low-toxicity to the host bacterium and other organisms, and must tightly regulate gene

expression. In light of the above discussion, there is a need in the art for novel promoters that are tightly regulated and are induced at low cost using low-toxicity inducers.

SUMMARY OF THE INVENTION

The present invention provides novel promoters that are useful for gene expression in commercial fermentation. In a more specific aspect, the invention provides benzoate-inducible promoters, anthranilate-inducible promoters, and tandem promoters that may be employed in bacterial commercial fermentation systems.

The present invention provides:

isolated and/or recombinant benzoate promoter nucleic acids comprising the –35 region of the *Pseudomonas fluorescens* native benzoate promoter attached upstream of the –10 region thereof, via a 15-20 nucleotide linker; and to the operative promoter nucleic acid segment(s) found in SEQ ID NO:1;

mutant and closely related promoter nucleic acids whose nucleotide sequences are at least 90% homologous to such promoter nucleic acids;

such promoter nucleic acids further comprising a benzoate promoter activator protein (BenR) binding site; and

such promoter nucleic acids further comprising a benzoate promoter activator protein coding sequence, and where such activator protein coding sequences encode a benzoate promoter activator protein having an amino acid sequence at least 90% homologous to any one of the native, mutant, and/or truncated activator protein amino acids sequences presented in SEQ ID NO:2.

The present invention also provides:

isolated and/or recombinant anthranilate promoter nucleic acids comprising the –35 region of the *Pseudomonas fluorescens* native anthranilate promoter attached upstream of the –10

region thereof, via a 15-20 nucleotide linker; and to the operative promoter nucleic acid segment(s) found in SEQ ID NO:7;

mutant and closely related promoter nucleic acids whose nucleotide sequences are at least 90% homologous to such promoter nucleic acids;

such promoter nucleic acids further comprising an anthranilate promoter activator protein (AntR) binding site; and

such promoter nucleic acids further comprising a anthranilate promoter activator protein coding sequence, and where such activator protein coding sequences encode an anthranilate promoter activator protein having an amino acid sequence at least 90% homologous to any one of the native, mutant, and/or truncated activator protein amino acids sequences presented in SEQ ID NO:9.

The present invention also provides:

tandem promoters comprising a non-catabolite-repressed promoter attached (*i.e.* covalently attached) to and upstream of a natively catabolite-repressed promoter, either directly or by means of an inter-promoter polynucleotide linker, in which the catabolite repression of the latter promoter is overcome and/or a different improved promoter property is exhibited;

tandem promoters prepared by a process comprising covalently attaching a prokaryotic non-catabolite-repressed promoter to and upstream of a prokaryotic natively catabolite-repressed promoter, either directly or by means of an inter-promoter polynucleotide linker;

such tandem promoters wherein the inter-promoter polynucleotide linker is about 100 or less than 100 nucleotides long;

such tandem promoters in which the component non-catabolite-repressed and natively catabolite-repressed promoters are prokaryotic promoters, or bacterial promoters; and to tandem promoters in which the component promoters are obtained from the same of different species of the Pseudomonads and closely related bacteria, and/or of the genus *Pseudomonas*, and/or from *Pseudomonas fluorescens*; and to tandem promoters in which

the component promoters are obtained from gene(s) or operon(s) encoding alternative carbon source utilization enzyme(s) or pathway(s); and to tandem promoters in which the non-catabolite-repressed promoter is obtained from an operon encoding an anthranilate degradation pathway and the natively catabolite-repressed promoter is obtained from an operon encoding a benzoate degradation pathway, and/or in which the anthranilate promoter and benzoate promoter are selected from among those summarized in the above paragraphs; and

the operative tandem promoter(s) found in, or constructed from the component promoters shown in, SEQ ID NO:13.

The present invention also provides:

altered promoters prepared by a process comprising obtaining at least one polynucleotide having a base sequence at least 90% identical to and heterologous to the base sequence of any one of the claimed promoters or the sequence of any one of at least bases 1275-1307 of SEQ ID NO:1, at least bases 1239-1274 of SEQ ID NO:7, and at least bases 1329-1509 of SEQ ID NO:13; screening the polynucleotide(s) for the ability to direct transcription in a prokaryotic host cell, and optionally for at least one promoter property; and identifying, based on the results, at least one promoter, optionally having at least one improved property; and

improved promoters prepared by a process of: utilizing a promoter polynucleotide, having a base sequence of any one of the claimed promoters or the sequence of any one of at least bases 1275-1307 of SEQ ID NO:1, at least bases 1239-1274 of SEQ ID NO:7, and at least bases 1329-1509 of SEQ ID NO:13, as a hybridization probe for sequence-altered polynucleotide(s) at least 90% homologous thereto, or of performing mutagenesis and/or recombination upon said promoter polynucleotide to generate said sequence-altered polynucleotide(s), or of utilizing an information string representing the base sequence of the promoter polynucleotide to perform a search for a heterologous string at least 90% homologous thereto and providing a sequence-altered polynucleotide having the base sequence represented by said heterologous string; or of modifying such an information string into such a heterologous string and utilizing said modified string to identify an

information string identical thereto and then providing a sequence-altered polynucleotide having the base sequence represented by said information string; followed by screening the sequence-altered polynucleotide(s) for the ability to direct transcription in a prokaryotic host cell, and for at least one promoter property; and identifying, based on the results, at least one promoter having at least one improved property.

The present invention also provides:
isolated nucleic acid molecules comprising a nucleic acid sequence whose complement hybridizes under stringent hybridization and wash conditions to a nucleobase polymer molecule having a base sequence of any one of the claimed promoters or of any one of at least bases 1275-1307 of SEQ ID NO:1, at least bases 1239-1274 of SEQ ID NO:7, and at least bases 1329-1509 of SEQ ID NO:13, wherein said isolated nucleic acid molecule can function as a promoter in a prokaryotic cell; and

isolated nucleobase polymer molecules having the base sequence of a prokaryotic promoter polynucleotide molecule having a base sequence at least 90% identical to the base sequence of any one of the claimed promoters or of any one of at least bases 1275-1307 of SEQ ID NO:1, at least bases 1239-1274 of SEQ ID NO:7, and at least bases 1329-1509 of SEQ ID NO:13

The present invention also provides:
recombinant nucleic acid molecules that can function as expression construct(s) in a prokaryotic cell, comprising a promoter containing a base sequence at least 90% identical to the base sequence of any one of the claimed promoters or of any one of at least bases 1275-1307 of SEQ ID NO:1, at least bases 1239-1274 of SEQ ID NO:7, and at least bases 1329-1509 of SEQ ID NO:13; such recombinant expression constructs comprising an mRNA-encoding sequence; such recombinant expression constructs wherein the expression construct is a vector; such recombinant expression constructs wherein the vector is a plasmid; genetically engineered prokaryotic host cells containing any such a recombinant expression construct, and preferably also at least one, and more preferably more than one, copy of a gene encoding the relevant activator protein for the promoter of said recombinant expression construct (and where said gene is expressed in the host cell); expression systems

comprising such a genetically engineered prokaryotic host cell that preferably contains at least one, and more preferably more than one, copy of a gene encoding the relevant activator protein for the promoter of said recombinant expression construct (and where said gene is expressed in the host cell); and such expression systems wherein the promoter is a benzoate-inducible promoter and the activator protein has an amino acid sequence of either any one of residues 1-335 or 21-335 of SEQ ID NO:2, optionally containing Asn152, ; and such expression systems wherein the promoter is an anthranilate-inducible promoter and the activator protein has an amino acid sequence of either residues 1-330 of SEQ ID NO:9or residues 1-330 of SEQ ID NO:9 containing Ala268.

The present invention also provides:
a process for preparing a transcription product comprising growing such a genetically engineered prokaryotic host cell, and inducing the recombinant expression construct therein, thereby expressing the transcription product encoded thereby; and a process for preparing a polypeptide comprising expressing an mRNA transcription product, by use of such a process for preparing a transcription product, and further permitting the host cell to translate the mRNA into the polypeptide encoded thereby.

The present invention also provides transcriptional activator proteins operative in prokaryotic cells. These include a transcriptional activator protein having an amino acid sequence at least 90% homologous to that of any one of residues 1-335 of SEQ ID NO:2, residues 1-335 of SEQ ID NO:2containing Asn152, residues 21-335 of SEQ ID NO:2, and residues 21-335 of SEQ ID NO:2 containing Asn152; and a transcriptional activator protein having an amino acid sequence at least 90% homologous to that of any one of residues 1-330 of SEQ ID NO:9, or of residues 1-330 of SEQ ID NO:9 containing Ala268. The present invention also provides polynucleotide molecules containing a base sequence encoding such transcriptional activator proteins.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 presents a plasmid map of pDOW1019, a vector comprising Pben509, a benzoate-inducible promoter according to the present invention.

Figure 2 presents a plasmid map of pDOW1028, a vector comprising Pben278, a benzoate-inducible promoter according to the present invention.

Figure 3 presents a plasmid map of pDOW1035, a vector comprising an anthranilate-inducible promoter according to the present invention, including the coding sequence of the activator protein therefor.

Figure 4 presents a plasmid map of pDOW1057, a vector comprising an anthranilate-benzoate tandem promoter according to the present invention.

Figure 5 presents a comparison of the nucleotide sequences of the Pben509 (nucleotides c994 - c1502 of SEQ ID NO:1) and Pben278 (nucleotides g1228 - c1502 of SEQ ID NO:1) benzoate-inducible promoters. The TTA translational stop site of the upstream ORF, and the putative –35 region (TTGACG), –10 region (TACGGT), and "C" transcription start site are underlined. The "C" double-underlined in the Pben278 sequence indicates a mutation in Pben278, which differs from Pben509 and the genomic sequence; this mutation was introduced during PCR amplification.

Figure 6 presents a bar chart showing β-galactosidase induction under the control of Pben509 (pDOW1019) or Pben278 (pDOW1028), in the presence of 0mm (□) or 10mM (■) sodium benzoate. pDOW1017 is a vector with no promoter. Results shown are average measurements from triplicate wells assayed at 24 hours post-induction (host cells were grown in a defined salts medium for 24 hours prior to induction).

Figure 7 presents a comparison of the nucleotide sequences of four Pant promoters: Pant+AntR (a1-c1395 of SEQ ID NO:13), Pant713 (nucleotides c592 - c1304 of SEQ ID NO:7), Pant705 (nucleotides c592 - g1288 of SEQ ID NO:7), and Pant311 (nucleotides c994 - c1304of SEQ ID NO:7). The double-underlined nucleotide triplets (TCA and CAT) respectively indicate the stop and start codons of the AntR ORF, which is encoded by the strand complementary to that shown. The lone double-underlined "A" indicates a mutation from the genomic sequence, confirmed on both strands, which lies within the AntR coding

sequence; this change results in an expressed change from alanine to serine in the AntR protein. The putative -35 region (ATAGCC), -10 region (CTTAAT), and transcription start site ("A"), and the complement of the putative ribosome binding site (GGAGG) are all underlined.

Figure 8 presents a graph comparing the activity of a Pant713 construct (pDOW1029) versus an *antR*/Pant construct (pDOW1035), when induced with either 5mM anthranilate or 2mM 6-chloroanthranilate. In all constructs, the promoter was fused to a β-galactosidase-encoding sequence. The activity of the β-galactosidase reporter was followed over an 8 hour time course post-induction. The results show the activity of pDOW1035 following induction occurs much faster and is much higher than that of pDOW1029.

Figure 9 presents bar charts showing induction of a Pant-Pben tandem promoter construct, and induction of component promoter constructs. In all constructs, the promoter was fused to a β-galactosidase-encoding sequence. Induction of a Pant-Pben tandem promoter construct (pDOW1057) with either benzoate or anthranilate is presented in comparison to that of constructs containing one of the component promoters: a Pben promoter construct (pDOW1028), containing Pben278, and a Pant promoter construct (pDOW1035), containing the antR/Pant. Figure 9A presents results at 5 hours post-induction and Figure 9B presents results at 24 hours post-induction.

Figure 10 presents a bar chart comparison of Pben509 activity to that of improved mutants 2d3 and 21b5, created by error prone PCR of Pben509. Expression constructs were formed by fusing promoter-containing fragments to a *phoA* reporter gene. Cultures containing the constructs were induced with 10mM benzoate and alkaline phosphatase activity was measured at 24 hours post-induction.

Figure 11 presents a comparison of the nucleotide sequences of Pben509 (nucleotides c994 - c1502 of SEQ ID NO:1) and two improved mutants thereof created by means of error prone PCR: mutant 2d3 (nucleotides c994 - c1502 of SEQ ID NO:1 with a1106→t1106) and mutant 21b5 (nucleotides c994 - c1502 of SEQ ID NO:1 with c1223→t1223 and g1302→a1302). The TTA translational stop site of the upstream ORF, and the putative -35 region (TTGACG), -10

region (TACGGT), and "C" transcription start site are underlined. Mutations are shown double-underlined.

Figure 12 presents a graph of anthranilate-induced expression in 20L fermentation conditions. An antR/Pant construct pDOW1035 (■) and a tandem promoter construct pDOW1057 (▲) were induced with 5mM sodium anthranilate plus a 1mM/hour anthranilate feed. The activity level for each 20L fermentation run is shown.

Figure 13 presents graphs demonstrating improved anthranilate-induced expression from a tandem promoter construct (pDOW1057) in an *antA* knockout host created by means of a single crossover knockout of the *antA* gene in the host cell (*Pseudomonas fluorescens*). Results are shown for 20L fermentations induced at a target level of about 5mM sodium anthranilate and followed over a 48 hour time course post-induction. No anthranilate feed was required. Figure 13A presents the activity of β-galactosidase expressed from the reporter gene construct; Figure 13B shows the maintenance of anthranilate concentration, demonstrating that the knockout host cell does not metabolize anthranilate.

Figure 14 presents a bar chart showing induction of Pben –10 mutants. The alkaline phosphatase activity of P. fluorescens containing the indicated Pben::*phoA* fusions is shown following 24-hour induction with 0mM (□) or 5mM (■) sodium benzoate. Representative experiment of triplicate samples is shown.

Figure 15 presents a bar chart showing induction of Pant -10 mutants. The alkaline phosphatase activity of P. fluorescens *antA*::kanR containing the indicated Pant::*phoA* fusions is shown following 24-hour induction with 0mM (□) or 5mM (■) anthranilate. Representative experiment of triplicate samples is shown.

Figure 16 presents bar charts indicating the effect of Pben88-10 mutations on the activity of the antR-Pant311-Pben tandem promoter. The β-galactosidase activity of P. fluorescens containing the indicated tandem promoter::*lacZ* fusions is shownfollowing 24-hour induction with 0mM (□) or 5mM (■) sodium benzoate (Figure 16A) or anthranilate (Figure 16B). Representative experiment of triplicate samples is shown.

Figure 17 presents a graph of benzoate consumption during 20L fermentation when *benR* is present in multicopy with pDOW1090. Data shown are HPLC analyses measuring the concentration of benzoate throughout the induction of two 20L fermentors in duplicate fermentations, labeled run "030211I" (●) and run "030211K" (■). The cultures were induced with 5mM sodium benzoate.

Figure 18 presents bar chart comparisons of tandem promoter construct activity. The β-galactosidase activity of P. fluorescens containing the indicated tandem promoter::*lacZ* fusions (except for pDOW1035, which is ant activator-Pant311::*lacZ*; and pDOW1126, which is ben activator Pben88 -10consensus::*lacZ*) is shown following induction with 0 or 5mM sodium benzoate (Figures 18A and 18B for 2, 4, and 24 hour time-points) or anthranilate (Figures 18C and 18D, for 0, 2, 6, and 24 hour time-points). Representative experiment of triplicate samples is shown. Time points are indicated as "post-induction" by the letter "I".

Figure 19 presents a bar chart demonstrating an analysis of Pben activity in the *benAB* knock-out strain. β-galactosidase activity of a P. fluorescens MB101 *benAB* knock-out strain carrying pDOW1019 (Pben278::*lacZ*) is shown as a result of induction with either 0mM (□) or 5mM (■) sodium benzoate for up to 24 hours. Cells were grown in LB medium.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

The present invention provides commercially useful benzoate-inducible promoters, anthranilate-inducible promoters, and tandem promoters that may be employed in bacterial commercial fermentation systems. Preferred bacterial host cells for use in such systems include Pseudomonads and closely related bacteria. The chemical inducers of these promoters include benzoic and anthranilic acids, their effective chemical analogs, and biologically acceptable salts thereof.

Benzoic and anthranilic acids and biologically acceptable salts, preferably sodium or potassium salts, thereof, are inexpensive chemicals with low toxicity that can be utilized as (alternative) carbon sources by bacterial host cells, including Pseudomonads and closely related bacteria. For example, these chemical inducers are available on the market at less than about US$0.15 per gram (versus about US$18 or US$36 per gram for IPTG).

The present inventors have isolated, sequenced, and characterized the native promoters responsible for expression of the *P. fluorescens* benzoate (*benABCD*) degradative genes, which may be induced with benzoate in the absence of glucose, and of the *P. fluorescens* (*antABC*) degradative genes, which may be induced with anthranilate. (The expression products of these operons are catabolic pathway enzymes responsible for degradation of benzoate and anthranilate, respectively, in *Pseudomonas fluorescens* biotype A.) These promoters have been found capable of inducing expression of exogenous genes about 250-fold, for the benzoate promoter, and about 25- to about 35-fold, for the anthranilate promoter, when induced with 5mM sodium benzoate and 5mM sodium anthranilate, respectively. The present inventors have found these promoters to be sufficiently inducible for use in commercial fermentation systems to produce proteins and chemicals in bacterial host cells, including Pseudomonads and closely related bacteria.

In addition, the present inventors have created tandem promoter constructs in which a non-catabolite-repressed promoter is linked upstream of a natively catabolite-repressed promoter, thereby surprisingly overcoming the catabolite repression of the latter promoter and/or thereby exhibiting a different improved property (*e.g.*, increased strength of induction or increased tightness of regulation). At least one example of a tandem promoter construct has been described for expression of foreign genes [6]. However, this example is a tandem arrangement of two copies of the same promoter, Plac, and the reference presents no evidence to suggest that the tandem Plac-Plac promoter has advantages over a single Plac promoter. Likewise, dual promoter constructs are known, *e.g.*, for use in shuttle vectors, in which two promoters operative in different species or genera are both operably attached to the same gene so that the gene can be expressed in either of the two different species or genera.

In contrast to these tandem and dual promoter constructs, the present creation of tandem promoter constructs, in which two non-identical promoters are placed in tandem arrangement, has surprisingly been found to retain advantageous features of both promoters. For example, the tandem arrangement of the anthranilate promoter, Pant, and the benzoate promoter, Pben, has resulted in formation of a tandem promoter that, when induced with anthranilate under fermentation conditions, exhibits both freedom from catabolite repression (a desirable feature of Pant, not shared by Pben) and improved strength of induction (a desirable feature of Pben, not shared by Pant). Thus, the tandem promoters of the present invention permit retention of desirable properties of the individual promoter elements, so that the resulting tandem promoter can exhibit improved properties: *e.g.*, increased strength of induction, or increased tightness of regulation (*i.e.* transcription only when contacted with the relevant inducer of the promoter's activator or repressor protein); and/or lack of catabolite repression.

GLOSSARY

A### (Absorption)

As used herein in regard to analytical detection, terms such as "A450" mean "absorption at a wavelength of 450 nm."

A and An (Indefinite Articles)

As used herein and in the appended claims, the singular forms "a", "an", and "the" include both singular and plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a host cell" literally defines both those embodiments employing only a single host cell, those employing a plurality of host cells of a single type, and those employing a plurality of host cells of a plurality of types.

* (Asterisk)

As used herein in regard to calculations, the "*" symbol (asterisk) indicates the mathematical multiplication function.

BCIP

5-Bromo-4-chloro-3-indolyl phosphate, *e.g.*, a divalent salt thereof, such as a disodium salt. This is used in conjunction with a, *e.g.*, tetrazolium salt, such as: a halide salt of Nitroblue Tetrazolium (NBT), *e.g.*, bis-[2-(4-yl-2-methoxyphenyl)-3-(4-nitrophenyl)-5-phenyl-tetrazolium chloride]; or of Iodo-Nitro-Tetrazolium (INT), also called Iodoblue Tetrazolium, *e.g.*, 2-(4-iodophenyl)-3-(4-nitrophenyl)-5-phenyl tetrazolium chloride.

Comprising

As used herein, the term "comprising" means that the subject contains the elements enumerated following the term "comprising" as well as any other elements not so enumerated. In this, the term "comprising" is to be construed as a broad and open-ended term; thus, a claim to a subject "comprising" enumerated elements is to be construed inclusively, *i.e.* construed as not limited to the enumerated elements. Therefore, the term "comprising" can be considered synonymous with terms such as, *e.g.*, "having," "containing," or "including."

The invention, as described herein, is spoken of using the terms "comprising" and "characterized in that." However, words and phrases having narrower meanings than these are also useful as substitutes for these open-ended terms in describing, defining, or claiming the invention more narrowly.

Corresponding

As used herein in reference to a sequence record's "corresponding to" a polynucleotide source, the term "corresponding" means that a given base sequence contained, as an information string, within the sequence record, is present in the form of a physical nucleobase sequence-containing molecule within the polynucleotide source.

ddH_2O

As used herein, ddH_2O refers to distilled, deionized water purified through a Milli-Q gradient system with Q-GARD purification pack (Millipore, Bedford, MA).

dNTPs

Except where otherwise indicated, as used herein in regard to reagents for polynucleotide synthesis reactions, the term "dNTPs" means an equimolar solution of each of the four deoxyribonucleotide triphosphates (dGTP, dCTP, dATP, dTTP). Thus, *e.g.*, reference to 10mM dNTPs indicates a solution containing 10mM each of dGTP, dCTP, dATP, and dTTP.

Exogenous and Foreign

The term "exogenous" means "from a source external to" a given cell or molecule. In the present application, as is common use in the art, this term is used interchangeably with the term "foreign," as synonyms. Both of these terms are used herein to indicate that a given object is foreign a given the cell or molecule (*e.g.*, a promoter polynucleotide), *i.e.* not found in nature in the cell and/or not found in nature with or connected to the molecule.

Heterologous

As used herein, the term "heterologous mean "non-identical" in sequence (not 100% identical in base sequence).

In and On

As used herein in regard to growing organisms by use of a growth medium, the organisms may be said to be grown "in" or "on" the medium. In the expression systems of the present invention, the medium is preferably a gel or liquid medium, more preferably an aqueous gel or liquid medium, and still more preferably an aqueous liquid medium. Thus, in this context, the terms "in" and "on" are used synonymously with one another to indicate growth of the host cells in contact with the medium.

Information String

As used herein, the phrase "information string" means a series of data elements (*e.g.*, bits, bytes, or alphanumeric characters), which series represents the information of a given series of nucleobases.

IPTG

Isopropyl-β-D-1-thiogalactopyranoside.

ONPG

O-Nitrophenyl-β-D-galactopyranoside, also known as 2-Nitrophenyl-β-D-galactopyranoside.

ORF

Open reading frame.

PNPP

para-Nitrophenyl phosphate, *e.g.*, a divalent salt thereof, such as a disodium salt. Also referred to as 4-nitrophenyl phosphate.

Polynucleotide Length

As used herein, the term "nucleotides" is used to describe the length of polynucleotides. However, in this context, the terminology is meant to refer both to length in nucleotides per se in regard to single stranded polynucleotides, and to length in base pairs in regard to double stranded polynucleotides.

Polynucleotide Source

As used herein, the phrase "polynucleotide source" means any source of a physical embodiment of a nucleic acid containing a given nucleobase sequence, such as a nucleic acid sample, clone, or native cell containing such a polynucleotide molecule.

Promoter Activator Protein Terminology

The native activator of a given promoter is designated with an "R" as, e.g., BenR, AntR for the native activators of the Pben and Pant promoters, respectively.

Promoter Terminology

"Pant": Promoter for the anthranilate operon of *P. fluorescens*.

"Pben": Promoter for the benzoate operon of *P. fluorescens*.

"Plac": Promoter for the lactose operon of *E. coli*.

"Ptandem" and "tandem promoter": a tandem arrangement of promoters in which a non-catabolite-repressed promoter is attached to and upstream of, by means of a sequence of 0 to about 100 nucleotides, a catabolite-repressed promoter. This is exemplified herein by Pant-Pben tandem promoter constructs.

"Promoter": a polynucleotide comprising at least 25 nucleotides, more commonly about 30 nucleotides, containing a prokaryotic "-35 region through -10 region." Preferably, this "-35 region through -10 region" is a "-35 region through -10 region" obtained from a single gene, or a combination of a -35 region and a -10 region obtained from cognate genes, the gene(s) being obtained from at least one prokaryote, more preferably at least one organism of the "Pseudomonads and closely related species." Preferably, the "-35 region through -10 region" is a σ70 "-35 region through -10 region," and the "-35 region" and the "-10 region" in the combination are, respectively, a σ70 -35 region and a σ70 -10 region.

The -35 region is linked upstream of the -10 region by an intra-promoter polynucleotide of preferably about 15 to about 20 nucleotides. More preferably, a promoter according to the present invention comprises about 35 nucleotides, which contains, in addition to the "-35 region through -10 region," a segment of about 5 to about 10 immediate downstream nucleotides, more preferably 6 to 7 immediate downstream nucleotides, terminating in a transcription start site nucleotide. In a preferred embodiment, this segment is obtained from the same gene as provides at least one of the -35 or -10 region. In a particularly preferred embodiment, the promoter will also contain an immediate upstream region of about 20 to about 250, more preferably about 40 to about 150, nucleotides comprising a promoter activator protein binding site, preferably an AraC/XylS-class binding site. In a preferred embodiment, the binding site region is obtained from the same gene as provides at least one of the -35 or -10 region.

As used herein, the term "-35 region" or "minus 35 region", indicates a 5-6 nucleotide sequence beginning approximately 35 nucleotides upstream (i.e. in a 5' direction from) a transcription start site, the transcription start site being numbered as "+1."

Likewise, the term "-10 region" or "minus 10" region" indicates a 5-6 nucleotide sequence beginning approximately 10 nucleotides upstream (i.e. in a 5' direction from) a transcription start site, the transcription start site being numbered as "+1."

RNA Terms

As used herein, the following RNA terms have the definitions recited below.

aRNA: anti-sense RNA

cRNA: cytoplasmic RNA

gRNA: guide RNA (for editing mitochondrial pre-mRNAs)

hnRNA: heteronuclear RNA

mRNA: messenger RNA

miRNA: microRNA (for regulating mRNA expression)

mtRNA: mitochondrial RNA

nRNA: nuclear RNA

ncRNA: non-coding RNA

pRNA: packaging RNA (for virus and phage particle assembly)

rRNA: ribosomal RNA

satRNA: satellite RNA

scRNA: small cytoplasmic RNA

siRNA: small interfering RNA

snRNA: small nuclear RNA

snoRNA: small nucleolar RNA

srpRNA: signal recognition particle RNA

stRNA: small temporal RNA (a subgroup of miRNA)

tRNA: transfer RNA

tmRNA: transfer-messenger-like RNA (for marking, for subsequent degradation, nascent polypeptides in stalled ribosomes)

vRNA: viral (and/or phage) RNA

Searching

As used herein in regard to seeking for information, the term "search" means performing (by manual, visual, or automated means) one or more comparisons between a known information string and other information strings in order to identify an identical or non-identical information string.

Sequence Record

As used herein, the phrase "sequence record" means a stored embodiment of one or more information strings, such as a computer readable record or a paper record.

-10 Variant Promoter Designations

"-10con": indicates a variant promoter in which the native -10 region has been substituted with the consensus -10 region sequence 'tataat.'
"-10ben": indicates a variant promoter in which the native -10 region has been substituted with the -10 region from *P. fluorescens* Pben 'tacggt.'
"-10benAc": indicates a variant promoter in which the native -10 region has been substituted with the -10 region from *Acinetobacter* Pben 'taaggt.'
"-10wt": indicates truncated native promoter retaining the wild-type -10 sequence.

~ (Tilde)

The ~ symbol (the tilde) is used herein to indicate "about".

× (Times)

The × symbol (the times symbol), as used herein in regard to the concentration of a solution, means, *e.g.*, that a 5× preparation is five times as concentrated as a 1× preparation, for the same solution composition (*i.e.* for the same relative amounts of all components therein).

Tris

The term "Tris" as used herein means Tris (hydroxymethyl) aminomethane (available from Fisher Scientific, Pittsburgh, PA).

X-gal

5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

General Materials & Methods

Unless otherwise noted, standard techniques, vectors, control sequence elements, and other expression system elements known in the field of molecular biology are used for nucleic acid manipulation, transformation, and expression. Such standard techniques, vectors, and elements can be found, for example, in: Ausubel *et al.* (eds.), *Current Protocols in Molecular Biology* (1995) (John Wiley & Sons); Sambrook, Fritsch, & Maniatis (eds.), *Molecular Cloning* (1989) (Cold Spring Harbor Laboratory Press, NY); Berger & Kimmel, *Methods in Enzymology* 152: Guide to Molecular Cloning Techniques (1987) (Academic Press); and Bukhari *et al.* (eds.), *DNA Insertion Elements, Plasmids and Episomes* (1977) (Cold Spring Harbor Laboratory Press, NY).

The promoters of the present invention include the benzoate promoter from *Pseudomonas fluorescens*, the anthranilate promoter from *Pseudomonas fluorescens*, and their derivatives. The promoters of the present invention also include tandem promoters having a non-catabolite-repressed promoter linked upstream of a natively catabolite-repressed promoter, in which the catabolite repression of the latter promoter is overcome and/or a different improved promoter property is exhibited.

The promoters of the present invention are typically in the form of DNA when in use in an expression system. However, the nucleobase sequence of the promoters may be present in the form of DNA, RNA, or any nucleic acid analog known in the art, *e.g.*, peptide nucleic acid (PNA).

Benzoate Promoters

In a preferred embodiment, a benzoate promoter of the present invention is the *Pseudomonas fluorescens* native benzoate promoter or an improved mutant thereof. The present inventors have found this promoter to be inducible with benzoic acid, benzoic acid analogs (*e.g.*, m-toluic acid), and biologically acceptable salts thereof (*e.g.*, sodium benzoate).

In a preferred embodiment, a benzoate promoter of the present invention comprises the –35 region of the *Pseudomonas fluorescens* native benzoate promoter attached upstream of the -10 region of this native promoter, via a 15-20 nucleotide linker. In a preferred embodiment, the linker is 15 nucleotides long.

In a preferred embodiment, a benzoate promoter of the present invention comprises nucleotides 1275-1280 of SEQ ID NO:1 attached upstream of nucleotides 1296-1301 of SEQ ID NO:1, via a 15-20 nucleotide linker. In a preferred embodiment, the linker is 15 nucleotides long. In a particularly preferred embodiment, the linker is nucleotides 1281-1295 of SEQ ID NO:1, a benzoate promoter of this preferred embodiment thereby comprising nucleotides 1275-1301 of SEQ ID NO:1. In a preferred embodiment, a benzoate promoter of the present invention comprises nucleotides 1275-1301 of SEQ ID NO:1 attached immediately upstream of a spacer segment of about 6 nucleotides, preferably of 6 nucleotides, in length, and terminating with a nucleotide that functions as a transcription start site. In a preferred embodiment, the spacer segment is nucleotides 1302-1307 of SEQ ID NO:1, a benzoate promoter of this preferred embodiment thereby comprising nucleotides 1275-1307 of SEQ ID NO:1.

In a preferred embodiment, a benzoate promoter of the present invention comprises both a "-35 to -10 region" and a benzoate promoter activator (or repressor) protein binding site, preferably an activator protein binding site. In a preferred embodiment, a benzoate promoter of the present invention comprises nucleotides 1275-1301 of SEQ ID NO:1 attached immediately downstream of a spacer region of about 50 nucleotides in length. In a preferred embodiment, a benzoate promoter of the present invention comprises nucleotides 1275-1301 of SEQ ID NO:1 attached immediately downstream of a spacer region of about 45 nucleotides in length. In a preferred embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:1, beginning about 50 nucleotides upstream of nucleotide 1275 and ending with nucleotide 1274. In a preferred embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:1, beginning about 45 nucleotides upstream of nucleotide 1275 and ending with nucleotide 1274. In a preferred embodiment, the spacer

region has the sequence of nucleotides 1228-1274 of SEQ ID NO:1, a benzoate promoter of this preferred embodiment thereby comprising the sequence of nucleotides 1228-1301.

In a preferred embodiment, a benzoate promoter of the present invention comprises nucleotides 1275-1301 of SEQ ID NO:1 attached immediately upstream of said spacer segment and attached immediately downstream of said spacer region. In a preferred embodiment, a benzoate promoter of the present invention comprises nucleotides 1228-1307 of SEQ ID NO:1.

In a preferred embodiment, in expression systems in which a benzoate promoter according to the present invention is used, the host cell will also contain and express at least one nucleic acid encoding a benzoate promoter activator protein. Even more preferred is the use therein of multiple expressed copies of such a Pben activator protein-encoding nucleic acid. In a preferred embodiment, the Pben activator protein will have an amino acid sequence of SEQ ID NO:2 or the residue 152 (Asn) variant thereof, or an amino acid sequence of residues 21-335 of SEQ ID NO:2 or the residue 152 (Asn) variant thereof. In a preferred embodiment, the nucleic acid encoding the Pben activator protein will contain the sequence of bases 285-1229 of SEQ ID NO:1 or the base 679 mutant variant thereof, or the sequence of bases 225-1229 of SEQ ID NO:1 or the base 679 mutant variant thereof; or the complement thereof of any of these; or the RNA equivalent of any of these.

Anthranilate Promoters

In a preferred embodiment, an anthranilate promoter of the present invention is the *Pseudomonas fluorescens* native anthranilate promoter or an improved mutant thereof. The present inventors have found this promoter to be inducible with anthranilic acid, anthranilic acid analogs (*e.g.*, haloanthranilic acids), and biologically acceptable salts thereof (*e.g.*, sodium anthranilate); and with o-toluate (o-toluate has been found to induce this promoter as well as does anthranilate).

In a preferred embodiment, an anthranilate promoter of the present invention comprises the –35 region of the *Pseudomonas fluorescens* native anthranilate promoter attached upstream

of the –10 region of this native promoter, via a 15-20 nucleotide linker, more preferably a 16-19 nucleotide linker. In a preferred embodiment, the linker is 19 nucleotides long.

In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1244 of SEQ ID NO:7 attached upstream of nucleotides 1264-1268 of SEQ ID NO:7, via a 15-20 nucleotide linker. In a preferred embodiment, the linker is 19 nucleotides long. In a particularly preferred embodiment, the linker is nucleotides 1245-1263 of SEQ ID NO:7, an anthranilate promoter of this preferred embodiment thereby comprising nucleotides 1239-1268 of SEQ ID NO:7. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately upstream of a spacer segment of about 6 nucleotides, preferably of 6 nucleotides, in length, and terminating with a nucleotide that functions as a transcription start site. In a preferred embodiment, the spacer segment is nucleotides 1269-1274 of SEQ ID NO:7, an anthranilate promoter of this preferred embodiment thereby comprising nucleotides 1239-1274 of SEQ ID NO:7.

In a preferred embodiment, an anthranilate promoter of the present invention comprises both a "-35 to -10 region" and an anthranilate promoter activator (or repressor) protein binding site, preferably an activator protein binding site. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 250 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 200 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 150 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 120 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 110 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises

nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 100 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 85 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 80 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 75 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 70 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 65 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 60 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 55 nucleotides in length. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1268 of SEQ ID NO:7 attached immediately downstream of a spacer region of about 50 nucleotides in length.

In a preferred embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:7, beginning about 100 nucleotides upstream of nucleotide 1239 and ending with nucleotide 1238. In a preferred embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:7, beginning about 85 nucleotides upstream of nucleotide 1239 and ending with nucleotide 1238. In a preferred embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:7, beginning about 80 nucleotides upstream of nucleotide 1239 and ending with nucleotide 1238. In a preferred embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:7, beginning about 75 nucleotides upstream of nucleotide 1239 and ending with nucleotide 1238. In a preferred

embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:7, beginning about 70 nucleotides upstream of nucleotide 1239 and ending with nucleotide 1238. In a preferred embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:7, beginning about 65 nucleotides upstream of nucleotide 1239 and ending with nucleotide 1238. In a preferred embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:7, beginning about 60 nucleotides upstream of nucleotide 1239 and ending with nucleotide 1238. In a preferred embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:7, beginning about 55 nucleotides upstream of nucleotide 1239 and ending with nucleotide 1238. In a preferred embodiment, the spacer region has the sequence of the region shown in SEQ ID NO:7, beginning about 50 nucleotides upstream of nucleotide 1239 and ending with nucleotide 1238. In a preferred embodiment, the spacer region has the sequence of nucleotides 1139-1238 of SEQ ID NO:7, an anthranilate promoter of this preferred embodiment comprising nucleotides 1139-1238 of SEQ ID NO:7.

In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1239-1274 of SEQ ID NO:7 attached immediately upstream of said spacer segment and attached immediately downstream of said spacer region. In a preferred embodiment, an anthranilate promoter of the present invention comprises nucleotides 1139-1274 of SEQ ID NO:7.

In a preferred embodiment, in expression systems in which an anthranilate promoter according to the present invention is used, the host cell will also contain and express at least one nucleic acid encoding an anthranilate promoter activator protein. Even more preferred is the use of multiple expressed copies of such a Pant activator protein-encoding nucleic acid. In a preferred embodiment, the Pant activator protein will have an amino acid sequence of SEQ ID NO:9 or the residue 268 (Ala) variant thereof. In a preferred embodiment, the nucleic acid encoding the Pant activator protein will contain the sequence of bases 1-990 of SEQ ID NO:8 or the base 802 variation thereof; or the complement thereof of any of these; or the RNA equivalent of any of these.

Mutant and Closely Related Activator Proteins and Polynucleotides Encoding Them

The same methods as described below for use in obtaining mutant promoters may similarly be used to obtain mutant activator proteins and the coding sequences and genes thereof. In this case, at least a portion of the gene encoding a given activator protein, *e.g.*, all or part of the coding sequence thereof, may be used as, or be used to form a probe for use in hybridization probing; or may provide a base sequence to be used in the form of an information string, identical or at least 90% identical thereto, to search a database for structurally related sequences for testing. Likewise all or part of the amino acid sequence of the activator protein may be used as an information string to perform such searching. The resulting sequences identified by hybridization or bioinformatic searching are then tested for promoter activation activity and/or for improved properties.

Thus, also included within the present invention are transcriptional activator proteins having an amino acid sequence at least 90% identical to and heterologous to that of: a Pben activator protein having an amino acid sequence of any one of residues 1-335 of SEQ ID NO:2, residues 1-335 of SEQ ID NO:2 containing Asn152, residues 21-335 of SEQ ID NO:2, and residues 21-335 of SEQ ID NO:7 containing Asn152; ; and a Pant activator protein having an amino acid sequence of any one of residues 1-330 of SEQ ID NO:9 and residues 1-330 of SEQ ID NO:9 containing Ala268. The present invention also includes polynucleotides encoding said mutant and closely related transcriptional activator proteins.

Tandem Promoters

In a preferred embodiment, a tandem promoter of the present invention comprises a (natively) non-catabolite-repressed promoter attached upstream of a natively catabolite-repressed promoter, in which the catabolite repression of the latter promoter is overcome and/or a different improved promoter property is exhibited.

In a preferred embodiment, both the non-catabolite-repressed promoter and the natively catabolite-repressed promoter are selected from the prokaryotes. In a preferred embodiment, both the non-catabolite-repressed promoter and the natively catabolite-repressed promoter are selected from the bacteria. In a preferred embodiment, both the non-catabolite-repressed promoter and the natively catabolite-repressed promoter are selected

from the Proteobacteria; preferably Gram negative Proteobacteria. In a preferred embodiment, both the non-catabolite-repressed promoter and the natively catabolite-repressed promoter are selected from the "Pseudomonads and closely related bacteria" or from a Subgroup thereof, as defined below.

In a preferred embodiment, both promoters are selected from the same species. In a preferred embodiment, both promoters are obtained from the same species in a genus selected from among the "Pseudomonads and closely related bacteria" or among a Subgroup thereof, as defined below. In a preferred embodiment, both promoters are selected from organisms of the genus *Pseudomonas*. In a preferred embodiment, both promoters are selected from the same species in the genus *Pseudomonas*. In a preferred embodiment, both promoters are selected from *Pseudomonas fluorescens*. In a preferred embodiment, both promoters are selected from *Pseudomonas fluorescens* biotype A.

The individual promoters selected for use in a tandem promoter according to the present invention may be activatable promoters, repressible promoters, or a combination thereof. In a preferred embodiment at least one of, and preferably both of, the individual promoters will be activatable promoters. Where a repressible promoter is present as a promoter element in such a tandem promoter, preferably the cell in which the tandem promoter is utilized will also contain at least one, and preferably more than, one copy of an expressible coding sequence for a repressor protein that mediates the regulation of the promoter. Where an activatable promoter is present as a promoter element in such a tandem promoter, preferably the cell in which the tandem promoter is utilized will also contain at least one, and preferably more than, one copy of an expressible coding sequence for an activator protein that mediates the regulation of the promoter.

In a preferred embodiment, both promoters are obtained as native promoters of genes or operons encoding enzyme(s) and/or pathway(s) capable of enabling a cell to utilize (*e.g.*, to import, export, transport, or metabolize) alternative carbon source(s). In a preferred embodiment, the non-catabolite-repressed promoter is a native promoter of a gene or operon encoding enzyme(s) and/or pathway(s) capable of biocatalytically degrading anthranilate, *i.e.* an "anthranilate promoter." In a preferred embodiment, the natively catabolite-repressed

promoter is a native promoter of a gene or operon encoding enzyme(s) and/or pathway(s) capable of biocatalytically degrading benzoate, *i.e.* a "benzoate promoter." In a preferred embodiment, the anthranilate promoter is an anthranilate promoter as described above. In a preferred embodiment, the benzoate promoter is a benzoate promoter as described above.

In a preferred embodiment, a tandem promoter of the present invention is a construct formed by linking the *Pseudomonas fluorescens* native anthranilate promoter to, and upstream of, the *Pseudomonas fluorescens* native benzoate promoter. The present inventors have found such a promoter arrangement to be inducible with anthranilic acid, anthranilic acid analogs (*e.g.*, haloanthranilic acids), and biologically acceptable salts thereof (*e.g.*, sodium anthranilate); with benzoic acid and biologically acceptable salts thereof; and with o-toluate (o-toluate has been found to induce this promoter as well as does anthranilate).

In a preferred embodiment, the non-catabolite-repressed promoter is attached immediately upstream of the natively catabolite-repressed promoter. This attachment may be made directly between the promoters (or directly between native nucleic acid segments containing the promoters) or by means of an, *e.g.*, polynucleotide linker connecting the promoters (or segments) to one another. In a preferred embodiment, the non-catabolite-repressed promoter is attached upstream of the natively catabolite-repressed promoter, via an inter-promoter linker. Preferably, the inter-promoter linker will be a polynucleotide, provided that that polynucleotide linker contains no sequence that functions as a transcription termination signal. In a preferred embodiment the inter-promoter linker is a polynucleotide of about 100 nucleotides in length. In a preferred embodiment, the inter-promoter linker is less than 100 nucleotides in length. In a preferred embodiment the inter-promoter linker is a polynucleotide of length equal to or less than 90 nucleotides. In a preferred embodiment the inter-promoter linker is a polynucleotide of length equal to or less than 80 nucleotides. In a preferred embodiment the inter-promoter linker is a polynucleotide of length equal to or less than 70 nucleotides. In a preferred embodiment the inter-promoter linker is a polynucleotide of length equal to or less than 60 nucleotides. In a preferred embodiment the inter-promoter linker is a polynucleotide of length equal to or less than 50 nucleotides. In a preferred embodiment the inter-promoter linker is a polynucleotide of length equal to or less than 40 nucleotides. In a preferred embodiment the inter-promoter linker is a polynucleotide of

length equal to or less than 30 nucleotides. In a preferred embodiment the inter-promoter linker is a polynucleotide of length equal to or less than 20 nucleotides. In a preferred embodiment, the inter-promoter linker is equal to or less than 10 nucleotides. In a preferred embodiment the inter-promoter linker is a polynucleotide at least about 5 nucleotides, or at least about 10 nucleotides, or at least about 20 nucleotides, or at least about 30 nucleotides, or at least about 40 nucleotides in length. In a preferred embodiment the inter-promoter linker is a polynucleotide about 5 to about 50 nucleotides, or about 10 to about 50 nucleotides, or about 20 to about 50 nucleotides, or about 30 to about 50 nucleotides in length. In a preferred embodiment the inter-promoter linker is a polynucleotide having a length of 43 nucleotides. In a preferred embodiment the inter-promoter linker has the sequence of SEQ ID NO:14.

In a preferred embodiment, a tandem promoter comprises an anthranilate promoter sequence selected from the group consisting of nucleotides 1221-1365, 1221-1371, 1329-1365, and 1329-1371 of SEQ ID NO:13 attached upstream of a benzoate promoter sequence selected from the group consisting of nucleotides 1430-1503, 1430-1509, 1477-1503, and 1477-1509 of SEQ ID NO:13. In a preferred embodiment, a tandem anthranilate-benzoate promoter of the present invention comprises, for the benzoate promoter portion, both a "-35 to -10 region" and a benzoate promoter activator (or repressor) protein binding site, preferably an activator protein binding site. In a preferred embodiment, a tandem promoter comprises nucleotides 1329-1503 of SEQ ID NO:13. In a preferred embodiment, a tandem promoter comprises nucleotides 1329-1509 of SEQ ID NO:13. In a preferred embodiment, a tandem promoter comprises nucleotides 1221-1503 of SEQ ID NO:13. In a preferred embodiment, a tandem promoter comprises nucleotides 1221-1509 of SEQ ID NO:13. In a preferred embodiment, a tandem promoter comprises nucleotides 1329-1544 of SEQ ID NO:13. In a preferred embodiment, a tandem promoter comprises nucleotides 1221-1544 of SEQ ID NO:13.

In a preferred embodiment, an anthranilate activator protein coding sequence or a benzoate activator protein coding sequence is included in, and expressed within, a system using, respectively, a Pant-containing or Pben-containing tandem promoter of the present invention. Where the tandem promoter contains both a Pant and a Pben, preferably an

anthranilate promoter activator protein coding sequence is selected, for example, the anthranilate activator protein (AntR) described above in regard to Pant promoters. Even more preferred in any expression system is the presence of such an expressed coding sequence in multiple copies.

Sources of Native Promoters for Use in Constructing Tandem Promoters

Tandem promoters according to the present invention may be constructed, *e.g.*, by obtaining from prokaryotic cells, preferably bacterial cells, native promoters from genes or operons encoding enzyme(s) responsible for utilization of alternative carbon sources, *i.e.* carbon sources other than glucose. In a preferred embodiment, the bacterial cells will be chosen from among the bacterial cells belonging to the "Pseudomonads and closely related bacteria," or any one of the 19 Subgroups thereof, as defined below.

Bacteria are known that are capable of utilizing a wide range of alternative carbon sources. In a preferred embodiment, a native promoter selected for use in constructing a tandem promoter will be obtained from a gene or operon from which is expressed an enzyme(s) having degradative activity toward at least one alternative carbon source chosen from among:

- Straight-chain, Branched-chain, Cyclic, and Alicyclic Homo- and Hetero-Hydrocarbons (saturated or unsaturated) and derivatives thereof;
- Aromatic and Alkylaryl Compounds and derivatives thereof, *e.g.*, benzene, naphthalene, anthracene, phenanthrene, toluene, xylene, biphenyl;
- Heterocyclic compounds and derivatives thereof, *e.g.*, steroids, sterols, allantoins, cyclic terpenes, yohimbines, indoles, imidazoles, oxazines, quinolines, phenazines, xanthenes;
- Alcohols & Polyols and derivatives thereof, *e.g.*, ethanol, phenol, naphthol, cresol, catechol, glycerol, benzyl alcohol, menthanol;
- Acids, Esters, Anhydrides, and derivatives thereof, *e.g.*, acetate, salicylate, benzoate, hydroxybenzoate, anthranilate, phthalate, benzylalkanoates, gentisate, amino acids (*e.g.*, glutamate); mono-, di-,and tri-carboxylic acids; fatty acids, lipids, and related compounds;
- Aldehydes, Ketones, Ethers and derivatives thereof;

- Halogenated Organic Compounds and derivatives thereof, *e.g.*, chlorobenzoate, chlorophenols, iodonaphthalene, bromoxylene, fluoropentane, trichloropropane;
- Organo-Phosphorus Compounds and derivatives thereof, *e.g.*, organophosphonates, organophosphates, organophosphites, phospha-compounds, phospho-compounds;
- Organo-Sulfur Compounds and derivatives thereof, *e.g.*, organosulfonates, organosulfates, organosulfites, thia-compounds, thio-compounds;
- Organo-Nitrogen Compounds and derivatives thereof, *e.g..*, organonitrates, nitro-organics (*e.g.*, nitrobenzoate, nitrophenol), cyano compounds, hydrazines, amines, imines, amides, imides, purines, pyrimidines, aza-compounds, azo-compounds;
- Other Hetero-Organic compounds, *e.g.*, organo-boron compounds, organo-silicon compounds, organometallic compounds, multi-heteroatom-organic compounds; and
- Multi-functional organic compounds.

The genes and operons encoding these biodegradative activities may be either catabolite-repressed or non-catabolite-repressed, as described above. The native promoters thereof may be readily obtained by one of ordinary skill in the art by methods well known in the art, *e.g.*, by isolating mRNA encoding such an enzyme and using the nucleic acid sequence of the mRNA or cDNA made therefrom, to probe the bacterial genome (or a record of the genomic sequence thereof) for occurrence(s) of the corresponding DNA gene. This is followed by identification of regulatory regions, including a transcription start site, located in the segment of DNA immediately upstream of (*i.e.* 5' to) the coding sequence. Expression constructs containing such regulatory region nucleic acid sequences are then formed and the expression construct(s) tested for induction in bacterial host cells by one or more alternative carbon source compounds, both in the presence and absence of glucose. This provides catabolite-repressed and non-catabolite-repressed promoters that may be used in constructing a tandem promoter according to the present invention.

A variety of catabolite repressed and non-catabolite-repressed genes and operons are known that either (a) encode enzymes that utilize (*e.g.*, that transport, anabolize, or catabolize) alternative carbon sources or (b) encode regulatory genes that control expression from such enzyme-encoding gene(s) and operon(s). The promoter of a typical gene or operon of this type is regulated in that transcription therefrom depends upon, *i.e.* the promoter is induced

or derepressed by, the presence of a relevant alternative carbon source or an analog compound thereof.

Examples of such catabolite repressed genes and operons include, *e.g.*:

- the *styABCD* operon of *Pseudomonas fluorescens* ST, which encodes enzymes required for the conversion of styrene to phenylacetate [Ref. 8];
- the *xylCMABN* operon of *P. putida* mt-2, which encodes enzymes required for the transformation of toluene to benzoate and the transformation of xylenes to toluates [Ref. 9];
- the *alkBFGHJKL* operon of *Burkholderia cepacia*, which encodes enzymes, including alkane hydroxylase, required for metabolism of alkanes and alkenes [Ref. 10];
- the *P. aeruginosa* gene, *oprD*, which encodes a specific porin that facilitates the uptake of basic amino acids, and of the carbapenem antibiotic, imipenem, a thienamycin derivative [Ref. 11];
- the *P. aeruginosa* gene, *aotJ*, which is part of an operon encoding enzymes required for the transport of arginine and ornithine [Ref. 12]; and
- the *P. putida* and *P. aeruginosa* genes, *bkdR*, which encode a protein regulating expression of an operon encoding a branched-chain keto acid dehydrogenase complex that is required for the metabolism of the branched-chain amino acids [Ref. 13].

Examples of such non-catabolite-repressed genes and operons include, *e.g.*:

- the *ttgDEF* operon of *P. putida* DOT-T1E, which encodes enzymes of a secondary toluene efflux system [Ref. 14];
- the *P. aeruginosa* gene, *gdhB*, which encodes an arginine-inducible NAD(+)-dependent glutamate dehydrogenase [Ref. 15]; and
- the *putA* and *putP* genes of *P. putida* mt-2, which encode enzymes necessary for proline utilization [Ref. 16].

Mutant and Closely Related Sequences of Promoters and Polypeptides

Mutant promoters made from a promoter(s) of a preferred embodiment hereof may also be created using any of the various random and/or directed, mutagenesis techniques known in the art. In a preferred embodiment, site-specific mutagenesis will be performed (*e.g.*, via mutagenic oligonucleotide-directed mutagenesis). In a preferred embodiment, an improved mutant promoter will be selected from a library of mutants made by an error-prone polymerase chain reaction (EP-PCR) performed on a promoter polynucleotide. Multiple rounds of mutagenesis may be performed either upon the pool of polynucleotides resulting from a previous round or upon one or more mutant promoters selected therefrom. Advantageous mutations identified in improved promoters may also be combined to obtain further increases in improvement (*e.g.*, cumulative improvements).

In addition to generating mutant tandem promoters by performing one or more of the techniques described above upon a non-mutant tandem promoter (*i.e.* a tandem promoter in which the individual promoter elements are themselves of native sequence), individual mutant promoters may be used in forming a tandem promoter(s) according to the present invention. For example, two mutant promoters may be linked together, or a mutant promoter and a non-mutant promoter may be linked together, to form a tandem promoter according to the present invention. In addition, directed mutagenesis and/or recombination may be performed (*e.g.*, using a technique such as is described in WO 91/16427) in order to create multiple promoter-promoter combinations in a given round.

Closely related promoters may be obtained by use of polynucleotides containing tandem and/or native promoter constructs and/or elements as hybridization probes, under stringent hybridization conditions, according to any of the various protocols known in the art. An exemplary stringent hybridization protocol is set forth below. Alternatively, a peptide nucleic acid (PNA), or other nucleic acid analog, having a base sequence of such a promoter may be used as a hybridization probe. Preferably the probe will contain a base sequence of at least about 6, at least about 8, at least about 10, at least about 15, at least about 20, at least about 25, at least about 30, at least about 35, at least about 40, at least about 45, or at least about 50 bases in length. In a preferred embodiment, the probe will contain a base sequence

of not more than about 100, about 80, about 60, or about 50 bases in length. In a preferred embodiment, the probe will contain a base sequence of about 20 to about 50, or about 25 to about 45, or about 30 to about 40 bases in length.

In order to perform hybridization probing of target nucleic acids, *e.g.*, target DNA at least suspected of containing a promoter, the target DNA to be probed is denatured, blotted crosslinked onto a nitrocellulose or nylon membrane according to standard protocols (see Sambrook *et al.* [Cold Spring Harbor Press], *Current Protocols in Molecular Biology* [John Wiley and Sons, Inc.]). The blot is then pre-hybridized using standard buffers as described in Sambrook *et al.* or Current Protocols or using a commercially available hybridization buffer such as EXPRESSHYB (BD Biosciences Clontech, Palo Alto, CA). Pre-hybridization may be performed at temperatures ranging from 50-65°C.

The probe to be used (*e.g.*, DNA representing or containing an, *e.g.*, anthranilate or benzoate, promoter fragment) may be labeled to identify specifically bound target DNA and/or the probe nucleic acid may be used as a primer to enzymatically copy specifically bound target DNA. The probe may be labeled according to any of the techniques known in the art. For example, the probe may be labeled with any detectable label, including, but not limited to a: peptide tag, an immunogenic moiety, avidin, biotin, a fluorescent or colored moiety, a detectable chelate, or a radionuclide moiety. In a preferred embodiment, a nucleic acid, preferably DNA, is used as the probe. In a preferred embodiment, the DNA of the probe is labeled. In a preferred embodiment, the label is a radioactive moiety, *e.g.*, a radionuclide-containing compound such as γ-^{32}P dATP. Kits for performing such labeling are commercially available: for example, the HIGH PRIME DNA labeling kit (Roche Molecular Biochemicals, Indianapolis, IN), in conjunction with a radionuclide-containing nucleoside-5'-triphosphate, such as γ-^{32}P dATP, may be used. The probe or labeled probe is then boiled and added to the pre-hybridization buffer. The blot is incubated with the probe at 50-65°C overnight, then washed twice with 2× SSC/0.5%SDS for 5 minutes per wash at room temperature. Then the blot is washed twice with 0.1× SSC/0.1%SDS for 15 minutes per wash at 50-65°C. The blot is then developed as appropriate for viewing the specifically bound labeled probe. For example, if a radionuclide moiety is used as the label on the probe, the blot is used to expose a film or a phosphor screen for viewing.

Alternately, an oligo or set of oligos may be designed that hybridize to known promoter elements (*i.e.*, the –35 and –10 sequences with intervening sequence), or to known activator protein binding sites; a set of degenerate oligos can be designed, at least one of which can hybridize to the target sequence of interest. These may be used as probes for Southern blot analysis as described above, or may be used to initiate synthesis of single (one oligo) or double (two oligos) stranded DNA that may be homologous to the promoter of interest. DNA synthesis may be carried out with, *e.g.*, *Taq* polymerase (with extension carried out at 72°C or as indicated in the manufacturer's protocol), or other polymerase, with buffers supplied by the manufacturer, 1-5mM concentration of primer, and 0.2-1mM final concentration dNTP mix. Annealing temperature can be varied to attain optimal amplification. Extension times for the polymerase may be 20-60 seconds, depending upon length of desired product. A linker could be added to the single-stranded fragment to allow for synthesis of a second strand and amplification, if necessary. Double-stranded fragments may then be sequenced using primers designed for extension/amplification. If restriction sites are also designed onto the oligo, this fragment could subsequently be directly cloned into a standard vector, such as pUC18, *e.g.*, for sequence analysis.

The target nucleic acid to which the probe has specifically bound is then selected by means of selecting probe-target hybrids that have been viewed. In a preferred embodiment, a selected target nucleic acid, will be at least 90% homologous, *i.e.* at least 90% identical in base sequence to the probe or the complementary base sequence thereof (wherein T and U are considered equivalent bases for these purposes). In a preferred embodiment, a selected target nucleic acid will be at least about 95% homologous thereto. In a preferred embodiment, a selected target nucleic acid will be at least about 98% homologous thereto. Where such a target nucleic acid is situated as a portion within a larger polynucleotide molecule, the target nucleic acid, or a fragment containing said target nucleic acid, may be recovered therefrom by any means known in the art, including, *e.g.*, endonuclease digestion and exonuclease digestion.

Alternatively, the base sequence of the probe may be used, in the form of an information string, to perform searching of a nucleotide sequence record, such as a paper or electronic

database record of nucleotide sequences present in polynucleotides containing with a polynucleotide source. The search parameters may specify that a successful match must be 100% identical (100% homologous) or less than 100% identical (heterologous) to the probe information string. Preferably, the search parameters will be selected so that a successful match must be at least 90% homologous, at least 95% homologous, or at least 98% homologous to the probe information string. Preferably, the search parameters will be selected so that a successful match must be heterologous to the probe information string. Once a successful match has been identified, the polynucleotide source corresponding thereto is selected.

Alternatively, a probe information string may be created by altering a first information string representing the nucleobase sequence of a given promoter from a modified information string representing a heterologous nucleobase sequence at least 90% homologous to that of said given promoter. This modified information string may then be used to synthesize a polynucleotide molecule containing the base sequence thereof or may be used to perform searching of a nucleotide sequence record for an identical information string as described above. Upon a successful match, the polynucleotide source corresponding thereto is selected.

Once a polynucleotide at least 90% homologous to the probe sequence is obtained, it is then tested, by forming an expression construct therewith, inserting the expression construct into a transcription system (or transcription and translation system), such as a prokaryotic host cell, and screening the resulting system, *e.g.*, the transformed cell, for the ability of the polynucleotide to direct transcription. Preferably, the screening also involves identifying at least one promoter property improved relative to that of the original promoter.

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 90% will identify those sequences with 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 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 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 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).

In a preferred embodiment, a nucleobase polymer (*e.g.*, a polynucleotide or polynucleotide analog) that is heterologous to, *i.e.* whose base sequence is heterologous to, the base sequence of a given promoter, promoter region, or other non-codon- or non-anti-codon-containing polynucleotide segment, will be at least 90% homologous thereto; preferably about or at least 93% homologous thereto; preferably about or at least 95% homologous thereto; preferably about or at least 96% homologous thereto; preferably about or at least 97% homologous thereto; preferably about or at least 98% homologous thereto; preferably about or at least 99% thereto.

In a preferred embodiment, a polypeptide (or segment thereof) that is heterologous to, *i.e.* whose amino acid sequence is heterologous to, the amino acid sequence of a given polypeptide (or segment thereof) will be at least 90% homologous thereto; preferably about or at least 93% homologous thereto; preferably about or at least 95% homologous thereto; preferably about or at least 96% homologous thereto; preferably about or at least 97% homologous thereto; preferably about or at least 98% homologous thereto; preferably about or at least 99% thereto.

In a preferred embodiment, a nucleobase polymer (or segment thereof) that is heterologous to, *i.e.* whose base sequence is heterologous to, the base sequence of a given codon- or anti-codon-containing polynucleotide (or segment thereof), will be at least 90% homologous thereto; preferably about or at least 93% homologous thereto; even more preferably about or at least 95% homologous thereto; still more preferably about or at least 96% homologous thereto. In a preferred embodiment, such a nucleobase polymer has such a degree of homology to the given codon- or anti-codon-containing polynucleotide that the amino acid sequence encoded by the nucleobase polymer will be at least 90% homologous to the amino acid sequence of the given polynucleotide; preferably about or at least 93% homologous thereto; preferably about or at least 95% homologous thereto; preferably about or at least 96% homologous thereto; preferably about or at least 97% homologous thereto; preferably about or at least 98% homologous thereto; preferably about or at least 99% thereto.

In a preferred embodiment, a nucleobase polymer (or segment thereof) that is heterologous to, *i.e.* whose base sequence is heterologous to, the base sequence of a given codon- or anti-codon-containing polynucleotide (or segment thereof), is about or at least 97% homologous thereto; preferably about or at least 98% homologous thereto; preferably about or at least 99% thereto.

Expression Constructs

In an expression construct, *e.g.*, a gene or operon, according to the present invention, a nucleic acid containing a transcription product-encoding sequence will be operably linked to a promoter according to the present invention, spacer. Where the transcription product is an

mRNA or a precursor molecule thereto, the spacer will be a ribosome-binding-site-containing spacer ("RBS spacer").

A "transcription product-encoding polynucleotide" is any polynucleotide that contains a transcription product-encoding sequence, wherein the transcription product is any functional or structural RNA molecule, including, but not limited to, *e.g.*, mRNA, rRNA, tRNA, cRNA, gRNA, hnRNA, miRNA, mtRNA, nRNA, ncRNA, pRNA, satRNA, scRNA, siRNA, snRNA, snoRNA, srpRNA, stRNA, tmRNA, vRNA, anti-sense RNA (also called "aRNA"), aptamer RNA, chromosomal RNA, enzyme-inhibitor RNA, genetic-control-element RNA, plastid RNA, ribozyme RNA, self-cleaving RNA, self-splicing RNA, telomerase RNA (TER or TERC), X-chromosome-inactivator RNA (XIST RNA), or a precursor RNA of any such RNA molecule. In a preferred embodiment, the transcription product will be an mRNA or a precursor RNA molecule thereto.

Other elements may be included in an expression construct. Such elements include, but are not limited to, *e.g.*: transcriptional enhancer sequences; translational enhancer sequences; leader peptide-encoding sequences, *e.g.*, for intra-cellular-targeting-peptides or secretion signal peptides; pro-peptide-, pre-peptide-, and pre-pro-peptide-coding sequences; other promoters; translational start and stop signals; polyadenylation signals; transcription terminators; introns; and tag sequences, such nucleotide sequence "tags" and "tag" peptide coding sequences (a "tag" facilitates identification, separation, purification, or isolation of an expressed polynucleotide, for which a nucleotide sequence tag is used, or of an expressed polypeptide, for which a "tag" peptide coding sequence is used).

At a minimum, an expression construct according to the present invention will include (in addition to a promoter and either a spacer or an RBS-spacer, operably linked to a transcription product-encoding sequence), a transcriptional terminator. Where the transcription product is an mRNA or pre-mRNA, the expression construct will, at minimum, further include translational start and stop signals operably linked to the transcription product-encoding sequence. The term "operably linked," 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 encoding

sequence, that in and by action of the host cell, the regulatory elements can direct the expression of the coding sequence. Every regulatory element in the expression construct must be "operably linked" to the transcription product-encoding sequence. In cases wherein the cell processes the expression construct before transcription or processes a precursor RNA transcribed from the expression construct, the regulatory element(s) must be so positioned that the cell's processing systems can manipulate the expression construct or the pre-RNA to operably link the regulatory element(s) therein. Likewise, in cases wherein the expression construct is present in the cell in distinct segments of polynucleotide(s), the segments, *i.e.* the polynucleotide molecules or regions collectively containing the regulatory element(s) and transcription product-encoding sequence(s), must be so positioned that the cell can manipulate the segments to create or to re-connect the expression construct wherein the regulatory elements are operatively linked to the transcription product-encoding sequence, and/or are so positioned that the cell's processing systems can manipulate a to-be-transcribed pre-RNA to operably link the regulatory element(s) thereto.

Any prokaryotic ribosome binding site (RBS) may be utilized in such an expression construct. Preferably a bacterial RBS is utilized. In preferred embodiment, an RBS operative in Gram positive bacteria is used; even more preferably an RBS operative in Gram negative bacteria is used. 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); or J.A. Wells *et al.*, Cloning, sequencing, and secretion of *Bacillus amyloliquefaciens* subtilisin in *Bacillus subtilis*, *Nucl. Acids Res.* 11(22):7911-25 (1983) (native RBS sequence of GAGAGG).

Furthermore, one or more marker genes or reporter genes may be used in an expression system to verify expression. Many such useful marker or reporter genes are known in the

art. See, *e.g.*, US Patent No. 4,753,876 to Hemming *et al.*, and DL Day *et al.*, in *J. Bact.* 157(3):937-39 (Mar 1984). In a preferred embodiment, the marker gene is selected from among the antibiotic resistance-conferring marker genes. In a preferred embodiment, the marker gene is selected from among the tetracycline and kanamycin resistance genes. In a preferred embodiment, a reporter gene is selected from among those encoding: (1) fluorescent proteins (*e.g.*, GFP); (2) colored proteins; and (3) fluorescence- or color-facilitating or -inducing proteins, the latter class (3) including, *e.g.*, luminases, alkaline phosphatases, and beta-galactosidases. Alkaline phosphatases hydrolyze BCIP to produce a blue color, and hydrolyze PNPP to produce a yellow color. Beta-galactosidases hydrolyze X-gal to create a blue-colored derivative, and hydrolyze ONPG to produce a yellow color. Fluorescent substrates are also available for alkaline phosphatase and β-galactosidase.

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

A great many bacterial vectors are known in the art as useful for expressing proteins in bacteria, and any of these may be 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 include, but are not limited to, the expression plasmids pMB9, pBR312, pBR322, pML122, RK2, RK6, and RSF1010. 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 useful *Pseudomonas* expression vectors include those listed in Table 3.

Table 3. Some Examples of Useful Expression Vectors

Replicon	Vector(s)
pPS10	PCN39, pCN51
RSF1010	PKT261-3
	PMMB66EH
	PEB8
	PPLGN1
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 a, *e.g.*, pMYC1803. Other exemplary useful vectors include those described in US Patent No. 4,680,264 to Puhler *et al.*

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

A vector can then be transformed into a bacterial host cell.

Transformation

Transformation of the 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.

In addition to the above elements of an expression construct, the bacterial host cell will also contain at least one, and preferably more than one, copy of a gene containing a coding sequence of an activator or repressor protein of the promoter. This gene may be attached to the expression construct, or it may be part of a separate nucleic acid. In a preferred embodiment, an anthranilate activator protein having the amino acid sequence encoded by the complement of the coding sequence shown at nucleotides 4-993 of SEQ ID NO:7 will be utilized; and a benzoate activator protein encoded by nucleotides 225-1229 or nucleotides 285-1229 of SEQ ID NO:1 will be utilized. Preferably, the activator- (or repressor-) encoding gene will be constitutively expressed in the bacterial host cell. When expression of the (*e.g.*, exogenous) coding sequence is desired, the host cell will be contacted with an activator (or de-repressor) compound to induce expression. In a preferred embodiment, the bacterial host cell will be contacted with anthranilic or benzoic acid or a biologically acceptable salt (preferably a sodium salt) thereof, in the case of anthranilate and benzoate promoters, respectively. In a preferred embodiment for a tandem promoter, the bacterial host cell will be contacted with an inducer compound that induces either the natively catabolite-repressed promoter element or the (natively) non-catabolite-repressed promoter element thereof. In a preferred embodiment of a tandem promoter, benzoic acid, anthranilic acid, or biologically acceptable salt(s) (preferably a sodium salt) thereof, will be used as the inducer (activator) compound.

Host Cells

In a preferred embodiment, the host cell in which the promoter is used will be selected from the prokaryotes. In a preferred embodiment, the host cell is selected from the bacteria. In a preferred embodiment, the host cell is selected from the Proteobacteria. In a preferred

embodiment, the host cell is selected from the "Pseudomonads and closely related bacteria" or from a Subgroup thereof, as defined below. In a preferred embodiment, the host cell is selected from the genus *Pseudomonas*. A particularly preferred species of *Pseudomonas* is *P. fluorescens*; even more preferred is *Pseudomonas fluorescens* biotype A.

In a preferred embodiment, both the organism from which the native promoter(s) are obtained and the host cells in which a promoter according to the present invention is utilized, will be selected from the prokaryotes. In a preferred embodiment, both the organism from which the native promoter(s) are obtained and the host cells in which a promoter according to the present invention is utilized, will be selected from the bacteria. In a preferred embodiment, both the bacteria from which the native promoter(s) are obtained and the bacterial host cells in which a promoter according to the present invention is utilized, will be selected from the Proteobacteria. In a preferred embodiment, both the bacteria from which the native promoter(s) are obtained and the bacterial host cells in which a promoter according to the present invention is utilized, will be selected from the Pseudomonads and closely related bacteria or from a Subgroup thereof, as defined below.

In a preferred embodiment, both the promoter source organism and the host cell will be selected from the same species. Preferably, the species will be a prokaryote; more preferably a bacterium, still more preferably a Proteobacterium. In a particularly preferred embodiment, both the promoter source organism and the host cell will 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; more preferably from the genus *Pseudomonas*. Especially preferred is the species *Pseudomonas fluorescens*; even more preferably, *Pseudomonas fluorescens* biotype A.

In a preferred embodiment, the host cells in which the promoter is used will lack biocatalyst(s) effective to degrade the inducer compound: *e.g.*, benzoate or anthranilate or an analog thereof; and/or the degradation product(s) thereof, if any, that is directly responsible for induction; and/or gratuitous inducer compounds. Such host cells are readily obtained as knock-out mutants. For example, the present inventors have found that, in the case of an anthranilate promoter, inactivation of at least the *antA* portion of the host cell's

antABC operon does inhibit the consumption of an anthranilate inducer and thereby permits the inducer to effect lasting induction. The *antA* open reading frame encodes the large subunit of the first enzyme utilized in the pathway for degradation of anthranilate. Similarly, in the case of a benzoate promoter, the inventors have found that inactivation of the *benAB* portion of the host cell's *benABCD* operon, *e.g.*, by deletion or mutation, does inhibit the consumption of a benzoate inducer, thereby improving the level of induction; inactivation of at least the *benA* portion would work similarly, as this encodes the large subunit of the first enzyme utilized in the pathway for degradation of benzoate.

Gene knock-outs may be constructed according to any method known effective in the art. Gene inactivation by insertion of a gene 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). Briefly, a portion of the gene to be disrupted in amplified and cloned into a vector containing a selectable marker, such as an antibiotic resistance gene, that is not able to replicate in the target host. Homologous recombination between the chromosomal copy of the gene and the portion of the target gene contained on the plasmid results in the disruption of the chomosomal copy of the gene and incorporation of the antibiotic resistance marker. Alternatively, transposon mutagenesis and selection for desired phenotype (such as the inability to metabolize benzoate or anthranilate) may 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, a 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.

Pseudomonads and Closely Related Bacteria

The "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)"), and the genus, *Acinetobacter*. Table 4 presents the families and genera of organisms listed in the Bergey taxonomic "Part."

Table 4. 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 thereunder, as well as all Proteobacteria that would be classified thereunder 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.

Also in accordance with this definition, "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.

Likewise in accordance with this definition, "Gram(-) Proteobacteria Subgroup 1" further 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 a preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 1," as defined above.

In a preferred embodiment, the bacteria are 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); *Oceanimonas* (4); and *Acinetobacter* (160).

Exemplary 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); *Oceanimonas doudoroffii* (ATCC 27123); and *Acinetobacter calcoaceticus* (ATCC 23055).

In a preferred embodiment, the bacteria are 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*; *Oceanimonas*; and *Acinetobacter*.

In a preferred embodiment, the bacteria are 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*; *Oceanimonas*; and *Acinetobacter*.

In a preferred embodiment, the bacteria are 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*; *Oceanimonas*; and *Acinetobacter*.

In a preferred embodiment, the host cell is 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*; *Oceanimonas*; and *Acinetobacter*.

In a preferred embodiment, the bacteria are 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*; *Oceanimonas*; and *Acinetobacter*.

In a preferred embodiment, the bacteria are 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*; *Oceanimonas*; and *Acinetobacter*.

In a preferred embodiment, the bacteria are 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*; *Oceanimonas*; and *Acinetobacter*.

In a preferred embodiment, the bacteria are 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*; *Xanthomonas*; and *Acinetobacter*.

In a preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 11." "Gram(-) Proteobacteria Subgroup 11" is defined as the group of Proteobacteria of the genera: *Pseudomonas*; *Stenotrophomonas*; *Xanthomonas*; and *Acinetobacter*.

In a preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 12." "Gram(-) Proteobacteria Subgroup 12" is defined as the group of Proteobacteria of the following genera: *Burkholderia*; *Ralstonia*; *Pseudomonas*.

In a preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 13." "Gram(-) Proteobacteria Subgroup 13" is defined as the group of Proteobacteria of the following genera: *Burkholderia*; *Ralstonia*; *Pseudomonas*; *Xanthomonas*; and *Acinetobacter*.

In a preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 14." "Gram(-) Proteobacteria Subgroup 14" is defined as the group of Proteobacteria of the following genera: *Pseudomonas* and *Xanthomonas*.

In a preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 15." "Gram(-) Proteobacteria Subgroup 15" is defined as the group of Proteobacteria of the genus *Pseudomonas*.

In a preferred embodiment, the bacteria are 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*.

In a preferred embodiment, the bacteria are 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*.

In a preferred embodiment, the bacteria are 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); and *Pseudomonas fluorescens* subsp. *cellulosa* (NCIMB 10462).

In a preferred embodiment, the bacteria are 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 preferred strain of this biotype is *P. fluorescens* strain MB101 (see US Patent No. 5,169,760 to Wilcox), and derivatives thereof.

In a particularly preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 1." In a particularly preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 2." In a particularly preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 3." In a particularly preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 5." In a particularly preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 7." In a particularly preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 12." In a particularly preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 15." In a particularly preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 17." In a particularly preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 18." In a particularly preferred embodiment, the bacteria are selected from "Gram(-) Proteobacteria Subgroup 19."

An expression system according to the present invention can be cultured in any fermentation format. For example, batch, fed-batch, semi-continuous, and continuous fermentation modes of any volume may be employed herein.

In the present invention, growth, culturing, and/or fermentation of the 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. Thus, *e.g.*, the terms "growth" (and "grow," "growing"), "culturing" (and "culture"), and "fermentation" (and "ferment," "fermenting"), as used herein in regard to the host cells of the present invention,

inherently and necessarily means "growth," "culturing," and "fermentation," within a temperature range of about 4°C to about 55°C, inclusive. In addition, "growth" is used to indicate both biological states of active cell division and/or enlargement, as well as biological states in which a non-dividing and/or non-enlarging cell is being metabolically sustained, the latter use of the term "growth" being synonymous with the term "maintenance."

In addition, growth "under conditions permitting expression" when used in regard to the bacterial host cells and expression systems of the present invention, is defined herein to mean: (1) growth of the recombinant bacterial host cells per se, where the promoter used in the control sequence operably linked to the coding sequence is a constitutive promoter; and (2) where the promoter used in the control sequence operably linked to the coding sequence is a regulated promoter, (a) growth of the recombinant bacterial host cells in the presence of (*i.e.* in contact with) an inducer thereof, and (b) growth of the recombinant bacterial host cells in the absence of an inducer thereof, followed by addition of such an inducer to the system, thereby causing contact between the cell and the inducer.

Biocatalyst Preparation

Once the coding sequence(s) under control of the promoter is expressed, the resulting gene product(s) and/or secondary products (*e.g.*, metabolites) resulting from expression of the gene product(s) can be separated, isolated, and/or purified using any recovery and/or purification methods known in the art as useful for such a product, *e.g.*, a protein, nucleic acid, or other molecule. Alternatively, the host cells themselves can be used, *e.g.*, in whole cell bioreactors or in other applications.

EXAMPLES

Materials & Methods

Promoters and Promoter-Plasmid Constructs

The following promoter nucleotide sequences are referred to herein,.

Pben509: nucleotides c994 - c1502 of SEQ ID NO:1.

Pben278: nucleotides g1228 - c1502 of SEQ ID NO:1.

Pben88: nucleotides g1228-c1316 of SEQ ID NO:1 with deletion of g1306.

Pant713: nucleotides c592 - c1304 of SEQ ID NO:7.

Pant705: nucleotides c592 - g1296 of SEQ ID NO:7.

Pant311: nucleotides c994 - c1304 of SEQ ID NO:7.

Pant289: nucleotides 994-1283 of SEQ ID NO:7 with deletion of g1269 and t1278.

AntR+Pant (also Pant+AntR): a1 - c1395 of SEQ ID NO:13.

Ptandem: SEQ ID NO:13.

The following promoterless plasmid constructs are referred to herein.

pDOW1017: No Promoter, but carrying a promoterless *lacZ* reporter gene.

pDOW1033: No Promoter, but carrying a promoterless *phoA* reporter gene.

The plasmid promoter constructs listed in Tables 5 and 6 are referred to herein.

Table 5. Plasmid Individual Promoter Constructs			
Promoter Type	**Plasmid Designation**	**Promoter or Activator-Promoter Sequence Identity**	**Construct Description**
Benzoate	pDOW1028	SEQ ID NO:1 N1228-N1502	Pben278::lacZ
	pDOW1041	SEQ ID NO:1 N1228-N1502	Pben278::phoA
	pDOW1019	SEQ ID NO:1 N994 -N1502	Pben509::lacZ
	pDOW1102	SEQ ID NO:1 N1228-N1316 with deletion of g1306	Pben88::lacZ
	pDOW1081	SEQ ID NO:1 N1228-N1316 with deletion of g1306	Pben88::phoA
	pDOW1083	SEQ ID NO:1 N1228-N1316 with deletion of g1306 and substitution of native –10 'tacggt'1296-1301 by 'tataat'	Pben88(–10con)::phoA
	pDOW1126	SEQ ID NO:1 N1-N1316 with deletion of g1306	BenR-Pben88(–10con)::lacZ
	pDOW1090	SEQ ID NO:1 N1-N1316 with deletion of g1306	BenR-Pben88(–10con)::phoA
	pDOW1100	SEQ ID NO:1 N1228-N1316 with deletion of g1306 and substitution of native –10 'tacggt'1296-1301 by 'taaggt'	Pben88(–10benAc)::lacZ
	pDOW1084	SEQ ID NO:1 N1228-N1316 with deletion of g1306 and substitution of native –10 'tacggt'1296-1301 by 'taaggt'	Pben88(–10benAc)::phoA
Anthranilate	pDOW1039	SEQ ID NO:7 N1-N1304	AntR-Pant
	pDOW1101	SEQ ID NO:7 N994-N1304	Pant311::lacZ
	pDOW1035	SEQ ID NO:7 N1-N1304	AntR-Pant::lacZ
	pDOW1056	SEQ ID NO:7 N1-N1304	AntR-Pant::phoA
	pDOW1029	SEQ ID NO:7 N592-N1304	Pant713::lacZ
	pDOW1095	SEQ ID NO:7 N1-N1283 with deletion of t1278	AntR-Pant(–10wt)::phoA
	pDOW1082	SEQ ID NO:7 N994-N1283 with deletion of g1269 and t1278 and substitution of native –10 region 'ttaat'1264-1268 by consensus -10 region 'tataat'	Pant289(–10con)::phoA
	pDOW1098	SEQ ID NO:7 N1-N1283 with deletion of g1269 and t1278 and substitution of native –10 region 'ttaat'1264-1268 by consensus -10 region 'tataat'	AntR-Pant289(–10con)::phoA
Mosaic	pDOW1099	SEQ ID NO:7 N1-N1283 with deletion of g1269 and t1278 and substitution of native Pant –10 region 'ttaat'1264-1268 by native Pben -10 region 'tacggt'	AntR-Pant289(–10ben)::phoA

Table 6. Plasmid Tandem Promoter Constructs		
Plasmid Designation	**Promoter or Activator-Promoter Sequence Identity**	**Construct Description**
pDOW1057	SEQ ID NO:13	AntR-Ptandem::lacZ
pDOW1111	SEQ ID NO:13 N1085-N1541	Pant311-Pben278::lacZ
pDOW1107	SEQ ID NO:13 N1-N1518 with deletion of g1508	AntR-Ptandem[Pben88(–10wt)]::lacZ
pDOW1108	SEQ ID NO:13 N1-N1518 with deletion of g1508 and substitution of Pben native –10 'tacggt' 1498-1503 by 'tataat'	AntR-Ptandem[Pben88(–10con)]::lacZ
pDOW1109	SEQ ID NO:13 N1-N1518 with deletion of g1508 and substitution of Pben native –10 'tacggt' 1498-1503 by 'taaggt'	AntR-Ptandem[Pben88(–10benAc)]::lacZ

The oligonucleotides listed in Table 7 are utilized in the following examples.

Table 7. Oligonucleotides Used Herein	
Primer Name	**Sequence (all listed 5' → 3')**
AntAKO5	GGAATTCTTCGTGACGATGCG (SEQ ID NO:16)
AntAKO3	CGGGATCCGCTCGCGATGCTGC (SEQ ID NO:17)
lacZPE	GGATGTGCTGCAAGGC (SEQ ID NO:18)
lacZPE2	GTAACCATGGTCATCGC (SEQ ID NO:19)
M13forward	GTAAAACGACGGCCAGT (SEQ ID NO:20)
M13reverse	AACAGCTATGACCATG (SEQ ID NO:21)
Bambenwtshort	CGGGATCCGTATCAGGCGCCTCACCGTACGTGCTC (SEQ ID NO:22)
Bambenconshort	CGGGATCCGTATCAGGCGCCTCATTATACGTGCTC (SEQ ID NO:23)
BambenAcshort	CGGGATCCGTATCAGGCGCCTCACCTTACGTGCTC (SEQ ID NO:24)
Bamantwtshort	CGGGATCCGCTAACGGTGAGCCATTAAGCGGCTGC (SEQ ID NO:25)
Bamantconshort	CGGGATCCGCTAACGGTGAGCATTATAGCGGCTGC (SEQ ID NO:26)
BenactKO-for	CGCGACACATTGCTGCCCAG (SEQ ID NO:27)
BenactKO-rev	AGTATCAGCCATCGCACCTT (SEQ ID NO:28)
1803H3seq	GTCCTGCAATTTCAGCCGA (SEQ ID NO:29)
BenL278	CCTTAATTAAGTTAAGCGACGTGCGC (SEQ ID NO:30)
3'Antactiv	CCCAAGCTTCTATCGAGGCAAGCCAG (SEQ ID NO:31)
Benact5'	AGCTTTGTTTAAACGCATGACGTTGTTGATTC (SEQ ID NO:32)
H3_5'BenAKOclean	CCCAAGCTTGCCATGAGGCGGAAAACGCTGC (SEQ ID NO:33)
H3_3'BenBKOclean	CCCAAGCTTCGGTGATCGCCACGCTGTCGC (SEQ ID NO:34)
BenKOmega	CATACGTCATGGCCCTCCGTTGTTC (SEQ ID NO:35)
InvbenKOmega	GAACAACGGAGGGCCATGACGTATG (SEQ ID NO:36)
5'BenA_Seq	CTGCTGGAAAACGCCTGCCTGGAG (SEQ ID NO:37)
Seq_3'BenB	GAGCACTTCAAGCATCGACAGGAAC (SEQ ID NO:38)
1261-8378F	CTTCAGATCCAGACTCACCAG (SEQ ID NO:39)
1261-103R	GACCATGATTACGCCAAGCGC (SEQ ID NO:40)
M13R21	CACACAGGAAACAGCTATGAC (SEQ ID NO:41)

Host Cells:

E. coli JM109 (obtained from Promega Corp.), E. coli TOP10 (obtained from Invitrogen Corp.), and *Pseudomonas fluorescens* biotype A (strains MB101 and MB214). *P. fluorescens* MB214 is a derivative of strain MB101 (a wild-type prototrophic *P. fluorescens* biovar A). MB214 had been prepared by integrating the *E. coli lacIZYA* operon (deleted of

the *lacZ* promoter region) into the chromosome of strain MB101 to provide a host cell wherein the lac promoter and its derivatives can be regulated by lactose or IPTG to drive inducible expression of transgenes of interest. The MB101 strain is Lac(-) whereas the MB214 strain is Lac(+).

Inducer Compounds

As used in the Examples below, an "anthranilate" inducer means sodium anthranilate, and a "benzoate" inducer means sodium benzoate.

Transformation Protocols

E. coli: Transformations of *E. coli* were performed as per the manufacturer's protocol, using strain JM109 chemically competent cells from Promega (Madison, Wis.).

P. fluorescens: Electroporation of *P. fluorescens* was performed by subculturing 1mL of an overnight culture (grown in rich medium; the present examples used Luria-Bertani Broth, Miller (*i.e.* LB Broth, Miller) (available from Difco, Detroit, Mich.) into 50 mL LB Broth, Miller and incubating at 30°C with shaking until an A600 measurement falls within the range of 0.4 - 0.6. The resulting cells were washed twice with 50 mL cold ddH2O and resuspend in 1 mL cold ddH2O. To 100 µL aliquots of competent cells were added approximately 10ng of a plasmid of interest, in a 0.2cm gap electroporation cuvette (Bio-Rad Laboratories, Inc., Hercules, Cal.). Electroporation was performed under the following conditions: 200 Ohms, 25 µF, 2.25kV. This was followed by the addition of 1 mL cold LB broth. Cells were permitted to recover on ice for 2 minutes, then incubated at 30°C, with no shaking, for 2 hours to overnight. Cells were then plated on selective medium; the present examples used LB agar Miller (Luria-Bertani) (available from Difco, Detroit, Mich.), supplemented with 15 µg/mL tetracycline (Fisher Scientific, Pittsburgh, PA) as the selective medium.

Cell Growth Protocols:

Cell growth for induction: Strains of interest were grown overnight (at 30°C with shaking at 250rpm) in 1× M9 minimal salts medium (diluted from a 5× preparation purchased from Fisher Scientific, Pittsburgh, Pennsylvania) supplemented with 0.5% or 1% (w/v) glucose,

1mM $MgSO_4$, and trace elements (for trace elements, the present examples used a solution containing salts of sodium, magnesium, manganese, iron, and cobalt, all at less than 0.5mg/mL final concentration). Strains were then subcultured 1:4 in the same medium to a volume of 10 or 20 mL and then induced with 0-10mM concentrations of anthranilate or benzoate, as indicated.

Cell growth for plasmid propagation: *E. coli* cells containing a plasmid of interest were grown overnight in 50-200 mL of LB Broth, Miller, supplemented with 15 μg/mL tetracycline or 100 μg/mL ampicillin (depending on the plasmid to be isolated) at 37°C, with shaking at 250rpm. Plasmids preparations were performed using the NUCLEOSPIN kit (plasmid DNA purification "miniprep" kit for use with culture volumes up to 5 mL; available from BD Biosciences Clontech, Palo Alto, Cal.) or the NUCLEOBOND kit (plasmid DNA purification "midiprep" kit for use with culture volumes up to 200ml; available from BD Biosciences Clontech, Palo Alto, Cal.).

Induction Protocols:

Strains of interest were grown overnight at 30°C in 1× M9 medium supplemented with 0.5% or 1% (w/v) glucose, 1mM $MgSO_4$, and 5L/L trace elements (as described above), and optionally tetracycline at 15 ug/mL. These were then subcultured 1:4 or 1:5 in the same medium and then induced with indicated concentrations of anthranilate, benzoate, or other inducer, for a desired amount of time (*e.g.*, for 2, 4, 6, 8, 12, or 24 hours, or overnight). Samples were taken at indicated times and those samples were assayed for reporter gene activity. Results are reported at time points taken at a given number of hours post-induction; time points are indicated by either a numeral for the number of hours, and in some cases this number is immediately preceded by the letter "I" indicating post-induction.

EP-PCR Protocol

The following protocol was used for error-prone PCR mutagenesis (see "Mutagenesis of Cloned DNA," in F.M. Ausubel, *Current Protocols in Molecular Biology* on CD-ROM (2002) (John Wiley & Sons, New York, NY)). The following reagents were combined: 63 μL water, 10 μL 0.1M Tris (pH 8.3), 5 μL 1M KCl, 0.7 μL 1M $MgCl_2$, 4 μL dNTP mix (either mix #1 [25mM dCTP, 25mM dTTP, 5mM dATP, 5mM dGTP] or mix #2 [20mM

dCTP, 20mM dTTP; 2mM dATP, 2mM dGTP]), 2 μL 100 μM M13forward primer (GTAAAACGACGGCCAGT) (SEQ ID NO:16), 2 μL 100μM M13reverse primer (AACAGCTATGACCATG) (SEQ ID NO:17), 1 μL (~5ng) template (consisting of a Pant or Pben promoter polynucleotide cloned into pNEB193, a plasmid available from New England BioLabs, Beverly, Mass.), 2 μL 25mM $MnCl_2$, and 1 μL *Taq* polymerase (5 Units/μL, obtained from Invitrogen Corp.). PCR conditions were as follows: 94°C for 3 min.; 30 cycles of 30 sec. at 94°C, 30 sec. at 50°C, and 90 sec. at 72°C; hold at 4°C. PCR products were purified using MICROCON YM-100 or MICROCON-PCR columns (nucleic acid purification columns, Millipore Corp., Bedford, Mass.) according to the manufacturer's instructions for AMICON devices. Products were digested with *Bam*HI and *Hin*dIII (New England BioLabs) in 1× NEBUFFER BAMH I + BSA (*Bam*H I restriction endonuclease buffer, available from New England BioLabs) and purified by gel extraction using either QIAEX II (gel extraction kit, from Qiagen, Valencia, Cal.), for fragments of 300bp and smaller, or PREP-A-GENE (DNA purification kit, from Bio-Rad Laboratories), for fragments larger than 300bp. The fragments were then cloned upstream of the *lacZ* or *phoA* reporter gene of pDOW1017 or pDOW1033, respectively.

Knock-Out Protocols

Construction of AntA Knock-Out. An internal fragment of the *antA* gene was amplified using primers AntAKO5 (G*GAATTC*TTCGTGACGATGCG) (SEQ ID NO:12) and AntAKO3 (CG*GGATCC*GCTCGCGATGCTGC) (SEQ ID NO:13) from *P. fluorescens* genomic DNA (*Eco*RI and *Bam*HI sites, respectively, shown in italics). The reaction mixture was formed by combining: 5 μL 10× buffer (*i.e.* the buffer supplied by Invitrogen with the *Taq* polymerase, which buffer contained 200mM Tris-HCl (pH 8.4) and 500mM KCl) , 2.5 μL 50mM $MgCl_2$, 1 μL 10mM dNTPs, 0.5μL 100μM AntKO3, 0.5 μL 100μM AntKO5, 1 μL (5 Units/μL) *Taq* polymerase (Invitrogen Corp., Carlsbad, Cal.), 0.5 μL *P. fluorescens* MB214 genomic DNA (~50ng), and 39 μL ddH_2O. The PCR cycle conditions used were: 2 min. at 96°C; 30 cycles of 30 sec. at 96°C, 30 sec. at 52°C, and 30 sec. at 72°C. The resulting PCR product was cloned into a plasmid unable to replicate in *P. fluorescens* (a pUC type plasmid was used, though, *e.g.*, pBR type plasmids will also work). The resulting plasmid was transformed into electrocompetent *P. fluorescens* cells, and the transformants were selected with the appropriate antibiotic. Since the plasmid cannot

replicate in *P. fluorescens*, only those bacteria which have the plasmid integrated at the *antA* locus, resulting in two truncated *antA* ORFs separated by the plasmid backbone, can be selected. Several transformants were cultured in M9 medium + 1.0% glucose, 5mM anthranilate, at 30°C with shaking for 24 hours, and culture supernatants were analyzed by HPLC for anthranilate concentration.

Construction of BenAB Knock-Out. Generally, following the above-described method for deletion of the AntA gene in *P. fluorescens*, the plasmid pDOW1139 was constructed to facilitate deletion of the *benAB* genes as follows. The 3' portion of the *benR* gene and the 5' portion of the *benC* gene were amplified using *P. fluorescens* MB214 genomic DNA as template. The *benR* region was amplified using primers H3_5'benAKOclean and BenKOmega. The *benC* region was amplified using primers H3_3'BenBKOclean and InvbenKOmega. For both reactions, the cycling conditions were 95°C for 5 minutes; (94°C, 30 seconds; 55°C, 30 seconds; 72°C, 1 minute) for 35 times; then 72°C for 5 minutes. This reaction was performed using *Taq* polymerase (Invitrogen) according to the manufacturer's protocol. The *benR* and *benC* fragments were fused using primers H3_5'benAKOclean and H3_3'benBKOclean, with both fragments as template. This fusion reaction employed KOD HOTSTART DNA polymerase (Novagen) under conditions of 94°C for 2 minutes; (94°C, 30 seconds; 50°C, 30 seconds; 68°C, 1.5 minutes) for 35 times; then 68°C for 5 minutes. The expected 1.1kb fragment was gel purified using QIAEX II (Qiagen) and cloned into *Srf*I-digested plasmid DNA to form plasmid, pDOW1139. pDOW1139 was then transformed into *P. fluorescens*). Transformants were selected by plating on LB medium with tetracycline for selection. Since the plasmid could not replicate in *P. fluorescens*, colonies resistant to tetracycline arose from the plasmid being integrated into the chromosome. The site of integration of the plasmid was analyzed by PCR. To obtain strains that lost the integrated plasmid by recombination between the homologous regions, single colonies of the first transformants were inoculated into liquid LB medium, grown overnight, and then plated onto selective medium to counterselect for loss of the plasmid (data not shown). Isolates having the expected phenotype were selected. DNA from the resulting strains was analyzed by PCR to confirm removal of the *benAB* region using primers 5'BenA_seq, Seq_3'BenB, M13R21, 1261-8378F and 1261-103R.

Inactivation of theP. fluorescens Chromosomal BenR Gene. The open reading frame (ORF) upstream of the *benA* gene e Fig.). A DNA fragment containing a portion of the ORF was amplified by PCR using the BenactKOfor and BenactKOrev primers, and *P. fluorescens* MB214 genomic DNA as template. Recombinant Taq polymerase (from Invitrogen Corp.) was used according to the manufacturer's protocol. The cycling profile [94°C for 2 min.; (94°C for 30sec, 62°C for 30sec, 72°C for 30sec) for 30 cycles; then 72°C for 7min.] was used. The resulting products were cloned into the pCR2.1 vector (form Invitrogen Corp.) and transformed into *E. coli* Top10 cells. Transformants were screened for insert by colony PCR using the above primers/ conditions, and the positive clones were further confirmed by DNA sequencing. The resulting plasmids were then used to insertionally inactivate the corresponding chromosomal ORFs. DNA samples were prepared using a NUCLEOBOND plasmid midiprep kit (from Clontech Corp.) and 4 μg of plasmid DNA was transformed into *P. fluorescens* strain MB101. The resulting transformants were screened again by colony PCR. To do this, putative knockout clones were picked into 20 μl H_2O and incubated at 100 °C for 10 min. PCR was performed on the DNA of the resulting lysed cells, using PCR reaction conditions of: 20 μl pre-incubated clone, 5 μl 10X buffer, 3 μl 25 mM $MgCl_2$, 1μl 10 mM dNTP, 5 μl 5 μM BenactKO-for, 5 μl 5 μM M13F (-40) and M13R (-21), 0.5 μl Taq polymerase (5U/μl; from Promega Corp.), and 5.5 μl H_2O. PCR reaction cycle conditions used were: 94 °C for 1 min; (94 °C, 1 min; 50 °C, 30 sec; 72 °C. 2 min.) for 30 times; then 72 °C for 10 min, followed by 4 °C hold. MB101 genomic DNA and pDOW1125 were used as controls. Inactivation of this BenR gene resulted in inability of the knock-out host cells to activate transgenic Pben-reporter gene constructs, as well as inability to metabolize benzoate.

Site-Directed Mutagenesis Protocol

Oligonucleotides used for site directed mutagenesis are found listed among SEQ ID NOs:16-41. Construction of the Pben $^{-}$10 promoter mutants was conducted as follows. The plasmid pDOW1022 was used as template for polymerase chain reaction (PCR) with 1uM primer benL278 and 1uM of bambenconshort, bambenwtshort, or bambenAcshort. Recombinant Taq polymerase (from Invitrogen Corp.) was used according to the manufacturer's instructions. The reaction cycling protocol was 94°C for 2 min.; (30 sec at 94°C, 30 sec at 62°C, and 30 sec at 72°C) for 25 times; then 72°C for 7 min. The resulting

products were cloned into the pCR2.1 vector (Invitrogen Corp.) and transformed into E. coli TOP10. The insert containing the mutated promoter was digested with *Bam*HI and *Pac*I and subsequently ligated to pDOW1033 digested with the same restriction enzymes yielding plasmids pDOW1081 and 1083-1084, which have a promoter::*phoA* transcriptional fusion. These plasmids were used as templates to re-amplify the mutant promoters using the primer 1803H3seq and either bambenconshort, bambenwtshort or bambenAcshort and using a recombinant Taq polymerase (from Promega Corp.), according to the manufacturer's instructions. Reaction cycling conditions were 94°C for 1min., (1 min at 94°C, 30 sec at 50°C, and 1 min. at 72°C) for 30 times; then 72°C for 7min. The resulting products were digested with *Hind*III and *Bam*HI, and subsequently ligated to pDOW1017 that had been digested with the same restriction enzymes. This resulted in formation of promoter::*lacZ* fusions pDOW1102, 1106 and 1100.

Construction of Pant-10 promoter mutants was conducted as follows. The plasmid pDOW1039 was used as template for PCR with 1uM primer 3'Antactiv and 1uM of primer bamantwtshort or bamantconshort. Recombinant Taq polymerase (from Invitrogen Corp.) was used according to the manufacturer's instructions. The reaction cycling protocol was 94°C for 2 min.; (30 sec at 94°C, 30 sec at 60°C, and 30 sec at 72°C) for 25 times; then 72°C for 7 min. The resulting products were digested with *Hind*III and *Bam*HI, and cloned into the same sites of pDOW1033: plasmids pDOW1095 and 1098 contain *antR*-Pant::*phoA* fusions, with variations of the –10 region of the promoter.

DNA Sequencing Protocol

Cloned inserts were sequenced using ABI PRISM BIGDYE V2.0 or V3.0 DNA sequencing kit from (Applied Biosystems, Inc., Foster City, Cal.) as follows: 4 μL of premix (containing buffer, *Taq* polymerase, and dye terminators, as supplied in the Applied Biosystems kit), 50 fmol of plasmid template, 3.2-5 pmol of desired sequencing primer, and 2 μL of 5× buffer (as supplied in the Applied Biosystems kit) were combined (to a final volume of 20 μL). The PCR cycling profile used was: 45 cycles of 30 sec. at 95°C, 20 sec. at 50°C, and 4 min. at 60°C. Samples were purified using SEPHADEX G-50 (a bead-form, dextran gel for chromatographic purification of nucleic acids, from Sigma Chemical Company, St. Louis, Missouri), dried, resuspended in formamide, and then run on an

ABI3100 automated DNA sequencer (a 16 capillary array, automated DNA sequencer, from Applied Biosystems, Inc.).

Primer Extension Protocol

RNA Isolation. An RNA isolation procedure was followed in order to identify the transcription start sites under the control of the *P. fluorescens* Pant and Pben promoters. The procedure used is as follows. An overnight culture of *P. fluorescens* MB101 carrying the appropriate plasmid was grown in 1× M9 medium supplemented with 1% glucose (w/v), 1mM $MgSO_4$, and trace elements (as described above) was subcultured 1:4 (v/v) in the same medium to a final volume of 50 mL. The culture was induced with 5mM benzoate or anthranilate as appropriate for 8 or 24 hours. Cells were pelleted and total RNA isolated using an RNEASY kit (a "maxi" bacterial RNA isolation kit from Qiagen, Valencia, Cal.). The RNA was resuspended to a final volume of 200 µL and treated with 10 Units of DNAse I (ribonuclease-free, from Ambion, Inc., Austin, Texas) according to manufacturer's protocol. Following DNAseI treatment, the RNA was purified using an RNEASY column (a "midi" or "mini" RNA purification column, from Qiagen) as appropriate (the RNEASY "midi" column was used for RNA amounts up to 1 mg; the RNEASY "mini" column was used for RNA amounts up to 100 µg). Once purified, the RNA concentration was determined using RIBOGREEN (RNA quantitation kit, from Molecular Probes, Inc., Eugene, Oregon), following the manufacturer's protocol.

Primer Labeling: This was performed by mixing 1 µL 10µM primer (either lacZPE, GGATGTGCTGCAAGGC (SEQ ID NO:14), or lacZPE2, GTAACCATGGTCATCGC (SEQ ID NO:15)), 1 µL 10× T4 polynucleotide kinase buffer (700mM Tris-HCl (pH 7.6), 100mM $MgCl_2$, 50mM dithiothreitol (DTT)), 5 µL ^{32}P-γATP (50µCi, Amersham-Pharmacia), 1 µL T4 kinase (New England BioLabs), and 2 µL ddH_2O; and incubating the resulting reaction mixture at 37°C for 30-60 min. Following incubation, 5 µL of the reaction mixture was reserved to use for a "sequencing ladder" analysis. 20 µL TE (10mM Tris, 1mM EDTA (pH8.0)) was added to the other 5 µL and mixed and the result was spun through a MICROSPIN G-25 column (Amersham-Pharmacia, Piscataway, New Jersey) to remove unincorporated nucleotides, thereby yielding a final concentration of 0.2µM labeled primer.

Sequencing ladder: This was performed according to the protocol that came with the FMOL kit (DNA sequencing kit from Promega Corp.), using 1 picomole (pmol) of the labeled primer described above. Plasmid template used corresponds with that contained in the strain from which RNA was isolated for the extension reaction.

Primer Extension reaction: Primer extension reactions were performed by mixing 10-20 μg of total RNA with 0.2 pmol primer to yield a final volume of 12 μL, followed by incubation at 70°C 10 min. Then, the following were added: 4 μL 5× SUPERSCRIPT II buffer (250mM Tris-HCl (pH 8.3), 375mM KCl, 15mM $MgCl_2$, available from Life Technologies, now Invitrogen Corp., Carlsbad, Cal.), 2 μL 1M DTT, 1 μL 10mM dNTPs, and 1 μL SUPERSCRIPT II (reverse transcriptase, from Life Technologies, now Invitrogen), followed by incubation at 42°C for 1hour. Then the resulting mixture was treated by either an addition of 5 μL sequencing stop solution (containing formamide and tracking dye, as supplied in the Promega FMOL kit) or, in those cases where the signal was weak, by: precipitation with 2 μL 3M sodium acetate/40 μL 100% ethanol, followed by centrifugation for 10 minutes to pellet suspended matter, drying of the pellet, and resuspension in 4 μL H_2O + 2 μL sequencing stop solution. The product mixture resulting from the primer extension reaction was then electrophoresed on a LONG RANGER gel (made from 6% pre-mixed gel solution, from Biowhittaker Molecular Applications, Rockland, Maine) containing 8M Urea and 1.2× TBE (*i.e.* Tris-Borate-EDTA, as diluted from 10× TBE obtained from Fisher Scientific, Pittsburgh, Pennsylvania) next to the sequencing ladder, with 0.6× TBE as an electrophoretic "running" buffer. The gel was dried and exposed to a phosphor screen (from Molecular Dynamics, now Amersham Biosciences, Inc., Piscataway, New Jersey) to detect radiolabeled DNA fragment, and imaged on the TYPHOON PHOSPHORIMAGER (Molecular Dynamics, now Amersham Biosciences, Inc., Piscataway, New Jersey).

Primer extension using Thermoscript reverse transcriptase: 30 ng total RNA, 1 μL 0.2μM primer, and ddH_2O to a final volume of 12 μL were mixed and then incubated at 70°C for 10 min. To this mixture were added 4 μL 5× cDNA synthesis buffer (250mM Tris acetate (pH 8.4), 375mM potassium acetate, 40mM magnesium acetate), 1 μL 0.1M DTT, 2 μL

10mM dNTPs, 1 μL THERMOSCRIPT RT (reverse transcriptase from Invitrogen Corp.), and the resulting mixture was incubated at 55°C for 1 hour. The reaction product was precipitated, dried, and resuspended in 4 μL ddH2O + 2 μL stop solution (described above). All reactions were heated at 70°C for 2 minutes immediately before being loaded onto the gel as described above. The gel was run as described above.

Microtiter β -Galactosidase Assay

We prepared enough of the following assay medium to provide for each sample well of a 96-well plate (*i.e.* for all those wells used, with at least one well being used for each time point measured during the reaction course for each sample): 152 μL Z buffer (0.06M $Na_2HPO_4 \cdot 7H_2O$, 0.04M $NaH_2PO_4 \cdot H_2O$, 0.01M KCl, 0.001M $MgSO_4 \cdot 7H_2O$) + 8μL 1M β-mercaptoethanol. For each 900 μL of the resulting mix, we added one drop of 0.1%SDS and two drops of $CHCl_3$, mixed (using a vortex-type mixer), and then added 144 μL thereof to each well. 16 μL of cells were then added to each well and the plate sealed with a plastic plate sealer. The plate was then mixed (by vortex) for 10 seconds, and then equilibrated to incubation temperature (room temperature) for 5 minutes. 50 μL 4mg/mL ONPG was then added. When a significant yellow color developed, 90 μL stop solution (1M Na_2CO_3) was added and the reaction time recorded. The resulting color intensity for each sample was then read at A420 and A550. In addition, the cell density of each culture providing the 16 μL of cells used in each sample was read at A600. Miller Units were calculated as follows:

1000 * ((A420 – (1.75*A550))/(time(in minutes) * 0.1 * A600)).

Alkaline Phosphatase Assay

For this assay we prepared SIGMA FAST (p-nitrophenyl phosphate (PNPP) substrate, from Sigma-Aldrich Corp., St. Louis, Missouri) by adding one of each tablet provided by the manufacturer (one table each PNPP and Tris; stored at –20°C) to 20 mL ddH_2O, giving a final concentration of 1 mg/mL PNPP and 0.2M Tris. At each time point, for each sample, 50 μL SIGMA FAST substrate was combined with 5 μL of cells. The result was then incubated at room temperature for 30 minutes. The resulting color intensity for each sample was then read at A410. In addition, the cell density of each culture providing the 5 μL of cells used in each tested sample was read at A600 (*i.e.* the cell cultures were read in a 96

well plate). The value of A410/(0.1 * A600) was then calculated to express alkaline phosphatase activity/cell.

Example 1. **Cloning and Analysis of a Benzoate-Inducible Promoter**

Benzoate is an inexpensive, essentially nontoxic compound, making it an ideal candidate for an inducer. A 509bp region of *P. fluorescens* DNA was cloned. This region was found located upstream of a putative *benA* translational start site (Figure 5), which was part of the coding sequence of a subunit of benzoate dioxygenase. The cloned region was found to contain a benzoate-inducible promoter (Pben), and was named "Pben509".

Benzoate-inducible promoter activity was tested by fusing the DNA fragment containing the putative promoter sequence of Pben509, or of Pben278 (described below), upstream of an easily assayable reporter gene (*i.e.* either *lacZ*, which encodes β-galactosidase and was used as the chief reporter gene, or *phoA*). The resulting plasmid was transformed into *P. fluorescens* MB101. Following addition of sodium benzoate, induction of β-galactosidase activity was measured using the chromogenic substrate o-nitrophenol-β-D-galactopyranoside (ONPG) (see Figure 6). Similar experiments were carried out using the *phoA* reporter gene and the chromogenic substrate p-nitrophenyl phosphate (PNPP). *P. fluorescens* strains carrying these constructs show β-galactosidase or alkaline phosphatase activity, respectively, upon addition of 1-10mM sodium benzoate. Varying the concentration of the inducer and/or the time of induction resulted in varying levels of reporter gene expression (data not shown).

A truncated version of the promoter-plus-reporter gene construct, containing a 275bp portion upstream of the predicted translational start site (Figure 5), which portion was named "Pben278", was found to retain activity similar to that of Pben509 (see Figure 6).

Both Pben 509 and Pben278 promoter activity was found to be inhibited during fermentation, due to the presence of a small, but significant concentration of glucose. Thus, these promoters are catabolite-repressed.

Northern analysis indicated that expression from Pben occurred only upon addition of the inducer compound (*e.g.*, sodium benzoate), demonstrating that inducible expression of Pben is not leaky like that of the lac family of promoters (data not shown). Primer extension analysis of total RNA isolated from induced cultures of MB101 carrying either Pben509 or Pben278 fused to *lacZ* indicated that the transcriptional start site was 196 nucleotides (nt) upstream of the predicted *benA* translational start site. This indicates that the promoter sequence and the positive regulatory *cis* acting elements are contained within 82bp upstream of the transcriptional start in the Pben278 clone.

The literature teaches that *cis,cis*-muconate, a benzoate metabolite, acts to induce the *benABCD* operon of other bacteria such as *Acinetobacter* sp. and *P. putida*. However, both cis,cis-muconate and the presumed preceding compound in the known metabolic pathway for benzoate degradation, *i.e.* catechol, fail to induce activity of either Pben509 or Pben278 (data not shown). As a result, either benzoate or an initial benzoate derivative, *e.g.*, 2-hydro-1,2-dihydroxybenzoate, may be directly responsible for inducing the benzoate promoter.

Example 2. **Cloning and Analysis of an Anthranilate-Inducible Promoter**

Anthranilate, like benzoate, is an inexpensive low toxic compound that can be utilized by *P. fluorescens*, making it an ideal compound to investigate as an inducer. Four promoter constructs have been cloned upstream of either a *lacZ* or *phoA* reporter and have been found to possess similar activity upon induction with anthranilate: Pant713, Pant705, Pant311, and Pant+antR coding sequence (CDS) (Figure 7). Pant713 and Pant705 have the same 5' end, but Pant713 contains the predicted ribosome binding site of the antA gene, whereas Pant705 does not (see the underlined CCTCC in the final octamer shown for Pant713). In an effort to determine the minimal region of DNA necessary for anthranilate-induced activation, the Pben713 construct was truncated on the 5' terminus to 311 base pairs (bp). The Pant311 construct was found to retain activity similar to that of Pant713 (data not shown). Expansion of the promoter clone to include the transcriptional activator gene 5' of the *antA* open reading frame (ORF) increased expression levels of the *lacZ* fusion. The transcriptional start site was mapped to 31 nucleotides upstream of the predicted *antA* translational start site, for both Pant713 and the expanded clone that includes the

transcriptional activator AntR (data not shown; also see Figure 7). The presence *antR* in multi-copy with the *lacZ* fusion was found to enable faster and stronger induction (see Figure 8).

In addition, further increasing expression of the AntR has been found to result in more improved anthranilate-inducible expression by Pant promoters. For example, as shown in Figure 8, both pDOW1029- and pDOW1035-constructs were induced in *P. fluorescens* host cells. Figure 8 demonstrates a substantial difference in the rates of induction and the maximum levels of induction achieved for each of these promoters during the 24 hour time course of the study. pDOW1029 contains the Pant713 promoter, which lacks the *antR* coding sequence; pDOW1035 contains the full activator CDS. The *P. fluorescens* host cell used contains an actively expressed, chromosomal copy of the *antR* CDS. Thus, the results shown in Figure 8 for pDOW1029 are for a system in which the *antR* gene is present in a single copy, while the results for pDOW1035 represent a two-copy system. These results demonstrate that the presence of an extra copy of the *antR* gene dramatically improves both the rate and level of response of the Pant promoter. Such improved expression can alternatively be obtained, or further enhanced, by driving *antR* expression with a very strong promoter. Improved induction/expression can also be obtained, as described below, by using a host cell in which a key gene (*e.g.*, antA) responsible for degradation of the inducer compound (*e.g.*, anthranilate) has been inactivated. Moreover, mutating the activator and/or promoter sequences (and selecting for mutants with increased activity) could also enhance the activator/promoter interaction and thereby allow for more improved anthranilate-inducible expression by Pant.

The anthranilate promoter was also found to be inducible by anthranilate analogs, including the halo-substituted anthranilic acid derivatives: 3-chloro-, 4-chloro-, 5-chloro- and 6-chloro-anthranilate. 6-chloroanthranilate is found to act as a gratuitous inducer of anthranilate metabolism, *i.e.* it is not metabolized by *P. fluorescens* yet induces expression from the anthranilate promoter. For example, 6-chloroanthranilate was found to induce the Pant713 and antR/Pant constructs (Figure 8). Taken together, these results indicate that anthranilate itself induces the metabolic pathway; and that it is possible to utilize substituted anthranilate compounds as gratuitous inducers as an alternative to inactivating the

anthranilate metabolic pathway of the host organism.

Example 3. **Construction and Testing of a Fused, Pant-Pben Tandem Promoter**

The relative strength of the Pant promoter with multi-copy *antR* was found to be approximately 1/5 that of the Pben278 promoter. However, unlike the catabolite-repressed Pben promoter, the activity of the Pant promoter was not inhibited during fermentation. A fusion of these two promoters was created by linking them together, as shown in SEQ ID NO:3, *i.e.* by cloning a fragment *antR* and Pant, upstream of the Pben278 promoter fused to *lacZ*.

The strength of the tandem "'*antR*/Pant' – 'Pben278'" construct, induced with anthranilate, was surprisingly found to be improved over that of "*antR*/Pant" alone upon induction with anthranilate. The strength of the tandem promoter upon induction with benzoate was found to be similar to that of Pben278 alone (Figure 9). The induction of greater β-galactosidase activity from the tandem promoter, upon addition of anthranilate in the presence of glucose, indicates that transcription from the tandem promoter, in which a natively catabolite-repressed Pben is located proximal to the coding sequence, surprisingly is not blocked by the catabolite repression of Pben. This is even more surprising in light of the fact that both (1) the bacterial source of the Pben and Pant elements in the tandem promoter and (2) the host cell in which induction was tested are the same: *Pseudomonas fluorescens* biotype A. Thus, even though the Pben element is native to the host cell, the upstream presence of the natively non-catabolite-repressed promoter (Pant) is able to overcome the catabolite repression of the natively catabolite-repressed promoter (Pben). Moreover, the presence of *antR* and Pant upstream of the Pben promoter appears to relieve the catabolite repression of Pben since the tandem promoter is active during fermentation, in the presence of glucose (see Figures 12 and 13).

Example 4. **Improved Mutants of Pben509**

In an effort to improve the Pben promoter, the Pben509 promoter was subjected to mutagenesis by error prone PCR. Mutants were screened for improved activity following induction with 10mM benzoate at the shake flask scale. The mutants identified showed

approximately 2-fold improvement over the wild type promoter (Figure 10). Positive hits were re-transformed into *P. fluorescens* and re-tested to ensure that the improved activity was in fact due to the new construct.

Sequence analysis (Figure 11) revealed one change in mutant 2d3 and two changes in mutant 21b5. As illustrated in the attached figure, the mutation in isolate 2d3 and one of the mutations in 21b5 fall within the coding region of the upstream ORF. This region is not contained within in the Pben278 construct. The fact that these mutations were isolated in improved promoter mutants indicates that upstream regions may affect transcription, although they are not necessary for activated transcription. The second mutation identified in 21b5 is located five base pairs upstream of the transcriptional start site.

Example 5. **Rationally Mutated Pben and Pant Promoters**

Construction and Analysis of Pben –10 Mutants. The native Pben predicted –10 region was mutated in an attempt to improve promoter activity. The promoter itself was truncated to 88 bp, and three derivatives of the $^{-}$10 were constructed: wild type (TACGGTT, consensus (TATAAT) , and *Acinetobacter* (Ac) Pben-10 (TAAGGT), as described in Materials and Methods. The primers were constructed such that one bp (G:C) upstream of the previously identified transcriptional start site was removed and 9bp downstream of the previously identified transcriptional start site are included. These promoters were fused to the *phoA* reporter gene and tested for activity in *P. fluorescens* MB101. Figure 14 shows that the truncation of Pben promoter to 88bp is sufficient to confer benzoate-activated expression, although altering the $^{-}$10 region either to the consensus TATAAT or to the $^{-}$10 sequence of the *Acinetobacter* Pben promoter did not appear to significantly improve benzoate induced promoter activity.

Construction and Analysis of Pant –10 Mutants. As described above for Pben, the predicted $^{-}$10 region of the Pant promoter was mutated in an attempt to improve promoter activity. In the construction of two Pant $^{-}$10 mutants, the promoter was truncated to 289bp. DNA and fragments containing the anthranilate transcriptional activator and the mutant promoter were fused to the *phoA* reporter gene. The resulting plasmids were transformed into a derivative

of MB101 in which the *antA* gene has been insertionally inactivated. Figure 15 shows that, following induction with anthranilate, the 3' truncation of the promoter to 289bp did not affect activity (pDOW1095). Altering the putative $^{-}$10 region to consensus $^{-}$10 (pDOW1098) resulted in the promoter becoming capable of expression even in the absence of inducing compound. Addition of anthranilate did result in higher expression, indicating that the promoter was still inducible.

Example 6. **Mutant Ptandem Promoters**

The tandem promoter having the sequence as shown in pDOW1057 (SEQ ID NO:13) was mutated in the Pben-10 region as follows to construct mutant Ptandem promoters. A 1.6 kb DNA fragment containing *antR* and Pant, obtained by digestion of pDOW1039 with *Hind*III and *Sma*I, was gel purified (using QIAEX II gel column, from Qiagen Corp.) and ligated into each of pDOW1102 (Pben88wt-10), pDOW1106 (Pben88con-10), and pDOW1100 (Pben88Ac-10), each of which had been digested with *Hind*III and *Pme*I. Following transformation into host cells, positive clones were identified for each plasmid by colony PCR and then confirmed by DNA sequencing. The resulting plasmids were named pDOW1107, pDOW1108, and pDOW1109, respectively.

The effect of the Pben mutants on Ptandem activity was assessed at the shake flask scale. As shown in Figure 16, the mutations did not have a significant effect on benzoate- or anthranilate-induced Ptandem activity. pDOW1107-1109 all showed β-galactosidase activity within 2 fold of that shown by the original construct, pDOW1057. One interesting finding was that the Pben88 -10consensus mutant alone or as part of a tandem promoter appeared to be expressed prior to induction with either benzoate or anthranilate. Addition of benzoate as an inducer resulted in an increased expression from Pben88 -10consensus alone (pDOW1106), or as part of a tandem (pDOW1108) (Figure 16A). However, addition of anthranilate did not result in an increase in *lacZ* expression from the Pben88 -10consensus tandem construct pDOW1108 (Figure 16B).

Analysis of pDOW1108 at the 20L scale revealed that MB101 carrying pDOW1108 induced with 2mM or 5mM benzoate pulses over a 24-hour period was not only active, but

also was able to metabolize benzoate (data not shown). A relatively high level of β-galactosidase activity was detected at I0, most likely a result of "leaky" expression, as had been detected at the shake flask scale (see Figure 16A). An initial decrease in activity was consistently detected upon induction with benzoate, but activity then rose to a level greater than that detected at I0. Induction of pDOW1108 at the 20L scale with 2mM anthranilate every 4 hours for a 24-hour period showed that although anthranilate was metabolized efficiently, cloned tandem promoter expression as measured by β-galactosidase activity was leaky, as observed at shake flask scale, but actually declined after addition of anthranilate (see Figure 18). A comparative induction of *P. fluorescens* carrying the original tandem promoter construct pDOW1057 with benzoate at the 20L scale shows that benzoate is metabolized, as with pDOW1108. However, induction of β-galactosidase activity seems to be delayed compared to similar 2mM dose inductions of the pDOW1108 construct, where increased activity was detected between 4 and 24 hours as opposed to between 30 and 48 hours post induction (data not shown).

Example 7. **Use of Benzoate- and Anthranilate-Induced Promoters for Controlled Gene Expression during Fermentation**

Testing of the Pben509 *lacZ* fusion at the 20L scale revealed transcriptional regulation issues not detected at the shake flask scale. Induction of the fusion with 5 or 10mM benzoate was not consistently observed (data not shown). A correlation between benzoate consumption and activation of Pben509 was also observed. The presence of glucose is thought to be responsible for the inhibition of reporter gene expression. Subsequent to these experiments, it has been observed in shake flask experiments that metabolism of benzoate follows the depletion of glucose. The benzoate-inducible system may be useful in fermentation processes that utilize carbon sources other than glucose. Shake flask experiments reveal that the highest levels of induction are observed when citrate is used as a carbon source. This observation should hold true for fermentation scale.

Testing of the *antR* Pant construct and of the tandem promoter construct at the 20L scale showed activity similar to that observed at the shake flask scale. Because the inducer is consumed by the culture, anthranilate was fed during the course of induction. Activity was

observed to increase over time. It is likely that higher activity will be observed in strains that are unable to metabolize the inducer. As observed in shake flask and 20L fermentation experiments, the tandem promoter construct is more active than the *antR* Pant construct (Figure 12). Inactivation of anthranilate metabolism by insertional inactivation of the *antA* gene allowed for greater expression of the tandem promoter:*lacZ* fusion at the 20L scale. As shown in Figure 13, anthranilate is not metabolized during the course of induction, and the level of β-galactosidase activity observe is much higher that that observed in a strain that does metabolize anthranilate.

Example 8. **Characterization of the BenAB Knock-Out Strain**

To verify whether benzoate is in fact the inducer of the Pben promoter, and not a downstream metabolite thereof, the benAB gene knock-out strain of P. fluorescens was further characterized. The *benA* and *benB* genes code for the large and small subunits of benzoate 1, 2 dioxygenase, respectively. Two isolated of the *benAB* knock-out strain were further tested for the ability to metabolize benzoate as follows. Cells were grown in LB-proline-uracil to high density; benzoate was then added to the cultures to a final concentration of ~5 mM before they were returned to incubate for 24 hr. The concentration of benzoate remaining in the cell-free broth, as measured by HPLC, showed that the *benAB* deletion mutants were unable to metabolize benzoate, while the parent, non-knock-out strain did metabolize benzoate efficiently. To assess whether the Pben promoter is still active in the *benAB* knockout strain, a plasmid containing a Pben278::*lacZ* construct was transformed into one of the strains, and transformants were grown in LB medium. Transformants were induced with 0 or 5mM benzoate and *lacZ* activity demonstrated that benzoate was indeed the inducer for Pben, rather than a downstream metabolite. See Figure 19.

Example 9. **Effect of Multi-Copy Expression of BenR**

A DNA fragment containing the BenR ORF upstream of *benA* along with Pben promoter was amplified from *P. fluorescens* MB214 genomic DNA using primers Benact5' and Bambenconshort under the following conditions: 94 °C for 1 min; (94 °C, 1 min; 50 °C, 30 sec; 72 °C, 90sec) for 30 cycles; then 72 °C for 10 min, and 4 °C hold. The PCR product

was ligated into the pCR2.1 vector, and the sequence verified. The insert fragment was digested with *Pme*I and *Bam*HI, and ligated to pDOW1033. The resulting plasmid was stocked as pDOW1090. The same promoter construct was fused to the *lacZ* reporter by digesting pDOW1090 with *Bam*HI and *Xho*I to remove the *phoA* reporter, and replacing it with the 3Kb *Bam*HI-*Xho*I fragment of pDOW1035, containing the *lacZ* reporter gene.

The *benR* ORF was cloned together with the Pben promoter upstream of the *phoA* reporter gene to determine whether expression of the transcriptional activator gene in multicopy would improve benzoate activated gene expression. At the shake flask scale, there was observed no significant difference in promoter activity with *benR* in multicopy. Since it has been shown in the literature that overexpressing the transcriptional activator can overcome catabolite repression, we tested 20L fermentations of P. fluorscens MB101 carrying pDOW1090. Previous studies showed that MB101 carrying a Pben::*lacZ* fusion was unable to metabolize benzoate during fermentation with a corn syrup feed. We found that MB101 carrying pDOW1090 is able to metabolize benzoate at the 20L scale. Benzoate was found to be consistently metabolized in triplicate 20L fermentations, indicating that the chromosomal Pben promoter was active. Thus, the presence of multi-copy expression of BenR overcame catabolite repression. See Figure 17.

As a result, we have found that overexpression of *benR* allows *P. fluorescens* to overcome catabolite repression observed for benzoate metabolism at the 20L scale when constructs containing Pben alone were tested. Demonstration of benzoate-induced promoter activity at the 20L scale is an important improvement, since benzoate-induced activation of tandem promoters is greater that that of anthranilate-induced activity at the shake flask scale, even though anthanilate-induced activity under control of Ptandem is already stronger than anthranilate-induced activity of Pant. Both pDOW1057 and pDOW1108 were found to be benzoate-inducible at the 20L scale. Although the pDOW1108 construct is "leaky", in that significant expression occurs prior to addition of the inducer, this should not present a large problem for its use in protein expression. In addition, because it has been found that Pben is active in the *benAB* knock-out strain, use of such a knock-out strain will improve benzoate-induced promoter activity for Pben, as well as Ptandem. Likewise, because it has now been shown that induction of the tandem promoter construct pDOW1057 with anthranilate is

improved in a strain carrying and insertionally inactivated chromosomal *antA* gene, improved anthranilate-induced promoter activity will be enhanced for Pant, as well as Ptandem.

Consequently, anthranilate- and benzoate-inducible promoters have now been developed for use in bacterial expression systems. These promoters have been found to permit tight regulation of transcription and are inducible with low-cost compounds such as benzoate and anthranilate; the presence of *antR* in multi-copy also now has been found to significantly improve the activity of the Pant promoter. In addition a new type of tandem promoters has now been developed for use in bacterial expression systems, exemplified by Pant-Pben tandem promoters that have been found to exhibit increased levels of anthranilate-induced gene expression, over Pant itself; were found to be benzoate-inducible, *i.e.* to the same level as Pben itself; and were found to surprisingly overcome the catabolite repression to which Pben alone was subject. Further, the present work has demonstrated that both the Pant promoter (with *antR*) and the tandem promoter constructs exhibit anthranilate-inducible gene expression under fermentation-scale conditions (*e.g.*, at the 20L scale); the tandem promoter constructs also exhibits benzoate inducible gene expresion under fermentations-scale condiitons.

It is to be understood that the preferred embodiments described above are merely exemplary of the present invention and that the terminology used therein is employed solely for the purpose of illustrating these preferred embodiments; thus, the preferred embodiments selected for the above description are not intended to limit the scope of the present invention. The invention being thus described, other embodiments, alternatives, variations, and obvious alterations will be apparent to those skilled in the art, using no more than routine experimentation, as equivalents to those preferred embodiments, methodologies, protocols, vectors, reagents, elements, and combinations particularly described herein. Such equivalents are to be considered within the scope of the present invention and are not to be regarded as a departure from the spirit and scope of the present invention. All such equivalents are intended to be included within the scope of the following claims, the true scope of the invention thus being defined by the following claims.

References

1. J. Brosius & A. Holy, Regulation of ribosomal RNA promoters with a synthetic lac operator, *Proc. Nat'l Acad. Sci. USA*, 81:6929-6933 (1984).

2. E. Amann *et al.*, Vectors bearing hybrid trp-lac promoter useful for regulated expression of cloned genes in Escherichia coli, *Genes* 25:167-178 (1983).

3. 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.).

4. H. Schweizer, Vectors to express foreign genes and techniques to monitor gene expression for Pseudomonads, *Current Opinion in Biotechnology*, 12:439-445 (2001).

5. 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).

6. D. Goeddel *et al.*, Direct expression in Escherichia coli of a DNA sequence coding for human growth hormone, *Nature* 281:544-48 (1979).

7. K.L. Hester *et al.*, Catabolite repression control by *crc* in 2xYT medium is mediated by post-transcriptional regulation of bkdR expression in *Pseudomonas putida*, *J. Bacteriol.* 182(4):1150-53 (2000).

8. P.M. Santos *et al.*, Physiological analysis of the expression of the styrene degradation gene cluster in *Pseudomonas fluorescens* ST, *Appl. & Environ. Microbiol.* 66(4):1305-10 (2000).

9. W.A. Duetz *et al.*, Catabolite repression of the toluene degradation pathway in *Pseudomonas putida* harboring pWW0 under various conditions of nutrient limitation in chemostat culture, *Appl. & Environ. Microbiol.* 62(2):601-06 (1996).

10. M.M. Marin *et al.*, The alkane hydroxylase gene of *Burkholderia cepacia* RR10 is under catabolite repression control, *J. Bacteriol.* 183(14):4202-09 (2001).

11. M.M. Ochs *et al.*, Amino acid-mediated induction of the basic amino acid-specific outer membrane porin *OprD* from *Pseudomonas aeruginosa*, *J. Bacteriol.* 181(17):5426-32 (1999).

12. T. Nishijyo *et al.*, Molecular characterization and regulation of an operon encoding a system for transport of arginine and ornithine and the ArgR regulatory protein in *Pseudomonas aeruginosa*, *J. Bacteriol.* 180(21):5559-66 (1998).

13. K.L. Hester *et al.*, *Crc* is involved in catabolite repression control of the *bkd* operons of *Pseudomonas putida* and *Pseudomonas aeruginosa*, *J. Bacteriol.* 182(4):1144-49 (2000).

14. G. Mosqueda & J.L. Ramos, A set of genes encoding a second toluene efflux system in *Pseudomonas putida* DOT-T1E is linked to the *tod* genes for toluene metabolism, *J. Bacteriol.* 182(4):937-43 (2000).

15. C.D. Lu & A.T. Abdelal, The *gdhB* gene of *Pseudomonas aeruginosa* encodes an arginine-inducible NAD(+)-dependent glutamate dehydrogenase which is subject to allosteric regulation, *J. Bacteriol.* 183(2):490-99 (2001).

16. S. Vilchez *et al.*, Control of expression of divergent *Pseudomonas putida* put promoters for proline catabolism. *Appl. & Environ. Microbiol.* 66(12):5221-25 (2000).

CLAIMS:

1. An isolated benzoate promoter nucleic acid comprising the –35 region of the *Pseudomonas fluorescens* native benzoate promoter attached upstream of the –10 region of this native promoter, via a 15-20 nucleotide linker.

2. The benzoate promoter nucleic acid according to Claim 1 wherein said promoter further comprises a benzoate promoter activator protein binding site.

3. The benzoate promoter nucleic acid according to Claim 1 wherein said promoter further comprises, attached immediately upstream, the native sequence of the approximately 50 nucleotide portion natively located immediately upstream of the native promoter.

4. The benzoate promoter nucleic acid according to Claim 1 wherein said promoter comprises nucleotides 1275-1280 of SEQ ID NO:1 attached upstream of nucleotides 1296-1301 of SEQ ID NO:1, via a 15-20 nucleotide linker.

5. The benzoate promoter nucleic acid according to Claim 1 wherein said promoter comprises nucleotides 1275-1301 of SEQ ID NO:1, and optionally at least one of the mutations listed therein.

6. The benzoate promoter nucleic acid according to Claim 1 wherein said promoter comprises nucleotides 1275-1307 of SEQ ID NO:1, and optionally at least one of the mutations listed therein.

7. The benzoate promoter nucleic acid according to Claim 1 wherein said promoter comprises nucleotides 1228-1301 of SEQ ID NO:1, and optionally at least one of the mutations listed therein.

8. The benzoate promoter nucleic acid according to Claim 1 wherein said promoter comprises nucleotides 1228-1307 of SEQ ID NO:1, and optionally at least one of the mutations listed therein.

9. A mutant promoter nucleic acid formed by mutation of a benzoate promoter nucleic acid according to any one of Claims 1-8 to obtain a mutant promoter whose nucleotide sequence is at least 90% homologous to and is heterologous to that of any one of the promoters according to Claims 1-8.

10. The benzoate promoter nucleic acid according to any one of Claims 1-9, further comprising a benzoate promoter activator protein coding sequence.

11. The benzoate promoter nucleic acid according to Claim 10, wherein said benzoate promoter activator protein coding sequence encodes a benzoate promoter activator protein having an amino acid sequence at least 90% homologous to any one of residues 1-335 of SEQ ID NO:2, residues 1-335 of SEQ ID NO:2 containing Asn152, residues 21-335 of SEQ ID NO:2, and residues 21-335 of SEQ ID NO:2 containing Asn152, and optionally at least one of the mutations listed therein.

12. An isolated anthranilate promoter nucleic acid comprising the –35 region of the *Pseudomonas fluorescens* native anthranilate promoter attached upstream of the –10 region of this native promoter, via a 15-20 nucleotide linker.

13. The anthranilate promoter nucleic acid according to Claim 12 wherein said promoter further comprises an anthranilate promoter activator protein binding site.

14. The anthranilate promoter nucleic acid according to Claim 12 wherein said promoter further comprises, attached immediately upstream, the native sequence of the approximately 110 nucleotide portion natively located immediately upstream of the native promoter.

15. The anthranilate promoter nucleic acid according to Claim 12 wherein said promoter comprises nucleotides 1239-1244 and 1264-1268 of SEQ ID NO:7, nucleotides 1239-1244 being attached upstream of nucleotides 1264-1268 via a 15-20 nucleotide linker, and optionally at least one of the mutations listed therein.

16. The anthranilate promoter nucleic acid according to Claim 12 wherein said promoter comprises nucleotides 1239-1268 of SEQ ID NO:7, and optionally at least one of the mutations listed therein.

17. The anthranilate promoter nucleic acid according to Claim 12 wherein said promoter comprises nucleotides 1239-1274 of SEQ ID NO:7, and optionally at least one of the mutations listed therein.

18. The anthranilate promoter nucleic acid according to Claim 12 wherein said promoter comprises nucleotides 1130-1268 of SEQ ID NO:7, and optionally at least one of the mutations listed therein.

19. The anthranilate promoter nucleic acid according to Claim 12 wherein said promoter comprises nucleotides 1130-1274 of SEQ ID NO:7, and optionally at least one of the mutations listed therein.

20. A mutant anthranilate promoter nucleic acid formed by mutation of an anthranilate promoter nucleic acid according to any one of Claims 12-19 to obtain a mutant promoter whose nucleotide sequence is at least 90% homologous to and heterologous to that of any one of the promoters according to Claims 12-19.

21. The anthranilate promoter nucleic acid according to any one of Claims 12-20, further comprising an anthranilate promoter activator protein coding sequence.

22. The anthranilate promoter nucleic acid according to Claim 21, wherein said anthranilate promoter activator protein coding sequence encodes an anthranilate promoter activator protein having an amino acid sequence at least 90% homologous to any one of residues 1-330 of SEQ ID NO:9, and residues 1-330 of SEQ ID NO:9 containing Ala268.

23. A tandem promoter comprising a non-catabolite-repressed promoter attached upstream of a natively catabolite-repressed promoter, in which the catabolite repression of the

latter promoter is overcome or a different improved promoter property is exhibited, or both.

24. The tandem promoter according to Claim 23 wherein both said non-catabolite-repressed promoter and said natively catabolite-repressed promoter are obtained from the Pseudomonads and closely related bacteria.

25. The tandem promoter according to Claim 24 wherein both said non-catabolite-repressed promoter and said natively catabolite-repressed promoter are obtained from the genus *Pseudomonas*.

26. The tandem promoter according to Claim 24 wherein both said non-catabolite-repressed promoter and said natively catabolite-repressed promoter are obtained from the same species.

27. The tandem promoter according to Claim 25 wherein both said non-catabolite-repressed promoter and said natively catabolite-repressed promoter are obtained from the same species.

28. The tandem promoter according to Claim 27 wherein both said non-catabolite-repressed promoter and said natively catabolite-repressed promoter are obtained from *Pseudomonas fluorescens*.

29. The tandem promoter according to any one of Claims 23-28 wherein both said non-catabolite-repressed promoter and said natively catabolite-repressed promoter are obtained from gene(s) or operon(s) encoding alternative carbon source utilization enzyme(s) or pathway(s).

30. The tandem promoter according to any one of Claims 23-28 wherein said non-catabolite-repressed promoter is obtained from an operon encoding an anthranilate degradation pathway and said natively catabolite-repressed promoter is obtained from an operon encoding a benzoate degradation pathway.

31. The tandem promoter according to any one of Claims 23-28, wherein said non-catabolite-repressed promoter is attached immediately upstream of said natively catabolite-repressed promoter.

32. The tandem promoter according to any one of Claims 23-28, wherein said non-catabolite-repressed promoter is attached upstream from said natively catabolite-repressed promoter via a polynucleotide linker.

33. The tandem promoter according to Claim 32, wherein said polynucleotide linker has the nucleotide sequence of g1396-a1429 of SEQ ID NO:13.

34. The tandem promoter according to Claim 28 wherein said non-catabolite-repressed promoter is an anthranilate promoter selected from the group consisting of nucleotides 1221-1359, 1221-1371, 1328-1359, and 1328-1371 of SEQ ID NO:13, and optionally at least one of the mutations listed therein, and said catabolite-repressed promoter is a benzoate promoter sequence selected from the group consisting of nucleotides 1430-1503, 1430-1509, 1477-1503, and 1477-1509 of SEQ ID NO:13, and optionally at least one of the mutations listed therein.

35. The tandem promoter according to any one Claim 23 wherein said promoter further comprises, upstream from and proximal to the benzoate "-35 to -10 region" a benzoate activator protein binding site, and upstream from and proximal to the anthranilate "-35 to -10 region" an anthranilate activator protein binding site.

36. The tandem promoter according to Claim 34 wherein said tandem promoter comprises nucleotides 1329-1503 of SEQ ID NO:13, and optionally at least one of the mutations listed therein.

37. The tandem promoter according to Claim 34 wherein said tandem promoter comprises nucleotides 1329-1509 of SEQ ID NO:13, and optionally at least one of the mutations listed therein.

38. The tandem promoter according to Claim 34 wherein said tandem promoter comprises nucleotides 1121-1503 of SEQ ID NO:13, and optionally at least one of the mutations listed therein.

39. The tandem promoter according to Claim 34 wherein said tandem promoter comprises nucleotides 1121-1509 of SEQ ID NO:13, and optionally at least one of the mutations listed therein.

40. The tandem promoter according to Claim 34 wherein said tandem promoter comprises nucleotides 1329-1541 of SEQ ID NO:13, and optionally at least one of the mutations listed therein.

41. The tandem promoter according to Claim 34 wherein said tandem promoter comprises nucleotides 1121-1541 of SEQ ID NO:13, and optionally at least one of the mutations listed therein.

42. An altered promoter prepared by a process comprising the steps of:

(A) obtaining at least one polynucleotide having a base sequence at least 90% identical to and heterologous to the base sequence of any one of:

(1) a promoter according to Claim 1, Claim 12, Claim 23, or Claim 30;

(2) at least bases 1275-1307 of SEQ ID NO:1;

(3) at least bases 1239-1274 of SEQ ID NO:7; or

(4) at least bases 1329-1509 of SEQ ID NO:13;

(B) screening said polynucleotide(s) for the ability, when operably attached to a transcription product-encoding polynucleotide in an expression construct, to direct transcription thereof in a prokaryotic host cell, and optionally for at least one promoter property; and

(C) identifying, based on the results of said screening, at least one polynucleotide having the ability, when operably attached to a transcription product-encoding polynucleotide in an expression construct, to direct transcription thereof in a prokaryotic host cell, and

optionally having at least one promoter property that is improved relative to that of the promoter of step (A)(1)-(A)(4).

43. An improved promoter prepared by a process comprising the steps of:

(A) providing

(1) a promoter polynucleotide

(a) having the base sequence of any one of

(i) a promoter according to Claim 1, Claim 12, Claim 23, or Claim 30;

(ii) at least bases 1275-1307 of SEQ ID NO:1;

(iii) at least bases 1239-1274 of SEQ ID NO:7; or

(iv) at least bases 1329-1509 of SEQ ID NO:13; and

(b) having the ability, when operably attached to a transcription product-encoding polynucleotide, to direct transcription thereof in a prokaryotic host cell, or

(2) an information string representing the base sequence of said promoter polynucleotide;

(B) obtaining at least one sequence-altered polynucleotide by

(1) performing at least one mutagenesis technique or at least one recombination technique or both upon said promoter polynucleotide to generate at least one sequence-altered polynucleotide, or

(2) modifying said information string to generate at least one modified string that is at least 90% identical to said information string and that represents at least one altered sequence, and

(a) synthesizing at least one sequence-altered polynucleotide having said altered sequence, or

(b) utilizing said modified string by

(i) searching at least one sequence record for at least one identical string that is identical to said modified string, said sequence record corresponding to at least one polynucleotide source,

(ii) identifying at least once occurrence of at least one said identical string therein, and

(iii) selecting the polynucleotide source to which the identified occurrence of said identical string corresponds, thereby obtaining at least one sequence-altered polynucleotide, or

(3) utilizing said information string by

(a) searching at least one sequence record for at least one heterologous string that is at least 90% identical to said information string or its complement, said heterologous string representing at least one altered sequence, and said sequence record corresponding to at least one polynucleotide source,

(b) identifying at least once occurrence of at least one said heterologous string therein, and

(c) selecting the polynucleotide source to which the identified occurrence of said heterologous string corresponds, thereby obtaining at least one sequence-altered polynucleotide, or

(4) utilizing said promoter polynucleotide as a probe by combining said probe with at least one heterologous test polynucleotide, under stringent conditions that permit hybridization of said probe to target sequences within said test polynucleotide that are at least 90% identical to the complement of said probe, and selecting the test polynucleotide to which said probe hybridizes, thereby obtaining at least one sequence-altered polynucleotide;

(C) screening at least one sequence-altered polynucleotide for at least one promoter property and for the ability, when operably attached to a transcription product-encoding polynucleotide in an expression construct, to direct transcription thereof in a prokaryotic host cell; and

(D) identifying, based on the results of said screening, at least one sequence-altered polynucleotide

(1) having at least one promoter property that is improved relative to that of the promoter polynucleotide of step (A)(1), and

(2) having the ability, when operably attached to a transcription product-encoding polynucleotide, to direct transcription thereof in a prokaryotic host cell,

thereby obtaining an improved promoter.

44. A process for utilizing the improved promoter according to Claim 43 wherein said process comprises a step of operably attaching said improved promoter to a transcription product-encoding polynucleotide to form an expression cassette.

45. The process according to Claim 44 wherein said process further comprises a step of including said expression cassette within a vector.

46. The process according to Claim 45 wherein said process further comprises a step of transforming, with said vector, a host cell capable of effecting expression from said expression cassette.

47. The process according to Claim 46 wherein said process further comprises a step of producing the expression product encoded by said expression cassette by maintaining said host cell under growth conditions that induce the improved promoter.

48. The improved promoter according to Claim 43 wherein said utilizing said promoter polynucleotide as a probe in step (B)(4) includes obtaining a polynucleotide or polynucleotide analog molecule containing a base sequence of said promoter polynucleotide that is at least 10 bases long, and labeling said molecule, thereby providing said probe.

49. The improved promoter according to Claim 43 wherein said utilizing said promoter polynucleotide as a probe in step (B)(4) includes obtaining a polynucleotide molecule containing a base sequence of said promoter polynucleotide that is at least 10 bases long, and optionally labeling said molecule, thereby providing said probe, and after said hybridization of the probe to target sequence(s), conducting nucleic acid polymerization in which said probe functions as a primer.

50. An isolated nucleic acid molecule comprising a nucleic acid sequence, the full-length complement of which hybridizes under stringent hybridization and wash conditions to a nucleobase polymer molecule having a base sequence of any one of:

(A) a promoter according to Claim 1, Claim 12, Claim 23, or Claim 30;

(B) at least bases 1275-1307 of SEQ ID NO:1;

(C) at least bases 1239-1274 of SEQ ID NO:7; or

(D) at least bases 1329-1509 of SEQ ID NO:13; and

wherein said isolated nucleic acid molecule can function as a promoter in a prokaryotic cell.

51. An isolated nucleobase polymer molecule having the base sequence of a polynucleotide molecule that can function as a promoter in a prokaryotic cell, wherein said nucleobase polymer molecule contains a base sequence at least 90% identical to the base sequence of any one of:

(A) a promoter according to Claim 1, Claim 12, Claim 23, or Claim 30;

(B) at least bases 1275-1307 of SEQ ID NO:1;

(C) at least bases 1239-1274 of SEQ ID NO:7; or

(D) at least bases 1329-1509 of SEQ ID NO:13.

52. A recombinant nucleic acid molecule that can function as an expression construct in a prokaryotic cell, molecule comprising a promoter containing a base sequence at least 90% identical to the base sequence of any one of:

(A) a promoter according to Claim 1, Claim 12, Claim 23, or Claim 30;

(B) at least bases 1275-1307 of SEQ ID NO:1;

(C) at least bases 1239-1274 of SEQ ID NO:7; or

(D) at least bases 1329-1509 of SEQ ID NO:13.

53. The recombinant expression construct according to Claim 52 wherein said expression construct comprises an mRNA-encoding sequence.

54. The recombinant expression construct according to Claim 52 wherein said expression construct is a vector

55. The recombinant expression construct according to Claim 54 wherein said vector is a plasmid, transposon, or artificial chromosome.

56. A genetically engineered prokaryotic host cell containing a recombinant expression construct according to any one of Claims 53-55.

57. An expression system comprising a genetically engineered prokaryotic host cell according to Claim 56.

58. The genetically engineered prokaryotic host cell according to Claim 56 wherein said host cell further comprises at least one copy, and preferably more than one copy, of a gene encoding the relevant activator protein for the promoter of said recombinant expression construct.

59. The expression system according to Claim 57 wherein said genetically engineered prokaryotic host cell further comprises at least one copy, and preferably more than one copy, of a gene encoding the relevant activator protein for the promoter of said recombinant expression construct.

60. A process for preparing a transcription product comprising growing a genetically engineered prokaryotic host cell according to Claim 58, inducing the recombinant expression construct therein, whereby the transcription product encoded by the recombinant expression product is expressed.

61. A process for preparing a polypeptide comprising expressing an mRNA transcription product by the process according to Claim 60, and further permitting said prokaryotic host cell to translate said mRNA into the polypeptide encoded thereby.

62. A polypeptide prepared by the process according to Claim 61.

63. A tandem promoter produced by a process comprising

(A) providing

(1) a non-catabolite-repressed prokaryotic promoter polynucleotide, and

(2) a catabolite-repressed prokaryotic promoter polynucleotide; and

(B) covalently attaching said non-catabolite-repressed promoter to and upstream of said catabolite-repressed promoter.

64. The tandem promoter according to Claim 63 wherein a polynucleotide inter-promoter linker covalently attaches said non-catabolite-repressed promoter to and upstream of said catabolite-repressed promoter.

65. The tandem promoter according to Claim 64 wherein said polynucleotide inter-promoter linker is about 100 or less than 100 nucleotides in length.

66. An isolated transcriptional activator protein having the amino acid sequence of residues 1-335 of SEQ ID NO:2, or of residues 1-335 of SEQ ID NO:2 containing Asn152.

67. An isolated transcriptional activator protein having the amino acid sequence of residues 21-335 of SEQ ID NO:2, or of residues 21-335 of SEQ ID NO:2 containing Asn152.

68. An isolated transcriptional activator protein having the amino acid sequence of residues 1-330 of SEQ ID NO:9, or of residues 1-330 of SEQ ID NO:9 containing Ala268.

69. An isolated transcriptional activator protein having an amino acid sequence at least 90% identical to and heterologous to that of the transcriptional activator protein according to any one of Claims 66-69.

70. An isolated polynucleotide containing a base sequence encoding a transcriptional activator protein according to any one of Claims 66-69.

71. An isolated polynucleotide containing a base sequence encoding a transcriptional activator protein according to Claim 70.

72. The expression system according to Claim 59 wherein said activator protein and said promoter are any one of the following:

(A) said activator protein is a benzoate promoter activator protein having an amino acid sequence of any one of

(1) residues 1-335 of SEQ ID NO:2,

(2) residues 1-335 of SEQ ID NO:2 containing Asn152,

(3) residues 21-335 of SEQ ID NO:2, and

(4) residues 21-335 of SEQ ID NO:2 containing Asn152,

and said promoter is a benzoate-inducible promoter;

or

(C) said activator protein is an anthranilate promoter activator protein having an amino acid sequence of any one of

(1) residues 1-330 of SEQ ID NO:9, and

(2) residues 1-330 of SEQ ID NO:9 containing Ala268,

and said promoter is an anthranilate-inducible promoter.

FIGURE 1

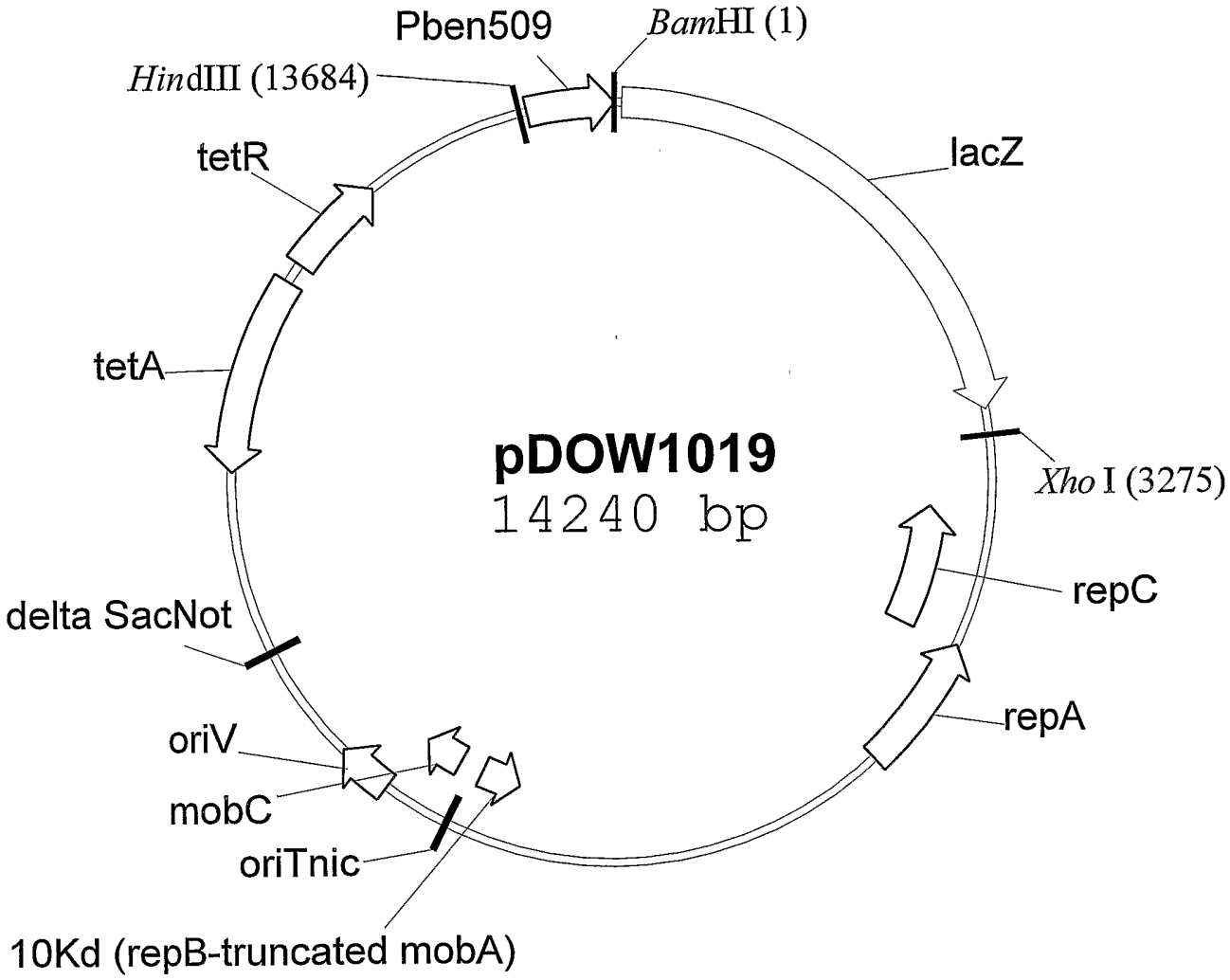

Pben509
*Bam*HI (1)
*Hin*dIII (13684)
tetR
lacZ
tetA
pDOW1019
14240 bp
Xho I (3275)
repC
delta SacNot
repA
oriV
mobC
oriTnic
10Kd (repB-truncated mobA)

FIGURE 2

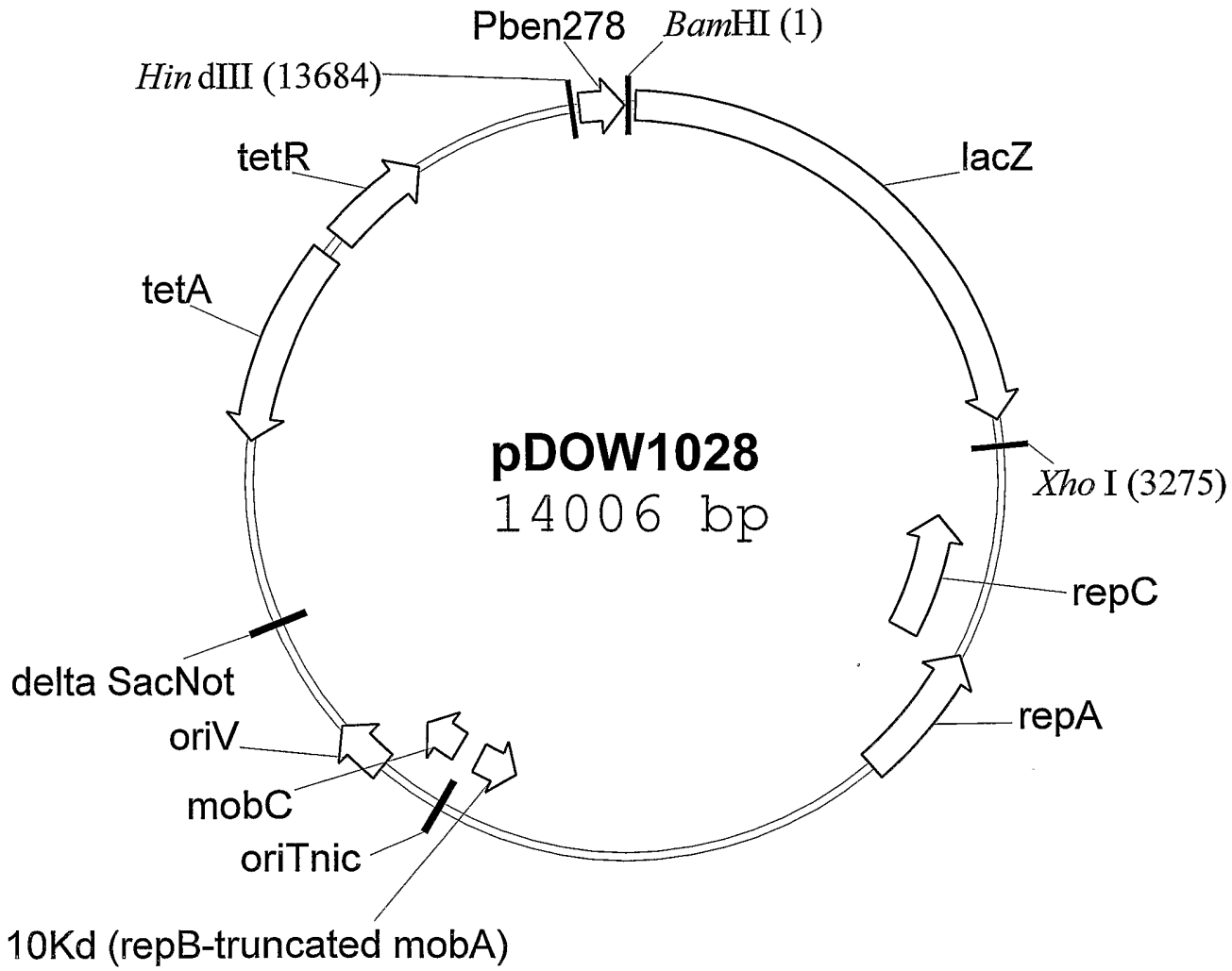

Pben278
BamHI (1)
Hin dIII (13684)
tetR
lacZ
tetA
pDOW1028
14006 bp
Xho I (3275)
repC
delta SacNot
repA
oriV
mobC
oriTnic
10Kd (repB-truncated mobA)

FIGURE 3

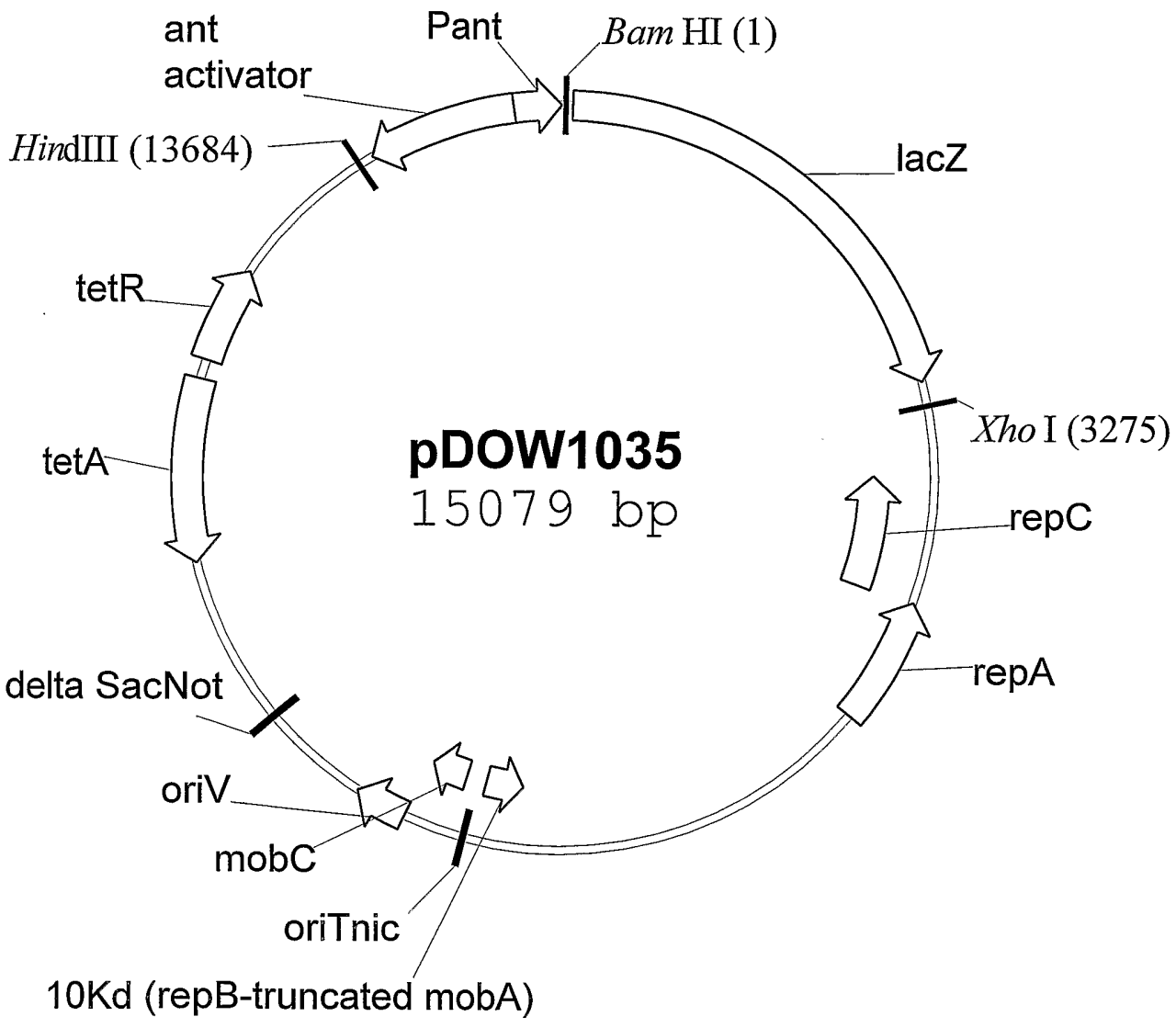

ant
activator
Pant
Bam HI (1)
HindIII (13684)
lacZ
tetR
Xho I (3275)
pDOW1035
15079 bp
tetA
repC
repA
delta SacNot
oriV
mobC
oriTnic
10Kd (repB-truncated mobA)

FIGURE 4

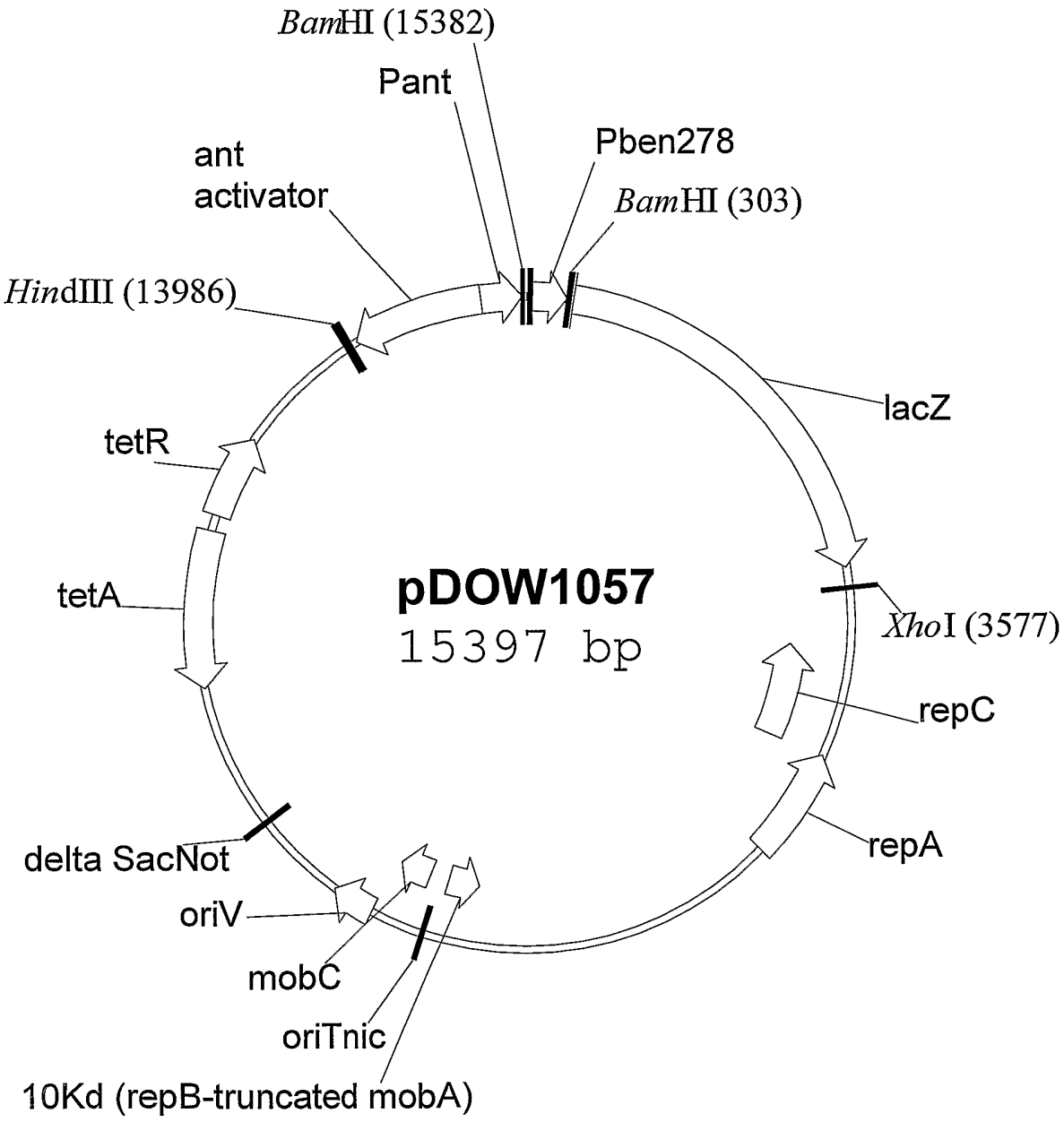

BamHI (15382)
Pant
Pben278
ant
activator
BamHI (303)
HindIII (13986)
lacZ
tetR
pDOW1057
15397 bp
tetA
XhoI (3577)
repC
repA
delta SacNot
oriV
mobC
oriTnic
10Kd (repB-truncated mobA)

FIGURE 5

```
Pben509  CCAACGTCAG TGAACGTTCG CTGTACAGCC TGTTTGAGCG CCAGGTGGGG CTGTCGCCGC GCGATTACGT

Pben509  ACGCCGCTGC AAGCTCGAAC GCGTACATGC ACGCTTGCAA CTAAGCAGCA CGCGCAGCGT GACCGAGGTG

Pben509  GCTTTGGACC ATGGGTTCAT GCACCTAGGG CGGTTTTCCG AAGCCTATCG CAAACGCTTC GGCGAACTGC

Pben509  CGTCGCAGAC CTGGAAACGC CATCGTTAAG CGACGTGCGC CTGGCGGATA GCGATGTGCA GGCAGCGGAT
Pben278                    GTTAAG CGACGTGCGC CTGGCGGATA GCGATGTGCA GGCAGCGGAT

Pben509  ATTGACGGGC AGGGCGAGCA CGTACGGTGA GGGCGCCTGA TACAAGAACA ACGGAGGGCC CGCCCCATGA
Pben278  ATTGACGGGC AGGGCGAGCA CGTACGGTGA GGGCGCCTGA TACAAGAACA ACGGAGGGCC CGCCCCATGA

Pben509  TCAGTACACT CGACCGACTC GCCTGCCAAT TGCGCGAGTC CGTACAGGAA GACCCCGCCA CTGGGGTGTT
Pben278  TCAGTACACC CGACCGACTC GCCTGCCAAT TGCGCGAGTC CGTACAGGAA GACCCCGCCA CTGGGGTGTT

Pben509  CCGCTGCCGC CGCGACATCT TCACCGACCC CGACCTGTTT GCCCTGGAGA TGAAACACAT CTTCGAAGGC
Pben278  CCGCTGCCGC CGCGACATCT TCACCGACCC CGACCTGTTT GCCCTGGAGA TGAAACACAT CTTCGAAGGC

Pben509  GGGTGGATCT ACCTGGCCCC
Pben278  GGGTGGATCT ACCTGGCCC
```

FIGURE 6

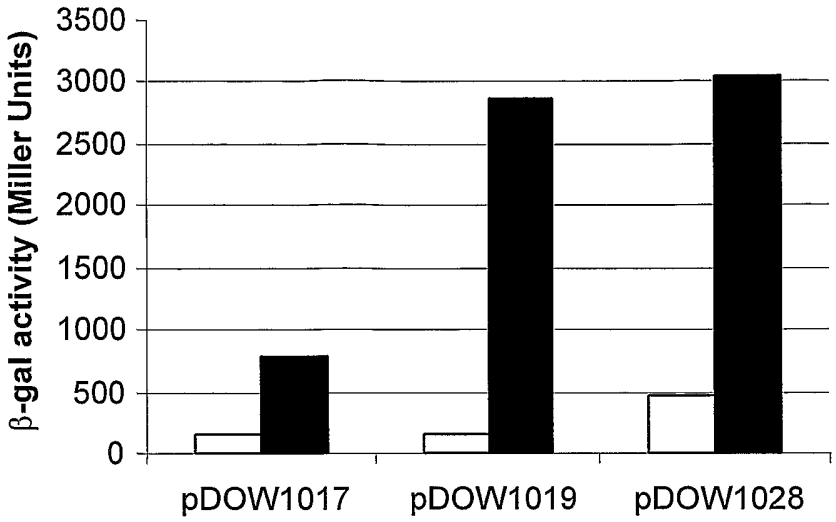

β-gal activity (Miller Units)
3500
3000
2500
2000
1500
1000
500
0
pDOW1017
pDOW1019
pDOW1028

FIGURE 7A

```
Pant+Act  TATCGCAGGC AAGCCAGCTC CCACAGATTG TTTTTCATCC AGTTCAAGTA ATGCGCAGGC GCTTGCGCTG

Pant+Act  CAATGTCTGG CTGGGCGACT CATCGAACAG CTTGCGGTAC TCCGCCGAAA ACCGCCCCAA ATGCGTAAAC

Pant+Act  CCCCAACCCA GGGCGATTTC AGAGATGGTG CGGATCGAGC CCTGCTCCAG AATTTCTTGG CGCACCGCCC

Pant+Act  CCAACCGATG CTTCTTCAAA TACGACATGG GCGACAGTGC GAAGTACTTG CGAAACGCAT CGAACAGTTT

Pant+Act  GAAACGCGAC ACGCCCGCCG CCGCTTCCAG GTCTTCCAGG TGCAGCGCTT CACGGGCGTT GTCGTGGATA

Pant+Act  AATTGCCGCG CGCGGATCAG GTAGTGCGGC AGTTTCACCC CCAGCACGTC GCGCAGTTCT TCGGAGTAGT

Pant+Act  TATTCGGTTG GGCCAGGATC AGGCCCTTGA TCAGCGAGCT TTCCAGGTCG CGAGTAAACG CCGCCTGCTC

Pant+Act  GTACAGTTCG CTGCTGCGCT CCAGTTCGGC GATGAAATAA CGCGCCATGC GCCACCACGA AGCCGGTGCT

Pant+Act  CCGTCCACAG CATCCATCAC CGACTCAAAG CGCAGCGGCG CATCAATGGG CCGTTGCAGC AAACCTTCCA

Pant+Act  GCGACTCGCT CATCGCCGCA CGGGTGATTA CCACCTGCAA CTTGCGGCAG TCACCGGAAA TCGCCAGCAC
Pant713       CTCGCT CATCGCCGCA CGGGTGATTA CCACCTGCAA CTTGCGGCAG TCACCGGAAA TCGCCAGCAC
Pant705       CTCGCT CATCGCCGCA CGGGTGATTA CCACCTGCAA CTTGCGGCAG TCACCGGAAA TCGCCAGCAC

Pant+Act  CTGATGCTCA TTGGGCGAAA TGATCACGCC TTGGTCGCGG TTGGAACTGA GACGTTCACC GTTCTTGCTC
Pant713   CTGATGCTCA TTGGGCGAAA TGATCACGCC TTGGTCGCGG TTGGAACTGA GACGTTCACC GTTCTTGCTC
Pant705   CTGATGCTCA TTGGGCGAAA TGATCACGCC TTGGTCGCGG TTGGAACTGA GACGTTCACC GTTCTTGCTC

Pant+Act  AGCTCCTGCT CGCCCACCAG TGGCAGGCTC AAGCTGTAGC TGCTGAAGTG CTCGGCGTCT TCGATGTCGA
Pant713   AGCTCCTGCT CGCCCACCAG TGGCAGGCTC AAGCTGTAGC TGCTGAAGTG CTCGGCGTCT TCGATGTCGA
Pant705   AGCTCCTGCT CGCCCACCAG TGGCAGGCTC AAGCTGTAGC TGCTGAAGTG CTCGGCGTCT TCGATGTCGA
```

FIGURE 7B

```
Pant+Act  TGGTCACATC AGTGCCGTAC TCGATCACGC CCAGGGTGGT GGCGCGGGAT TTGAACACGT TGGCGCTGTG
Pant713   TGGTCACATC AGTGCCGTAC TCGATCACGC CCAGGGTGGT GGCGCGGGAT TTGAACACGT TGGCGCTGTG
Pant705   TGGTCACATC AGTGCCGTAC TCGATCACGC CCAGGGTGGT GGCGCGGGAT TTGAACACGT TGGCGCTGTG

Pant+Act  GTGAAAGCGC AGGCGCTCGG GGGTTGCCGT CGCCAGGCGA TGGGGCCCGC AGATGCCGGA CATCCAGCTG
Pant713   GTGAAAGCGC AGGCGCTCGG GGGTTGCCGT CGCCAGGCGA TGGGGCCCGC AGATGCCGGA CATCCAGCTG
Pant705   GTGAAAGCGC AGGCGCTCGG GGGTTGCCGT CGCCAGGCGA TGGGGCCCGC AGATGCCGGA CATCCAGCTG

Pant+Act  CGCGCGCCTT CCAGGTCGAA GCGTTGAATA TGAATATCGC GTGTCTGACT AGTCATCAGG GTGCACCCAC
Pant713   CGCGCGCCTT CCAGGTCGAA GCGTTGAATA TGAATATCGC GTGTCTGACT AGTCATCAGG GTGCACCCAC
Pant705   CGCGCGCCTT CCAGGTCGAA GCGTTGAATA TGAATATCGC GTGTCTGACT AGTCATCAGG GTGCACCCAC
Pant311                                                       CAGG GTGCACCCAC

Pant+Act  GGCGGTTAGG CGTTTGCGCG CTCTGACGGC GCGTCGTTGA ACCTCGACAG CAAGTTCCAG GCCACGCCAG
Pant713   GGCGGTTAGG CGTTTGCGCG CTCTGACGGC GCGTCGTTGA ACCTCGACAG CAAGTTCCAG GCCACGCCAG
Pant705   GGCGGTTAGG CGTTTGCGCG CTCTGACGGC GCGTCGTTGA ACCTCGACAG CAAGTTCCAG GCCACGCCAG
Pant311   GGCGGTTAGG CGTTTGCGCG CTCTGACGGC GCGTCGTTGA ACCTCGACAG CAAGTTCCAG GCCACGCCAG

Pant+Act  TGCAGTTCTC ACTGGGTGGA TAGCAACGGT CGACTATGTG GATAAACCCC AGAGTTTTGC GACCATCGCC
Pant713   TGCAGTTCTC ACTGGGTGGA TAGCAACGGT CGACTATGTG GATAAACCCC AGAGTTTTGC GACCATCGCC
Pant705   TGCAGTTCTC ACTGGGTGGA TAGCAACGGT CGACTATGTG GATAAACCCC AGAGTTTTGC GACCATCGCC
Pant311   TGCAGTTCTC ACTGGGTGGA TAGCAACGGT CGACTATGTG GATAAACCCC AGAGTTTTGC GACCATCGCC

Pant+Act  CGCCATCACA GTAGCGCATG CCGTCACCGG CGCGCACCGT CATGGGTATT TGCCGCCCAA CTTTGCGGCC
Pant713   CGCCATCACA GTAGCGCATG CCGTCACCGG CGCGCACCGT CATGGGTATT TGCCGCCCAA CTTTGCGGCC
Pant705   CGCCATCACA GTAGCGCATG CCGTCACCGG CGCGCACCGT CATGGGTATT TGCCGCCCAA CTTTGCGGCC
Pant311   CGCCATCACA GTAGCGCATG CCGTCACCGG CGCGCACCGT CATGGGTATT TGCCGCCCAA CTTTGCGGCC
```

FIGURE 7C

```
Pant+Act  TACGTTCCCC CATTAAGCGG ATAGCCCGCC ACCGCATCGC AGCCGCTTAA TGGCTCACCG TTTAGCCATG
Pant713   TACGTTCCCC CATTAAGCGG ATAGCCCGCC ACCGCATCGC AGCCGCTTAA TGGCTCACCG TTTAGCCATG
Pant705   TACGTTCCCC CATTAAGCGG ATAGCCCGCC ACCGCATCGC AGCCGCTTAA TGGCTCACCG TTTAGCCATG
Pant311   TACGTTCCCC CATTAAGCGG ATAGCCCGCC ACCGCATCGC AGCCGCTTAA TGGCTCACCG TTTAGCCATG

Pant+Act  ATCAAAAGGT GCCTCCC
Pant713   ATCAAAAGGT GCCTCCC
Pant705   ATCAAAAGG
Pant311   ATCAAAAGGT GCCTCCC
```

FIGURE 8

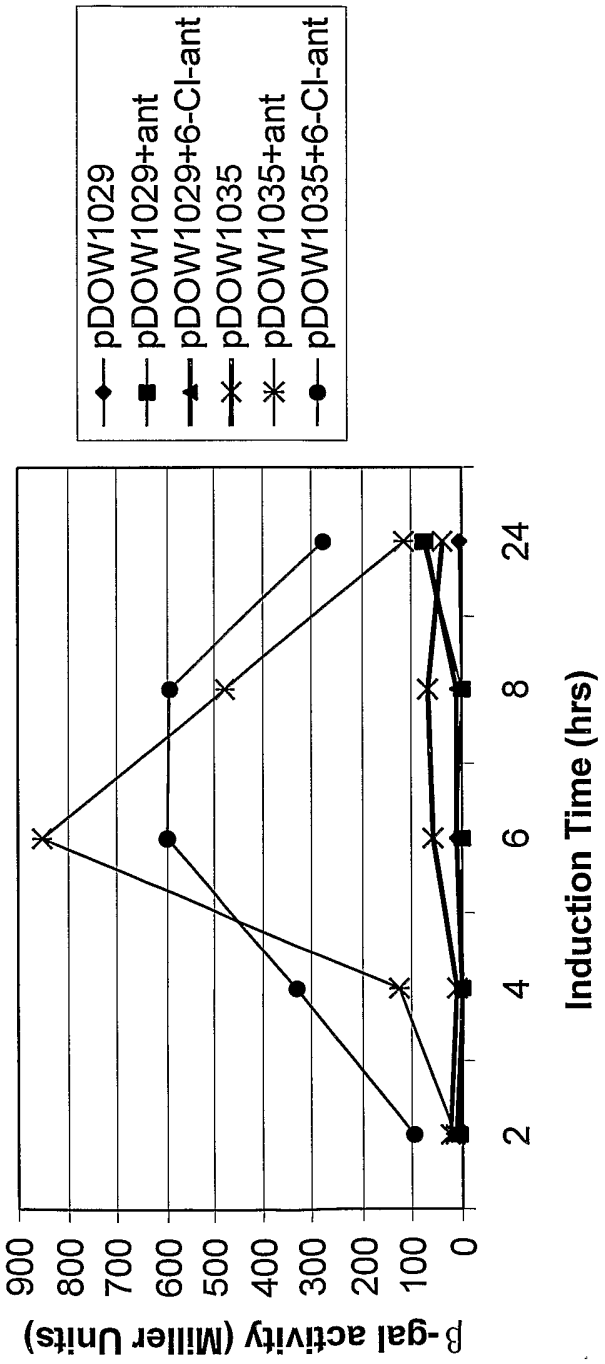

pDOW1029
pDOW1029+ant
pDOW1029+6-Cl-ant
pDOW1035
pDOW1035+ant
pDOW1035+6-Cl-ant
β-gal activity (Miller Units)
0
100
200
300
400
500
600
700
800
900
2
4
6
8
24
Induction Time (hrs)

FIGURE 9A

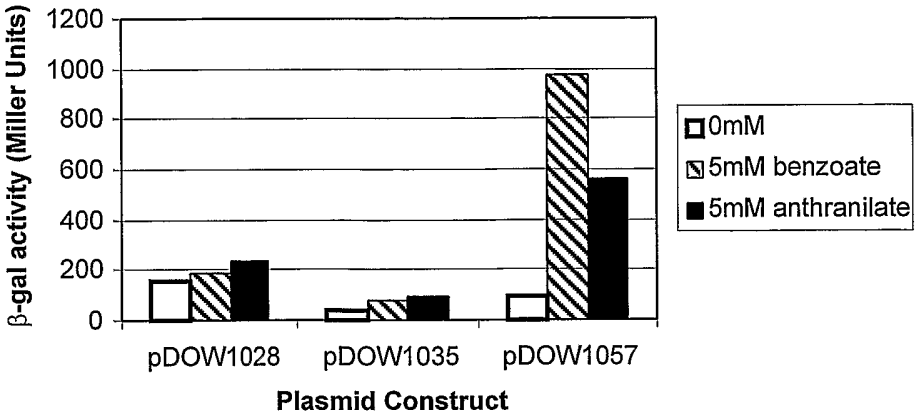


FIGURE 9B

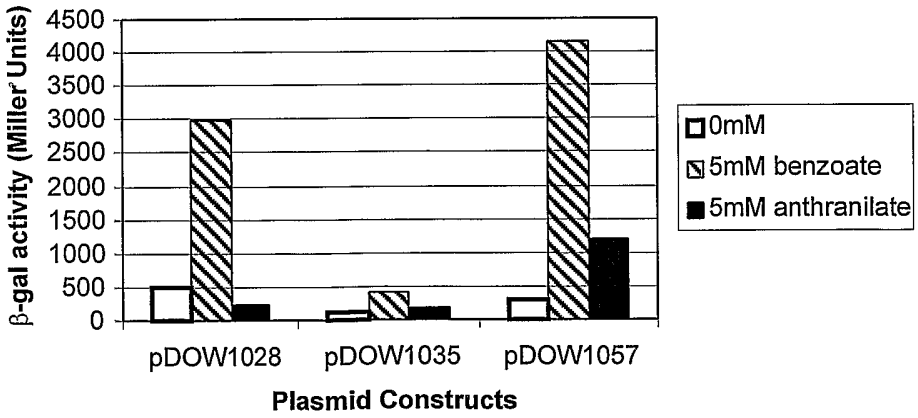


FIGURE 10

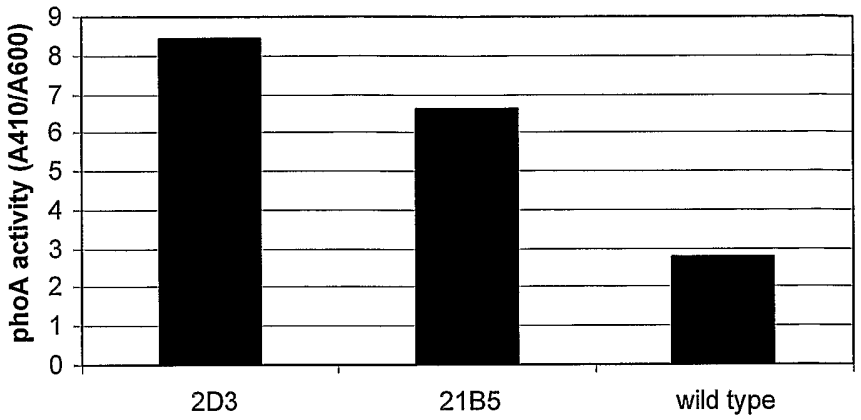


FIGURE 11

```
Pben509   CCAACGTCAG TGAACGTTCG CTGTACAGCC TGTTTGAGCG CCAGGTGGGG CTGTCGCCGC GCGATTACGT
2d3       CCAACGTCAG TGAACGTTCG CTGTACAGCC TGTTTGAGCG CCAGGTGGGG CTGTCGCCGC GCGATTACGT
21b5      CCAACGTCAG TGAACGTTCG CTGTACAGCC TGTTTGAGCG CCAGGTGGGG CTGTCGCCGC GCGATTACGT

Pben509   ACGCCGCTGC AAGCTCGAAC GCGTACATGC ACGCTTGCAA CTAAGCAGCA CGCGCAGCGT GACCGAGGTG
2d3       ACGCCGCTGC AAGCTCGAAC GCGTACATGC ACGCTTGCAA CTTAGCAGCA CGCGCAGCGT GACCGAGGTG
21b5      ACGCCGCTGC AAGCTCGAAC GCGTACATGC ACGCTTGCAA CTAAGCAGCA CGCGCAGCGT GACCGAGGTG

Pben509   GCTTTGGACC ATGGGTTCAT GCACCTAGGG CGGTTTTCCG AAGCCTATCG CAAACGCTTC GGCGAACTGC
2d3       GCTTTGGACC ATGGGTTCAT GCACCTAGGG CGGTTTTCCG AAGCCTATCG CAAACGCTTC GGCGAACTGC
21b5      GCTTTGGACC ATGGGTTCAT GCACCTAGGG CGGTTTTCCG AAGCCTATCG CAAACGCTTC GGCGAACTGC

Pben509   CGTCGCAGAC CTGGAAACGC CATCGTTAAG CGACGTGCGC CTGGCGGATA GCGATGTGCA GGCAGCGGAT
2d3       CGTCGCAGAC CTGGAAACGC CATCGTTAAG CGACGTGCGC CTGGCGGATA GCGATGTGCA GGCAGCGGAT
21b5      CGTCGCAGAC CTGGAAACGT CATCGTTAAG CGACGTGCGC CTGGCGGATA GCGATGTGCA GGCAGCGGAT

Pben509   ATTGACGGGC AGGGCGAGCA CGTACGGTGA GGGCGCCTGA TACAAGAACA ACGGAGGGCC CGCCCCATGA
2d3       ATTGACGGGC AGGGCGAGCA CGTACGGTGA GGGCGCCTGA TACAAGAACA ACGGAGGGCC CGCCCCATGA
21b5      ATTGACGGGC AGGGCGAGCA CGTACGGTAA GGGCGCCTGA TACAAGAACA ACGGAGGGCC CGCCCCATGA

Pben509   TCAGTACACT CGACCGACTC GCCTGCCAAT TGCGCGAGTC CGTACAGGAA GACCCCGCCA CTGGGGTGTT
2d3       TCAGTACACT CGACCGACTC GCCTGCCAAT TGCGCGAGTC CGTACAGGAA GACCCCGCCA CTGGGGTGTT
21b5      TCAGTACACT CGACCGACTC GCCTGCCAAT TGCGCGAGTC CGTACAGGAA GACCCCGCCA CTGGGGTGTT

Pben509   CCGCTGCCGC CGCGACATCT TCACCGACCC CGACCTGTTT GCCCTGGAGA TGAAACACAT CTTCGAAGGC
2d3       CCGCTGCCGC CGCGACATCT TCACCGACCC CGACCTGTTT GCCCTGGAGA TGAAACACAT CTTCGAAGGC
21b5      CCGCTGCCGC CGCGACATCT TCACCGACCC CGACCTGTTT GCCCTGGAGA TGAAACACAT CTTCGAAGGC

Pben509   GGGTGGATCT ACCTGGCCC
2d3       GGGTGGATCT ACCTGGCCC
21b5      GGGTGGATCT ACCTGGCCC
```

FIGURE 12

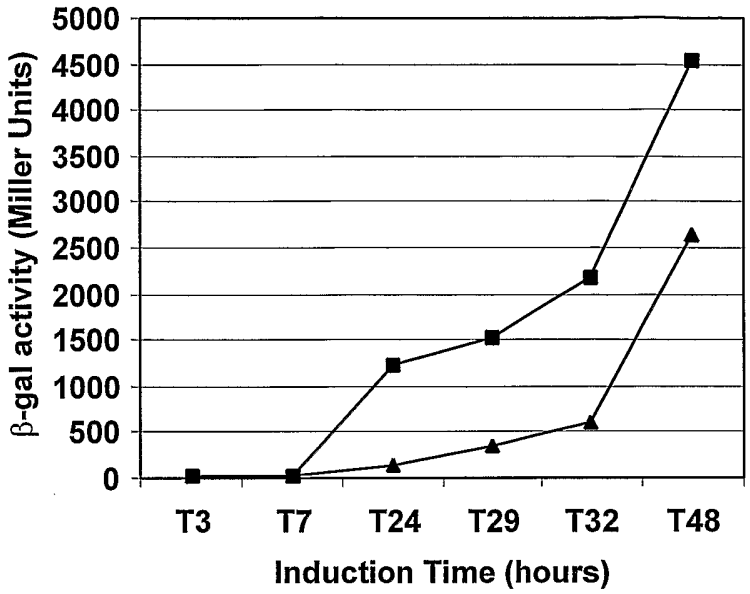

5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
β-gal activity (Miller Units)
T3
T7
T24
T29
T32
T48
Induction Time (hours)

FIGURE 13A

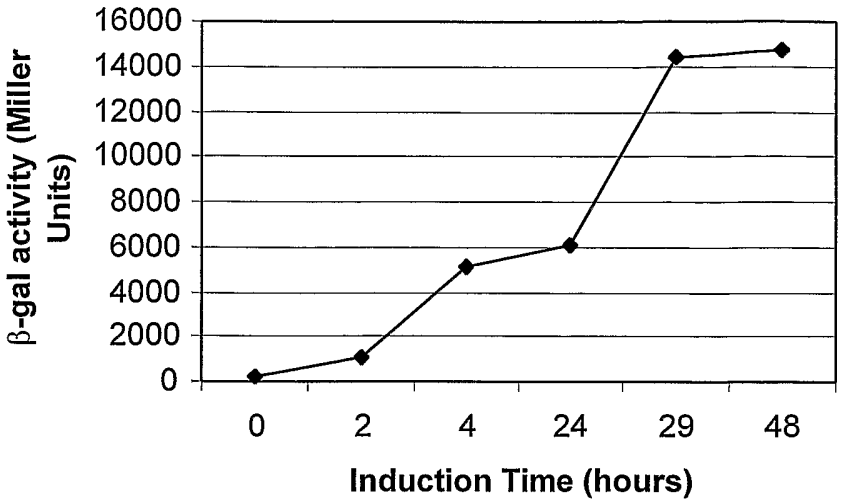


FIGURE 13B

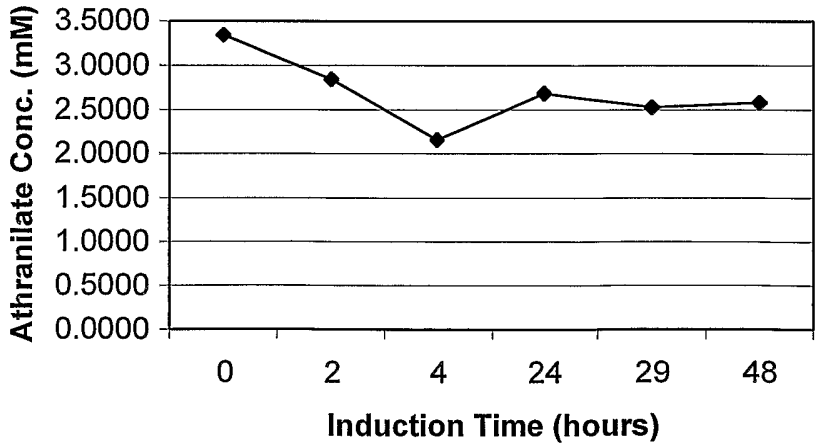


FIGURE 14

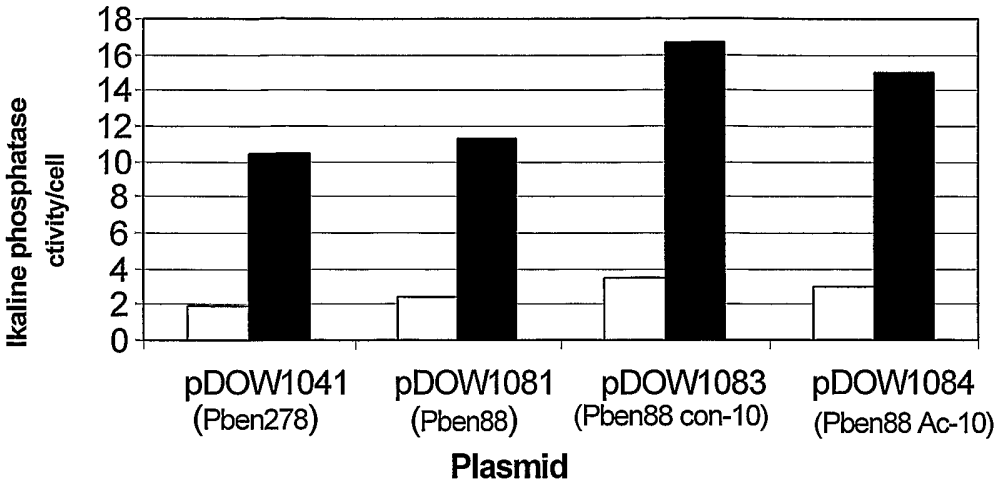


FIGURE 15

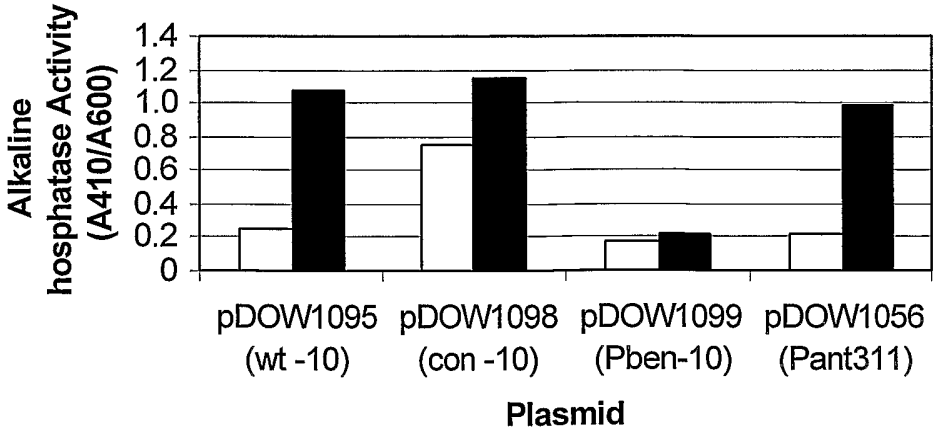


FIGURE 16A

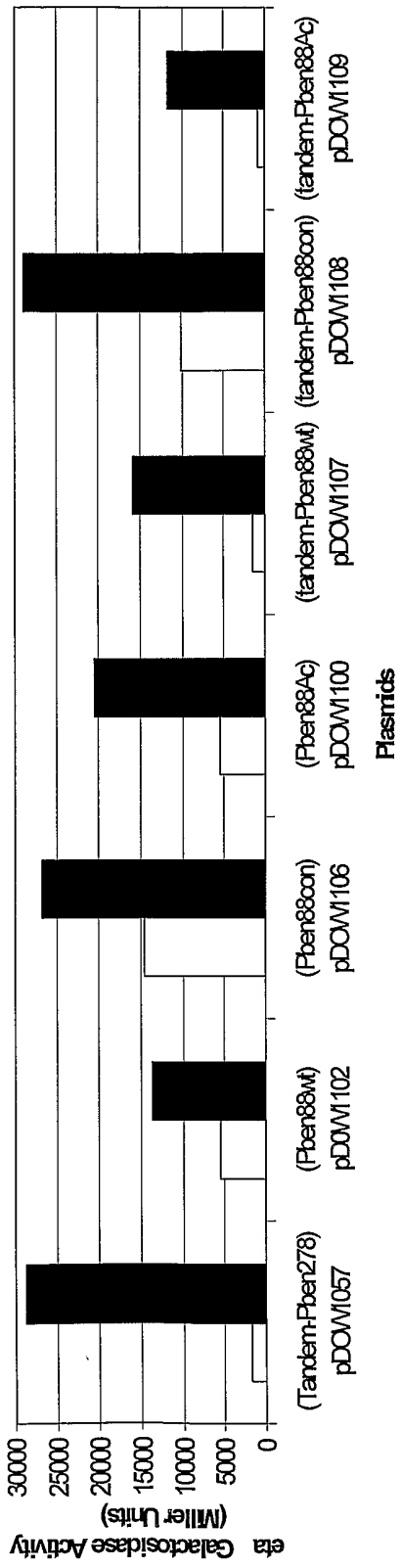


FIGURE 16B

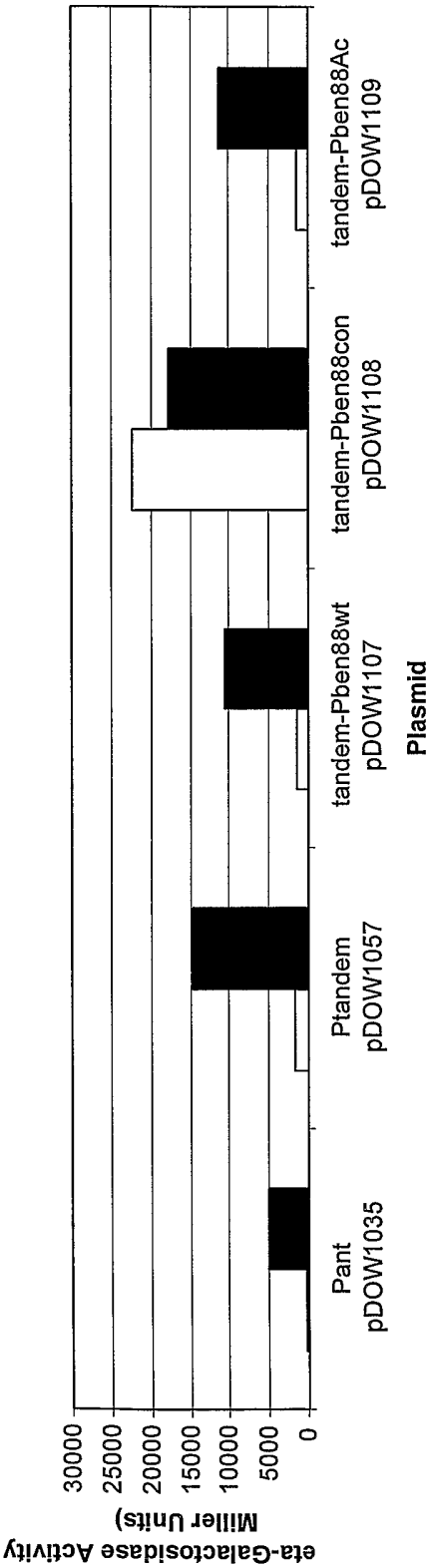


FIGURE 17

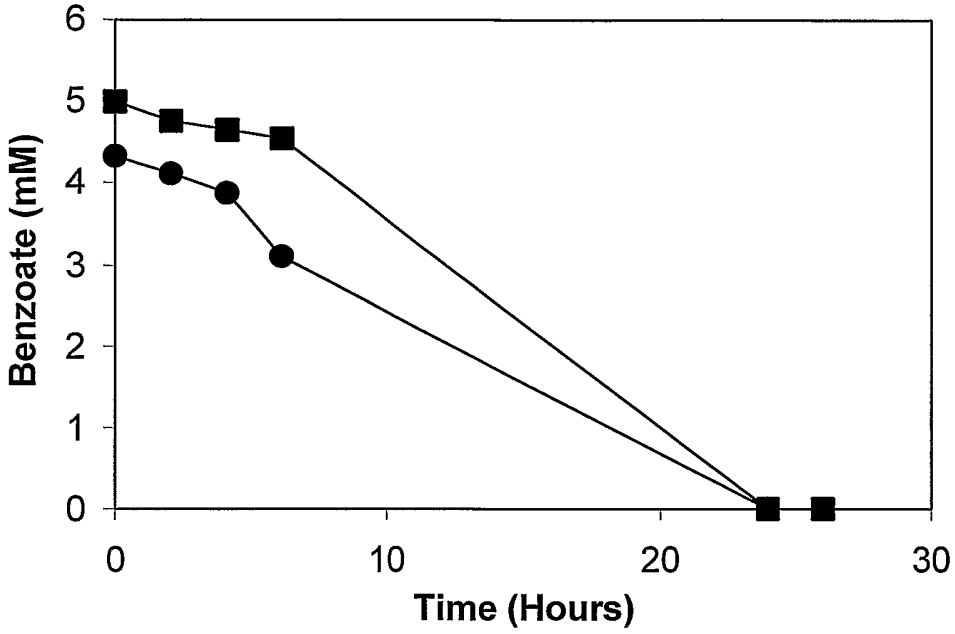


FIGURE 18A

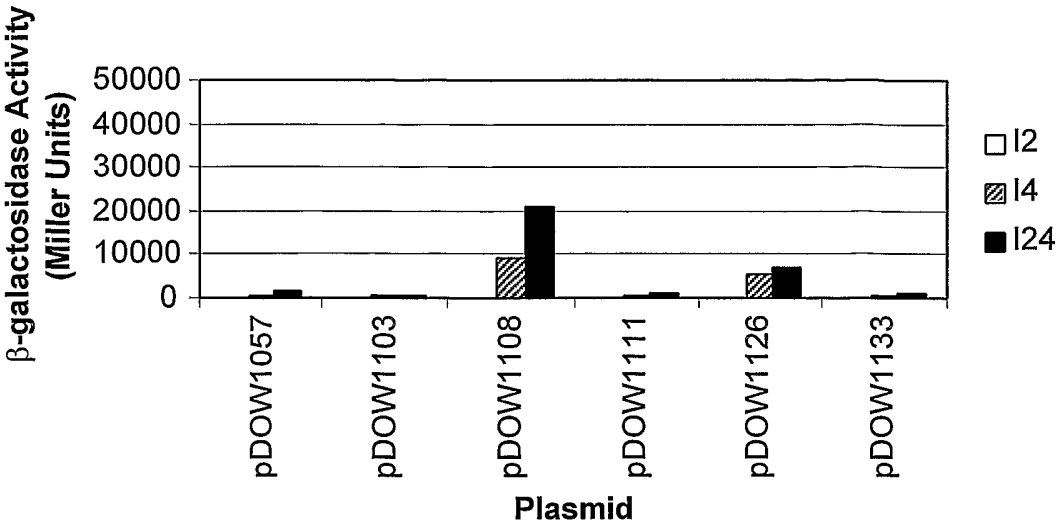


FIGURE 18B

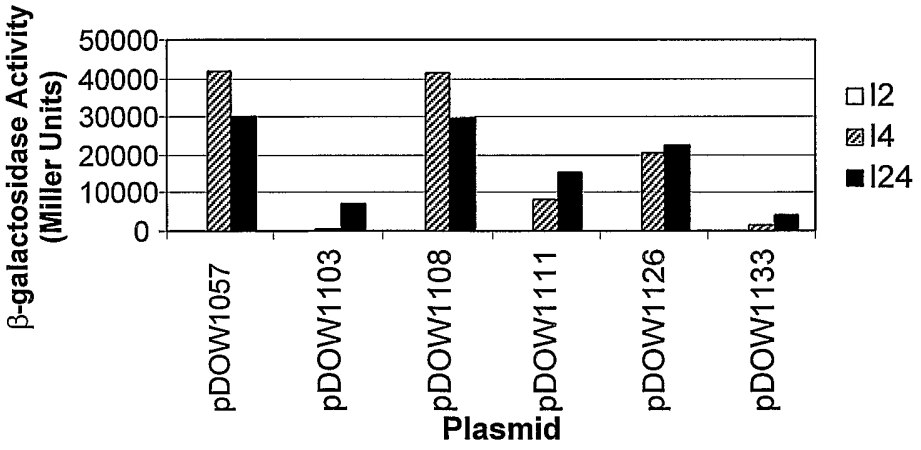


FIGURE 18C

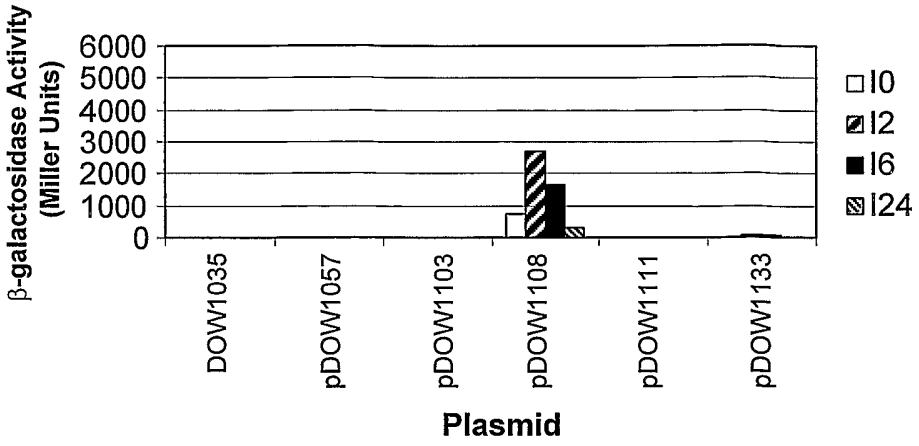


FIGURE 18D

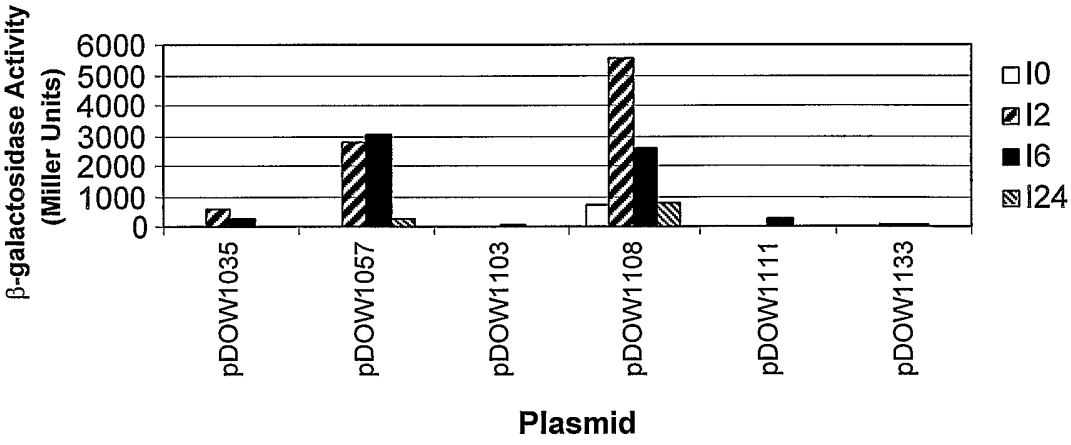
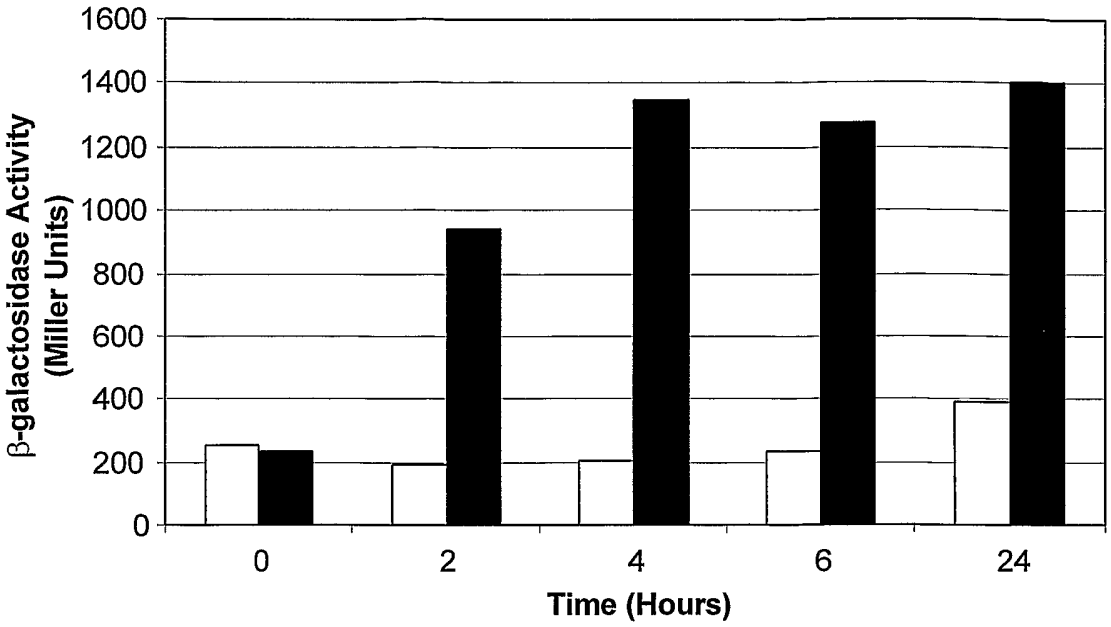


FIGURE 19


1600
1400
1200
1000
800
600
400
200
0
β-galactosidase Activity (Miller Units)
0
2
4
6
24
Time (Hours)

SEQUENCE LISTING

<110> Retallack, Diane M.
Subramanian, Venkiteswaran

<120> Benzoate- and Anthranilate-Inducible Promoters

<130> 62033A WO

<150> US 60/393,422
<151> 2002-07-03

<160> 41

<170> Microsoft Word 97 SR-2

<210> 1
<211> 5006
<212> DNA
<213> Pseudomonas fluorescens

<220>
<221> misc_feature
<222> 1..5006
<223> Benzoate Operon controlling expression of benABCD

<220>
<221> CDS
<222> 225..284
<223> Alternative amino-terminal-portion-encoding CDS, starting from alternative putative initator codon (ttg225-227) of CDS encoding BenR, giving Met1->Arg335 as the amino acid sequence of the Pben activator protein

<220>
<221> CDS
<222> 285..1229
<223> Sense strand of ORF encoding the putative Pben activator protein, starting from putative initiator codon (atg285-287) of CDS encoding BenR, giving Met21->Arg335 as the full amino acid sequence of the Pben activator protein

<220>
<221> mutation
<222> 679..679
<223> An expressed mutation of g679->a679, changing agc->aac and giving Ser152->Asn152 upon expression

<220>
<221> misc_feature
<222> 1106..1106
<223> A mutation of a1106->t1106, found in Pben509 mutant 2d3

<220>
<221> misc_feature
<222> 1223..1223
<223> A mutation of c1223->t1223, found in Pben509 mutant 21b5

<220>
<221> misc_signal
<222> 1228..1274
<223> Approximate region estimated to contain the BenR binding site

<220>
<221> promoter
<222> 1275..1307
<223> Putative promoter (Pben) from benzoate operon (benABCD)

<220>
<221> -35_signal
<222> 1275..1280
<223> Putative -35 region of Pben promoter

<220>
<221> -10_signal
<222> 1296..1301
<223> Putative -10 region of Pben promoter

<220>
<221> misc_feature
<222> 1296..1301
<223> Substitution mutation of Pben -10 region by tataat to form -10con mutants, and by taaggt to form -10benAc mutants

<220>
<221> misc_feature
<222> 1302..1302
<223> A mutation of g1302->a1302, found in Pben509 mutant 21b5

<220>
<221> misc_feature
<222> 1306..1306
<223> A deletion of g1306 found in mutant promoter variants herein

<220>
<221> misc_signal
<222> 1307..1307
<223> Putative transcription start site under control of Pben

<220>
<221> misc_signal
<222> 1340..1342
<223> Putative native translation initiator codon

<220>
<221> CDS
<222> 1340..2713
<223> BenA open reading frame encoding benzoate 1,2-dioxygenase alpha subunit

<220>
<221> misc_signal
<222> 2714..2716
<223> BenA stop codon

<220>
<221> CDS
<222> 2713..3198
<223> BenB open reading frame encoding benzoate 1,2-dioxygenase beta subunit

<220>
<221> misc_signal
<222> 3199..3201

<223> BenB stop codon

<220>
<221> CDS
<222> 3212..4231
<223> BenC open reading frame encoding benzoate 1,2-dioxygenase electron transfer component

<220>
<221> misc_signal
<222> 4232..4234
<223> BenC stop codon

<220>
<221> CDS
<222> 4224..5003
<223> BenD open reading frame encoding cis-1,2-dihydroxycyclohexa-3,5-diene-1-carboxylate dehydrogenase

<220>
<221> misc_signal
<222> 5004..5006
<223> BenD stop codon

<400> 1
gcatgacgtt gttgatttct catcgattct ccagcgtgcg caatgccgac cacatcatcg 60

tgctggacgg cgggcgaatc ctggaggagg gcagtcatca ccaactgatg gcggcgggtg 120

gacgctatgc cgagttgttc gatgttcagg cgcgggggta tcgctaggtc ccgagcgcca 180

cgcccgacga tacctccagg gtatccctcc agactttatc cata ttg gac gcc ccc 236
Met Asp Ala Pro
1

cta ccc aag cgt cag cct gag ccc aat aac gat aaa aga gtc gcg acc 284
Leu Pro Lys Arg Gln Pro Glu Pro Asn Asn Asp Lys Arg Val Ala Thr
5 10 15 20

atg acc gtg cta ttg agt gag cgc agc cag att ttc cag ggc gcc gat 332
Met Thr Val Leu Leu Ser Glu Arg Ser Gln Ile Phe Gln Gly Ala Asp
25 30 35

gcc tac gcg gtg tcg gac tac gtc aac cag cat gtg ggc agc cac tgc 380
Ala Tyr Ala Val Ser Asp Tyr Val Asn Gln His Val Gly Ser His Cys
40 45 50

att cgc ctg cct ccc agg ggc cag ccc cgg gca agt atc agc cat cgc 428
Ile Arg Leu Pro Pro Arg Gly Gln Pro Arg Ala Ser Ile Ser His Arg
55 60 65

acc ttc gcc agc ctg gac ctg tgc cgc atc agc tac ggc gca ccg gtg 476
Thr Phe Ala Ser Leu Asp Leu Cys Arg Ile Ser Tyr Gly Ala Pro Val
70 75 80

cgg gtc acg tcg gtg gcg ctg gag acc atc tac cac ctg cag atc ctc 524
Arg Val Thr Ser Val Ala Leu Glu Thr Ile Tyr His Leu Gln Ile Leu
85 90 95 100

ttg agc ggg cat tgc cgc tcc aac tcc cgt ggc gag gat gat gtg ttc 572
Leu Ser Gly His Cys Arg Ser Asn Ser Arg Gly Glu Asp Asp Val Phe
105 110 115

```
ggg ccg ggg gaa atc ctg ctg atc aat ccg gac gac ccg gta gac ctg    620
Gly Pro Gly Glu Ile Leu Leu Ile Asn Pro Asp Asp Pro Val Asp Leu
            120                 125                 130

acc tat tcc gcc gac tgc gaa aaa ttc atc atc aaa ctg ccg gtg cgc    668
Thr Tyr Ser Ala Asp Cys Glu Lys Phe Ile Ile Lys Leu Pro Val Arg
        135                 140                 145

ctg ctg gaa agc gcc tgc ctg gag cag cac tgg agc ctg ccg cgg gcg    716
Leu Leu Glu Ser Ala Cys Leu Glu Gln His Trp Ser Leu Pro Arg Ala
    150                 155                 160

ggg gtc cgc ttc acg acg gcc cgc cac gcg ctc agt gaa atg ggc ggc    764
Gly Val Arg Phe Thr Thr Ala Arg His Ala Leu Ser Glu Met Gly Gly
165                 170                 175                 180

ttc ctg ccg ttg ctc ggg ttg atc tgc cat gag gcg gaa aac gct gcc    812
Phe Leu Pro Leu Leu Gly Leu Ile Cys His Glu Ala Glu Asn Ala Ala
                185                 190                 195

gag ccc cac atg caa ggc ctg tac gaa cgc atc gtg gcc aac aag ctg    860
Glu Pro His Met Gln Gly Leu Tyr Glu Arg Ile Val Ala Asn Lys Leu
            200                 205                 210

ctg gca ttg ctg ggc agc aat gtg tcg cgg gtg acc ccc cgg gct gcc    908
Leu Ala Leu Leu Gly Ser Asn Val Ser Arg Val Thr Pro Arg Ala Ala
        215                 220                 225

cac ggc ggt ggg ttt gaa gcg gtg cac gaa ttt atc cag cag cac ctg    956
His Gly Gly Gly Phe Glu Ala Val His Glu Phe Ile Gln Gln His Leu
    230                 235                 240

ggc gat gac atc agc gtc gag cag ttg atg gcc gtg gcc aac gtc agt   1004
Gly Asp Asp Ile Ser Val Glu Gln Leu Met Ala Val Ala Asn Val Ser
245                 250                 255                 260

gaa cgt tcg ctg tac agc ctg ttt gag cgc cag gtg ggg ctg tcg ccg   1052
Glu Arg Ser Leu Tyr Ser Leu Phe Glu Arg Gln Val Gly Leu Ser Pro
                265                 270                 275

cgc gat tac gta tgc cgc tgc aag ctc gaa cgc gta cat gca cgc ttg   1100
Arg Asp Tyr Val Cys Arg Cys Lys Leu Glu Arg Val His Ala Arg Leu
            280                 285                 290

caa cta agc agc acg cgc agc gtg acc gag gtg gct ttg gac cat ggg   1148
Gln Leu Ser Ser Thr Arg Ser Val Thr Glu Val Ala Leu Asp His Gly
        295                 300                 305

ttc atg cac cta ggg cgg ttt tcc gaa gcc tat cgc aaa cgc ttc ggc   1196
Phe Met His Leu Gly Arg Phe Ser Glu Ala Tyr Arg Lys Arg Phe Gly
    310                 315                 320

gaa ctg ccg tcg cag acc tgg aaa cgc cat cgt taagcgacgt gcgcctggcg 1249
Glu Leu Pro Ser Gln Thr Trp Lys Arg His Arg
325                 330                 335

gatagcgatg tgcaggcagc ggatattgac gggcagggcg agcacgtacg gtgagggcgc 1309

ctgatacaag aacaacggag ggcccgcccc atg atc agt aca ccc gac cga ctc  1363
                                  Met Ile Ser Thr Pro Asp Arg Leu
                                  1               5
```

```
gcc tgc caa ttg cgc gag tcc gta cag gaa gac ccc gcc act ggg gtg   1411
Ala Cys Gln Leu Arg Glu Ser Val Gln Glu Asp Pro Ala Thr Gly Val
    10                  15                  20

ttc cgc tgc cgc cgc gac atc ttc acc gac ccc gac ctg ttt gcc ctg   1459
Phe Arg Cys Arg Arg Asp Ile Phe Thr Asp Pro Asp Leu Phe Ala Leu
25                  30                  35                  40

gag atg aaa cac atc ttc gaa ggc ggg tgg atc tac ctg gcc cat gaa   1507
Glu Met Lys His Ile Phe Glu Gly Gly Trp Ile Tyr Leu Ala His Glu
                45                  50                  55

agc cag gtg ccg cag atc aac gat tac ttc acc acc tgg atc ggc cgc   1555
Ser Gln Val Pro Gln Ile Asn Asp Tyr Phe Thr Thr Trp Ile Gly Arg
            60                  65                  70

cag ccg gtg gtc atc acc cgt gac aag cac ggc gcg ctg cat ggc ctg   1603
Gln Pro Val Val Ile Thr Arg Asp Lys His Gly Ala Leu His Gly Leu
        75                  80                  85

gtc aac gcc tgc gcg cat cgc ggc gcc atg ttg tgc cgg cgc aaa caa   1651
Val Asn Ala Cys Ala His Arg Gly Ala Met Leu Cys Arg Arg Lys Gln
    90                  95                  100

ggc aac aag ggc tca ttc act tgc ccc ttc cat ggc tgg acg ttc agc   1699
Gly Asn Lys Gly Ser Phe The Cys Pro Phe His Gly Trp Thr Phe Ser
105                 110                 115                 120

aac gcc ggc aag ctg ctc aag gta aag gac gca aag acc ggc gcc tac   1747
Asn Ala Gly Lys Leu Leu Lys Val Lys Asp Ala Lys Thr Gly Ala Tyr
                125                 130                 135

ccg gac agc ttc gac tgc gac ggc tcc cat gac ctc aag cgc ctg gcg   1795
Pro Asp Ser Phe Asp Cys Asp Gly Ser His Asp Leu Lys Arg Leu Ala
            140                 145                 150

cgc ttt gaa aac tac cgc ggt ttc ctg ttc gcc agc ctc agc gat gcg   1843
Arg Phe Glu Asn Tyr Arg Gly Phe Leu Phe Ala Ser Leu Ser Asp Ala
        155                 160                 165

gtg ccg gaa ctc agc gat tac ttg ggt gaa acc cgc gtc atc atc gac   1891
Val Pro Glu Leu Ser Asp Tyr Leu Gly Glu Thr Arg Val Ile Ile Asp
    170                 175                 180

cag atg gtc gac cag gcc cct ttg ggc ctg gag gtg ctg cgc ggc agc   1939
Gln Met Val Asp Gln Ala Pro Leu Gly Leu Glu Val Leu Arg Gly Ser
185                 190                 195                 200

tct tcc tat gtc tat gac ggc aac tgg aag ctg caa atc gaa aac ggc   1987
Ser Ser Tyr Val Tyr Asp Gly Asn Trp Lys Leu Gln Ile Glu Asn Gly
                205                 210                 215

gcc gac ggt tac cac gtc agc tcc gtg cac tgg aac tac tcg gcg acc   2035
Ala Asp Gly Tyr His Val Ser Ser Val His Trp Asn Tyr Ser Ala Thr
            220                 225                 230

atg ggc cgg cgc aac tac gac gcc gaa ggc acg cgc acc gtc gac gcc   2083
Met Gly Arg Arg Asn Tyr Asp Ala Glu Gly Thr Arg Thr Val Asp Ala
        235                 240                 245

aat ggc tgg tcg aaa agc ctg ggc ggc gtc tac gcc ttc gac cac ggg   2131
```

```
Asn Gly trp Ser Lys Ser Leu Gly Gly Val Tyr Ala Phe Asp His Gly
    250                 255                 260

cat atc ctg ctg tgg acg cgc ctg ctt aac ccc caa gtg cgc ccg gtg   2179
His Ile Leu Leu Trp Thr Arg Leu Leu Asn Pro Gln Val Arg Pro Val
265                 270                 275                 280

cac gct cac cgc gag gcc ttg gcc gaa cgc ctg ggc caa gcg cgc gcc   2227
His Ala His Arg Glu Ala Leu Ala Glu Arg Leu Gly Gln Ala Arg Ala
                285                 290                 295

gac ttt atc gtc gac cag acc cgc aac ctc tgt ctc tac ccc aat gtg   2275
Asp Phe Ile Val Asp Gln Thr Arg Asn Leu Cys Leu Tyr Pro Asn Val
            300                 305                 310

tac ctg atg gac cag ttc tcg acc cag atc cgc gtg gtg cgg ccc ctc   2323
Tyr Leu Met Asp Gln Phe Ser Thr Gln Ile Arg Val Val Arg Pro Leu
        315                 320                 325

gcc gtg gat aaa acc gaa gtg aca atc tat tgc atg gcg ccc atc ggc   2371
Ala Val Asp Lys The Glu Val The Ile Tyr Cys Met Ala Pro Ile Gly
    330                 335                 340

gaa agc gcc cag gag cgc gcc acg cgg att cgc cag tac gaa gac ttc   2419
Glu Ser Ala Gln Glu Arg Ala Thr Arg Ile Arg Gln Tyr Glu Asp Phe
345                 350                 355                 360

ttc aat gtc agc ggc atg ggc acc ccg gat gac ctc gag gag ttc cgc   2467
Phe Asn Val Ser Gly Met Gly Thr Pro Asp Asp Leu Glu Glu Phe Arg
                365                 370                 375

gcc tgc cag acc ggt tac cag ggc gcg agc acc ctg tgg aat gac ttg   2515
Ala Cys Gln Thr Gly Tyr Gln Gly Ala Ser Thr Leu Trp Asn Asp Leu
            380                 385                 390

agc cgt ggc gcc aag cag tgg gtc gag ggt gcg gac gaa aat gcc ttg   2563
Ser Arg Gly Ala Lys Gln Trp Val Glu Gly Ala Asp Glu Asn Ala Leu
        395                 400                 405

gcc atg ggt atg caa ccg cag ctc agc ggg gtc aag acc gag gac gag   2611
Ala Met Gly Met Gln Pro Gln Leu Ser Gly Val Lys The Glu Asp Glu
    410                 415                 420

ggc ttg ttt gtg cgc cag cat gcg cac tgg gcc caa agc ctg cag cgt   2659
Gly Leu Phe Val Arg Gln His Ala His Trp Ala Gln Ser Leu Gln Arg
425                 430                 435                 440

gca atc gag cgc gaa cag caa ggg ctg ata gcc agc gac tgt gag gtg   2707
Ala Ile Glu Arg Glu Gln Gln Gly Leu Ile Ala Ser Asp Cys Glu Val
                445                 450                 455

ctg cca tg agc ctt gcc cgg gac cac ctg ctg gat ttt ctt tac cgt   2754
Leu Pro    Ser Leu Ala Arg Asp His Leu Leu Asp Phe Leu Tyr Arg
                       5                  10

gaa gcg cgc ctg ctc gac gac cgc caa tgg gat gaa tgg ctg gcc tgc   2802
Glu Ala Arg Leu Leu Asp Asp Arg Gln Trp Asp Glu Trp Leu Ala Cys
15                  20                  25                  30

tat tcg ccc aag gcc gag ttc tgg atg ccc gcc tgg gac gat cac gac   2850
Tyr Ser Pro Lys Ala Glu Phe Trp Met Pro Ala Trp Asp Asp His Asp
                35                  40                  45
```

```
act ctt acc gaa gac ccg cag cgc gaa atc tcg ctg atc tac tac ccc   2898
Thr Leu Thr Glu Asp Pro Gln Arg Glu Ile Ser Leu Ile Tyr Tyr Pro
            50                  55                  60

aac cgt gac ggc ctg gaa gac cgc atc ttt cgc atc aag act gag cgc   2946
Asn Arg Asp Gly Leu Glu Asp Arg Ile Phe Arg Ile Lys Thr Glu Arg
        65                  70                  75

tcc agc gcc agc acg ccc gag ccg cgc acc gtg cac atg ctg tgc aac   2994
Ser Ser Ala Ser Thr Pro Glu Pro Arg Thr Val His Met Leu Cys Asn
    80                  85                  90

ctc gaa gtg ctg gcc gac gac ggc gcg cag gtg gac ctg cgt ttc aac   3042
Leu Glu Val Leu Ala Asp Asp Gly Ala Gln Val Asp Leu Arg Phe Asn
95                  100                 105                 110

tgg cac acc ctc agc cac cgc tac aaa acc acc gac agt tat ttc ggt   3090
Trp His Thr Leu Ser His Arg Tyr Lys Thr Thr Asp Ser Tyr Phe Gly
                115                 120                 125

acc tcc ttc tat cgc ctc gac atc cgt gcc gag cag ccg ttg ata acg   3138
Thr Ser Phe Tyr Arg Leu Asp Ile Arg Ala Glu Gln Pro Leu Ile Thr
            130                 135                 140

cgc aag aag gtg gtg ctg aaa aac gat tac atc cac cag gtc atc gac   3186
Arg Lys Lys Val Val Leu Lys Asn Asp Tyr Ile His Gln Val Ile Asp
        145                 150                 155

atc tac cat atc tgaggacacc gcc atg acg tat gcc atc gcc ttg aac   3235
Ile Tyr His Ile               Met Thr Tyr Ala Ile Ala Leu Asn
    160                       1               5

ttc gag gat gga gtg acc cgc ttc atc gac tgc aag gtg gga gaa aag   3283
Phe Glu Asp Gly Val Thr Arg Phe Ile Asp Cys Lys Val Gly Glu Lys
    10                  15                  20

gtg ctc gat gcg gcc ttc cgc caa cgc atc aac ctg ccc atg gac tgc   3331
Val Leu Asp Ala Ala Phe Arg Gln Arg Ile Asn Leu Pro Met Asp Cys
25                  30                  35                  40

tcg gac ggc gtg tgc ggc acc tgc aaa tgc cgc tgt gaa acc ggc gcc   3379
Ser Asp Gly Val Cys Gly Thr Cys Lys Cys Arg Cys Glu Thr Gly Ala
                45                  50                  55

tac gac ctg ggc gac gac ttt atc gac gac gcc ctg agc gcc gac gaa   3427
Tyr Asp Leu Gly Asp Asp Phe Ile Asp Asp Ala Leu Ser Ala Asp Glu
            60                  65                  70

gcg cag gcg cgc cgg gtg ctg acc tgc caa atg gtg ccg cag tcc gac   3475
Ala Gln Ala Arg Arg Val Leu Thr Cys Gln Met Val Pro Gln Ser Asp
        75                  80                  85

tgc gtg atc gcc gtg ccg gtg ccg tcc agc gcc tgc aag acc ggc acc   3523
Cys Val Ile Ala Val Pro Val Pro Ser Ser Ala Cys Lys Thr Gly Thr
    90                  95                  100

acg cac ttt gcc gcg acg ctg gcc ggc atc acc cga cat gcc gat gcg   3571
Thr His Phe Ala Ala Thr Leu Ala Gly Ile Thr Arg His Ala Asp Ala
105                 110                 115                 120

gcg ctg gag gtg agt ttc gaa ctg gac cag gcg ccg gta ttc ctg ccc   3619
```

```
Ala Leu Glu Val Ser Phe Glu Leu Asp Gln Ala Pro Val Phe Leu Pro
                125                 130                 135

ggc cag tac gtg aat atc agc gtg ccc gac agt ggg cag act cgt gct   3667
Gly Gln Tyr Val Asp Ile Ser Val Pro Asp Ser Gly Gln Thr Arg Ala
            140                 145                 150

tac tcc ttc agc agt ccc ccg ggc gac ccg cgc gcc agc ttc ctg atc   3715
Tyr Ser Phe Ser Ser Pro Pro Gly Asp Pro Arg Ala Ser Phe Leu Ile
        155                 160                 165

aag cac gtg ccc ggc ggg ttg atg agc ggt tgg ctc gag cgc gcc cag   3763
Lys His Val Pro Gly Gly Leu Mey Ser Glu Trp Leu Glu Arg Ala Gln
    170                 175                 180

ccg ggc gac agc gtg gcg atc acc ggc cca ctg ggg agt ttc tac ctg   3811
Pro Gly Asp Ser Val Ala Ile Thr Gly Pro Leu Gly Ser Phe Tyr Leu
185                 190                 195                 200

cgt gag gtg gcg cgg ccg ctg ctg tta ctg gcc ggt ggt acc ggc ctg   3859
Arg Glu Val Ala Arg Pro Leu Leu Leu Leu Ala Gly Gly Thr Gly Leu
                205                 210                 215

gcg ccg ttc ctg tcg atg ctt gaa gtg ctc gcg cag cgc cag gaa acc   3907
Ala Pro Phe Leu Ser Met Leu Glu Val Leu Ala Gln Arg Gln Glu Thr
            220                 225                 230

cgc ccg atc cgg ttg atc tac ggc gta acg cgg gat cag gac ctg gtg   3955
Arg Pro Ile Arg Leu Ile Tyr Gly Val Thr Arg Asp Gln Asp Leu Val
        235                 240                 245

atg att gag gcg ttg cag gct ttt acc gcg cgt ttg ccc gac ttc aac   4003
Met Ile Glu Ala Leu Gln Ala Phe Thr Ala Arg Leu Pro Asp Phe Asn
    250                 255                 260

ctg gtg acc tgc gtg gct gat ccg cac acc act cac ccg cgc cag ggc   4051
Leu Val Thr Cys Val Ala Asp Pro His Thr Thr His Pro Arg Gln Gly
265                 270                 275                 280

tat gtg acc cag cac atg gcc gac gaa gcc ctc aat ggc ggc gat gtc   4099
Tyr Val Thr Gln His Met Ala Asp Glu Ala Leu Asn Gly Gly Asp Val
                285                 290                 295

gac gtg tac ctg tgc ggc ccg ccg ccg atg gtc gat gcg gtg cgc gag   4147
Asp Val Tyr Leu Cys Gly Pro Pro Pro Met Val Asp Ala Val Arg Glu
            300                 305                 310

cac ttc aag cag caa agc gtg acc ccg gcc agc ttc cat tac gag aaa   4195
His Phe Lys Gln Gln Ser Val Thr Pro Ala Ser Phe His Tyr Glu Lys
        315                 320                 325

ttc acc cct aac gcc gtc gcc acg tgc gat gcc gcc t gag gac tgc cgc 4244
Phe Thr Pro Asn Ala Val Ala Thr Cys Asp Ala Ala   Glu Asp Cys Arg
    330                 335                 340       5

atg act caa cgg ttt aac aac aag gtc gcg ctg gtt acc ggc gct gcg   4292
Met Thr Gln Arg Phe Asn Asn Lys Val Ala Leu Val Thr Gly Ala Ala
        10                  15                  20

caa ggc atc ggc cga cgt gtc gcc gaa cgc ttg ctg gag gag ggg gcc   4340
Gln Gly Ile Gly Arg Arg Val Ala Glu Arg Leu Leu Glu Glu Gly Ala
    25                  30                  35
```

```
tgg ctg gtc gcg gtg gat cgc tcc gag ctc gtg cat gaa ttg cag cat   4388
Trp Leu Val Ala Val Asp Arg Ser Glu Leu Val His Glu Leu Gln His
40                  45                  50                  55

gag cga gcg cta ctg ctg acc gcc gac ctg gaa caa tac agc gag tgc   4436
Glu Arg Ala Leu Leu Leu Thr Ala Asp Leu Glu Gln Tyr Ser Glu Cys
                60                  65                  70

gca cgg gta atg gcc gcc gcc acg gcg cgt ttc ggg cgc ata gac gtg   4484
Ala Arg Val Met Ala Ala Ala Thr Ala Arg Phe Gly Arg Ile Asp Val
            75                  80                  85

ctg gtc aat aac gtc ggc ggg acc atc tgg gcc aag cct ttt gag cat   4532
Leu Val Asn Asn Val Gly Gly The Ile Trp Ala Lys Pro Phe Glu His
        90                  95                  100

tat gcc gag gct gaa atc gag gcc gaa gtg cgc cgc tcg ctg ttc cct   4580
Tyr Ala Glu Ala Glu Ile Glu Ala Glu Val Arg Arg Ser Leu Phe Pro
    105                 110                 115

acg ttg tgg tgc tgc cat tgc gtg ctg ccc tat atg ctg gag cag ggc   4628
Thr Leu Trp Cys Cys His Cys Val Leu Pro Tyr Met Leu Glu Gln Gly
120                 125                 130                 135

gcg ggc gcg atc gtc aac gtg tct tcc gtg gcc acg cgc ggg gtc aat   4676
Ala Gly Ala Ile Val Asn Val Ser Ser Val Ala Thr Arg Gly Val Asn
                140                 145                 150

cgc gtg ccc tat ggc gca gcc aag ggc ggc gtg aat gcc ttg acg gcc   4724
Arg Val Pro Tyr Gly Ala Ala Lys Gly Gly Val Asn Ala Leu Thr Ala
            155                 160                 165

tgc ctg gcc ctg gag act gca ggc agc ggg att cgc gtc aac gcc acc   4772
Cys Leu Ala Leu Glu Thr Ala Gly Ser Gly Ile Arg Val Asn Ala Thr
        170                 175                 180

gcg ccc ggc ggc acc gag gca ccg cca cgg cgc atc ccg cgc aac agc   4820
Ala Pro Gly Gly Thr Glu Ala Pro Pro Arg Arg Ile Pro Arg Asn Ser
    185                 190                 195

cag ccg cag agc gag cag gaa cgt gtg tgg tac cag cag atc gtc gac   4868
Gln Pro Gln Ser Glu Gln Glu Arg Val Trp Tyr Gln Gln Ile Val Asp
200                 205                 210                 215

cag acc ctc gag agc agc tcg atg aaa cgc tac ggc agc atc gac gaa   4916
Gln Thr Leu Glu Ser Ser Ser Met Lys Arg Tyr Gly Ser Ile Asp Glu
                220                 225                 230

caa gct ggc gca att ctg ttc ctg gcc tgc gac gag gcc tcc tac atc   4964
Gln Ala Gly Ala Ile Leu Phe Leu Ala Cys Asp Glu Ala Ser Tyr Ile
            235                 240                 245

acc ggc gtg acc ttg ccg gtg ggc ggc ggc gac ctc ggc taa           5006
Thr Gly Val Thr Leu Pro Val Gly Gly Gly Asp Leu Gly
        250                 255                 260
```

<210> 2
<211> 335
<212> PRT
<213> Pseudomonas fluorescens

<220>
<221> INIT_MET
<222> 1..1
<223> Alternative putative initiator methionine for BenR, giving Met1->Arg335 as the amino acid sequence of the Pben activator protein

<220>
<221> INIT_MET
<222> 21..21
<223> Putative initiator methionine of BenR, giving Met21->Arg335 as the amino acid sequence of the Pben activator protein

<220>
<221> VARIANT
<222> 152..152
<223> A mutation of Ser152->Asn152

<400> 2

```
Met Asp Ala Pro Leu Pro Lys Arg Gln Pro Glu Pro Asn Asn Asp Lys
1               5                  10                  15

Arg Val Ala Thr Met Thr Val Leu Leu Ser Glu Arg Ser Gln Ile Phe
            20                  25                  30

Gln Gly Ala Asp Ala Tyr Ala Val Ser Asp Tyr Val Asn Gln His Val
        35                  40                  45

Gly Ser His Cys Ile Arg Leu Pro Pro Arg Gly Gln Pro Arg Ala Ser
    50                  55                  60

Ile Ser His Arg Thr Phe Ala Ser Leu Asp Leu Cys Arg Ile Ser Tyr
65                  70                  75                  80

Gly Ala Pro Val Arg Val Thr Ser Val Ala Leu Glu Thr Ile Tyr His
                85                  90                  95

Leu Gln Ile Leu Leu Ser Gly His Cys Arg Ser Asn Ser Arg Gly Glu
            100                 105                 110

Asp Asp Val Phe Gly Pro Gly Glu Ile Leu Leu Ile Asn Pro Asp Asp
        115                 120                 125

Pro Val Asp Leu Thr Tyr Ser Ala Asp Cys Glu Lys Phe Ile Ile Lys
    130                 135                 140

Leu Pro Val Arg Leu Leu Glu Ser Ala Cys Leu Glu Gln His Trp Ser
145                 150                 155                 160

Leu Pro Arg Ala Gly Val Arg Phe Thr Thr Ala Arg His Ala Leu Ser
                165                 170                 175

Glu Met Gly Gly Phe Leu Pro Leu Leu Gly Leu Ile Cys His Glu Ala
            180                 185                 190

Glu Asn Ala Ala Glu Pro His Met Gln Gly Leu Tyr Glu Arg Ile Val
        195                 200                 205

Ala Asn Lys Leu Leu Ala Leu Leu Gly Ser Asn Val Ser Arg Val Thr
    210                 215                 220

Pro Arg Ala Ala His Gly Gly Gly Phe Glu Ala Val His Glu Phe Ile
225                 230                 235                 240
```

```
Gln Gln His Leu Gly Asp Asp Ile Ser Val Glu Gln Leu Met Ala Val
                245                 250                 255

Ala Asn Val Ser Glu Arg Ser Leu Tyr Ser Leu Phe Glu Arg Gln Val
            260                 265                 270

Gly Leu Ser Pro Arg Asp Tyr Val Cys Arg Cys Lys Leu Glu Arg Val
        275                 280                 285

His Ala Arg Leu Gln Leu Ser Ser Thr Arg Ser Val Thr Glu Val Ala
    290                 295                 300

Leu Asp His Gly Phe Met His Leu Gly Arg Phe Ser Glu Ala Tyr Arg
305                 310                 315                 320

Lys Arg Phe Gly Glu Leu Pro Ser Gln Thr Trp Lys Arg His Arg
                325                 330                 335

<210> 3
<211> 458
<212> PRT
<213> Pseudomonas fluorescens

<220>
<221> misc_feature
<222> 1..458
<223> BenA expression product, benzoate 1,2-dioxygenase alpha subunit

<400> 3
Met Ile Ser Thr Pro Asp Arg Leu Ala Cys Gln Leu Arg Glu Ser Val
1               5                   10                  15

Gln Glu Asp Pro Ala Thr Gly Val Phe Arg Cys Arg Arg Asp Ile Phe
            20                  25                  30

Thr Asp Pro Asp Leu Phe Ala Leu Glu Met Lys His Ile Phe Glu Gly
        35                  40                  45

Gly Trp Ile Tyr Leu Ala His Glu Ser Gln Val Pro Gln Ile Asn Asp
    50                  55                  60

Tyr Phe Thr Thr Trp Ile Gly Arg Gln Pro Val Val Ile Thr Arg Asp
65                  70                  75                  80

Lys His Gly Ala Leu His Gly Leu Val Asn Ala Cys Ala His Arg Gly
                85                  90                  95

Ala Met Leu Cys Arg Arg Lys Gln Gly Asn Lys Gly Ser Phe The Cys
            100                 105                 110

Pro Phe His Gly Trp Thr Phe Ser Asn Ala Gly Lys Leu Leu Lys Val
        115                 120                 125

Lys Asp Ala Lys Thr Gly Ala Tyr Pro Asp Ser Phe Asp Cys Asp Gly
    130                 135                 140

Ser His Asp Leu Lys Arg Leu Ala Arg Phe Glu Asn Tyr Arg Gly Phe
145                 150                 155                 160

Leu Phe Ala Ser Leu Ser Asp Ala Val Pro Glu Leu Ser Asp Tyr Leu
                165                 170                 175
```

```
Gly Glu Thr Arg Val Ile Ile Asp Gln Met Val Asp Gln Ala Pro Leu
            180                 185                 190

Gly Leu Glu Val Leu Arg Gly Ser Ser Ser Tyr Val Tyr Asp Gly Asn
        195                 200                 205

Trp Lys Leu Gln Ile Glu Asn Gly Ala Asp Gly Tyr His Val Ser Ser
    210                 215                 220

Val His Trp Asn Tyr Ser Ala Thr Met Gly Arg Arg Asn Tyr Asp Ala
225                 230                 235                 240

Glu Gly Thr Arg Thr Val Asp Ala Asn Gly trp Ser Lys Ser Leu Gly
                245                 250                 255

Gly Val Tyr Ala Phe Asp His Gly His Ile Leu Leu Trp Thr Arg Leu
            260                 265                 270

Leu Asn Pro Gln Val Arg Pro Val His Ala His Arg Glu Ala Leu Ala
        275                 280                 285
Glu Arg Leu Gly Gln Ala Arg Ala Asp Phe Ile Val Asp Gln Thr Arg
    290                 295                 300

Asn Leu Cys Leu Tyr Pro Asn Val Tyr Leu Met Asp Gln Phe Ser Thr
305                 310                 315                 320

Gln Ile Arg Val Val Arg Pro Leu Ala Val Asp Lys The Glu Val The
                325                 330                 335

Ile Tyr Cys Met Ala Pro Ile Gly Glu Ser Ala Gln Glu Arg Ala Thr
            340                 345                 350

Arg Ile Arg Gln Tyr Glu Asp Phe Phe Asn Val Ser Gly Met Gly Thr
        355                 360                 365

Pro Asp Asp Leu Glu Glu Phe Arg Ala Cys Gln Thr Gly Tyr Gln Gly
    370                 375                 380

Ala Ser Thr Leu Trp Asn Asp Leu Ser Arg Gly Ala Lys Gln Trp Val
385                 390                 395                 400

Glu Gly Ala Asp Glu Asn Ala Leu Ala Met Gly Met Gln Pro Gln Leu
                405                 410                 415

Ser Gly Val Lys The Glu Asp Glu Gly Leu Phe Val Arg Gln His Ala
            420                 425                 430

His Trp Ala Gln Ser Leu Gln Arg Ala Ile Glu Arg Glu Gln Gln Gly
        435                 440                 445

Leu Ile Ala Ser Asp Cys Glu Val Leu Pro
    450                 455

<210> 4
<211> 162
<212> PRT
<213> Pseudomonas fluorescens

<220>
<221> misc_feature
<222> 1..162
```

<223> BenB expression product, benzoate 1,2-dioxygenase beta subunit

<400> 4
Met Ser Leu Ala Arg Asp His Leu Leu Asp Phe Leu Tyr Arg Glu Ala
1 5 10 15

Arg Leu Leu Asp Asp Arg Gln Trp Asp Glu Trp Leu Ala Cys Tyr Ser
20 25 30

Pro Lys Ala Glu Phe Trp Met Pro Ala Trp Asp Asp His Asp Thr Leu
35 40 45

Thr Glu Asp Pro Gln Arg Glu Ile Ser Leu Ile Tyr Tyr Pro Asn Arg
50 55 60

Asp Gly Leu Glu Asp Arg Ile Phe Arg Ile Lys Thr Glu Arg Ser Ser
65 70 75 80

Ala Ser Thr Pro Glu Pro Arg Thr Val His Met Leu Cys Asn Leu Glu
85 90 95

Val Leu Ala Asp Asp Gly Ala Gln Val Asp Leu Arg Phe Asn Trp His
100 105 110

Thr Leu Ser His Arg Tyr Lys Thr Thr Asp Ser Tyr Phe Gly Thr Ser
115 120 125

Phe Tyr Arg Leu Asp Ile Arg Ala Glu Gln Pro Leu Ile Thr Arg Lys
130 135 140

Lys Val Val Leu Lys Asn Asp Tyr Ile His Gln Val Ile Asp Ile Tyr
145 150 155 160

His Ile

<210> 5
<211> 340
<212> PRT
<213> Pseudomonas fluorescens

<220>
<221> misc_feature
<222> 1..340
<223> BenC expression product, benzoate 1,2-dioxygenase electron transfer component

<400> 5
Met Thr Tyr Ala Ile Ala Leu Asn Phe Glu Asp Gly Val Thr Arg Phe
1 5 10 15

Ile Asp Cys Lys Val Gly Glu Lys Val Leu Asp Ala Ala Phe Arg Gln
20 25 30

Arg Ile Asn Leu Pro Met Asp Cys Ser Asp Gly Val Cys Gly Thr Cys
35 40 45

Lys Cys Arg Cys Glu Thr Gly Ala Tyr Asp Leu Gly Asp Asp Phe Ile
50 55 60

Asp Asp Ala Leu Ser Ala Asp Glu Ala Gln Ala Arg Arg Val Leu Thr
65 70 75 80

```
Cys Gln Met Val Pro Gln Ser Asp Cys Val Ile Ala Val Pro Val Pro
                85                  90                  95

Ser Ser Ala Cys Lys Thr Gly Thr Thr His Phe Ala Ala Thr Leu Ala
            100                 105                 110

Gly Ile Thr Arg His Ala Asp Ala Ala Leu Glu Val Ser Phe Glu Leu
        115                 120                 125

Asp Gln Ala Pro Val Phe Leu Pro Gly Gln Tyr Val Asp Ile Ser Val
    130                 135                 140

Pro Asp Ser Gly Gln Thr Arg Ala Tyr Ser Phe Ser Ser Pro Pro Gly
145                 150                 155                 160

Asp Pro Arg Ala Ser Phe Leu Ile Lys His Val Pro Gly Gly Leu Mey
                165                 170                 175

Ser Glu Trp Leu Glu Arg Ala Gln Pro Gly Asp Ser Val Ala Ile Thr
            180                 185                 190

Gly Pro Leu Gly Ser Phe Tyr Leu Arg Glu Val Ala Arg Pro Leu Leu
        195                 200                 205

Leu Leu Ala Gly Gly Thr Gly Leu Ala Pro Phe Leu Ser Met Leu Glu
    210                 215                 220

Val Leu Ala Gln Arg Gln Glu Thr Arg Pro Ile Arg Leu Ile Tyr Gly
225                 230                 235                 240

Val Thr Arg Asp Gln Asp Leu Val Met Ile Glu Ala Leu Gln Ala Phe
                245                 250                 255

Thr Ala Arg Leu Pro Asp Phe Asn Leu Val Thr Cys Val Ala Asp Pro
            260                 265                 270

His Thr Thr His Pro Arg Gln Gly Tyr Val Thr Gln His Met Ala Asp
        275                 280                 285

Glu Ala Leu Asn Gly Gly Asp Val Asp Val Tyr Leu Cys Gly Pro Pro
    290                 295                 300

Pro Met Val Asp Ala Val Arg Glu His Phe Lys Gln Gln Ser Val Thr
305                 310                 315                 320

Pro Ala Ser Phe His Tyr Glu Lys Phe Thr Pro Asn Ala Val Ala Thr
                325                 330                 335

Cys Asp Ala Ala
            340
```

<210> 6
<211> 260
<212> PRT
<213> Pseudomonas fluorescens

<220>
<221> misc_feature
<222> 1..260
<223> BenD expression product, cis-1,2-dihydroxycyclohexa-3,5-diene-1-carboxylate dehydrogenase

<400> 6
Met Pro Pro Glu Asp Cys Arg Met Thr Gln Arg Phe Asn Asn Lys Val
1 5 10 15

Ala Leu Val Thr Gly Ala Ala Gln Gly Ile Gly Arg Arg Val Ala Glu
20 25 30

Arg Leu Leu Glu Glu Gly Ala Trp Leu Val Ala Val Asp Arg Ser Glu
35 40 45

Leu Val His Glu Leu Gln His Glu Arg Ala Leu Leu Leu Thr Ala Asp
50 55 60

Leu Glu Gln Tyr Ser Glu Cys Ala Arg Val Met Ala Ala Ala Thr Ala
65 70 75 80

Arg Phe Gly Arg Ile Asp Val Leu Val Asn Asn Val Gly Gly The Ile
85 90 95

Trp Ala Lys Pro Phe Glu His Tyr Ala Glu Ala Glu Ile Glu Ala Glu
100 105 110

Val Arg Arg Ser Leu Phe Pro Thr Leu Trp Cys Cys His Cys Val Leu
115 120 125

Pro Tyr Met Leu Glu Gln Gly Ala Gly Ala Ile Val Asn Val Ser Ser
130 135 140

Val Ala Thr Arg Gly Val Asn Arg Val Pro Tyr Gly Ala Ala Lys Gly
145 150 155 160

Gly Val Asn Ala Leu Thr Ala Cys Leu Ala Leu Glu Thr Ala Gly Ser
165 170 175

Gly Ile Arg Val Asn Ala Thr Ala Pro Gly Gly Thr Glu Ala Pro Pro
180 185 190

Arg Arg Ile Pro Arg Asn Ser Gln Pro Gln Ser Glu Gln Glu Arg Val
195 200 205

Trp Tyr Gln Gln Ile Val Asp Gln Thr Leu Glu Ser Ser Ser Met Lys
210 215 220

Arg Tyr Gly Ser Ile Asp Glu Gln Ala Gly Ala Ile Leu Phe Leu Ala
225 230 235 240

Cys Asp Glu Ala Ser Tyr Ile Thr Gly Val Thr Leu Pro Val Gly Gly
245 250 255

Gly Asp Leu Gly
260

<210> 7
<211> 4330
<212> DNA
<213> Pseudomonas fluorescens

<220>
<221> misc_feature
<222> 1..4330
<223> Anthranilate Operon controlling expression of antABC

<220>
<221> misc_signal
<222> 1..3
<223> Anticodon of stop codon for the CDS of AntR

<220>
<221> CDS
<222> 4..993
<223> Antisense strand of ORF encoding AntR

<220>
<221> variation
<222> 192..192
<223> Mutation to A235 from native C235; resulting in the anticodon mutation CGC->CGA shown at 233..235, thus codon mutation GCG->TCG and amino acid mutation of Ala268 to Ser268

<220>
<221> misc_signal
<222> 1130..1237
<223> Approximate region estimated to contain the AntR binding site

<220>
<221> promoter
<222> 1239..1274
<223> Putative promoter (Pant) from anthranilate operon (antABC)

<220>
<221> -35_signal
<222> 1239..1244
<223> Putative -35 region of Pant promoter

<220>
<221> -10_signal
<222> 1264..1268
<223> Putative -10 region of Pant promoter

<220>
<221> misc_feature
<222> 1264..1268
<223> Substitution mutation of Pant -10 region by tataat to form -10con mutants

<220>
<221> misc_feature
<222> 1269..1269
<223> A deletion of g1269 found in mutant promoter variants herein

<220>
<221> misc_signal
<222> 1274..1274
<223> Putative transcription start site under control of Pant

<220>
<221> misc_feature
<222> 1278..1278
<223> A deletion of t1278 found in mutant promoter variants herein

<220>
<221> misc_signal
<222> 1305..1307
<223> Putative native translation initiator codon

<220>
<221> CDS
<222> 1305..2693
<223> AntA open reading frame encoding anthranilate dioxygenase large subunit

<220>
<221> misc_signal
<222> 2694..2696
<223> AntA stop codon

<220>
<221> CDS
<222> 2696..3184
<223> AntB open reading frame encoding anthranilate dioxygenase small subunit

<220>
<221> misc_signal
<222> 3185..3187
<223> AntB stop codon

<220>
<221> CDS
<222> 3323..4327
<223> AntC open reading frame encoding anthranilate dioxygenase reductase

<220>
<221> misc_signal
<222> 4328..4330
<223> AntC stop codon

<400> 7
tca agt aat gcg cag gcg ctt gcg ctg caa tgt ctg gct ggg cga ctc 48
Thr Ile Arg Leu Arg Lys Arg Gln Leu Thr Gln Ser Pro Ser Glu
-330 -325 -320

atc gaa cag ctt gcg gta ctc cgc cga aaa ccg ccc caa atg cgt aaa 96
Asp Phe Leu Lys Arg Tyr Glu Ala Ser Phe Arg Gly Leu His Thr Phe
-315 -310 -305 -300

ccc cca acc cag ggc gat ttc aga gat ggt gcg gat cga gcc ctg ctc 144
Gly Trp Gly Leu Ala Ile Glu Ser Ile Thr Arg Ile Ser Gly Gln Glu
-295 -290 -285

cag aat ttc ttg gcg cac cgc ccc caa ccg atg ctt ctt caa ata cgc 192
Leu Ile Glu Gln Arg Val Ala Gly Leu Arg His Lys Lys Leu Tyr Ala
-280 -275 -270

cat ggg cga cag tgc gaa gta ctt gcg aaa cgc atc gaa cag ttt gaa 240
Met Pro Ser Leu Ala Phe Tyr Lys Arg Phe Ala Asp Phe Leu Lys Phe
-265 -260 -255

acg cga cac gcc cgc cgc cgc ttc cag gtc ttc cag gtg cag cgc ttc 288
Arg Ser Val Gly Ala Ala Ala Glu Leu Asp Glu Leu His Leu Ala Glu
-250 -245 -240

acg ggc gtt gtc gtg gat aaa ttg ccg cgc gcg gat cag gta gtg cgg 336
Arg Ala Asn Asp His Ile Phe Gln Arg Ala Arg Ile Leu Tyr His Pro
-235 -230 -225 -220

```
cag ttt cac ccc cag cac gtc gcg cag ttc ttc gga gta gtt att cgg    384
Leu Lys Val Gly Leu Val Asp Arg Leu Glu Glu Ser Tyr Asn Asn Pro
                -215                -210                -205

ttg ggc cag gat cag gcc ctt gat cag cga gct ttc cag gtc gcg agt    432
Gln Ala Leu Ile Leu Gly Lys Ile Leu Ser Ser Glu Leu Asp Arg Thr
            -200                -195                -190

aaa cgc cgc ctg ctc gta cag ttc gct gct gcg ctc cag ttc ggc gat    480
Phe Ala Ala Gln Glu Tyr Leu Glu Ser Ser Arg Glu Leu Glu Ala Ile
        -185                -180                -175

gaa ata acg cgc cat gcg cca cca cga agc cgg tgc tcc gtc cac agc    528
Phe Tyr Arg Ala Met Arg Trp Trp Ser Ala Pro Ala Gly Asp Val Ala
    -170                -165                -160

atc cat cac cga ctc aaa gcg cag cgg cgc atc aat ggg ccg ttg cag    576
Asp Met Val Ser Glu Phe Arg Leu Pro Ala Asp Ile Pro Arg Gln Leu
-155                -150                -145                -140

caa acc ttc cag cga ctc gct cat cgc cgc acg ggt gat tac cac ctg    624
Leu Gly Glu Leu Ser Glu Ser Met Ala Ala Arg Thr Ile Val Val Gln
                -135                -130                -125

caa ctt gcg gca gtc acc gga aat cgc cag cac ctg atg ctc att ggg    672
Leu Lys Arg Cys Asp Gly Ser Ile Ala Leu Val Gln His Glu Asn Pro
            -120                -115                -110

cga aat gat cac gcc ttg gtc gcg gtt gga act gag acg ttc acc gtt    720
Ser Ile Ile Val Gly Gln Asp Arg Asn Ser Ser Leu Arg Glu Gly Asn
        -105                -100                -95

ctt gct cag ctc ctg ctc gcc cac cag tgg cag gct caa gct gta gct    768
Lys Ser Leu Glu Gln Glu Gly Val Leu Pro Leu Ser Leu Ser Tyr Ser
    -90                 -85                 -80

gct gaa gtg ctc ggc gtc ttc gat gtc gat ggt cac atc agt gcc gta    816
Ser Phe His Glu Ala Asp Glu Ile Asp Ile Thr Val Asp Thr Gly Tyr
-75                 -70                 -65                 -60

ctc gat cac gcc cag ggt ggt ggc gcg gga ttt gaa cac gtt ggc gct    864
Glu Ile Val Gly Leu Thr Thr Ala Arg Ser Lys Phe Val Asn Ala Ser
                -55                 -50                 -45

gtg gtg aaa gcg cag gcg ctc ggg ggt tgc cgt cgc cag gcg atg ggg    912
His His Phe Arg Leu Arg Glu Pro Thr Ala Thr Ala Leu Arg His Pro
            -40                 -35                 -30

ccc gca gat gcc gga cat cca gct gcg cgc gcc ttc cag gtc gaa gcg    960
Gly Cys Ile Gly Ser Met Trp Ser Arg Ala Gly Glu Leu Asp Phe Arg
        -25                 -20                 -15

ttg aat atg aat atc gcg tgt ctg act agt cat cagggtgcac ccacggcggt 1013
Gln Ile His Ile Asp Arg Thr Gln Ser Thr Met
    -10                 -5                  -1

taggcgtttg cgcgctctga cggcgcgtcg ttgaacctcg acagcaagtt ccaggccacg 1073

ccagtgcagt tctcactggg tggatagcaa cggtcgacta tgtggataaa ccccagagtt 1133

ttgcgaccat cgcccgccat cacagtagcg catgccgtca ccggcgcgca ccgtcatggg 1193
```

```
tatttgccgc ccaactttgc ggcctacgtt cccccattaa gcggatagcc cgccaccgca 1253

tcgcagccgc ttaatggctc accgtttagc catgatcaaa aggtgcctcc c atg agt  1310
                                                          Met Ser
                                                          1

ggt gca aga acc gtc gag caa tgg aaa tcc ttt atc gaa agc tgc ctg   1358
Gly Ala Arg Thr Val Glu Gln Trp Lys Ser Phe Ile Glu Ser Cys Leu
        5                  10                  15

gac ttt cgc ccg gcg gat gaa gtg ttc cgc atc gcc cgc gac atg ttc   1406
Asp Phe Arg Pro Ala Asp Glu Val Phe Arg Ile Ala Arg Asp Met Phe
    20                  25                  30

acc gag ccc gag ttg ttc gac ctg gag atg gag ctg atc ttc gag aag   1454
Thr Glu Pro Glu Leu Phe Asp Leu Glu Met Glu Leu Ile Phe Glu Lys
35                  40                  45                  50

aac tgg atc tac gcc tgc cac gaa agc gaa ctg gcc aat aac cac gac   1502
Asn Trp Ile Tyr Ala Cys His Glu Ser Glu Leu Ala Asn Asn His Asp
                55                  60                  65

ttc gtg acg atg cgc gcc ggc cgc cag ccg atg atc atc acc cgt gac   1550
Phe Val Thr Met Arg Ala Gly Arg Gln Pro Met Ile Ile Thr Arg Asp
            70                  75                  80

ggc gaa ggc cga ctc aac gcg ttg atc aac gcc tgc cag cat cgc ggt   1598
Gly Glu Gly Arg Leu Asn Ala Leu Ile Asn Ala Cys Gln His Arg Gly
        85                  90                  95

acc acc ctc acc cgc gtg ggc aag ggt aac cag tcc acc ttc acc tgc   1646
Thr Thr Leu Thr Arg Val Gly Lys Gly Asn Gln Ser Thr Phe Thr Cys
    100                 105                 110

ccg ttc cac gcc tgg tgc tac aag agc gat ggc cga ctg gta aag gtc   1694
Pro Phe His Ala Trp Cys Tyr Lys Ser Asp Gly Arg Leu Val Lys Val
115                 120                 125                 130

aag gcg ccg ggg gaa tac ccg gaa ggt ttc gac aag gcc acc cgc ggc   1742
Lys Ala Pro Gly Glu Tyr Pro Glu Gly Phe Asp Lys Ala Thr Arg Gly
                135                 140                 145

ctg aaa aaa gcg cgc atc gaa agc tac agg ggc ttt gtg ttt atc agc   1790
Leu Lys Lys Ala Arg Ile Glu Ser Tyr Arg Gly Phe Val Phe Ile Ser
            150                 155                 160

ctg gac gtg aac ggc acc aac agc ctg gag gac ttc ctg ggc gat gcc   1838
Leu Asp Val Asn Gly Thr Asn Ser Leu Glu Asp Phe Leu Gly Asp Ala
        165                 170                 175

aaa gtg ttc ttc gac atg atg gtg gcg caa tcg gcc acc ggt gag ctg   1886
Lys Val Phe Phe Asp Met Met Val Ala Gln Ser Ala Thr Gly Glu Leu
    180                 185                 190

gaa gtg ctg ccg ggc aag tcc gcc tac acc tac gac ggc aac tgg aag   1934
Glu Val Leu Pro Gly Lys Ser Ala Tyr Thr Tyr Asp Gly Asn Trp Lys
195                 200                 205                 210

ctg caa aac gaa aac ggc ctg gac ggt tat cac gtc agc acc gtg cac   1982
Leu Gln Asn Glu Asn Gly Leu Asp Gly Tyr His Val Ser Thr Val His
                215                 220                 225
```

```
tac aac tac gtg gcc acc gtg cag cat cgc gag cag gtc aac acc gaa   2030
Tyr Asn Tyr Val Ala Thr Val Gln His Arg Glu Gln Val Asn Thr Glu
            230                 235                 240

aac ggc gca ggt tcc agc acg acg ttg gac tac agc aag ctc ggc gcc   2078
Asn Gly Ala Gly Ser Ser Thr Thr Leu Asp Tyr Ser Lys Leu Gly Ala
        245                 250                 255

ggc gac gcc aat acc gac gac ggc tgg ttc gcc ttc aac aac ggc cac   2126
Gly Asp Ala Asn Thr Asp Asp Gly Trp Phe Ala Phe Asn Asn Gly His
    260                 265                 270

agc gtg ttg ttt agc gac atg ccc aac ccc agc gtg cgc tcc ggc tac   2174
Ser Val Leu Phe Ser Asp Met Pro Asn Pro Ser Val Arg Ser Gly Tyr
275                 280                 285                 290

gcc acc atc atg ccg cgc ctg gta gaa gaa cac ggc cag cag aag gcc   2222
Ala Thr Ile Met Pro Arg Leu Val Glu Glu His Gly Gln Gln Lys Ala
                295                 300                 305

gag tgg atg atg cac cgc ctg cgc aac ctg aat atc tac ccc agc ctg   2270
Glu Trp Met Met His Arg Leu Arg Asn Leu Asn Ile Tyr Pro Ser Leu
            310                 315                 320

ttt ttc ctc gac cag atc agc tcg cag ttg cgc atc atc cgc ccg gtg   2318
Phe Phe Leu Asp Gln Ile Ser Ser Gln Leu Arg Ile Ile Arg Pro Val
        325                 330                 335

gcc tgg aac aag acc gag atc atc agc cag tgc ctg ggg gtt aag ggc   2366
Ala Trp Asn Lys Thr Glu Ile Ile Ser Gln Cys Leu Gly Val Lys Gly
    340                 345                 350

gag tcc gac gcc gac cgc gaa aac cgg att cgt cag ttc gaa gac ttc   2414
Glu Ser Asp Ala Asp Arg Glu Asn Arg Ile Arg Gln Phe Glu Asp Phe
355                 360                 365                 370

ttc aac gtt tca ggc atg ggc acg ccc gat gac ctg gtg gag ttt cgc   2462
Phe Asn Val Ser Gly Met Gly Thr Pro Asp Asp Leu Val Glu Phe Arg
                375                 380                 385

gaa gcc cag cgt ggc ttt cag ggc cgc ctg gaa cgc tgg agc gac atc   2510
Glu Ala Gln Arg Gly Phe Gln Gly Arg Leu Glu Arg Trp Ser Asp Ile
            390                 395                 400

tca cgg ggc agc cat cgc tgg gag acc ggg ccg acg cca aac agc gag   2558
Ser Arg Gly Ser His Arg Trp Glu Thr Gly Pro Trp Pro Asn Ser Glu
        405                 410                 415

gcc atc ggc atc caa ccg gcg atg acc ggt acc gaa ttc acc cat gaa   2606
Ala Ile Gly Ile Gln Pro Ala Met Thr Gly Thr Glu Phe Thr His Glu
    420                 425                 430

ggc ctg tac gtc aac cag cat cgc aac tgg cag cag ttc ctg cta aag   2654
Gly Leu Tyr Val Asn Gln His Arg Asn Trp Gln Gln Phe Leu Leu Lys
435                 440                 445                 450

ggt ttg gac cag cga gcc ctg gca ctg cgg gag gtg aag tg atg aat    2701
Gly Leu Asp Gln Arg Ala Leu Ala Leu Arg Glu Val Lys    Met Asn
                455                 460                 1

gcg caa ttg cag tac cag atc gag cag ttc ttc tat cgc aag tcc gag   2749
Ala Gln Leu Gln Tyr Gln Ile Glu Gln Phe Phe Tyr Arg Lys Ser Glu
```

```
        5                   10                  15

ctg tgc gac gcc cag gac tgg gac gcc tac gtg cag ttg ttc gac ccg  2797
Leu Cys Asp Ala Gln Asp Trp Asp Ala Tyr Val Gln Leu Phe Asp Pro
    20                  25                  30

cag agt gaa ttc cac ctg ccg caa tgg gac tcc gaa cac gtc tac acc  2845
Gln Ser Glu Phe His Leu Pro Gln Trp Asl Ser Glu His Val Tyr Thr
35                  40                  45                  50

caa gac ccc aag cgc gag atg tca ttg atc tac tac gcc aac cgt tcg  2893
Gln Asp Pro Lys Arg Glu Met Ser Leu Ile Tyr Tyr Ala Asn Arg Ser
                55                  60                  65

ggc ctg gaa gac cgt gtg ttc cgc ctg cgc acc ggc aaa gcc gcc tct  2941
Gly Leu Glu Asp Arg Val Phe Arg Leu Arg Thr Glu Lys Ala Ala Ser
            70                  75                  80

gcc acg ccg atg ccg cgc act ttg cac ctg atc aat aac gta cgc att  2989
Ala Thr Pro Met Pro Arg Thr Leu His Leu Ile Asn Asn Val Arg Ile
        85                  90                  95

gcc gag cag gcc gat ggc acg ttg gag gtg cgt ttg aac tgg cac aca  3037
Ala Glu Gln Ala Asp Gly Thr Leu Glu Val Arg Leu Asn Trp His Thr
    100                 105                 110

ttg ttt tat cgc ctg gcc acg tcc gag cag ttt tac ggg cat gcc acg  3085
Leu Phe Tyr Arg Lau Ala Thr Ser Glu Gln Phe Tyr Gly His Ala Thr
115                 120                 125                 130

tac cgc ctc aag cct gcg ggc gac agc tgg ttg atc atg cgc aag cac  3133
Tyr Arg Leu Lys Pro Ala Gly Asp Ser Trp Leu Ile Met Arg Lys His
                135                 140                 145

gcc ttg ttg ctc aac gac acc atc aac tcg gtg ctg gat ttc tac cac  3181
Ala Leu Leu Leu Asn Asp Thr Ile Asn Ser Val Leu Asp Phe Tyr His
            150                 155                 160

ctg taacggtggt gcatcgccct gtaggagcga gcttgctcgc gaaaaacgta       3234
Leu

agtacgccgc gttcattcag gatgtcccgc gtcatcgttg acgtttttcg cgagcaaggg 3294

gttcatacct attcacggag ttatgtga atg aat cac aaa gtg gcc ttc agc    3346
                              Met Asn His Lys Val Ala Phe Ser
                              1               5

ttt gcc gat ggc aag acc ctg ttc ttc ccg gtg ggc gcc cat gaa atc  3394
Phe Ala Asp Gly Lys Thr Leu Phe Phe Pro Val Gly Ala His Glu Ile
    10                  15                  20

ctc ctg gac gcg gcc ctg cgc aac ggc atc aag atc ccg ctc gat tgc  3442
Leu Leu Asp Ala Ala Leu Arg Asn Gly Ile Lys Ile Pro Leu Asp Cys
25                  30                  35                  40

cgc gaa ggc gtg tgc ggc acc tgc cag ggg cgc tgt gag tcc ggc gag  3490
Arg Glu Gly Val Cys Gly Thr Cys Gln Gly Arg Cys Glu Ser Gly Glu
                45                  50                  55

tac acc cag gac tat gtc gat gag gaa gcc ctc tcc agc ctc gac ctg  3538
Tyr Thr Gln Asp Tyr Val Asp Glu Glu Ala Leu Ser Ser Leu Asp Leu
            60                  65                  70
```

```
caa caa cgc aag atg ctc agt tgc caa acc cgg gtg aag tcc gac gcc   3586
Gln Gln Arg Lys Met Leu Ser Cys Gln Thr Arg Val Lys Ser Asp Ala
        75                  80                  85

acg ttt tat ttc gac ttt gac tca agc ctg tgc aac gcc cca ggc ccc   3634
Thr Phe Tyr Phe Asp Phe Asp Ser Ser Leu Cys Asn Ala Pro Gly Pro
    90                  95                  100

gtg cag gtg cgc ggc act gtg agc gcg gtg cag cag gta tcg acc agc   3682
Val Gln Val Arg Gly Thr Val Ser Ala Val Gln Gln Val Ser Thr Ser
105                 110                 115                 120

acc gcc att ttg cag gtg caa ctg gac cag cct ctg gat ttt ttg ccg   3730
Thr Ala Ile Leu Gln Val Gln Leu Asp Gln Pro Leu Asp Phe Leu Pro
                125                 130                 135

ggc caa tac gcg cgt ctg tcg gtg ccc ggc acc gat agc tgg cgc tcc   3778
Gly Gln Tyr Ala Arg Leu Ser Val Pro Gly Thr Asp Ser Trp Arg Ser
            140                 145                 150

tac tcc ttc gcc aac cgg ccg ggt aat cag ttg cag ttc ctg gta cgc   3826
Tyr Ser Phe Ala Asn Arg Pro Gly Asn Gln Leu Gln Phe Leu Val Arg
        155                 160                 165

ctg ctg ccc gac gga gtc atg agc aac tac ctg cgt gaa cgc tgc cag   3874
Leu Leu Pro Asp Gly Val Met Ser Asn Tyr Leu Arg Glu Arg Cys Gln
    170                 175                 180

gtg ggt gat gaa atg ctg atg gag gcg ccc ttg ggt gcg ttt tat ctg   3922
Val Gly Asp Glu Met Leu Met Glu Ala Pro Leu Gly Ala Phe Tyr Leu
185                 190                 195                 200

cgg cac gtc acc caa ccg ctg gta ctg gtg gcg ggc ggc acc ggg ttg   3970
Arg His Val Thr Gln Pro Leu Val Leu Val Ala Gly Glu Thr Gly Leu
                205                 210                 215

tcg gcg ttg ttg ggc atg ctc gat gag ctg gtc gtc aac ggc tgc aca   4018
Ser Ala Leu Leu Gly Met Leu Asp Glu Leu Val Val Asn Glu Cys Thr
            220                 225                 230

caa cct gtg cac ctg tac tac ggc gtg cgc ggc gcc gaa gac tta tgt   4066
Gln Pro Val His Leu Tyr Tyr Gly Val Arg Gly Ala Glu Asp Leu Cys
        235                 240                 245

gaa gcg gca cgt atc cac gcc tac gcg acg aaa atc ccg aac ttt cgc   4114
Glu Ala Ala Arg Ile His Ala Tyr Ala Thr Lys Ile Pro Asn Phe Arg
    250                 255                 260

tac acc gaa gtg ctg agc gac gcc tca gtc gag tgg acg ggc aaa cgc   4162
Tyr Thr Glu Val Leu Ser Asp Ala Ser Val Glu Trp Thr Gly Lys Arg
265                 270                 275                 280

ggc tac ctg acc gaa cat ttt gac ctg gcc gaa ttg cgg gac aga tcg   4210
Gly Tyr Leu Thr Glu His Phe Asp Leu Ala Glu Leu Arg Asp Arg Ser
                285                 290                 295

gcg gat atg tac gtg tgc ggc ccc cct cca atg gtc gaa tcc atc caa   4258
Ala Asp Met Tyr Val Cys Gly Pro Pro Pro Met Val Glu Ser Ile Gln
            300                 305                 310

caa tgg ctg gcg gat cag aca ctt gat ggc gtt cag ttg tat tac gaa   4306
```

```
Gln Trp Leu Ala Asp Gln Thr Leu Asp Gly Val Gln Lys Tyr Tyr Glu
        315                 320                 325

aag ttt acc cag agt aat atc tga                                    4330
Lys Phe Thr Gln Ser Asn Ile
    330                 335

<210> 8
<211> 993
<212> DNA
<213> Pseudomonas fluorescens

<220>
<221> CDS
<222> 1..990
<223> Coding sequence for AntR

<220>
<221> variation
<222> 802..802
<223> Mutation to T802 from native G802; resulting in the codon mutation
GCG->TCG shown at 802..804, and the amino acid mutation of Ala268 to
Ser268

<220>
<221> misc_signal
<222> 991..993
<223> Stop codon for the CDS of AntR

<400> 8
atg act agt cag aca cgc gat att cat att caa cgc ttc gac ctg gaa     48
Met Thr Ser Gln Thr Arg Asp Ile His Ile Gln Arg Phe Asp Leu Glu
1               5                   10                  15

ggc gcg cgc agc tgg atg tcc ggc atc tgc ggg ccc cat cgc ctg gcg     96
Gly Ala Arg Ser Trp Met Ser Gly Ile Cys Gly Pro His Arg Leu Ala
            20                  25                  30

acg gca acc ccc gag cgc ctg cgc ttt cac cac agc gcc aac gtg ttc    144
Thr Ala Thr Pro Glu Arg Leu Arg Phe His His Ser Ala Asn Val Phe
        35                  40                  45

aaa tcc cgc gcc acc acc ctg ggc gtg atc gag tac ggc act gat gtg    192
Lys Ser Arg Ala Thr Thr Leu Gly Val Ile Glu Tyr Gly Thr Asp Val
    50                  55                  60

acc atc gac atc gaa gac gcc gag cac ttc agc agc tac agc ttg atc    240
Thr Ile Asp Ile Glu Asp Ala Glu His Phe Ser Ser Tyr Ser Leu Ser
65                  70                  75                  80

ctg cca ctg gtg ggc gag cag gag ctg agc aag aac ggt gaa cgt ctc    288
Leu Pro Leu Val Gly Glu Gln Glu Leu Ser Lys Asn Gly Glu Arg Leu
                85                  90                  95

agt tcc aac cgc gac caa ggc gtg atc att tcg ccc aat gag cat cag    336
Ser Ser Asn Arg Asp Gln Gly Val Ile Ile Ser Pro Asn Glu His Gln
            100                 105                 110

gtg ctg gcg att tcc ggt gac tgc cgc aag ttg cag gtg gta atc acc    384
Val Leu Ala Ile Ser Gly Asp Cys Arg Lys Leu Gln Val Val Ile Thr
        115                 120                 125
```

```
tgc gcg gcg atg agc gag tcg ctg gaa ggt ttg ctg caa cgg ccc att    432
Arg Ala Ala Met Ser Glu Ser Leu Glu Gly Leu Leu Gln Arg Pro Ile
    130                 135                 140

gat gcg ccg ctg cgc ttt gag tcg gtg atg gat gct gtg gac gga gca    480
Asp Ala Pro Leu Arg Phe Glu Ser Val Met Asp Ala Val Asp Gly Ala
145                 150                 155                 160

ccg gct tcg tgg tgg cgc atg gcg cgt tat ttc atc gcc gaa ctg gag    527
Pro Ala Ser Trp Trp Arg Met Ala Arg Tyr Phe Ile Ala Glu Leu Glu
                165                 170                 175

cgc agc agc gaa ctg tac gag cag gcg gcg ttt act cgc gac ctg gaa    576
Arg Ser Ser Glu Leu Tyr Glu Gln Ala Ala Phe Thr Arg Asp Leu Glu
            180                 185                 190

agc tcg ctg atc aag ggc ctg atc ctg gcc caa ccg aat aac tac tcc    624
Ser Ser Leu Ile Lys Gly Leu Ile Leu Ala Gln Pro Asn Asn Tyr Ser
        195                 200                 205

gaa gaa ctg cgc gac gtg ctg ggg gtg aaa ctg ccg cac tac ctg atc    672
Glu Glu Leu Arg Asp Val Leu Gly Val Lys Leu Pro His Tyr Leu Ile
    210                 215                 220

cgc gcg cgg caa ttt atc cac gac aac gcc cgt gaa gcg ctg cac ctg    720
Arg Ala Arg Gln Phe Ile His Asp Asn Ala Arg Glu Ala Leu His Leu
225                 230                 235                 240

gaa gac ctg gaa gcg gcg gcg ggc gtg tcg cgt ttc aaa ctg ttc gat    768
Glu Asp Leu Glu Ala Ala Ala Gly Val Ser Arg Phe Lys Leu Phe Asp
                245                 250                 255

gcg ttt cgc aag tac ttc gca ctg tcg ccc atg tcg tat ttg aag aag    816
Ala Phe Arg Lys Tyr Phe Ala Leu Ser Pro Met Ser Tyr Leu Lys Lys
            260                 265                 270

cat cgg ttg ggg gcg gtg cgc caa gaa att ctg gag cag ggc tcg atc    864
His Arg Leu Gly Ala Val Arg Gln Glu Ile Leu Glu Gln Gly Ser Ile
        275                 280                 285

cgc acc tac tct gaa atc gcc ctg ggt tgg ggg ttt acg cat ttg ggg    912
Arg Thr Ile Ser Glu Ile Ala Leu Gly Trp Gly Phe Thr His Leu Gly
    290                 295                 300

cgg ttt tcg gcg gag tac cgc aag ctg ttc gat gag tcg ccc agc cag    960
Arg Phe Ser Ala Glu Tyr Arg Lys Leu Phe Asp Glu Ser Pro Ser Gln
305                 310                 315                 320

aca ttg cag cgc aag cgc ctg cgc att act tga                        993
Thr Leu Gln Arg Lys Arg Leu Arg Ile Thr
                325                 330
```

<210> 9
<211> 330
<212> PRT
<213> Pseudomonas fluorescens

<220>
<221> INIT_MET
<222> 1..1
<223> Putative initiator methionine for AntR, giving Met1->Thr330 as the amino acid sequence of the putative Pant activator protein

<220>
<221> VARIANT
<222> 268..268
<223> Mutation to Ser268 from Ala268

<400> 9

```
Met Thr Ser Gln Thr Arg Asp Ile His Ile Gln Arg Phe Asp Leu Glu
1               5                  10                  15

Gly Ala Arg Ser Trp Met Ser Gly Ile Cys Gly Pro His Arg Leu Ala
            20                  25                  30

Thr Ala Thr Pro Glu Arg Leu Arg Phe His His Ser Ala Asn Val Phe
        35                  40                  45

Lys Ser Arg Ala Thr Thr Leu Gly Val Ile Glu Tyr Gly Thr Asp Val
    50                  55                  60

Thr Ile Asp Ile Glu Asp Ala Glu His Phe Ser Ser Tyr Ser Leu Ser
65                  70                  75                  80

Leu Pro Leu Val Gly Glu Gln Glu Leu Ser Lys Asn Gly Glu Arg Leu
                85                  90                  95

Ser Ser Asn Arg Asp Gln Gly Val Ile Ile Ser Pro Asn Glu His Gln
            100                 105                 110

Val Leu Ala Ile Ser Gly Asp Cys Arg Lys Leu Gln Val Val Ile Thr
        115                 120                 125

Arg Ala Ala Met Ser Glu Ser Leu Glu Gly Leu Leu Gln Arg Pro Ile
    130                 135                 140

Asp Ala Pro Leu Arg Phe Glu Ser Val Met Asp Ala Val Asp Gly Ala
145                 150                 155                 160

Pro Ala Ser Trp Trp Arg Met Ala Arg Tyr Phe Ile Ala Glu Leu Glu
                165                 170                 175

Arg Ser Ser Glu Leu Tyr Glu Gln Ala Ala Phe Thr Arg Asp Leu Glu
            180                 185                 190

Ser Ser Leu Ile Lys Gly Leu Ile Leu Ala Gln Pro Asn Asn Tyr Ser
        195                 200                 205

Glu Glu Leu Arg Asp Val Leu Gly Val Lys Leu Pro His Tyr Leu Ile
    210                 215                 220

Arg Ala Arg Gln Phe Ile His Asp Asn Ala Arg Glu Ala Leu His Leu
225                 230                 235                 240

Glu Asp Leu Glu Ala Ala Ala Gly Val Ser Arg Phe Lys Leu Phe Asp
                245                 250                 255

Ala Phe Arg Lys Tyr Phe Ala Leu Ser Pro Met Ser Tyr Leu Lys Lys
            260                 265                 270

His Arg Leu Gly Ala Val Arg Gln Glu Ile Leu Glu Gln Gly Ser Ile
        275                 280                 285

Arg Thr Ile Ser Glu Ile Ala Leu Gly Trp Gly Phe Thr His Leu Gly
```

```
    290                 295                 300

Arg Phe Ser Ala Glu Tyr Arg Lys Leu Phe Asp Glu Ser Pro Ser Gln
305                 310                 315                 320

Thr Leu Gln Arg Lys Arg Leu Arg Ile Thr
                325                 330

<210> 10
<211> 463
<212> PRT
<213> Pseudomonas fluorescens

<220>
<221> misc_feature
<222> 1..463
<223> AntA expression product, anthranilate dioxygenase large subunit

<400> 10
Met Ser Gly Ala Arg Thr Val Glu Gln Trp Lys Ser Phe Ile Glu Ser
1               5                   10                  15

Cys Leu Asp Phe Arg Pro Ala Asp Glu Val Phe Arg Ile Ala Arg Asp
            20                  25                  30

Met Phe Thr Glu Pro Glu Leu Phe Asp Leu Glu Met Glu Leu Ile Phe
        35                  40                  45

Glu Lys Asn Trp Ile Tyr Ala Cys His Glu Ser Glu Leu Ala Asn Asn
    50                  55                  60

His Asp Phe Val Thr Met Arg Ala Gly Arg Gln Pro Met Ile Ile Thr
65                  70                  75                  80

Arg Asp Gly Glu Gly Arg Leu Asn Ala Leu Ile Asn Ala Cys Gln His
                85                  90                  95

Arg Gly Thr Thr Leu Thr Arg Val Gly Lys Gly Asn Gln Ser Thr Phe
            100                 105                 110

Thr Cys Pro Phe His Ala Trp Cys Tyr Lys Ser Asp Gly Arg Leu Val
        115                 120                 125

Lys Val Lys Ala Pro Gly Glu Tyr Pro Glu Gly Phe Asp Lys Ala Thr
    130                 135                 140

Arg Gly Leu Lys Lys Ala Arg Ile Glu Ser Tyr Arg Gly Phe Val Phe
145                 150                 155                 160

Ile Ser Leu Asp Val Asn Gly Thr Asn Ser Leu Glu Asp Phe Leu Gly
                165                 170                 175

Asp Ala Lys Val Phe Phe Asp Met Met Val Ala Gln Ser Ala Thr Gly
            180                 185                 190

Glu Leu Glu Val Leu Pro Gly Lys Ser Ala Tyr Thr Tyr Asp Gly Asn
        195                 200                 205

Trp Lys Leu Gln Asn Glu Asn Gly Leu Asp Gly Tyr His Val Ser Thr
    210                 215                 220

Val His Tyr Asn Tyr Val Ala Thr Val Gln His Arg Glu Gln Val Asn
```

225 230 235 240

Thr Glu Asn Gly Ala Gly Ser Ser Thr Thr Leu Asp Tyr Ser Lys Leu
245 250 255

Gly Ala Gly Asp Ala Asn Thr Asp Asp Gly Trp Phe Ala Phe Asn Asn
260 265 270

Gly His Ser Val Leu Phe Ser Asp Met Pro Asn Pro Ser Val Arg Ser
275 280 285

Gly Tyr Ala Thr Ile Met Pro Arg Leu Val Glu Glu His Gly Gln Gln
290 295 300

Lys Ala Glu Trp Met Met His Arg Leu Arg Asn Leu Asn Ile Tyr Pro
305 310 315 320

Ser Leu Phe Phe Leu Asp Gln Ile Ser Ser Gln Leu Arg Ile Ile Arg
325 330 335

Pro Val Ala Trp Asn Lys Thr Glu Ile Ile Ser Gln Cys Leu Gly Val
340 345 350

Lys Gly Glu Ser Asp Ala Asp Arg Glu Asn Arg Ile Arg Gln Phe Glu
355 360 365

Asp Phe Phe Asn Val Ser Gly Met Gly Thr Pro Asp Asp Leu Val Glu
370 375 380

Phe Arg Glu Ala Gln Arg Gly Phe Gln Gly Arg Leu Glu Arg Trp Ser
385 390 395 400

Asp Ile Ser Arg Gly Ser His Arg Trp Glu Thr Gly Pro Trp Pro Asn
405 410 415

Ser Glu Ala Ile Gly Ile Gln Pro Ala Met Thr Gly Thr Glu Phe Thr
420 425 430

His Glu Gly Leu Tyr Val Asn Gln His Arg Asn Trp Gln Gln Phe Leu
435 440 445

Leu Lys Gly Leu Asp Gln Arg Ala Leu Ala Leu Arg Glu Val Lys
450 455 460

<210> 11
<211> 163
<212> PRT
<213> Pseudomonas fluorescens

<220>
<221> misc_feature
<222> 1..163
<223> AntB expression product, anthranilate dioxygenase small subunit

<400> 11
Met Asn Ala Gln Leu Gln Tyr Gln Ile Glu Gln Phe Phe Tyr Arg Lys
1 5 10 15

Ser Glu Leu Cys Asp Ala Gln Asp Trp Asp Ala Tyr Val Gln Leu Phe
20 25 30

Asp Pro Gln Ser Glu Phe His Leu Pro Gln Trp Asl Ser Glu His Val

35 40 45

Tyr Thr Gln Asp Pro Lys Arg Glu Met Ser Leu Ile Tyr Tyr Ala Asn
50 55 60

Arg Ser Gly Leu Glu Asp Arg Val Phe Arg Leu Arg Thr Glu Lys Ala
65 70 75 80

Ala Ser Ala Thr Pro Met Pro Arg Thr Leu His Leu Ile Asn Asn Val
85 90 95

Arg Ile Ala Glu Gln Ala Asp Gly Thr Leu Glu Val Arg Leu Asn Trp
100 105 110

His Thr Leu Phe Tyr Arg Lau Ala Thr Ser Glu Gln Phe Tyr Gly His
115 120 125

Ala Thr Tyr Arg Leu Lys Pro Ala Gly Asp Ser Trp Leu Ile Met Arg
130 135 140

Lys His Ala Leu Leu Leu Asn Asp Thr Ile Asn Ser Val Leu Asp Phe
145 150 155 160

Tyr His Leu

<210> 12
<211> 335
<212> PRT
<213> Pseudomonas fluorescens

<220>
<221> misc_feature
<222> 1..335
<223> AntC expression product, anthranilate dioxygenase reductase

<400> 12
Met Asn His Lys Val Ala Phe Ser Phe Ala Asp Gly Lys Thr Leu Phe
1 5 10 15

Phe Pro Val Gly Ala His Glu Ile Leu Leu Asp Ala Ala Leu Arg Asn
20 25 30

Gly Ile Lys Ile Pro Leu Asp Cys Arg Glu Gly Val Cys Gly Thr Cys
35 40 45

Gln Gly Arg Cys Glu Ser Gly Glu Tyr Thr Gln Asp Tyr Val Asp Glu
50 55 60

Glu Ala Leu Ser Ser Leu Asp Leu Gln Gln Arg Lys Met Leu Ser Cys
65 70 75 80

Gln Thr Arg Val Lys Ser Asp Ala Thr Phe Tyr Phe Asp Phe Asp Ser
85 90 95

Ser Leu Cys Asn Ala Pro Gly Pro Val Gln Val Arg Gly Thr Val Ser
100 105 110

Ala Val Gln Gln Val Ser Thr Ser Thr Ala Ile Leu Gln Val Gln Leu
115 120 125

Asp Gln Pro Leu Asp Phe Leu Pro Gly Gln Tyr Ala Arg Leu Ser Val
130 135 140

```
Pro Gly Thr Asp Ser Trp Arg Ser Tyr Ser Phe Ala Asn Arg Pro Gly
145                 150                 155                 160

Asn Gln Leu Gln Phe Leu Val Arg Leu Leu Pro Asp Gly Val Met Ser
                165                 170                 175

Asn Tyr Leu Arg Glu Arg Cys Gln Val Gly Asp Glu Met Leu Met Glu
            180                 185                 190

Ala Pro Leu Gly Ala Phe Tyr Leu Arg His Val Thr Gln Pro Leu Val
        195                 200                 205

Leu Val Ala Gly Glu Thr Gly Leu Ser Ala Leu Leu Gly Met Leu Asp
    210                 215                 220

Glu Leu Val Val Asn Glu Cys Thr Gln Pro Val His Leu Tyr Tyr Gly
225                 230                 235                 240

Val Arg Gly Ala Glu Asp Leu Cys Glu Ala Ala Arg Ile His Ala Tyr
                245                 250                 255

Ala Thr Lys Ile Pro Asn Phe Arg Tyr Thr Glu Val Leu Ser Asp Ala
            260                 265                 270

Ser Val Glu Trp Thr Gly Lys Arg Gly Tyr Leu Thr Glu His Phe Asp
        275                 280                 285

Leu Ala Glu Leu Arg Asp Arg Ser Ala Asp Met Tyr Val Cys Gly Pro
    290                 295                 300

Pro Pro Met Val Glu Ser Ile Gln Gln Trp Leu Ala Asp Gln Thr Leu
305                 310                 315                 320

Asp Gly Val Gln Lys Tyr Tyr Glu Lys Phe Thr Gln Ser Asn Ile
                325                 330                 335

<210> 13
<211> 1544
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..1544
<223> Construct containing the pDOW1057 Pant-Pben tandem promoter

<220>
<221> misc_feature
<222> 95..1395
<223> Portion containing Pant with activator CDS

<220>
<221> misc_signal
<222> 92..94
<223> Anticodon of stop codon for the CDS of AntR

<220>
<221> CDS
<222> 95..1084
<223> Antisense strand of ORF encoding AntR
```

<220>
<221> variation
<222> 283..283
<223> Mutation to A235 from native C235; resulting in anticodon mutation CGC->CGA, thus codon mutation GCG->TCG and amino acid mutation of Ala268 to Ser268

<220>
<221> misc_signal
<222> 1221..1327
<223> Approximate region estimated to contain the AntR binding site

<220>
<221> misc_feature
<222> 1329..1509
<223> Pant-Pben tandem promoter construct

<220>
<221> promoter
<222> 1329..1365
<223> Putative promoter (Pant) from anthranilate operon (antABC)

<220>
<221> -35_signal
<222> 1329..1333
<223> Putative -35 region of Pant promoter

<220>
<221> -10_signal
<222> 1355..1359
<223> Putative -10 region of Pant promoter

<220>
<221> misc_signal
<222> 1371..1371
<223> Putative native transcription start site under control of Pant

<220>
<221> misc_feature
<222> 1396..1429
<223> linker

<220>
<221> misc_feature
<222> 1430..1541
<223> Portion containing Pben

<220>
<221> misc_signal
<222> 1430-1476
<223> Approximate region estimated to contain the BenR binding site

<220>
<221> promoter
<222> 1477-1509
<223> Putative promoter (Pben) from benzoate operon (benABCD)

<220>
<221> -35_signal
<222> 1477-1482
<223> Putative -35 region of Pben promoter

<220>
<221> -10_signal
<222> 1498-1503
<223> Putative -10 region of Pben promoter

<220>
<221> misc_feature
<222> 1498-1503
<223> Substitution mutation of Pben -10 region by tataat to form -10con mutants, and by taaggt to form -10benAc mutants

<220>
<221> misc_feature
<222> 1508..1508
<223> A deletion of g1508 found in mutant promoter variants herein

<220>
<221> misc_signal
<222> 1509..1509
<223> Putative transcription start site

<220>
<221> misc_signal
<222> 1542..1544
<223> Putative native translation initiator codon attached to Pben

<400> 13
agcttgcatg cctgcaggtt taaacagtcg actctagact taattaacta tcgcaggcaa 60

gccagctccc acagattgtt tttcatccag ttca agt aat gcg cag gcg ctt gcg 115
Thr Ile Arg Leu Arg Lys Arg
-330 -325

ctg caa tgt ctg gct ggg cga ctc atc gaa cag ctt gcg gta ctc cgc 163
Gln Leu Thr Gln Ser Pro Ser Glu Asp Phe Leu Lys Arg Tyr Glu Ala
-320 -315 -310

cga aaa ccg ccc caa atg cgt aaa ccc cca acc cag ggc gat ttc aga 211
Ser Phe Arg Gly Leu His Thr Phe Gly Trp Gly Leu Ala Ile Glu Ser
-305 -300 -295

gat ggt gcg gat cga gcc ctg ctc cag aat ttc ttg gcg cac cgc ccc 259
Ile Thr Arg Ile Ser Gly Gln Glu Leu Ile Glu Gln Arg Val Ala Gly
-290 -285 -280

caa ccg atg ctt ctt caa ata cga cat ggg cga cag tgc gaa gta ctt 307
Leu Arg His Lys Lys Leu Tyr Ser Met Pro Ser Leu Ala Phe Tyr Lys
-275 -270 -265 -260

gcg aaa cgc atc gaa cag ttt gaa acg cga cac gcc cgc cgc cgc ttc 355
Arg Phe Ala Asp Phe Leu Lys Phe Arg Ser Val Gly Ala Ala Ala Glu
-255 -250 -245

cag gtc ttc cag gtg cag cgc ttc acg ggc gtt gtc gtg gat aaa ttg 403
Leu Asp Glu Leu His Leu Ala Glu Arg Ala Asn Asp His Ile Phe Gln
-240 -235 -230

ccg cgc gcg gat cag gta gtg cgg cag ttt cac ccc cag cac gtc gcg 451
Arg Ala Arg Ile Leu Tyr His Pro Leu Lys Val Gly Leu Val Asp Arg
-225 -220 -215

cag ttc ttc gga gta gtt att cgg ttg ggc cag gat cag gcc ctt gat 499

```
Leu Glu Glu Ser Tyr Asn Asn Pro Gln Ala Leu Ile Leu Gly Lys Ile
    -210                -205                -200

cag cga gct ttc cag gtc gcg agt aaa cgc cgc ctg ctc gta cag ttc    547
Leu Ser Ser Glu Leu Asp Arg Thr Phe Ala Ala Gln Glu Tyr Leu Glu
-195                -190                -185                -180

gct gct gcg ctc cag ttc ggc gat gaa ata acg cgc cat gcg cca cca    595
Ser Ser Arg Glu Leu Glu Ala Ile Phe Tyr Arg Ala Met Arg Trp Trp
                -175                -170                -165

cga agc cgg tgc tcc gtc cac agc atc cat cac cga ctc aaa gcg cag    643
Ser Ala Pro Ala Gly Asp Val Ala Asp Met Val Ser Glu Phe Arg Leu
            -160                -155                -150

cgg cgc atc aat ggg ccg ttg cag caa acc ttc cag cga ctc gct cat    691
Pro Ala Asp Ile Pro Arg Gln Leu Leu Gly Glu Leu Ser Glu Ser Met
        -145                -140                -135

cgc cgc acg ggt gat tac cac ctg caa ctt gcg gca gtc acc gga aat    739
Ala Ala Arg Thr Ile Val Val Gln Leu Lys Arg Cys Asp Gly Ser Ile
    -130                -125                -120

cgc cag cac ctg atg ctc att ggg cga aat gat cac gcc ttg gtc gcg    787
Ala Leu Val Gln His Glu Asn Pro Ser Ile Ile Val Gly Gln Asp Arg
-115                -110                -105                -100

gtt gga act gag acg ttc acc gtt ctt gct cag ctc ctg ctc gcc cac    835
Asn Ser Ser Leu Arg Glu Gly Asn Lys Ser Leu Glu Gln Glu Gly Val
                -95                 -90                 -85

cag tgg cag gct caa gct gta gct gct gaa gtg ctc ggc gtc ttc gat    883
Leu Pro Leu Ser Leu Ser Tyr Ser Ser Phe His Glu Ala Asp Glu Ile
            -80                 -75                 -70

gtc gat ggt cac atc agt gcc gta ctc gat cac gcc cag ggt ggt ggc    931
Asp Ile Thr Val Asp Thr Gly Tyr Glu Ile Val Gly Leu Thr Thr Ala
        -65                 -60                 -55

gcg gga ttt gaa cac gtt ggc gct gtg gtg aaa gcg cag gcg ctc ggg    979
Arg Ser Lys Phe Val Asn Ala Ser His His Phe Arg Leu Arg Glu Pro
    -50                 -45                 -40

ggt tgc cgt cgc cag gcg atg ggg ccc gca gat gcc gga cat cca gct   1027
Thr Ala Thr Ala Leu Arg His Pro Gly Cys Ile Gly Ser Met Trp Ser
-35                 -30                 -25                 -20

gcg cgc gcc ttc cag gtc gaa gcg ttg aat atg aat atc gcg tgt ctg   1075
Arg Ala Gly Glu Leu Asp Phe Arg Gln Ile His Ile Asp Arg Thr Gln
                -15                 -10                 -5

act agt cat cagggtgcac ccacggcggt taggcgtttg cgcgctctga          1124
Ser Thr Met
        -1

cggcgcgtcg ttgaacctcg acagcaagtt ccaggccacg ccagtgcagt tctcactggg 1184

tggatagcaa cggtcgacta tgtggataaa ccccagagtt ttgcgaccat cgcccgccat 1244

cacagtagcg catgccgtca ccggcgcgca ccgtcatggg tatttgccgc ccaactttgc 1304

ggcctacgtt cccccattaa gcggatagcc cgccaccgca tcgcagccgc ttaatggctc 1364
```

accgtttagc catgatcaaa aggtgcctcc cggatcccca aacagtcgac tctagactta 1424

attaagttaa gcgacgtgcg cctggcggat agcgatgtgc aggcagcgga tattgacggg 1484

cagggcgagc acgtacggtg agggcgcctg atacaagaac aacggagggc ccgcccc 1541

atg 1544
Met
1

<210> 14
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..20
<223> Alternative Inter-promoter linker for tandem promoter

<400> 14
ggatccggcg cgcccccatc 20

<210> 15
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..18
<223> Linker used to connect promoters to reporter genes

<400> 15
actagtagga ggtaactt 18

<210> 16
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..21
<223> Knock-out primer AntAKO5

<220>
<221> misc_feature
<222> 2..7
<223> EcoRI recognition site

<400> 16
ggaattcttc gtgacgatgc g 21

<210> 17
<211> 22
<212> DNA
<213> Artificial Sequence

<220>

<221> misc_feature
<222> 1..22
<223> Knock-out primer AntAKO3

<220>
<221> misc_feature
<222> 3..8
<223> BamHI recognition site

<400> 17
cgggatccgc tcgcgatgct gc 22

<210> 18
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..16
<223> Labeling primer lacZPE

<400> 18
ggatgtgctg caaggc 16

<210> 19
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..17
<223> Labeling primer lacZPE2

<400> 19
gtaaccatgg tcatcgc 17

<210> 20
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..17
<223> M13forward primer

<400> 20
gtaaaacgac ggccagt 17

<210> 21
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..16
<223> M13reverse primer

<400> 21
aacagctatg accatg 16

<210> 22
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..35
<223> Oligonucleotide Bambenwtshort

<400> 22
cgggatccgt atcaggcgcc tcaccgtacg tgctc 35

<210> 23
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..35
<223> Oligonucleotide Bambenconshort

<400> 23
cgggatccgt atcaggcgcc tcattatacg tgctc 35

<210> 24
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..35
<223> Oligonucleotide BambenAcshort

<400> 24
cgggatccgt atcaggcgcc tcaccttacg tgctc 35

<210> 25
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..35
<223> Oligonucleotide Bamantwtshort

<400> 25
cgggatccgc taacggtgag ccattaagcg gctgc 35

<210> 26
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature

<222> 1..35
<223> Oligonucleotide Bamantconshort

<400> 26
cgggatccgc taacggtgag cattatagcg gctgc 35

<210> 27
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..20
<223> Primer BenactKO-for

<400> 27
cgcgacacat tgctgcccag 20

<210> 28
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..20
<223> Primer BenactKO-rev

<400> 28
agtatcagcc atcgcacctt 20

<210> 29
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..19
<223> Primer 1803H3seq

<400> 29
gtcctgcaat ttcagccga 19

<210> 30
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..26
<223> Oligonucleotide BenL278

<400> 30
ccttaattaa gttaagcgac gtgcgc 26

<210> 31
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..26
<223> Oligonucleotide 3'Antactiv

<400> 31
cccaagcttc tatcgaggca agccag 26

<210> 32
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..32
<223> Oligonucleotide Benact5'

<400> 32
agctttgttt aaacgcatga cgttgttgat tc 32

<210> 33
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..31
<223> Oligonucleotide H3_5'BenAKOclean

<400> 33
cccaagcttg ccatgaggcg gaaaacgctg c 31

<210> 34
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..30
<223> Oligonucleotide H3_3'BenBKOclean

<400> 34
cccaagcttc ggtgatcgcc acgctgtcgc 30

<210> 35
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..25
<223> Oligonucleotide BenKOmega

<400> 35
catacgtcat ggccctccgt tgttc 25

<210> 36

<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..25
<223> Oligonucleotide InvbenKOmega

<400> 36
gaacaacgga gggccatgac gtatg 25

<210> 37
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..24
<223> Oligonucleotide 5'BenA_seq

<400> 37
ctgctggaaa acgcctgcct ggag 24

<210> 38
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..25
<223> Oligonucleotide Seq_3'BenB

<400> 38
gagcacttca agcatcgaca ggaac 25

<210> 39
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..21
<223> Primer 1261-8378F

<400> 39
cttcagatcc agactcacca g 21

<210> 40
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..21
<223> Primer 1261-103R

<400> 40

gaccatgatt acgccaagcg c 21

<210> 41
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 1..21
<223> Primer M13R21

<400> 41
cacacaggaa acagctatga c 21