



Office de la Propriété
Intellectuelle
du Canada

Un organisme
d'Industrie Canada

Canadian
Intellectual Property
Office

An agency of
Industry Canada

CA 2176236 C 2001/10/16

(11)(21) **2 176 236**

(12) **BREVET CANADIEN
CANADIAN PATENT**

(13) **C**

(86) Date de dépôt PCT/PCT Filing Date: 1994/11/08
(87) Date publication PCT/PCT Publication Date: 1995/05/18
(45) Date de délivrance/Issue Date: 2001/10/16
(85) Entrée phase nationale/National Entry: 1996/05/09
(86) N° demande PCT/PCT Application No.: US 94/12857
(87) N° publication PCT/PCT Publication No.: WO 95/13386
(30) Priorité/Priority: 1993/11/12 (152,483) US

(51) Cl.Int.⁶/Int.Cl.⁶ C12N 15/62, C12N 15/53, C12P 17/16,
C12N 15/12, C12N 9/02
(72) Inventeurs/Inventors:
Kumagai, Monto H., US;
Della-Cioppa, Guy, US
(73) Propriétaire/Owner:
BIOSOURCE TECHNOLOGIES INCORPORATED, US
(74) Agent: CARTON, JOHN K.

(54) Titre : ENZYME DE FUSION DE LA PROTEINE ACTIVATRICE DE LA TYROSINASE
(54) Title: TYROSINASE-ACTIVATOR PROTEIN FUSION ENZYME

(57) **Abrégé/Abstract:**

The present invention relates to a nucleic acid sequence encoding a fusion enzyme comprising a nucleic acid sequence encoding for a tyrosinase and a nucleic acid sequence encoding for a tyrosinase activator protein. Activator proteins used in the present invention are ORF438 and URF402. Prokaryotic tyrosinases that require a separate activator protein, particularly tyrosinases derived from *Streptomyces* are preferred. It is also preferred that the activator protein sequence be positioned (5') relative to the tyrosinase sequence. The present invention also relates to a fusion enzyme which comprises an amino acid sequence for tyrosinase and an amino acid sequence for a tyrosinase activator protein, which may also contain a linker positioned between the amino acid sequences of the activator protein and the tyrosinase. The present invention also relates to melanin produced by a fusion enzyme of the present invention.



ABSTRACT

The present invention relates to a nucleic acid sequence encoding a fusion enzyme comprising a nucleic acid sequence encoding for a tyrosinase and a nucleic acid sequence encoding for a tyrosinase activator protein. Activator proteins used in the present invention are ORF438 and URF402. Prokaryotic tyrosinases that require a separate activator protein, particularly tyrosinases derived from *Streptomyces* are preferred. It is also preferred that the activator protein sequence be positioned (5') relative to the tyrosinase sequence. The present invention also relates to a fusion enzyme which comprises an amino acid sequence for tyrosinase and an amino acid sequence for a tyrosinase activator protein, which may also contain a linker positioned between the amino acid sequences of the activator protein and the tyrosinase. The present invention also relates to melanin produced by a fusion enzyme of the present invention.

- 1 -

TITLE OF THE INVENTION

TYROSINASE-ACTIVATOR PROTEIN FUSION ENZYME

5

BACKGROUND OF INVENTION

10 Melanogenesis, production of the biological
polymer melanin, is a widespread phenomena in nature
occurring in most phyla from fungi to mammals. The
black, brown, buff and Tyndall-blue pigments found in
feathers, hairs, eyes, insect cuticle, fruit and seeds
15 are usually melanins. Melanins have been assigned a
photoprotective role in the skin, their role in the eye
and inner ear is unknown. Melanins

-2-

are also found in the mammalian brain where they are referred to as neuromelanins. The biological function of neuromelanins is unknown.

5 The stepwise biosynthesis of melanins is depicted in Figure 1. Tyrosinase (E.C. 1.14.18.1) catalyzes two types of reactions involved in the biosynthesis of melanin: the orthohydroxylation of
10 monophenols to catechols, which is referred to as cresolase activity, and the dehydrogenation of catechols to o-quinones, designated as catecholase activity. Molecular oxygen is used for the hydroxylation reaction. For this reason, tyrosinase acting on a monophenol is referred to as a "mixed
15 function oxidase". Hayaishi, in "Biological Oxidation" (Singer, ed.) p.581, Interscience Publishers, New York (1968). As used herein tyrosinase refers to all enzymes possessing the above described enzymatic activity. Tyrosinases are also
20 sometimes referred to as polyphenol oxidases.

As can be seen from the biosynthetic pathway of melanin depicted in Figure 1, tyrosinase is essential to the production of melanin. Therefore, the ability of an organism to express significant quantities of
25 tyrosinase activity is essential to the organism's in vivo production of melanin.

Tyrosinase is not naturally found in all organisms. Rather, tyrosinase has been found to occur in a relatively limited number of prokaryotes, is absent in a variety of higher plants and is
30 generally confined to specific cells of the skin in higher animals but may also occur in interior tissue, such as the substantia nigra, eye and inner ear. Given the limited number of organisms that produce melanin, one objective of the present invention is to
35 provide a means for introducing tyrosinase activity

- 3 -

into organisms that are otherwise unable to produce melanin.

5 Even in those organisms where tyrosinase activity occurs naturally, such activity is generally present at low levels. As a result, melanin is generally produced in small quantities by those organisms that possess tyrosinase activity. It is therefore a further
10 objective of the present invention to provide a means for enhancing the level of tyrosinase activity in organisms in order to enhance the in vivo production of melanin.

Several prior art references teach how to genetically engineer an organism to possess tyrosinase
15 activity. In PCT application WO 92/00373, published January 9, 1992, Applicants teach a method for producing melanin from transformed microorganisms wherein a sequence encoding for tyrosinase is introduced into the organism. Microorganisms that have
20 be genetically engineered to enhance their abilities to produce tyrosinases include, but are not limited to species of Streptomyces, Escherichia, Bacillus, Streptococcus, Salmonella, Staphylococcus, and Vibrio. For example, cloned tyrosinase genes from Streptomyces
25 sp. have been shown to produce melanin pigments in culture. J. Gen. Microbiol. 129:2703-2714 (1983); Gene 37:101-110 (1985). The cloned genes have also been expressed in Streptomyces and E. coli. della-Cioppa, Bio/Technology 8:634-638 (1990). In each case, both
30 tyrosinase and ORF438 were required for melanin production.

U.S. Patent No. 4,898,814 issued to Kwon discloses a cDNA clone of human tyrosinase and claims

-4-

a method of making human tyrosinase by expressing the cDNA in E. coli.

Many forms of tyrosinase from bacteria, such as Streptomyces, require an activator protein. For example, the mel locus of S. antibioticus has been shown to contain two open reading frames (ORF's) that encode a putative ORF438 protein ($M_r=14,754$) and tyrosinase ($M_r=30,612$). ORF438 and tyrosinase are thought to be transcribed from the same promoter in S. antibioticus. Bernan, et al., Gene 37:101 (1985). Both genes are required for melanin production. Bernan, et al., Gene 37:101 (1985). Based on genetic evidence, ORF438 protein has been shown to function as a trans-activator of tyrosinase. Lee, et al., Gene 65:71 (1988). It has been suggested that the ORF438 protein is involved in tyrosinase secretion, or it may function as a metallothionein-like protein that delivers copper to tyrosinase, Bernan, et al., Gene 37:101 (1985); Lee, et al., Gene 65:71 (1988). The mel locus of S. glaucescens has a nearly identical ORF sequence upstream of tyrosinase that probably serves a similar function. Huber, et al. Biochemistry 24:6038 (1985); Huber, et al., Nucleic Acids Res. 15:8106 (1987). The existence of an ORF438 protein, however, has never been confirmed in vivo.

The melanin operons of S. antibioticus and S. glaucescens have been isolated and sequenced, and both share sequence homology and similar gene arrangement. The polypeptide sequence encoded by ORF438 (146 amino acids) in S. antibioticus is structurally and functionally equivalent to URF402 (134 amino acids) from S. glaucescens Huber, et al. Nucleic Acids Res. 15:8106 (1987). Disruption of the URF402 coding sequence abolishes the melanin

- 5 -

phenotype similar to that already known for ORF438. Recent evidence suggest that the ORF438 protein (and by analogy the URF402 protein equivalent) functions as a molecular chaperone for tyrosinase. Chen, et al., J. Biol Chem. 268:18710 (1993).

Tyrosinase has also been isolated and employed to produce melanin in vitro. For example, in U.S. Patent No. 5,340,734, issued August 23, 1994, Applicants teach the production of melanin in vitro using a tyrosinase which is excreted from the microorganism during fermentation. In vitro production of melanin is dependant on the production of significant quantities of tyrosinase activity. Tyrosinase activity may be lost during isolation of the secreted tyrosinase due to a disruption of the tyrosinase - activator protein complex. It is therefore an objective of the present invention to stabilize the tyrosinase - activator protein complex in order to reduce tyrosinase activity loss during isolation and purification.

Fusion proteins have been synthesized to overcome instability and proteolytic degradation of a polypeptide of interest. Many eucaryotic proteins have been produced in E. coli as fusion proteins with E. coli polypeptides such as Beta-galactosidase. Beta-galactosidase can be used to protect the foreign passenger protein from degradation in E. coli. Somatostatin, the first eucaryotic protein to be produced in E. coli was produced by fusing a synthetic gene to the entire Beta-galactosidase coding sequence (Itakura, et al., Science 198:1056 (1977)) and somatostatin was subsequently isolated by chemical cleavage of the fusion protein. When two independently functioning polypeptides are fused into

-6-

5 a single polyprotein by way of a synthetic gene
fusion, the resulting gene construct can be
engineered for expression behind a single promoter,
ribosome binding site, and initiation codon. A gene
fusion constructed in such a way can be placed in a
single location within a chromosome or episomal
element for subsequent expression of the fusion
protein. Such fusion genes are inherently more
stable because both gene sequences are coordinately
10 expressed at high levels from a single promoter.
Undesirable effects such as differential down
regulation or rearrangement and deletion of one of
the genes can thus be avoided.

15 Unfortunately, however, most fusion protein
constructs are not functional. Proteins fold into
conformationally active states as dictated by their
primary amino acid sequence. Additional polypeptide
sequences at the C- or N- terminus can interfere with
proper folding, thereby preventing the formation of a
20 biologically active protein. Enzymes typically fold
into well defined three-dimensional conformations to
allow the formation of a catalytic pocket that
excludes potential substrate molecules of abnormal
size or conformation, but allows access of substrate
25 molecules with the correct three-dimensional
structure. Many enzymes will not function as fusion
polypeptides because the sequence extensions
interfere with proper folding, or interfere by
sterically blocking the substrate's access to the
30 catalytic site.

The present invention relates to Applicants'
recognition that it might be possible to construct a
biologically active fusion enzyme coupling tyrosinase
with an activator protein. It has been shown that
35 histidine residues #102 and # 117 of the ORF438

-7-

activator protein are critical for copper binding and delivery to the active site of tyrosinase. Chen, et al., J. Biol. Chem. 268:18710 (1993). In order to form a biologically active fusion enzyme between tyrosinase and an activator protein, both functional domains of tyrosinase and the activator protein must be correctly folded. Further, the fusion enzyme must be able to assume a structural conformation enabling the activator protein to intermolecularly deliver copper to the active site of the tyrosinase to form a biologically active tyrosinase. Based on the existing prior art, it is unclear whether such a fusion enzyme would be biologically active.

SUMMARY OF INVENTION

The present invention relates to a nucleic acid sequence encoding a fusion enzyme comprising a nucleic acid sequence encoding for a tyrosinase and a nucleic acid sequence encoding for a tyrosinase activator protein. Activator proteins used in the present invention include but are not limited to ORF438 and URF402. Tyrosinases used in the present invention include both prokaryotic and eucaryotic tyrosinases. Prokaryotic tyrosinases that require a separate activator protein, particularly tyrosinases derived from Streptomyces are preferred. It is also preferred that the activator protein sequence be positioned 5' relative to the tyrosinase sequence.

The present invention also relates to a vector useful for introducing a nucleic acid sequence encoding a fusion enzyme into an organism. The vector comprises a nucleic acid sequence encoding for a fusion enzyme of the present invention and a promoter sequence that regulates the transcription of fusion enzyme. The present invention also relates to

WO 95/13386

- 8a -

PCT/US94/12857

organisms, such as bacteria, yeast fungi, plants and animals which have been transformed by the vector of the present invention.

5 The present invention also relates a fusion enzyme which comprises an amino acid sequence for tyrosinase and an amino acid sequence for a tyrosinase activator protein. The fusion enzyme may also contain a linker positioned between the amino acid sequences of the activator protein
10 and the tyrosinase.

The present invention also relates to melanin produced by a fusion enzyme of the present invention. The melanin may be produced *in vitro* by contacting a fusion enzyme with an
15 enzyme substrate under suitable reaction conditions to form melanin. Melanin may also be produced in vivo by an organism transformed with a vector of the present invention.

The present invention also relates to a
20 nucleic acid sequence encoding a fusion enzyme comprising a nucleic acid sequence encoding for a tyrosinase and a nucleic acid sequence encoding for a tyrosinase activator protein positioned 5' relative to the tyrosinase
25 sequence.

The present invention also relates to a vector for transforming an organism, the vector comprising: a nucleic acid sequence encoding for a fusion enzyme, the fusion enzyme comprising an
30 amino acid sequence encoding for a tyrosinase and an amino acid sequence encoding for a tyrosinase activator protein that is positioned on the N-terminus end of the tyrosinase

WO 95/13386

- 8b -

PCT/US94/12857

sequence; and a promoter sequence that regulates the transcription of fusion enzyme.

5 The present invention also relates to a fusion enzyme comprising an amino acid sequence for tyrosinase and an amino acid sequence for a tyrosinase activator protein that is positioned on the N-terminus end of the tyrosinase sequence.

10 The present invention also relates to a melanin wherein the melanin is produced *in vitro* by a fusion enzyme comprising an amino acid sequence for tyrosinase and an amino acid sequence for a tyrosinase activator protein that is positioned on the N-terminus end of the tyrosinase sequence.

15 The present invention also relates to an *in vitro* method of producing a melanin comprising the steps of contacting a fusion enzyme comprising an amino acid sequence for tyrosinase and an amino acid sequence for a tyrosinase
20 activator protein and a reaction substrate selected from the group consisting of L-tyrosine, X/L-tyrosine, L-tyrosine/X and X/L-tyrosine/X where X is a single amino acid, a
25 dipeptide or an oligopeptide bound to L-tyrosine, under suitable reaction conditions, to form melanin.

30 The present invention also relates to a method of producing melanins, comprising the step of growing an organism which has been transformed with a vector comprising a nucleic acid sequence encoding for a fusion enzyme, the fusion enzyme comprising an amino acid sequence encoding for a tyrosinase and an amino acid

WO 95/13386

- 8c -

PCT/US94/12857

sequence encoding for a tyrosinase activator
protein that is positioned on the N-terminus end
of the tyrosinase sequence; and a promoter
sequence that regulates the transcription of
5 fusion enzyme.

The present invention also relates to a
nucleic acid sequence encoding a fusion enzyme
comprising a nucleic acid sequence encoding for
a protein that has cresolase activity and
10 catecholase activity; and a nucleic acid
encoding for a gene product that functions as a
tyrosinase activator protein.

The present invention also relates to an
amino acid sequence for a fusion enzyme
15 comprising: an amino acid sequence for a protein
that has cresolase activity and catecholase
activity; and an amino acid sequence for a
protein that has a tyrosinase activator
function.

20

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 depicts the stepwise biosynthesis
of melanins.

25 Figure 2 provides the plasmid map of
pBGC623.

Figure 3 provides the plasmid map of
pBGC635.

30 Figure 4 provides the plasmid map of
pBGC636.

Figure 5 provides the plasmid map of
pBGC646.

Figure 6 provides the plasmid map of
pBGC648.

WO 95/13386

- 8d -

PCT/US94/12857

Figure 7 provides a comparison of the melanin production capabilities of E. coli when transformed by plasmids pBGC623, pBGC635, pBGC636, pBGC646 and pBGC648.

5

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a nucleic acid sequence encoding for a fusion enzyme of tyrosinase and an activator protein for tyrosinase. The present

10

2176236

-9-

invention also relates to a vector useful for transforming a host organism that contains the nucleic acid sequence of the present invention.

5 The present invention also relates to the expression of a fusion enzyme of a tyrosinase and an activator protein by the transformed organism. Host organisms that may be transformed to express the fusion enzyme of the present invention include but are not limited to bacteria, yeast fungi, plants and
10 animals. Expression of the fusion enzyme by the host organism enables and/or enhances tyrosinase activity in the host thereby enabling and/or enhancing melanin production by the host.

15 The present invention also relates to a fusion enzyme comprising an amino acid sequence for tyrosinase and an amino acid sequence for an activator protein of tyrosinase.

20 Finally, the present invention relates to the in vivo and in vitro production of melanin using the fusion enzyme of the present invention.

25 The fusion enzyme of the present invention comprises an amino acid sequence for a tyrosinase and an amino acid sequence for an activator protein. The nucleic acid sequences for both proteins are under the same promoter control and are thus expressed as a single peptide.

30 Tyrosinases used in the present invention include both prokaryotic and eucaryotic tyrosinases. In some organisms, tyrosinases are referred to as polyphenol oxidases. Tyrosinases from prokaryotes are preferred, particularly those tyrosinases that require a separate activator protein such as those that have been obtained from Streptomyces.

35 Activator proteins used in the present invention include but are not limited to ORF438 and URF402.

-10-

5 The activator protein employed in the present invention need not naturally occur in the organism from which the tyrosinase is derived. Rather, activator proteins from a variety of sources should function with a given tyrosinase in view of the similarity of the copper binding sites of different tyrosinases. See Chen, et al., J. Biol. Chem. 268 18710 (1993).

10 The fusion enzymes of the present invention are preferably constructed such that the amino acid sequence for the activator protein is positioned on the N- terminus of the tyrosinase amino acid sequence. As can be seen from Example 5, fusion enzymes where the activator protein is positioned on the N- terminus of tyrosinase exhibit equivalent melanin production capabilities as where tyrosinase and the activator protein are expressed separately.

15 Eucaryotic tyrosinases tend to comprise higher molecular weight (50-70 kD) single polypeptide chains without defined activator proteins. Applicants speculate that the activator equivalent of the ORF438 protein is built into the polypeptide backbone of eucaryotic tyrosinases. Based on the amino acid homology at the active sites between eucaryotic and Streptomyces tyrosinases, it appears that excess polypeptide sequences occur in the C- terminal region of the larger eucaryotic enzymes. Based on this observation, it might be predicted that ORF438 would function best if fused to the C- terminus rather than the N- terminus of the 30 kD Streptomyces tyrosinase since the fusion protein would then mimic the eucaryotic tyrosinases with respect to placement of the catalytic site. However, since many newly synthesized polypeptides undergo three dimensional folding into their preferred active conformation

20
25
30
35

-11-

beginning with the free C- terminus, it might also be expected that C- terminal additions are disruptive to proper folding and, hence, normal catalytic activity. Hence, prior to preparing the fusion enzymes of the present invention, it was unclear whether a fusion enzyme at either the N- or C- terminus would be functional.

The amino acid sequence encoding the activator protein need not be directly attached to the tyrosinase sequence. In plasmid pBGC648, a His residue has been inserted between the activator protein and the tyrosinase sequence. A single amino acid and a repeating Pro-Thr amino acid sequence, which behaves as a natural hinge, are preferred as linking sequences between tyrosinase and the activator protein.

Determination of the wide variety of amino acid sequences that may be interposed between the tyrosinase and activator protein sequences can be routinely determined by one of ordinary skill in the art in view of the teachings of the present invention.

Other polypeptide linkers that are known from the literature to provide a high degree of flexibility between two polypeptide sub-domains might work equally as well as linking sequences between the tyrosinase sequence and the activator protein sequence. One example of this type of hinge polypeptide is a proline - threonine repeating unit such as those known from cellulose binding proteins. Ong, et al., Bio/Technology 7 604(1989). Construction of synthetic oligonucleotide linkers that encode hinge polypeptides of known flexibility is within the level of ordinary skill. In addition, synthetic oligonucleotide linkers could be used to

-12-

construct random polypeptide linker sequences to join the tyrosinase - activator protein domains together. Linkers with a high degree of flexibility and sufficient length might be expected to work best for intermolecular cis-activation of tyrosinase. Inflexible linker polypeptides, or extremely short linker polypeptides, might be expected to permit only intramolecular trans-activation of neighboring fusion enzymes.

In order to provide a clear and consistent understanding of the specification and the claims, including the scope given to such terms, the following definitions are provided:

Activator protein: a gene product that alters, activates or enhances the activity of tyrosinase. The activator protein may function as a trans-activator, as a metallothionein-like protein that delivers an ion to a tyrosinase apoenzyme or it may function in assisting the secretion of tyrosinase. ORF438 gene, URF402 and ORF(s) 3' to the tyrosinase coding sequence code for activator proteins that enhance melanogenesis in all of the ways described.

Melanin: Melanins are polymers produced by polymerization of reactive intermediates. The polymerization mechanisms include but are not limited to autoxidation, enzyme catalyzed oxidation and free radical initiated polymerization. The reactive intermediates are produced chemically or enzymatically from precursors. Suitable enzymes include, but are not limited to

-13-

peroxidase and catalases, polyphenol oxidases, tyrosinases, tyrosine hydroxylases or laccases. The precursors which are converted to the reactive intermediates are hydroxylated aromatic compounds. Suitable hydroxylated aromatic compounds include, but are not limited to

5 1) phenols, polyphenols, aminophenols and thiophenols of aromatic or polycyclic aromatic hydrocarbons, including but not limited to phenol, tyrosine, pyrogallol, 3-aminotyrosine, thiophenol and α -naphthol;

10 2) phenols, polyphenols, aminophenols, and thiophenols of aromatic heterocyclic or heteropolycyclic hydrocarbons such as but not limited to 2-hydroxypyrrole, 4-hydroxy-1,2-pyrazole, 4-hydroxypyridine, 8-hydroxyquinoline, and 4,5-dihydroxybenzothiazole. Suitable

15 hydroxylated aromatic compounds also include X/L-tyrosine, L-tyrosine/X and X/L-tyrosine/X where X is a single amino acid, a dipeptide or an oligopeptide bound to L-tyrosine.

20

25 The nucleic acid sequences encoding for the fusion enzymes of the present invention may be inserted into a wide variety of vector constructs known in the art for transforming a host organism. Suitable techniques include those described in

30 Maniatis, et al., Molecular Cloning, 1st Ed., Cold Spring Harbor Laboratory, New York (1982); Molecular Cloning, 2nd Ed., Cold Spring Harbor Laboratory, New York (1989); Methods in Enzymology, Vols. 68 (1979), 100 (1983), 101 (1983), 118 (1986) and Vols. 152-154

- 14 -

(1987) DNA Cloning, Glover, Ed., IRL Press, Oxford
(1985); and Plant Molecular Biology: Manual, Gelvin, et
al., Eds., Kluwer Academic Publishers, Podrecht (1988).

5 Medium compositions have been described in Miller,
Experiments in Molecular Genetics, Cold Spring Harbor
Laboratory, New York (1972), as well as the references
previously identified. Hopewood, et al., "Genetic
10 Manipulation of Streptomyces: A Laboratory Manual", The
John Innes Foundation, Norwich, England (1985).

With regard to the expression of the tyrosinase-
activator protein fusion enzyme of the present
invention in a transformed microorganism, it is
preferred that the transformed organism be grown under
15 the conditions described in PCT application WO
92/00373, published January 9, 1992.

With regard to the expression of the tyrosinase-
activator protein fusion enzyme of the present
invention in plants, it is preferred that the nucleic
20 acid cassette encoding the fusion enzyme be inserted
into one of the viral constructs described in U.S.
Patent No. 5,316,931, issued May 31, 1994, or PCT
application WO 94/16089, published July 21, 1994.

Vectors encoding the tyrosinase-activator protein
25 fusion enzyme of the present invention may be produced
by standard techniques. Appropriate vectors which can
be utilized as starting materials are known in the art.

The DNA sequence coding for the fusion enzyme is
inserted into an appropriate vector in such a manner
30 that the enzyme is correctly expressed. In other words,
the DNA sequence is positioned in the proper

-15-

orientation and reading frame so that the correct amino acid sequence is produced upon expression of the DNA sequence in the host. In accordance with conventional techniques, a chimeric DNA sequence is generally constructed which contains a promoter operable in the specific host and the DNA sequence coding for the desired enzyme. The chimeric DNA sequence may further contain 3' non-coding sequences operable in the host. The chimeric DNA sequence can be prepared in situ within a suitable vector by inserting the DNA sequence coding for the enzyme into a restriction site of a known host transformation vector. Alternatively, the chimeric gene could be first constructed and then inserted into a vector to produce a transformation vector. The vector can be further modified by utilizing an enhancer sequence and/or a strong promoter, which leads to an increased production of the fusion enzyme.

The typical vector is a plasmid having one or more marker genes for antibiotic resistance, an origin of replication, and a variety of useful restriction sites for cloning or subcloning restriction fragments. A large number of vectors have been described which are useful for transforming many microorganisms including but not limited to Streptomyces and E. coli. See, for example, Cloning Vectors, Pouwels, et al. ed. Elsevier Science Publishers Amsterdam (1985).

A large number of naturally occurring Streptomyces plasmids have been described, many of which are conjugally proficient. Two such isolates, SLP1.2 and pIJ101, have formed the basis of a series of useful plasmid vectors. Thompson, et al., Gene 20:51 (1982). The plasmids of the SLP1 family, of which SLP1.2 is the largest detected member, were

-16-

discovered as autonomous replicons in S. lividans 66 after interspecific matings with S. coelicolor A3(2). The SLP1 replicon is integrated in the S. coelicolor genome but can be excised together with various lengths of neighboring DNA to become autonomous in S. lividans. The SLP1 plasmids exist stably at a copy number of 4-5 per chromosome in S. lividans and have a narrow host range.

The 8.9 kb plasmid pIJ101 was discovered in S. lividans ISP5434 (Kieser, et al., Mol. Gen. Genet. 185:223 (1982)) but can be conjugally transferred to a wide variety of Streptomyces species. Derivatives (e.g. pIJ102) have been isolated from the plasmid which have similar properties but are smaller. Kieser, et al. (1982), supra. Plasmid pIJ101 has a copy number of 100-300 per chromosome equivalent in most hosts and a minimum replicon of less than 2.1 kb. Derivatives carrying drug-resistance determinants have been constructed to act as vectors, and a chimeric plasmid which can be used as a shuttle vector between E. coli and Streptomyces is available.

The temperate phage ϕ C31 has a wide host range within the Streptomyces and lysogenizes S. coelicolor A-3(2) via a site-specific integration event. Lomovshaya, et al., Bacteriol Rev. 44, 206 (1980). Up to 42.4 kb of DNA can be packaged within a viable phage particle, but only 32 kb (at the most) of the DNA contains the genetic information essential for plaque formation. Derivatives of ϕ C31 containing deletions can be used as vectors, and recombinant phages can either be grown lytically or used to lysogenize suitable streptomyces strains.

pBR322-derived plasmids are very common for use in E. coli transformation. They possess a pair of antibiotic resistance genes which confer antibiotic

2176236

-17-

resistance when Escherichia coli are successfully transformed. Typically, the insertion of a DNA segment is made so that one of the antibiotic resistance genes is inactivated. Selection then is accomplished by selecting for E. coli exhibiting antibiotic resistance conferred by the second gene. Bolivar, et al., Gene 2:95 (1977); and Sutcliff, J., Proc. Natl. Acad. Sci., USA 75:3737 (1978).

Another example of transforming vectors is the bacteriophage. The M13 series are modified filamentous E. coli bacteriophage containing single stranded circular DNA. The M13 series carry the lacZ gene for β -galactosidase and will metabolize the galactose analog Xgal to produce a blue color. Placing a cloned insert into the polylinker sequence located in the amino terminus of the lacZ gene inactivates the gene. Microorganisms carrying an M13 with an inactivated lacZ (representing a cloned insert) are distinguishable from those carrying an M13 with an active lacZ gene by their lack of blue color. Messing, et al., Proc. Natl. Acad. Sci., USA 74:3642 (1977); and Messing, Methods in Enzymology 101:20 (1983).

Other transforming vectors are the pUC series of plasmids. They contain the ampicillin resistance gene and origin of replication from pBR322, and a portion of the lacZ gene of E. coli. The lac region contains a polylinker sequence of restriction endonuclease recognition sites identical to those in the M13 series. The pUC series have the advantage that they can be amplified by chloramphenicol. When a DNA fragment is cloned into the lac region the lac gene is inactivated. When E. coli containing a pUC plasmid with an inactivated lacZ gene is grown in the presence of isopropylthiogalactoside (IPTG) and

- 18 -

5-bromo-4-chloro-3-indolyl β -D-galactopyranoside (Xgal) its colonies are white. If it carries a pUC plasmid with an active lacZ gene its colonies are blue. Vieira, et al., Gene 19:259 (1982). Bacteria are transformed by means conventional in the art.

The genus Streptomyces is one of three aerobic genera of bacteria of the order Actinomycetales. Streptomyces are Gram-positive, mycelial, spore-forming bacteria. Several naturally occurring Streptomyces plasmids have been described. Streptomyces lividans TK64 has no tyrosinase gene and produces no melanin. Applicants teach the transformation of Streptomyces lividans TK64 with plasmid pIJ702 which encodes for the tyrosinase gene in PCT application WO 92/00373, published January 9, 1992. Transformation is carried out by means standard in the art. Similarly, transformation of Streptomyces can be performed using a plasmid encoding the fusion enzyme of the present invention. Transformation vectors of the present invention may also be used to transform a variety of microorganisms after insertion into vectors which are useful for transforming the corresponding host microorganism.

Bluescript (obtained from Stratagene, LaJolla, CA) is a pUC derivative having a β -galactosidase color indicator and a lac promoter. In PCT application WO 92/00373, published January 9, 1992, Applicants teach modification of the Bluescript plasmid by inserting a tyrosinase gene. This modified plasmid was used to successfully transform E. coli which formed pigmented colonies. Similarly, the Bluescript may be modified by inserting a nucleic acid sequence encoding for the fusion enzyme of the present invention.

-19-

Melanin has been purified from bacterial cells with 0.5N NaOH at room temperature and at 100°C. Pigmented fractions were found to be: (1) soluble in acid and base; (2) soluble in ethyl alcohol and base; and (3) soluble base only. Pavlenko, et al., Microbiology USSR 50 539 (1981)

Soluble melanin can be extracted from the medium and purified. This is done by first removing cells and particulate matter using, for example, filtration or centrifugation. A variety of filtration methods are known in the art including filtration through glass wool. If centrifugation is used, 5,000 X gravity is usually sufficient. The melanin is then precipitated at between pH 2-4, preferably about 3. Precipitated melanin is removed by either filtration or centrifugation. The melanin is washed by successive resolubilization at high pH, i.e. about pH 7.0 to about pH 9.0, preferably about pH 8.0, and precipitation at low pH followed by filtration or centrifugation. The melanin may also be concentrated using molecular weight filtration, such as reverse osmosis. Salt precipitation can be as effective in precipitating the melanin as low pH.

The invention is further illustrated by the following non-limiting examples.

EXAMPLES

1. Preparation of Plasmids pBGC623, pBGC635 and pBGC636

Plasmid pBGC623 is a plasmid containing the nucleic acid sequence for ORF438. Figure 2 provides the plasmid map of pBGC623. Plasmid pBGC635 is a plasmid containing a nucleic acid sequence for tyrosinase. Figure 3 provides the plasmid map of pBGC635. Plasmid pBGC636 is a plasmid encoding for

- 20 -

both ORF438 and tyrosinase under separate promoters. Figure 4 provides the plasmid map of pBGC636. Preparation of plasmids pBGC623, pBGC635 and pBGC636 are taught in PCT application WO 92/00373, published January 9, 1992.

2. Preparation of pBGC646

In order to prepare pBGC646, the Bcl I site at the stop codon of the tyrosinase gene (in pBGC188Nde) was blunted with mung bean nuclease and ligated with a synthetic Eco RI linker (dGGAATTCC; SEQ. NO. 1). The tyrosinase gene was removed from plasmid pBGC188Nde as a 822bp Nde I/Eco RI fragment and gel purified in low melt agarose. The purified fragment was ligated into an ORF438 containing plasmid (pBGC623) that was modified as follows. Plasmid pBGC623 was cleaved with Nco I, blunted with mung bean nuclease, and ligated with a synthetic Eco RI I linker (dGGAATTCC). The plasmid was gel purified in low melt agarose and ligated with the Nde I/Eco RI fragment containing the modified tyrosinase gene. Figure 5 provides the plasmid map of pBGC646 [SEQ. NO. 2]. The translated amino acid sequence for the fusion enzyme encoded for by plasmid pBGC646 is provided as SEQ. NO. 3.

Transformants were screened in HB101 and two independent transformants were identified as having the correct orientation. These plasmids are pBS646.9 and pBS646.19, and both are identical. Both plasmids were transformed into *E. coli* strain K38 that harbors plasmid PGP1-2 and plated on agar containing tyrosine and copper. Both transformants gave a melanin phenotype as described previously (della, Cioppa, et al. Bio/Technology 8:634-638 (1990)), and were shown by SDS-PAGE to give a ~45kD band by SDS-PAGE. The

-21-

hybrid tyrosinase/ORF438 fusion enzyme is predicted to encode a 45,806 MW protein with four additional amino acids linking the two functional domains (NH2-tyrosinase-trp-asn-ser-ala-ORF438-COOH).

5 3. Preparation of pBGC648

To construct a second hybrid fusion gene between ORF438 and tyrosinase, the two synthetic oligonucleotides shown below

5'-dCCAGGGCGCCCGGCTCCTCOCCTTCCCCTCCAACCA-3' [SEQ. NO. 4]

10 3'-dTCGAGGTCCCGCGGGCCGAGGAGGGGAAGGGGAGGTTGGTAT-5' [SEQ. NO. 5]

were annealed, kinased, and used to clone into the Sac I site of the ORF438 gene. This replacement oligonucleotide sequence results in the destruction of the TGA stop codon of ORF438 and creates an in-frame NdeI site in its place for insertion of tyrosinase. A triple ligation reaction was set up that included the annealed oligonucleotide shown above, a 2,934 bp Hind III/Sac I fragment from pBGC623, and a 1,107 bp NdeI/Hind III fragment from pBGC635. The new plasmid (pBGC648) creates an ORF438/tyrosinase in-frame gene fusion that encodes a single polypeptide chain of 46,230 Daltons. The introduction of the new NdeI site introduces a single histidine residue between the two coding sequences upon translation. Figure 6 provides the plasmid map of pBGC648. The nucleic acid sequence for pBGC648 is provided as SEQ. NO. 6. The translated amino acid sequence for the fusion enzyme encoded for by plasmid pBGC648 is provided as SEQ. NO. 7.

30 The DNA at the junction of the ORF438 and tyrosinase genes in plasmid pBGC648 were sequenced. Approximately 180 bp 3' of the Sac I site in the fusion were found to be correct, and the translated amino acid sequence is as predicted (See SEQ. No. 7).

-22-

Plasmid pBGC648 was transformed into *E. coli* strain K38 that harbors plasmid pGP1-2 and plated on agar containing tyrosine and copper. The transformants gave a black melanin phenotype as described previously (della-Cioppa, et al., Bio/Technology 8:634-638 (1990)), and were shown by SDS-PAGE to give a ~46kD band by SDS-PAGE. Transformants harboring pBGC648 give rise to the black melanin phenotype as rapidly as that seen when ORF438 and tyrosinase are expressed as single polypeptide chains from single genes. Phenotypically, the ORF438/tyrosinase gene fusion retains full catalytic activity similar to the wild type ~30kD tyrosinase holoenzyme.

4. Transformation of *E. coli* with pBGC623, pBGC635, pBGC636, pBGC646 and pBGC648

Plasmids pBGC623, pBGC635, pBGC636, pBGC646 and pBGC648 were introduced into *E. coli* by pretreating exponentially growing cultures of the *E. coli* with CaCl_2 as described in Maniatis, et al., Molecular Cloning (1982).

5. Comparison of Melanin Production By *E. coli* transformed with pBGC623, pBGC635, pBGC636, pBGC646 and pBGC648

Plasmids pBGC623, pBGC635, pBGC636, pBGC646 and pBGC648 were each introduced into *E. coli*. The transformed strains of *E. coli* were then grown on agar plates containing tyrosine and copper as described previously in della-Cioppa, et al., Bio/Technology 8 634 (1990) in order to evaluate each strains ability to produce melanin. Figure 7 provides a comparison of the melanin production capabilities of *E. coli* when transformed by these five different plasmids.

-23-

5 The amount of melanin formed was determined by the size and color intensities of black melanin halos that formed around the transformed E. coli colonies. The rate of color development on agar plates, and the intensity of the black halo formation, is directly proportional to the level of tyrosinase enzymatic activity in each of the different plasmid bearing E. coli colonies. Melanin formation was quantitated as (-) none, (+) very weak, (++) weak, (+++) moderate, 10 (+++++) strong and (+++++) very strong.

E. coli transformed with either pBGC623 or pBGC635 do not produce a positive black melanin phenotype. E. coli transformed with pBGC636 produced a positive black melanin phenotype.

15 E. coli transformed with pBGC646 gave rise to a positive black melanin phenotype. However, E. coli transformed with pBGC646 produced melanin at a slower rate than E. coli transformed with pBGC636 in which ORF438 and tyrosinase are expressed as single 20 polypeptide chains from single genes.

E. coli transformed with pBGC648 produced a positive black melanin phenotype at a comparable rate as pBGC636, indicating that the fusion enzyme produced by pBGC648 functions equally well as when 25 ORF438 and tyrosinase are expressed independently.

6. Expression Of Tyrosinase/ORF438 Fusion Enzyme Containing Chloroplast Targeting Sequence In Plants

30 For expression of a tyrosinase - activator protein fusion enzyme in higher plants, it may be advantageous to target the fusion enzyme to the chloroplast. Chloroplasts are known to contain the enzymatic pathway for production of L-tyrosine (the

- 24 -

primary substrate for tyrosinase), and the oxidative environment inside the chloroplast may be well suited for achieving optimal enzymatic activity.

5 In order to target a tyrosinase - activator protein fusion enzyme to chloroplasts, the nucleotide sequence encoding the chloroplast transit peptide (CTP) from ribulose biphosphate carboxylase small subunit (RuBPCase SSU) from Nicotiana tabacum was cloned (as an
10 Nco I/Sph I fragment) and fused by way of its naturally occurring Sph I site (at the cys-met cleavage site of the CTP) to the naturally occurring Sph I site at the N-terminus of ORF438. The CTP/ORF438 nucleotide fusion was then exchanged as an Nco I/Sac I fragment in the
15 plasmid BlueScript™ that contained the ORF438/tyrosinase sequence as described in plasmid pBS648. The resulting nucleotide sequence and translated amino acid sequence of the CTP/ORF438/tyrosinase fusion are shown in SEQ. No. 8
20 and SEQ. No. 9 respectively. The CTP/ORF438 tyrosinase fusion gene encodes 478 amino acids residues of which the 57 at the N-terminus direct the ORF438/tyrosinase fusion into the chloroplast. Upon import into the chloroplast compartment, the 57 amino acid CTP is
25 proteolytically removed thus resulting in the identical ORF438/tyrosinase fusion enzyme of ~45 kD as previously produced in E. Coli. The CTP/ORF438/tyrosinase fusion enzyme has a deduced molecular weight of 51,461 Daltons.

30 The CTP/ORF438/tyrosinase nucleotide sequence may then be inserted into a viral construct such as those described in U.S. Patent No. 5,316,931, issued May 31, 1994, or PCT application WO 94/16089, published July 21, 1994 and used to systemically infect higher plants.

-25-

5 While the invention has been disclosed by reference to the details of preferred embodiments, the disclosure is intended in an illustrative rather than in a limiting sense, as it is contemplated that modifications will readily occur to those skilled in the art, within the spirit of the invention and the scope of the appended claims.

-26-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Della-Cioppa, Guy
Kumagai, Monto
- (ii) TITLE OF INVENTION: TYROSINASE-ACTIVATOR
PROTEIN FUSION ENZYME
- (iii) NUMBER OF SEQUENCES: 9
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Pennie & Edmonds
 - (B) STREET: 2730 Sand Hill Road
 - (C) CITY: Menlo Park
 - (D) STATE: California
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 94025
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: Patent in Release #1.0,
Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/152,483
 - (B) FILING DATE: November 12, 1993
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 7/857,602
 - (B) FILING DATE: March 30, 1992
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 923,692
 - (B) FILING DATE: July 31, 1992
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 600,244
 - (B) FILING DATE: October 22, 1990
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 641,617
 - (B) FILING DATE: January 16, 1991
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 737,899
 - (B) FILING DATE: July 26, 1991
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Halluin, Albert P.
 - (B) REGISTRATION NUMBER: 25,227
 - (C) REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (415) 854-3660
 - (B) TELEFAX: (415) 854-3694
 - (C) TELEX: 66141 PENNIE

-27-

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 8
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 - (A) DESCRIPTION: DNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM:
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE:
- (ix) FEATURE:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGAATTCC

8

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4294
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 - (A) DESCRIPTION: DNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM:
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE:
- (ix) FEATURE:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

AAATCAATCT AAAGTATATA TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT60
GAGGCACCTA TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC120
GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CAGTGCTGCA ATGATACCGC180
GAGACCCACG CTGACCGGCT CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG240
AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA GTCTATTAAT TGTTGCCGGG300
AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG360

-28-

GCATCGTGGT GTCACGCTCG GCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT420
CAAGGCGAGT TACATGATCC CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCGGTCCTC480
CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTATCACT CATGGTTATG GCAGCACTGC540
ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA600
CCAAGTATTT GGAAGATGCG CGACCGAGTT GCTCTTGCCC GCGGTCAACA CGGGATAATA660
CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT TCGGGGCGAA720
AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA780
ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC840
AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC900
TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG960
AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC1020
CTGACGTCTA AGAAACCAT ATTATCATGA CATTAACTA TAAAAATAGG CGTATCACGA1080
GGCCCTTTTCG TCTTCAAGAA ttaaaaggat ctaggtgaag atccttttttg ataatctcat1140
gaccaaaatc ccttaacgtg agttttcggt ccactgagcg tcagaccccg tagaaaagat1200
caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa1260
accaccgcta ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc tttttccgaal320
ggtaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt agccgtagtt1380
aggccaccac ttcaagaact ctgtagcacc gcctacatac ctcgctctgc taatcctggt1440
accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact caagacgata1500
gttaccggat aaggcgcagc ggtcgggctg aacggggggg tcgtgcacac agcccagctt1560
ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag aaagcgccac1620
gcttcccga gggagaaagg cggacaggta tccggtaagc ggcagggtcg gaacaggaga1680
gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg tcgggtttcgl1740
ccacctctga cttgagcgtc gattttttgtg atgctcgtca ggggggcgga gcctatggaa1800
aaacgccagc aacgcggcct ttttacgggt cctggccttt tgctggcctt ttgctcacat1860
gttcttttct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc1920
tgataccgct cgcgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga1980
agagcgctg atgcggtatt ttctccttac gcatctgtgc ggtattttcac accgcaCAGA2040
TCTGtggtgc actctcagta caatctgctc tgatgccgca tagttaagcc agtatacact2100
ccgctatcgc tacgtgactg ggtcatggct gcgccccgac acccgccaac acccgctgac2160
gcgcctgac gggcttgtct gctcccggca tccgcttaca gacaagctgt gaccgtctcc2220
gggagctgca tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag gCCcagctgC2280

-29-

GATTCGAAct tctcgattcg aacttctgat agacttcgaa attaatagca ctcactatag2340
ggagaccaca acggtttccc tctagaaata attttgttta actttaagaa ggagatatatac2400
atatgACCGT CCGCAAGAAC CAGGCGTCCC TGACCGCCGA GGAGAAGCGC CGCTTCGTTCG2460
CCGCCCTGCT CGAACTCAAG CGCACCGGCC GCTACGACGC CTTCGTCACC ACGCACAACG2520
CGTTCATCCT GGGCGACACC GACAACGGCG AGCGCACCGG CCACCGTTCG CCGTCCTTCC2580
TGCCCTGGCA CCGCAGATTT CTGCTGGAGT TCGAGCGGGC GCTCCAGTCG GTGGACGCGT2640
CGGTGGCGCT GCCGTACTGG GACTGGTCCG CCGACCGGTC CACCCGGTCC TCGCTGTGGG2700
CGCCGGACTT CCTCGGCGGC ACCGGGCGCA GCCGGGACGG CCAGGTGATG GACGGGCCGT2760
TCGCCGCGTC GGCCGGCAAC TGGCCGATCA ATGTGCGGGT GGACGGCCGT ACGTTCCTGC2820
GGCGGGCGCT CGGCGCGGGC GTGAGCGAAC TGCCACGCG TGCCGAGGTC GACTCGGTGC2880
TGGCGATGGC GACGTACGAC ATGGCGCCCT GGAACAGCG CTCCGACGGC TTCCGCAACC2940
ATCTCGAAGG GTGGCGCGGG GTCAATCTGC ACAACCGGGT GCATGTCTGG GTCGGCGGCC3000
AGATGGCGAC CGGGGTCTCC CCCAACGACC CGGTGTTCTG GCTGCACCAC GCCTACATCG3060
ACAAGCTGTG GGCCGAGTGG CAGCGGCGGC ACCCCTCGTC CCCGTATCTG CCGGGCGGCG3120
GCACGCCGAA CGTCGTGAC CTCAACGAGA CGATGAAGCC GTGGAACGAC ACCACCCCGG3180
CGGCCCTGCT GGACCACACC CGGCACTACA CCTTCGACGT Ctggaaattcc GCGGAACTCA3240
CCCGTCGTG CGCGCTCGGC GCCGCAGCCG TCGTCGCCGC CGGTGTCCCG CTGGTCGCCC3300
TTCCCGCCGC CCGCGCGGAC GATCGGGGGC ACCACACCCC CGAGGTCCCC GGGAAACCCGG3360
CCGCGTCCGG CGCCCCCGCC GCCTTCGACG AGATCTACAA GGGCCGCCGG ATACAGGGCC3420
GGACGGTCAC CGACGGCGGG GGCCACCACG GCGGCGGTCA CGGCGGTGAC GGTCACGGCG3480
GCGGCCATCA CGGCGGCGGT TACGCCGTGT TCGTGGACGG CGTCGAACTG CATGTGATGC3540
GCAACGCCGA CGGCTCGTGG ATCAGCGTCG TCAGCCACTA CGAGCCGGTG GACACCCCGC3600
GCGCCGCGGC CCGCGCTGCG GTCGACGAGC TCCAGGGCGC CCGGCTCCTC CCCTTCCCCT3660
CCAActgaCC TTCTCCCCCG CACTTTTGGA GCACCCGCAC atgACCGTCC GCAAGAACCA3720
GGCGTCCCTG ACCGCCGAGG AGAAGCGCCG CTTCGTGCGC GCCCTGCTCG AACTCAAGCG3780
CACCGGCCGC TACGACGCCT TCGTCACCAC GCACAACGCG TTCATCCTGG GCGACACCGA3840
CAACGGCGAG CGCACCGGCC ACCGTTGCGC GTCCTTCCTG CCCTGGCACC GCAGATTTCT3900
GCTGGAGTTC GAGCGGGCGC TCCAGTCGGT GGACGCGTCG GTGGCGCTGC CGTACTGGGA3960
CTGGTCCGCC GACCGGTCCA CCCGGTCCTC GCTGTGGGCG CCGGACTTCC TCGGCGGCAC4020
CGGGCGCAGC CGGGACGGCC AGGTGATGGA CGGGCCGTTT GCCGCGTCGG CCGGCAACTG4080
GCCGATCAAT GTGCGGGTGG ACGGCCGTAC GTTCCTGCGG CGGGCGCTCG GCGCGGGCGT4140
GAGCGAACTG CCCACGCGTG CCGAGgtcga cCTCAACGAG ACGATGAAGC CGTGGAAACGA4200

-30-

CACCACCCCG GCGGCCCTGC TGGACCACAC CCGGCACTAC ACCTTCGACG TctgaTCcaa4260
gcttATCGAT GATAAGCTGT CAAACATGAG AATT

4294

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 422
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 - (A) DESCRIPTION: protein
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM:
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE:
- (ix) FEATURE:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met	Thr	Val	Arg	Lys	Asn	Gln	Ala	Ser	Leu	Thr	Ala	Glu	Glu	Lys	
				5					10					15	
Arg	Arg	Phe	Val	Ala	Ala	Leu	Leu	Glu	Leu	Lys	Arg	Thr	Gly	Arg	
				20					25					30	
Tyr	Asp	Ala	Phe	Val	Thr	Thr	His	Asn	Ala	Phe	Ile	Leu	Gly	Asp	
				35					40					45	
Thr	Asp	Asn	Gly	Glu	Arg	Thr	Gly	His	Arg	Ser	Pro	Ser	Phe	Leu	
				50					55					60	
Pro	Trp	His	Arg	Arg	Phe	Leu	Leu	Glu	Phe	Glu	Arg	Ala	Leu	Gln	
				65					70					75	
Ser	Val	Asp	Ala	Ser	Val	Ala	Leu	Pro	Tyr	Trp	Asp	Trp	Ser	Ala	
				80					85					90	
Asp	Arg	Ser	Thr	Arg	Ser	Ser	Leu	Trp	Ala	Pro	Asp	Phe	Leu	Gly	
				95					100					105	
Gly	Thr	Gly	Arg	Ser	Arg	Asp	Gly	Gln	Val	Met	Asp	Gly	Pro	Phe	
				110					115					120	
Ala	Ala	Ser	Ala	Gly	Asn	Trp	Pro	Ile	Asn	Val	Arg	Val	Asp	Gly	
				125					130					135	
Arg	Thr	Phe	Leu	Arg	Arg	Ala	Leu	Gly	Ala	Gly	Val	Ser	Glu	Leu	
				140					145					150	
Pro	Thr	Arg	Ala	Glu	Val	Asp	Ser	Val	Leu	Ala	Met	Ala	Thr	Tyr	
				155					160					165	
Asp	Met	Ala	Pro	Trp	Asn	Ser	Gly	Ser	Asp	Gly	Phe	Arg	Asn	His	
				170					175					180	
Leu	Glu	Gly	Trp	Arg	Gly	Val	Asn	Leu	His	Asn	Arg	Val	His	Val	
				185					190					195	
Trp	Val	Gly	Gly	Gln	Met	Ala	Thr	Gly	Val	Ser	Pro	Asn	Asp	Pro	
				200					205					210	
Val	Phe	Trp	Leu	His	His	Ala	Tyr	Ile	Asp	Lys	Leu	Trp	Ala	Glu	

-31-

Trp	Gln	Arg	Arg	His	Pro	Ser	Ser	Pro	Tyr	Leu	Pro	Gly	Gly	Gly	225
Thr	Pro	Asn	Val	Val	Asp	Leu	Asn	Glu	Thr	Met	Lys	Pro	Trp	Asn	240
Asp	Thr	Thr	Pro	Ala	Ala	Leu	Leu	Asp	His	Thr	Arg	His	Tyr	Thr	255
Phe	Asp	Val	Trp	Asn	Ser	Ala	Glu	Leu	Thr	Arg	Arg	Arg	Ala	Leu	270
Gly	Ala	Ala	Ala	Val	Val	Ala	Ala	Gly	Val	Pro	Leu	Val	Ala	Leu	285
Pro	Ala	Ala	Arg	Ala	Asp	Asp	Arg	Gly	His	His	Thr	Pro	Glu	Val	300
Pro	Gly	Asn	Pro	Ala	Ala	Ser	Gly	Ala	Pro	Ala	Ala	Phe	Asp	Glu	315
Ile	Tyr	Lys	Gly	Arg	Arg	Ile	Gln	Gly	Arg	Thr	Val	Thr	Asp	Gly	330
Gly	Gly	His	His	Gly	Gly	Gly	His	Gly	Gly	Asp	Gly	His	Gly	Gly	345
Gly	His	His	Gly	Gly	Gly	Tyr	Ala	Val	Phe	Val	Asp	Gly	Val	Glu	360
Leu	His	Val	Met	Arg	Asn	Ala	Asp	Gly	Ser	Trp	Ile	Ser	Val	Val	375
Ser	His	Tyr	Glu	Pro	Val	Asp	Thr	Pro	Arg	Ala	Ala	Ala	Arg	Ala	390
Ala	Val	Asp	Glu	Leu	Gln	Gly	Ala	Arg	Leu	Leu	Pro	Phe	Pro	Ser	405
Asn	Gly														420

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 - (A) DESCRIPTION: DNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM:
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE:
- (ix) FEATURE:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CCAGGGCGCC CGGCTCCTCC CCTTCCCCTC CAACCA

36

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 42
 - (B) TYPE: nucleic acid

-32-

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM:

(vii) IMMEDIATE SOURCE:
(B) CLONE:

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TCGAGGTCCC GCGGGCCGAG GAGGGGAAGG GGAGGTTGGT AT 42

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4009
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE:
(A) DESCRIPTION: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM:

(vii) IMMEDIATE SOURCE:
(B) CLONE:

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AAATCAATCT AAAGTATATA TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT60
GAGGCACCTA TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC120
GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CAGTGCTGCA ATGATACCGC180
GAGACCCACG CTGACCGGCT CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG240
AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA GTCTATTAAT TGGTGCCGGG300
AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG360
GCATCGTGGT GTCACGCTCG GCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT420
CAAGGCGAGT TACATGATCC CCCATGTTGT GCAAAAAGC GGTTAGCTCC TTCGGTCCTC480

-33-

CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTATCACT CATGGTTATG GCAGCACTGC540
ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA600
CCAAGTATTT GGAAGATGCG CGACCGAGTT GCTCTTGCCC GCGTCAACA CGGGATAATA660
CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT TCGGGGCGAA720
AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA780
ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC840
AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC900
TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG960
AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC1020
CTGACGTCTA AGAAACCATT ATTATCATGA CATTAACTA TAAAAATAGG CGTATCACGA1080
GGCCCTTTTCG TCTTCAAGAA ttaaaaggat ctaggtgaag atcctttttg ataatctcat1140
gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg tcagaccccg tagaaaagat1200
caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa1260
accaccgcta ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc tttttccgaa1320
ggtaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt agccgtagtt1380
aggccaccac ttcaagaact ctgtagcacc gcctacatac ctgcctctgc taatcctgtt1440
accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact caagacgata1500
gttaccggat aaggcgcagc ggtcgggctg aacggggggg tctgtcacac agcccagctt1560
ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag aaagcgccac1620
gcttcccga gggagaaagg cggacaggta tccggtaagc ggcagggtcg gaacaggaga1680
gcgcacgagg gagcttccag ggggaaacgc ctgggtatctt tatagtcctg tccgggtttcg1740
ccacctctga cttgagcgtc gattttttgtg atgctcgtca ggggggcgga gcctatggaa1800
aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt ttgctcacat1860
gttcttttct gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc1920
tgataccgct cgcgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga1980
agagcgctg atgcggtatt ttctccttac gcatctgtgc ggtattttcac accgcaCAGA2040
TCTGtggtgc actctcagta caatctgctc tgatgccgca tagttaagcc agtatacact2100
ccgctatcgc tacgtgactg ggtcatggct gcgccccgac acccgccaac acccgctgac2160
gcgcccctgac gggcttgtct gctcccggca tccgcttaca gacaagctgt gaccgtctcc2220
gggagctgca tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag gCCcagctgC2280
GATTCGAAct tctcgattcg aacttctgat agacttcgaa attaatacga ctcactatag2340
ggagaccaca acggtttccc TCTAGAAata attttgttta actttaagaa ggagatatatac2400

-34-

atATGGCTAG AATTGCCatg GCGGAACTCA CCCGTCGTCG CGCGCTCGGC GCCGCAGCCG2460
TCGTGCGCCG CGGTGTCCCG CTGGTCGCCC TTCCCGCCGC CCGCGCGGAC GATCGGGGGC2520
ACCACACCCC CGAGGTCCCC GGGAAACCCGG CCGCGTCCGG CGCCCCCGCC GCCTTCGACG2580
AGATCTACAA GGGCCGCCGG ATACAGGGCC GGACGGTCAC CGACGGCGGG GGCCACCACG2640
GCGGCGGTCA CGGCGGTGAC GGTCACGGCG GCGGCCATCA CGGCGGCGGT TACGCCGTGT2700
TCGTGGACGG CGTCGAACTG CATGTGATGC GCAACGCCGA CGGCTCGTGG ATCAGCGTCG2760
TCAGCCACTA CGAGCCGGTG GACACCCCGC GCGCCGCGGC CCGCGCTGCG GTCGACGAGC2820
TCCAGGGCGC CCGGCTCCTC CCCTTCCCCT CCAACcatAT GACCGTCCGC AAGAACCAGG2880
CGTCCCTGAC CGCCGAGGAG AAGCGCCGCT TCGTCCCGC CCTGCTCGAA CTCAAGCGCA2940
CCGGCCGCTA CGACGCCTTC GTCACCACGC ACAACGCGTT CATCCTGGGC GACACCGACA3000
ACGGCGAGCG CACCGGCCAC CGTTCGCCGT CCTTCCTGCC CTGGCACCGC AGATTTCTGC3060
TGGAGTTCGA GCGGGCGCTC CAGTCGGTGG ACGCGTCGGT GGCGCTGCCG TACTGGGACT3120
GGTCCGCCGA CCGGTCCACC CGGTCCTCGC TGTGGGCGCC GGACTTCCTC GCGGGCACCG3180
GGCGCAGCCG GGACGGCCAG GTGATGGACG GGCCGTTCGC CGCGTCGGCC GGCAACTGGC3240
CGATCAATGT GCGGGTGGAC GGCCGTACGT TCCTGCGGCG GGCGCTCGGC GCGGGCGTGA3300
GCGAACTGCC CACGCGTGCC GAGGTCGACT CGGTGCTGGC GATGGCGACG TACGACATGG3360
CGCCCTGGAA CAGCGGCTCC GACGGCTTCC GCAACCATCT CGAAGGGTGG CGCGGGGTCA3420
ATCTGCACAA CCGGGTG CAT GTCTGGGTG GCGGCCAGAT GGCGACCGGG GTCTCCCCCA3480
ACGACCCGGT GTTCTGGCTG CACCACGCCT ACATCGACAA GCTGTGGGCC GAGTGGCAGC3540
GGCGGCACCC CTCGTCCCCG TATCTGCCGG GCGGCGGCAC GCCGAACGTC GTCGACCTCA3600
ACGAGACGAT GAAGCCGTGG AACGACACCA CCCC GGCGGC CCTGCTGGAC CACACCCGGC3660
ACTACACCTT CGACGTctga tcatcactga cgaatcgagg tcgaggaacc gagcgtccga3720
ggaacagagg cgcttatcgg ttggccgcga gattcctgtc gatcctctcg tgcagcgcga3780
ttccgagggg aacggaaacg ttgagagact cggctctggt catcatgggg atggaaaccg3840
aggcggaaga cgctcctcg aacaggtcgg aaggcccacc cttttcgtg ccgaacagca3900
aggccagccg atccggattg tccccgagtt ccttcacgga aatgtcgcca tccgccttga3960
gcgtcatcag ATCaagcttA TCGATGATAA GCTGTCAAAC ATGAGAATT 4009

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 426
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: circular

2176236

-35-

(ii) MOLECULE TYPE:
(A) DESCRIPTION: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM:

(vii) IMMEDIATE SOURCE:
(B) CLONE:

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

MET	ALA	ARG	ILE	ALA	MET	ALA	GLU	LEU	THR	ARG	ARG	ARG	ALA	LEU	5	10	15
GLY	ALA	ALA	ALA	VAL	VAL	ALA	ALA	GLY	VAL	PRO	LEU	VAL	ALA	LEU	20	25	30
PRO	ALA	ALA	ARG	ALA	ASP	ASP	ARG	GLY	HIS	HIS	THR	PRO	GLU	VAL	35	40	45
PRO	GLY	ASN	PRO	ALA	ALA	SER	GLY	ALA	PRO	ALA	ALA	PHE	ASP	GLU	50	55	60
ILE	TYR	LYS	GLY	ARG	ARG	ILE	GLN	GLY	ARG	THR	VAL	THR	ASP	GLY	65	70	75
GLY	GLY	HIS	HIS	GLY	GLY	GLY	HIS	GLY	GLY	ASP	GLY	HIS	GLY	GLY	80	85	90
GLY	HIS	HIS	GLY	GLY	GLY	TYR	ALA	VAL	PHE	VAL	ASP	GLY	VAL	GLU	95	100	105
LEU	HIS	VAL	MET	ARG	ASN	ALA	ASP	GLY	SER	TRP	ILE	SER	VAL	VAL	110	115	120
SER	HIS	TYR	GLU	PRO	VAL	ASP	THR	PRO	ARG	ALA	ALA	ALA	ARG	ALA	125	130	135
ALA	VAL	ASP	GLU	LEU	GLN	GLY	ALA	ARG	LEU	LEU	PRO	PHE	PRO	SER	140	145	150
ASN	HIS	MET	THR	VAL	ARG	LYS	ASN	GLN	ALA	SER	LEU	THR	ALA	GLU	155	160	165
GLU	LYS	ARG	ARG	PHE	VAL	ALA	ALA	LEU	LEU	GLU	LEU	LYS	ARG	THR	170	175	180
GLY	ARG	TYR	ASP	ALA	PHE	VAL	THR	THR	HIS	ASN	ALA	PHE	ILE	LEU	185	190	195
GLY	ASP	THR	ASP	ASN	GLY	GLU	ARG	THR	GLY	HIS	ARG	SER	PRO	SER	200	205	210
PHE	LEU	PRO	TRP	HIS	ARG	ARG	PHE	LEU	LEU	GLU	PHE	GLU	ARG	ALA	215	220	225
LEU	GLN	SER	VAL	ASP	ALA	SER	VAL	ALA	LEU	PRO	TYR	TRP	ASP	TRP	230	235	240
SER	ALA	ASP	ARG	SER	THR	ARG	SER	SER	LEU	TRP	ALA	PRO	ASP	PHE	245	250	255
LEU	GLY	GLY	THR	GLY	ARG	SER	ARG	ASP	GLY	GLN	VAL	MET	ASP	GLY	260	265	270
PRO	PHE	ALA	ALA	SER	ALA	GLY	ASN	TRP	PRO	ILE	ASN	VAL	ARG	VAL	275	280	285
ASP	GLY	ARG	THR	PHE	LEU	ARG	ARG	ALA	LEU	GLY	ALA	GLY	VAL	SER	290	295	300
GLU	LEU	PRO	THR	ARG	ALA	GLU	VAL	ASP	SER	VAL	LEU	ALA	MET	ALA	305	310	315
THR	TYR	ASP	MET	ALA	PRO	TRP	ASN	SER	GLY	SER	ASP	GLY	PHE	ARG	320	325	330
ASN	HIS	LEU	GLU	GLY	TRP	ARG	GLY	VAL	ASN	LEU	HIS	ASN	ARG	VAL			

-36-

HIS VAL TRP VAL	330	GLY GLY GLN MET ALA THR	335	GLY VAL SER PRO	340
ASP PRO VAL PHE	345	TRP LEU HIS HIS ALA TYR	350	ILE ASP LYS LEU	355
ALA GLU TRP GLN	360	ARG ARG HIS PRO SER	370	SER PRO TYR LEU	375
GLY GLY THR PRO	380	ASN VAL VAL ASP LEU	385	ASN GLU THR MET	390
TRP ASN ASP THR	395	THR PRO ALA ALA LEU	400	LEU ASP HIS THR	405
TYR THR PHE ASP	410	VAL GLY	415		420
	425				

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1442
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 - (A) DESCRIPTION: DNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM:
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE:
- (ix) FEATURE:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

ctcgagccAT GGCTTCCTCA GTTCTTTCCT CTGCAGCAGT TGCCACCCGC AGCAATGTTG60
 CTCAAGCTAA CATGGTTGCA CCTTTCACCTG GCCTTAAGTC AGCTGCCTCA TTCCCTGTTT120
 CAAGGAAGCA AAACCTTGAC ATCACTTCCA TTGCCAGCAA CGGCGGAAGA GTGCAATGCA180
 TGCCGGAACT CACCCGTCGT CGCGCGCTCG GCGCCGCAGC CGTCGTCGCC GCCGGTGTCC240
 CGCTGGTTCGC CCTTCCCGCC GCCCGCGCGG ACGATCGGGG GCACCACACC CCCGAGGTCC300
 CCGGGAACCC GGCCGCGTCC GCGCCCCCG CCGCCTTCGA CGAGATCTAC AAGGGCCGCC360
 GGATACAGGG CCGGACGGTC ACCGACGGCG GGGGCCACCA CGGCGGCGGT CACGGCGGTG420
 ACGGTCACGG CGGCGGCCAT CACGGCGGCG GTTACGCCGT GTTCGTGGAC GGCGTCGAAC480
 TGCATGTGAT GCGCAACGCC GACGGCTCGT GGATCAGCGT CGTCAGCCAC TACGAGCCGG540
 TGGACACCCC GCGCGCCGCG GCCCGCGCTG CGGTGACGA GCTCCAGGGC GCCCGGCTCC600
 TCCCCTTCCC CTCCAACCAT ATGACCGTCC GCAAGAACCA GGCGTCCCTG ACCGCCGAGG660
 AGAAGCGCCG CTTCGTCGCC GCCCTGCTCG AACTCAAGCG CACCGGCCGC TACGACGCCT720

-37-

TCGTCACCAC GCACAACGCG TTCATCCTGG GCGACACCGA CAACGGCGAG CGCACC GGCC780
 ACCGTTTCGCC GTCCTTCCTG CCCTGGCACC GCAGATTCT GCTGGAGTTC GAGCGGGCGC840
 TCCAGTCGGT GGACGCGTCG GTGGCGCTGC CGTACTGGGA CTGGTCCGCC GACCGGTCCA900
 CCCGGTCCTC GCTGTGGGCG CCGGACTTCC TCGGCGGCAC CGGGCGCAGC CGGGACGGCC960
 AGGTGATGGA CGGGCCGTTT GCCGCGTCGG CCGGCAACTG GCCGATCAAT GTGCGGGTGG1020
 ACGGCCGTAC GTTCCTGCGG CGGGCGCTCG GCGCGGGCGT GAGCGAACTG CCCACGCGTG1080
 CCCAGGTCGA CTCGGTGCTG GCGATGGCGA CGTACGACAT GGCGCCCTGG AACAGCGGCT1140
 CCGACGGCTT CCGCAACCAT CTCGAAGGGT GGCGCGGGGT CAATCTGCAC AACCGGGTGC1200
 ATGTCTGGGT CGGCGGCCAG ATGGCGACCG GGGTCTCCCC CAACGACCCG GTGTTCTGGC1260
 TGCACCACGC CTACATCGAC AAGCTGTGGG CCGAGTGGCA GCGGCGGCAC CCCTCGTCCC1320
 CGTATCTGCC GGGCGGCGGC ACGCCGAACG TCGTCGACCT CAACGAGACG ATGAAGCCGT1380
 GGAACCACAC CACCCCGGCG GCCCTGCTGG ACCACACCCG GCACTACACC TTCGACGTCT1440
 GA

1442

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 478
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 - (A) DESCRIPTION: DNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM:
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE:
- (ix) FEATURE:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Met	Ala	Ser	Ser	Val	Leu	Ser	Ser	Ala	Ala	Val	Ala	Thr	Arg	Ser	5	10	15
Asn	Val	Ala	Gln	Ala	Asn	Met	Val	Ala	Pro	Phe	Thr	Gly	Leu	Lys	20	25	30
Ser	Ala	Ala	Ser	Phe	Pro	Val	Ser	Arg	Lys	Gln	Asn	Leu	Asp	Ile	35	40	45
Thr	Ser	Ile	Ala	Ser	Asn	Gly	Gly	Arg	Val	Gln	Cys	Met	Pro	Glu	50	55	60
Leu	Thr	Arg	Arg	Arg	Ala	Leu	Gly	Ala	Ala	Ala	Val	Val	Ala	Ala	65	70	75

-38-

Gly Val Pro Leu Val Ala Leu Pro Ala Ala Arg Ala Asp Asp Arg	80	85	90
Gly His His Thr Pro Glu Val Pro Gly Asn Pro Ala Ala Ser Gly	95	100	105
Ala Pro Ala Ala Phe Asp Glu Ile Tyr Lys Gly Arg Arg Ile Gln	110	115	120
Gly Arg Thr Val Thr Asp Gly Gly Gly Asp His Gly Gly Gly His	125	130	135
Gly Gly Asp Gly His Gly Gly Gly His His Gly Gly Gly Tyr Ala	140	145	150
Val Phe Val Asp Gly Val Glu Leu His Val Met Arg Asn Ala Asp	155	160	165
Gly Ser Trp Ile Ser Val Val Ser His Tyr Glu Pro Val Asp Thr	170	175	180
Pro Arg Ala Ala Ala Arg Ala Ala Val Asp Glu Leu Gln Gly Ala	185	190	195
Arg Leu Leu Pro Phe Pro Ser Asn His Met Thr Val Arg Lys Asn	200	205	210
Gln Ala Ser Leu Thr Ala Glu Glu Lys Arg Arg Phe Val Ala Ala	215	220	225
Leu Leu Glu Leu Lys Arg Thr Gly Arg Tyr Asp Ala Phe Val Thr	230	235	240
Thr His Asn Ala Phe Ile Leu Gly Asp Thr Asp Asn Gly Glu Arg	245	250	255
Thr Gly His Arg Ser Pro Ser Phe Leu Pro Trp His Arg Arg Phe	260	265	270
Leu Leu Glu Phe Glu Arg Ala Leu Gln Ser Val Asp Ala Ser Val	275	280	285
Ala Leu Pro Tyr Trp Asp Trp Ser Ala Asp Arg Ser Thr Arg Ser	290	295	300
Ser Leu Trp Ala Pro Asp Phe Leu Gly Gly Thr Gly Arg Ser Arg	305	310	315
Asp Gly Gln Val Met Asp Gly Pro Phe Ala Ala Ser Ala Gly Asn	320	325	330
Trp Pro Ile Asn Val Arg Val Asp Gly Arg Thr Phe Leu Arg Arg	335	340	345
Ala Leu Gly Ala Gly Val Ser Glu Leu Pro Thr Arg Ala Glu Val	350	355	360
Asp Ser Val Leu Ala Met Ala Thr Tyr Asp Met Ala Pro Trp Asn	365	370	375
Ser Gly Ser Asp Gly Phe Arg Asn His Leu Glu Gly Trp Arg Gly	380	385	390
Val Asn Leu His Asn Arg Val His Val Trp Val Gly Gly Gln Met	395	400	405
Ala Thr Gly Val Ser Pro Asn Asp Pro Val Phe Trp Leu His His	410	415	420
Ala Tyr Ile Asp Lys Leu Trp Ala Glu Trp Gln Arg Arg His Pro	425	430	435
Ser Ser Pro Tyr Leu Pro Gly Gly Gly Thr Pro Asn Val Val Asp	440	445	450
Leu Asn Glu Thr Met Lys Pro Trp Asn Asp Thr Thr Pro Ala Ala	455	460	465
Leu Leu Asp His Thr Arg His Tyr Thr Phe Asp Val Gly	470	475	

- 39 -

International Application No: PCT/

/

MICROORGANISMS	
Optional Sheet in connection with the microorganism referred to on page __, lines __ of the description *	
A. IDENTIFICATION OF DEPOSIT * Further deposits are identified on an additional sheet *	
Name of depositary institution * American Type Culture Collection	
Address of depositary institution (including postal code and country) * 12301 Parklawn Drive Rockville, MD 20852 US	
Date of deposit * <u>December 28, 1993</u> Accession Number * <u>75632</u>	
B. ADDITIONAL INDICATIONS * (leave blank if not applicable). This information is continued on a separate attached sheet	
C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE * (if the indications are not all designated States)	
D. SEPARATE FURNISHING OF INDICATIONS * (leave blank if not applicable) The indications listed below will be submitted to the International Bureau later * (Specify the general nature of the indications e.g., "Accession Number of Deposit")	
E. ☐ This sheet was received with the International application when filed (to be checked by the receiving Office) _____ (Authorized Officer) ☐ The date of receipt (from the applicant) by the International Bureau * was _____ (Authorized Officer)	

- 40 -

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

- 5 1. A nucleic acid sequence encoding a fusion enzyme comprising a nucleic acid sequence encoding for a tyrosinase and a nucleic acid sequence encoding for a tyrosinase activator protein positioned 5' relative to the tyrosinase sequence.
- 10 2. A nucleic acid sequence encoding a fusion enzyme according to claim 1 wherein the tyrosinase is derived from a prokaryote.
- 15 3. A nucleic acid sequence encoding a fusion enzyme according to claim 2 wherein the tyrosinase is derived from Streptomyces.
- 20 4. A nucleic acid sequence encoding a fusion enzyme according to claim 1 wherein the activator protein is selected from the group consisting of ORF438 and URF402.
- 25 5. A nucleic acid sequence encoding a fusion enzyme according to claim 1 wherein the activator protein is ORF438 and the tyrosinase is derived from S.antibioticus.
- 30 6. A nucleic acid sequence encoding a fusion enzyme according to claim 1 which further comprises a linker sequence positioned between the nucleic acid sequences encoding the activator protein and the tyrosinase.
- 35 7. A nucleic acid sequence encoding a fusion enzyme according to claim 6 wherein linker sequence encodes for a single amino acid.
8. A nucleic acid sequence encoding a fusion enzyme according to claim 7 wherein linker sequence encodes for a histidine residue.

- 41 -

9. A nucleic acid sequence encoding a fusion enzyme according to claim 7 wherein linker sequence encodes for a repeating Pro-Thr amino acid sequence.
- 5 10. A nucleic acid sequence encoding a fusion enzyme according to claim 6 wherein the activator protein is selected from the group consisting of ORF438 and URF402.
- 10 11. A nucleic acid sequence encoding a fusion enzyme according to claim 6 wherein the tyrosinase is derived from a prokaryote.
- 15 12. A nucleic acid sequence encoding a fusion enzyme according to claim 11 wherein the tyrosinase is derived from Streptomyces.
13. A vector for transforming an organism, said vector comprising:
a nucleic acid sequence encoding for a fusion enzyme,
20 said fusion enzyme comprising an amino acid sequence encoding for a tyrosinase and an amino acid sequence encoding for a tyrosinase activator protein that is positioned on the N-terminus end of the tyrosinase sequence; and
25 a promoter sequence that regulates the transcription of fusion enzyme.
- 30 14. A vector according to claim 13 wherein the activator protein is selected from the group consisting of ORF438 and URF402.
15. A vector according to claim 13 wherein the tyrosinase is derived from a prokaryote.
- 35 16. A vector according to claim 15 wherein the tyrosinase is derived from Streptomyces.

- 42 -

17. A vector according to claim 13 wherein the activator protein is ORF438 and the tyrosinase is derived from S.antibioticus.

5

18. A vector according to claim 13 which further comprises a linker sequence positioned between the nucleic acid sequences encoding the activator protein and the tyrosinase.

10

19. A vector according to claim 18 wherein linker sequence encodes for a single amino acid.

15

20. A vector according to claim 19 wherein linker sequence encodes for a histidine residue.

21. A vector according to claim 18 wherein linker sequence encodes for a repeating Pro-Thr amino acid sequence.

20

22. A vector according to claim 13 wherein the activator protein is selected from the group consisting of ORF438 and URF402.

25

23. A vector according to claim 13 wherein the tyrosinase is derived from S. antibioticus.

30

24. A fusion enzyme comprising an amino acid sequence for tyrosinase and an amino acid sequence for a tyrosinase activator protein that is positioned on the N-terminus end of the tyrosinase sequence.

25. A fusion enzyme according to claim 24 wherein the activator protein is selected from the group consisting of ORF438 and URF402.

35

26. A fusion enzyme according to claim 24 wherein the tyrosinase is derived from a prokaryote.

- 43 -

27. A fusion enzyme according to claim 26 wherein the tyrosinase is derived from Streptomyces.

5 28. A fusion enzyme according to claim 24 wherein the activator protein is ORF438 and the tyrosinase is derived from S. antibioticus.

10 29. A fusion enzyme according to claim 24 which further comprises a linker positioned between the amino acid sequences of the activator protein and the tyrosinase.

15 30. A fusion enzyme according to claim 29 wherein the linker sequence comprises a single amino acid.

31. A fusion enzyme according to claim 30 wherein the linker sequence comprises a histidine residue.

20 32. A fusion enzyme according to claim 29 wherein the linker sequence comprises a repeating Pro-Thr amino acid sequence.

25 33. An *in vitro* method of producing a melanin comprising the steps of contacting
a fusion enzyme comprising an amino acid sequence for tyrosinase and an amino acid sequence for a tyrosinase activator protein that is positioned on the N-terminus end of the tyrosinase sequence and a reaction substrate selected from the group consisting of L-tyrosine,
30 X/L-tyrosine, L-tyrosine/X and X/L-tyrosine/X where X is a single amino acid, a dipeptide or an oligopeptide bound to L-tyrosine, under suitable reaction conditions, to form melanin.

- 44 -

34. A method of producing melanins, comprising the step of growing an organism which has been transformed with a vector comprising

- 5 a nucleic acid sequence encoding for a fusion enzyme, said fusion enzyme comprising an amino acid sequence encoding for a tyrosinase and an amino acid sequence encoding for a tyrosinase activator protein that is positioned on the N-terminus end of the tyrosinase
10 sequence; and
a promoter sequence that regulates the transcription of fusion enzyme.

- 15 35. A method according to claim 34 comprising the further step of isolating the melanin produced.

36. The method of claim 34 wherein the organism is prokaryotic.

- 20 37. The method of claim 36 wherein said prokaryotic organism is selected from the group consisting of *Escherichia coli* and *Streptomyces*.

38. A nucleic acid sequence encoding a fusion enzyme comprising

- 25 a nucleic acid sequence encoding a tyrosinase with cresolase activity and catecholase activity; and
a nucleic acid encoding a tyrosinase activator protein that is positioned 5' relative to the nucleic acid sequence encoding the tyrosinase with cresolase activity and catecholase activity.

- 30 39. The nucleic acid sequence encoding a fusion enzyme of claim 38 further comprising a linker sequence positioned between said nucleic acid sequence encoding a tyrosinase with cresolase activity and catecholase activity and said nucleotide sequence encoding a tyrosinase activator
35 protein.

WO 95/13386

PCT/US94/12857

- 45 -

40. The nucleic acid sequence encoding a fusion enzyme of claim 39 wherein said linker sequence encodes for a His residue.

5 41. An amino acid sequence for a fusion enzyme comprising:
an amino acid sequence for a tyrosinase; and
an amino acid sequence for a tyrosinase activator
10 protein that is positioned on the N-terminus end of the amino acid sequence for the tyrosinase.

15 42. The amino acid sequence of claim 41 further comprising a linker amino acid sequence positioned between said amino acid sequence for a tyrosinase and said amino acid sequence for a tyrosinase activator protein.

43. The amino acid sequence encoding a fusion enzyme of claim 42 wherein said linker amino acid sequence comprises a His residue.

2176236

1/7

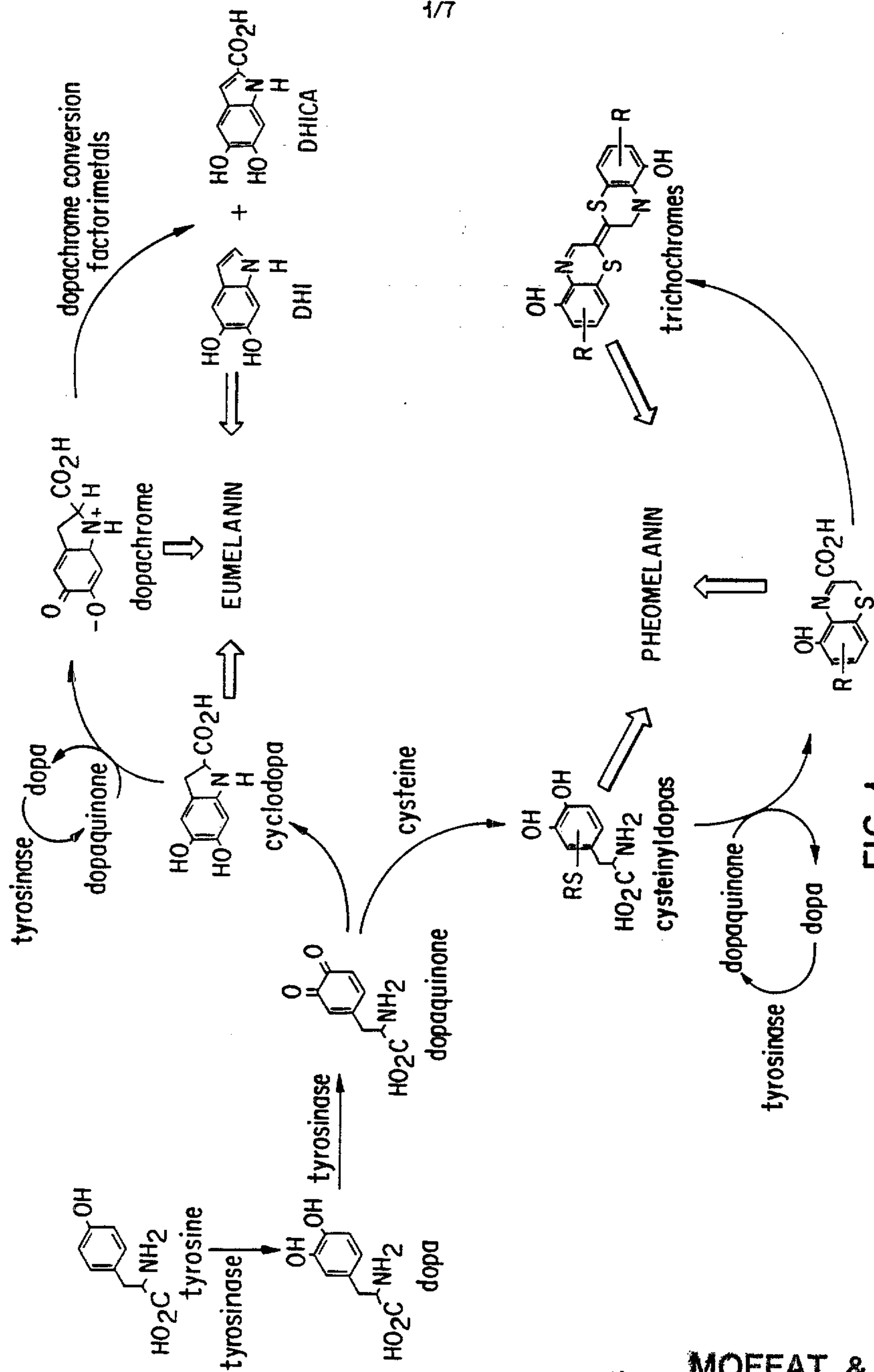


FIG. 1

2/7

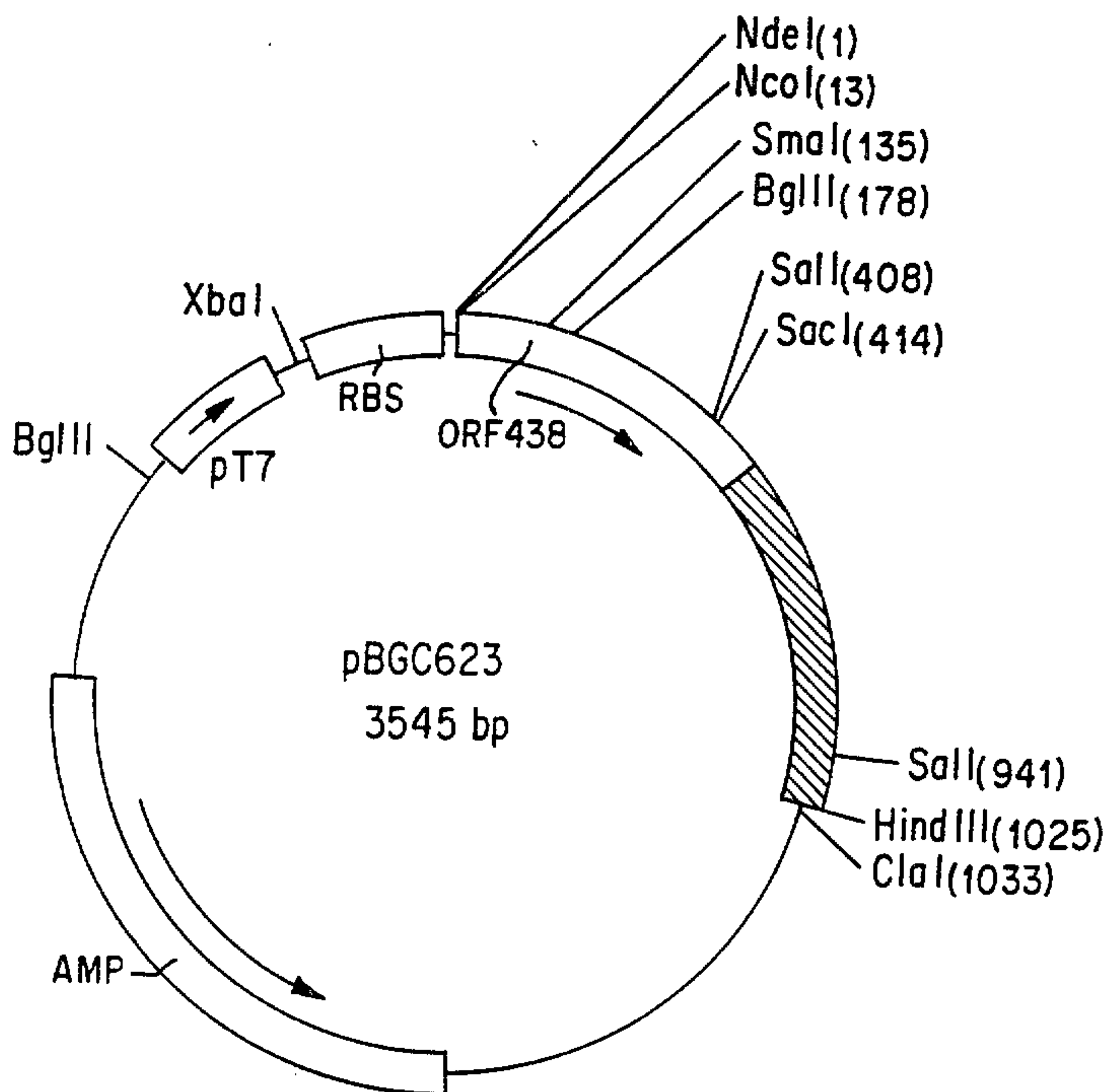


FIG. 2

2176236

3/7

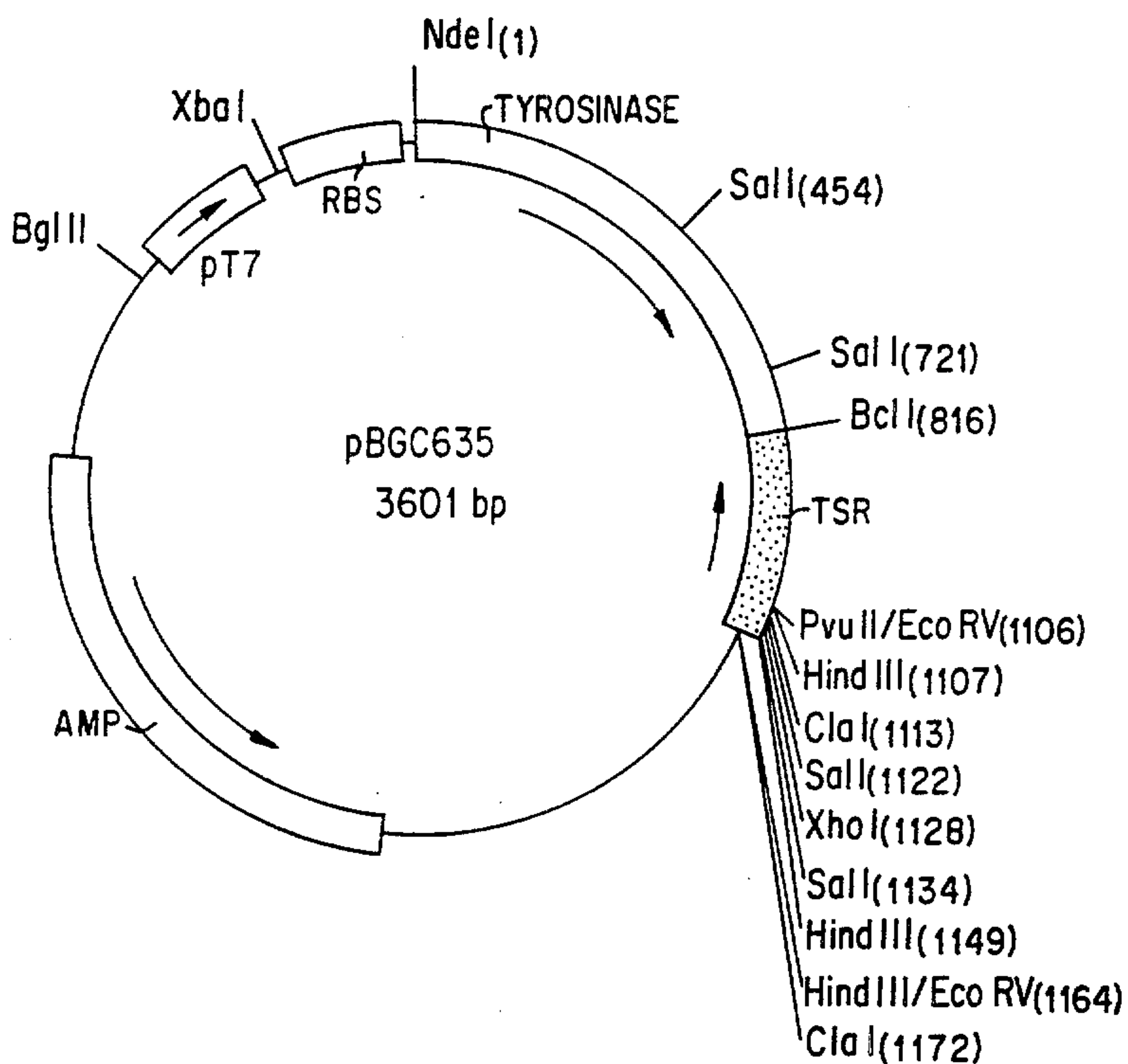


FIG. 3

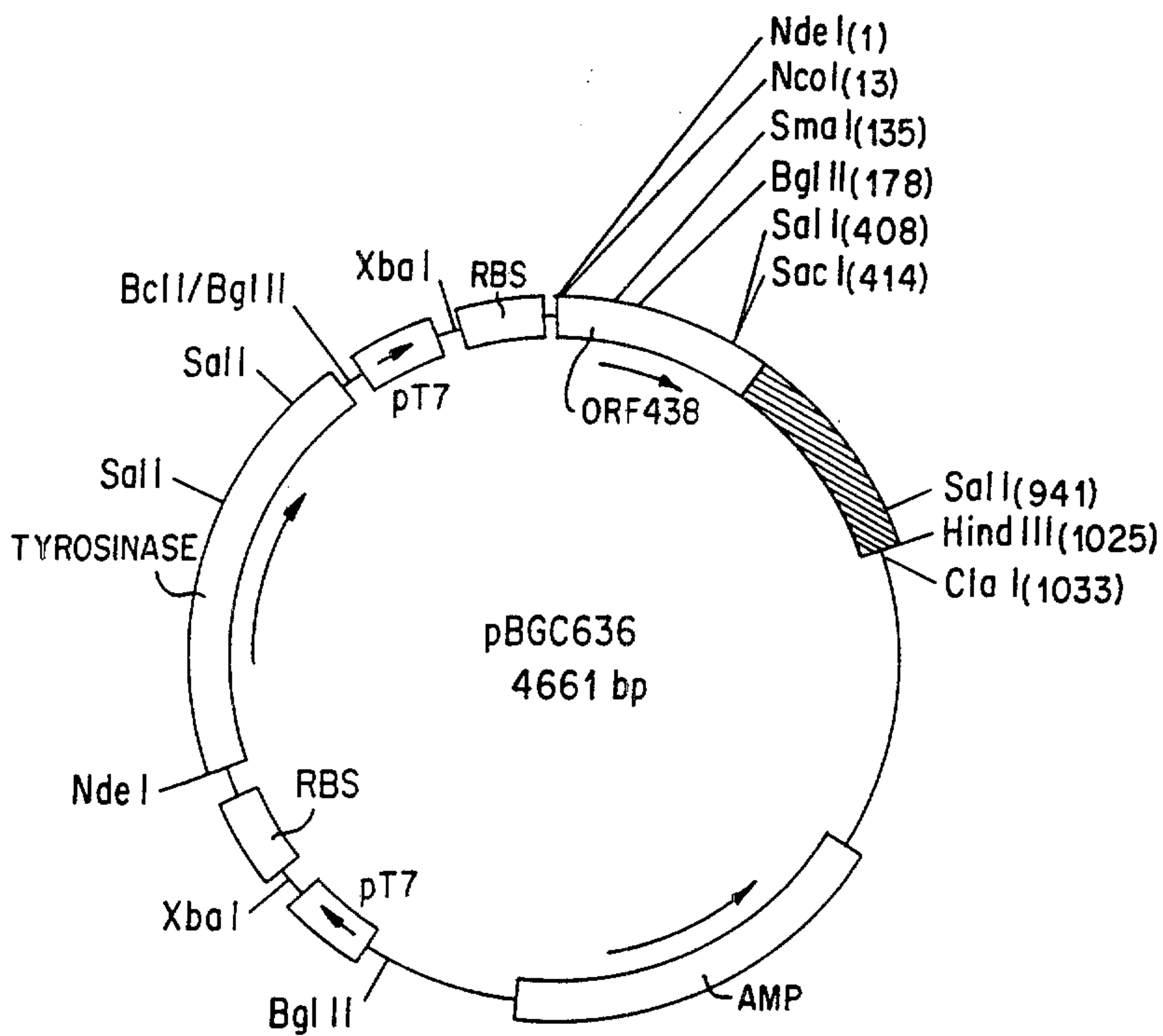


FIG. 4

5/7

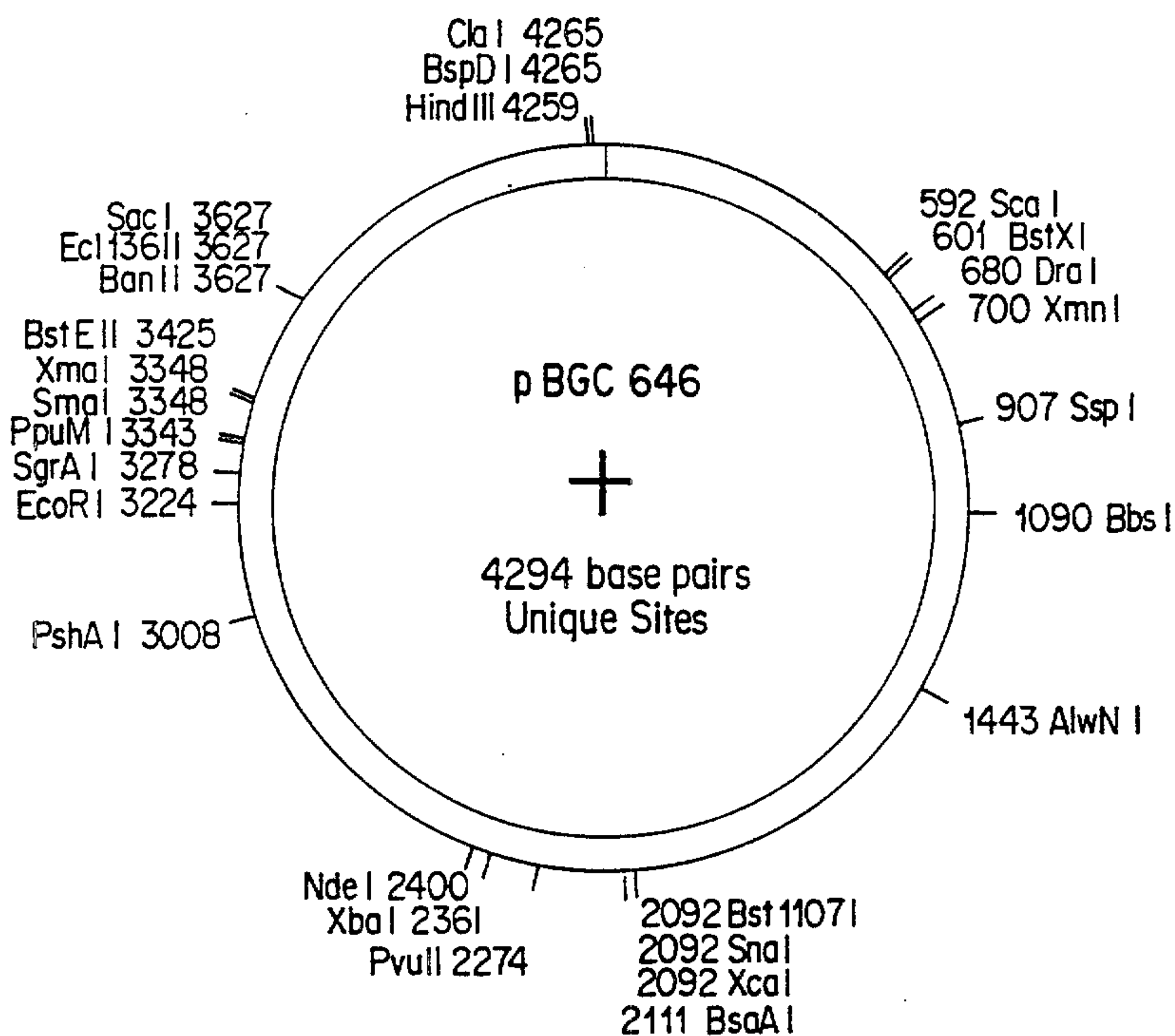


FIG. 5

6/7

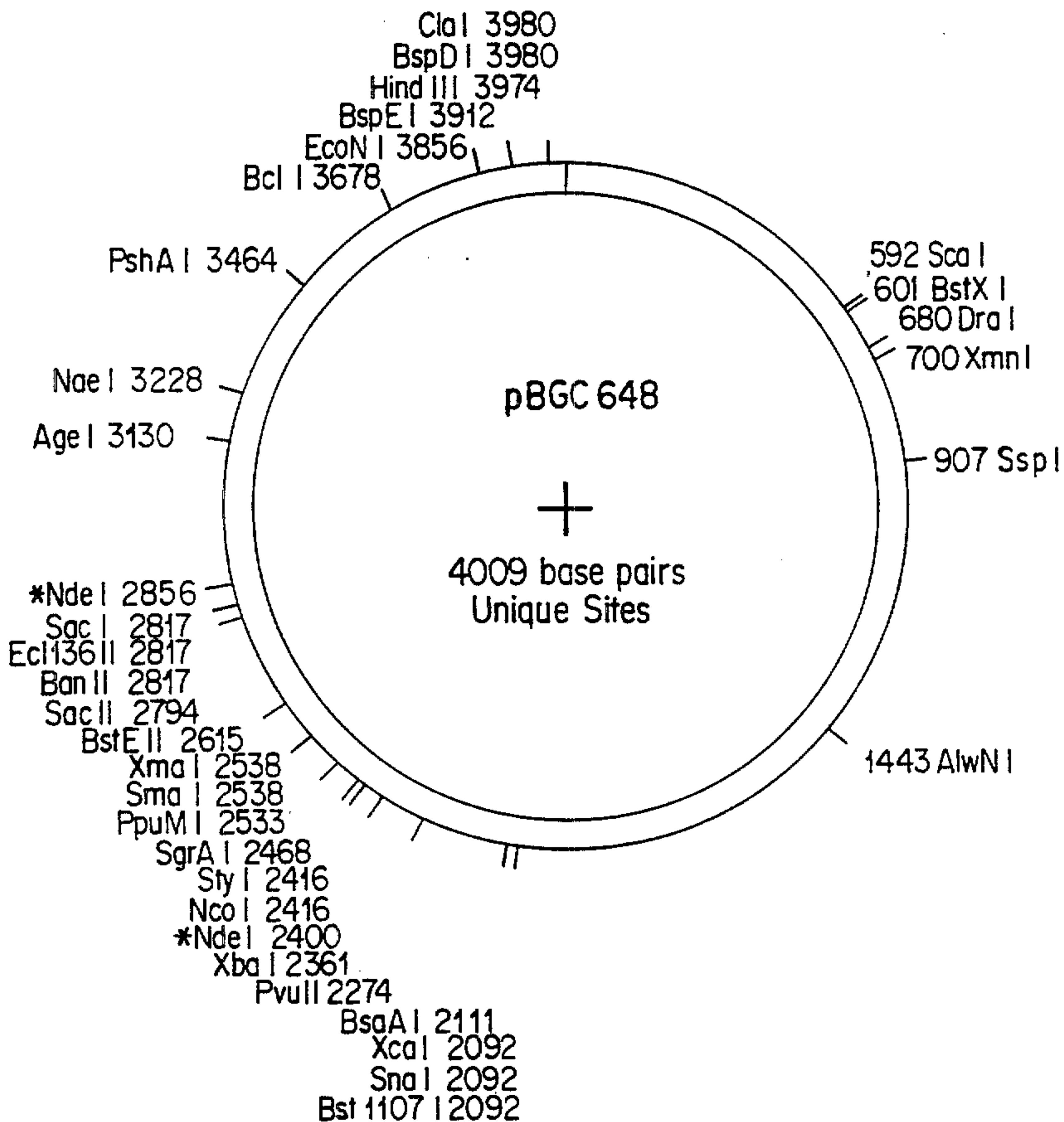


FIG. 6

Melanin Phenotype

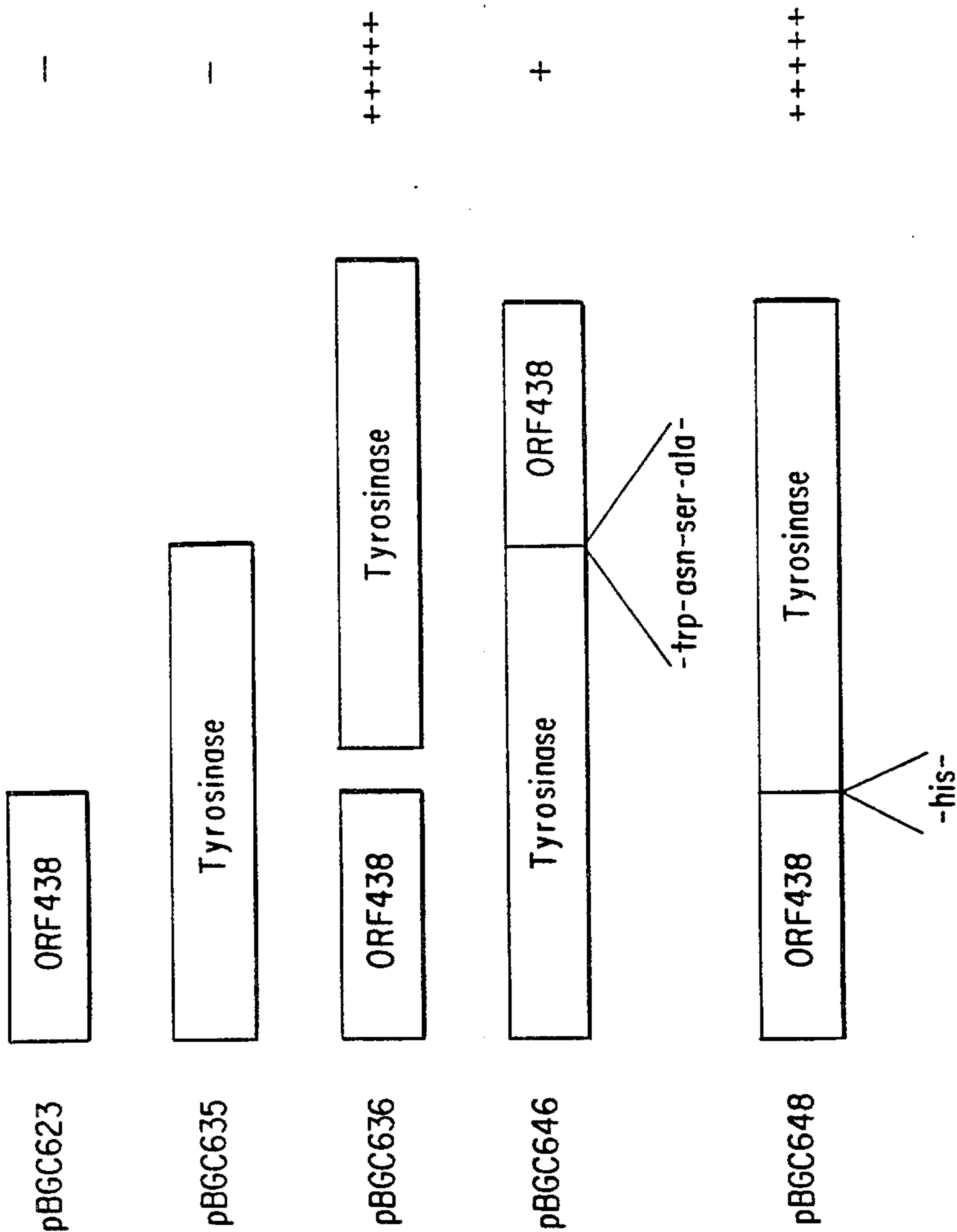


FIG. 7