Title: HERBICIDE-TOLENT PRETORPORPHYRINOGEN OXIDASE

Abstract: The present invention provides novel DNA sequences coding for protoporphyrinogen oxidase (protox) enzymes from soybeans, wheat, cotton, sugar beet, oilseeds rape, rice, sorghum, and sugar cane. In addition, the present invention teaches modified forms of protox enzymes that are herbicide tolerant. Plants expressing herbicide tolerant protox enzymes taught herein are also provided. These plants may be engineered for resistance to protox inhibitors via mutation of the native protox gene to a resistant form or they may be transformed with a gene encoding a herbicide tolerant form of a plant protox enzyme.
HERBICIDE-TOLERANT PROTOPORPHYRINOGEN OXIDASE

The present invention relates to DNA molecules encoding herbicide-tolerant forms of the enzyme protoporphyrinogen oxidase ("protox"). The invention further relates to herbicide-tolerant plants as well as methods for tissue culture selection and herbicide application based on these herbicide-tolerant forms of protox.

The present invention provides modified forms of plant protoporphyrinogen oxidase (protox) enzymes that are resistant to compounds that inhibit unmodified naturally occurring plant protox enzymes, and DNA molecules coding for such inhibitor-resistant plant protox enzymes.

For clarity, certain terms used in the specification are defined and presented as follows:

Associated With / Operatively Linked: refers to two DNA sequences that are related physically or functionally. For example, a promoter or regulatory DNA sequence is said to be "associated with" a DNA sequence that codes for an RNA or a protein if the two sequences are operatively linked, or situated such that the regulator DNA sequence will affect the expression level of the coding or structural DNA sequence.

Chimeric Gene: a recombinant DNA sequence in which a promoter or regulatory DNA sequence is operatively linked to, or associated with, a DNA sequence that codes for an mRNA or which is expressed as a protein, such that the regulator DNA sequence is able to regulate transcription or expression of the associated DNA sequence. The regulator DNA sequence of the chimeric gene is not normally operatively linked to the associated DNA sequence as found in nature.

Coding DNA Sequence: a DNA sequence that is translated in an organism to produce a protein.

Corresponding To: in the context of the present invention, "corresponding to" means that when the amino acid sequences of various protox enzymes are aligned with each other, such as in Table 1A, the amino acids that "correspond to" certain enumerated positions in Table 1A are those that align with these positions in Table 1A, but that are not necessarily in these exact numerical positions relative to the particular protox enzyme's amino acid sequence. Likewise, when the amino acid sequence of a particular protox enzyme (for example, the soybean protox enzyme) is aligned with the amino acid sequence
of a reference protox enzyme (for example, the Arabidopsis protox-1 sequence given in SEQ ID NO:2), the amino acids in the soybean protox sequence that "correspond to" certain enumerated positions of SEQ ID NO:2 are those that align with these positions of SEQ ID NO:2, but are not necessarily in these exact numerical positions of the soybean protox enzyme's amino acid sequence.

Herbicide: a chemical substance used to kill or suppress the growth of plants, plant cells, plant seeds, or plant tissues.

Heterologous DNA Sequence: a DNA sequence not naturally associated with a host cell into which it is introduced, including non-naturally occurring multiple copies of a naturally occurring DNA sequence.

Homologous DNA Sequence: a DNA sequence naturally associated with a host cell into which it is introduced.

Homoplasmonic: refers to a plant, plant tissue or plant cell, wherein all of the plastids are genetically identical. In different tissues or stages of development, the plastids may take different forms, e.g., chloroplasts, proplastids, etioplasts, amyloplasts, chromoplasts, and so forth.

Inhibitor: a chemical substance that inactivates the enzymatic activity of a protein such as a biosynthetic enzyme, receptor, signal transduction protein, structural gene product, or transport protein that is essential to the growth or survival of the plant. In the context of the instant invention, an inhibitor is a chemical substance that inactivates the enzymatic activity of protox. The term "herbicide" is used herein to define an inhibitor when applied to plants, plant cells, plant seeds, or plant tissues.

Isolated: in the context of the present invention, an isolated nucleic acid molecule or an isolated enzyme is a nucleic acid molecule or enzyme that, by the hand of man, exists apart from its native environment and is therefore not a product of nature. An isolated nucleic acid molecule or enzyme may exist in a purified form or may exist in a non-native environment such as, for example, a transgenic host cell.

Minimal Promoter: promoter elements, particularly a TATA element, that are inactive or that have greatly reduced promoter activity in the absence of upstream activation. In the presence of a suitable transcription factor, the minimal promoter functions to permit transcription.

Modified Enzyme Activity: enzyme activity different from that which naturally occurs in a plant (i.e. enzyme activity that occurs naturally in the absence of direct or indirect
manipulation of such activity by man), which is tolerant to inhibitors that inhibit the naturally occurring enzyme activity.

**Nucleic Acid Molecule:** a linear segment of single- or double-stranded DNA or RNA that can be isolated from any source. In the context of the present invention, the nucleic acid molecule is preferably a segment of DNA.

**Plant:** refers to any plant or part of a plant at any stage of development. Therein are also included cuttings, cell or tissue cultures and seeds. As used in conjunction with the present invention, the term "plant tissue" includes, but is not limited to, whole plants, plant cells, plant organs, plant seeds, protoplasts, callus, cell cultures, and any groups of plant cells organized into structural and/or functional units.

**Plastome:** the genome of a plastid.

**Protox-1:** chloroplast protox.

**Protox-2:** mitochondrial protox.

**Significant Increase:** an increase in enzymatic activity that is larger than the margin of error inherent in the measurement technique, preferably an increase by about 2-fold or greater of the activity of the wild-type enzyme in the presence of the inhibitor, more preferably an increase by about 5-fold or greater, and most preferably an increase by about 10-fold or greater.

**Substantially Similar:** with respect to nucleic acids, a nucleic acid molecule that has at least 60 percent sequence identity with a reference nucleic acid molecule. In a preferred embodiment, a substantially similar DNA sequence is at least 80% identical to a reference DNA sequence; in a more preferred embodiment, a substantially similar DNA sequence is at least 90% identical to a reference DNA sequence; and in a most preferred embodiment, a substantially similar DNA sequence is at least 95% identical to a reference DNA sequence. A substantially similar nucleotide sequence typically hybridizes to a reference nucleic acid molecule, or fragments thereof, under the following conditions: hybridization at 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50°C; wash with 2X SSC, 1% SDS, at 50°C. With respect to proteins or peptides, a substantially similar amino acid sequence is an amino acid sequence that is at least 90% identical to the amino acid sequence of a reference protein or peptide and has substantially the same activity as the reference protein or peptide.

**Tolerance / Resistance:** the ability to continue normal growth or function when exposed to an inhibitor or herbicide.
Transformation: a process for introducing heterologous DNA into a cell, tissue, or plant. Transformed cells, tissues, or plants are understood to encompass not only the end product of a transformation process, but also transgenic progeny thereof.

Transit Peptide: a signal polypeptide that is translated in conjunction with a protein encoded by a DNA molecule, forming a polypeptide precursor. In the process of transport to a selected site within the cell, a chloroplast for example, the transit peptide can be cleaved from the remainder of the polypeptide precursor to provide an active or mature protein.

Transformed: refers to an organism such as a plant into which a heterologous DNA molecule has been introduced. The DNA molecule can be stably integrated into the genome of the plant, wherein the genome of the plant encompasses the nuclear genome, the plastid genome and the mitochondrial genome. In a transformed plant, the DNA molecule can also be present as an extrachromosomal molecule. Such an extrachromosomal molecule can be auto-replicating. A “non-transformed” plant refers to a wild-type organism, i.e., a plant, which does not contain the heterologous DNA molecule.

Transplastome: a transformed plastid genome.

Nucleotides are indicated by their bases by the following standard abbreviations: adenine (A), cytosine (C), thymine (T), and guanine (G). Amino acids are likewise indicated by the following standard abbreviations: alanine (ala; A), arginine (Arg; R), asparagine (Asn; N), aspartic acid (Asp; D), cysteine (Cys; C), glutamine (Gln; Q), glutamic acid (Glu; E), glycine (Gly; G), histidine (His; H), isoleucine (Ile; I), leucine (Leu; L), lysine (lys; K), methionine (Met; M), phenylalanine (Phe; F), proline (Pro; P), serine (Ser; S), threonine (Thr; T), tryptophan (Trp; W), tyrosine (Tyr; Y), and valine (Val; V). Furthermore (Xaa; X) represents any amino acid.

The present invention discloses modified forms of plant protoporphyrinogen oxidase (protox) enzymes that are resistant to compounds that inhibit unmodified naturally occurring plant protox enzymes, and DNA molecules coding for such inhibitor-resistant plant protox enzymes.

Thus, in one aspect the present invention discloses a DNA molecule encoding a plant protox enzyme that is capable of being incorporated into a DNA construct used to transform a plant containing wild-type, herbicide-sensitive protox, wherein the DNA molecule has at least one point mutation relative to a wild-type DNA molecule encoding plant protox such that upon transformation with the DNA construct the plant contains the DNA molecule, which renders the plant resistant to the application of a herbicide that inhibits naturally occurring plant protox.
In particular, the present invention relates to a nucleic acid molecule comprising a nucleotide sequence that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein said modified enzyme is resistant to an inhibitor of the naturally occurring form of said enzyme, wherein said modified enzyme comprises at least one of the following amino acid sub-sequences:

(a) $\text{KA} \Delta_{18}\text{F}$, wherein $\Delta_{18}$ is an amino acid other than alanine,
(b) $\text{Q} \Delta_{19}\text{H}$, wherein $\Delta_{19}$ is an amino acid other than leucine,
(c) $\text{AP} \Delta_{1}\text{F}$, wherein $\Delta_{1}$ is an amino acid other than arginine, but especially leucine or cysteine,
(d) $\text{F} \Delta_{2}\text{S}$, wherein $\Delta_{2}$ is leucine,
(e) $\text{Y} \Delta_{3}\text{G}$, wherein $\Delta_{3}$ is isoleucine,
(f) $\Delta_{7}\text{IG}$, wherein $\Delta_{7}$ is histidine or alanine,
(g) $\text{T} \Delta_{16}\text{G}$, wherein $\Delta_{16}$ is an amino acid other than leucine and $\text{YV} \Delta_{17}\text{G}$, wherein $\Delta_{17}$ is an amino acid other than alanine,

wherein the nucleotide sequence that encodes the naturally occurring form of said enzyme is derived from a plant.

In a preferred embodiment the invention relates to a nucleic acid molecule comprising a nucleotide sequence that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein said modified enzyme is resistant to an inhibitor of the naturally occurring form of said enzyme, wherein said modified enzyme comprises at least one amino acid sub-sequence selected from the group consisting of:

(a) $\text{KA} \Delta_{18}\text{F}$, wherein $\Delta_{18}$ is an amino acid other than alanine,
(b) $\text{Q} \Delta_{19}\text{H}$, wherein $\Delta_{19}$ is an amino acid other than leucine,
(c) $\text{AP} \Delta_{1}\text{F}$, wherein $\Delta_{1}$ is an amino acid other than arginine, but especially leucine,
(d) $\text{F} \Delta_{2}\text{S}$, wherein $\Delta_{2}$ is leucine,
(e) $\text{Y} \Delta_{3}\text{G}$, wherein $\Delta_{3}$ is isoleucine,
(f) $\Delta_{7}\text{IG}$, wherein $\Delta_{7}$ is histidine or alanine,
(g) $\text{T} \Delta_{16}\text{G}$, wherein $\Delta_{16}$ is an amino acid other than leucine and $\text{YV} \Delta_{17}\text{G}$, wherein $\Delta_{17}$ is an amino acid other than alanine,

wherein said nucleotide sequence that encodes said modified enzyme hybridizes to any one of the nucleotide sequences set forth in SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23 or SEQ ID NO:36 under the following conditions:
(i) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO₄ pH 7.0, 1 mM EDTA at 50°C; and
(ii) wash in 2X SSC, 1% SDS at 50°C.

Further preferred is a nucleic acid molecule comprising a nucleotide sequence that encodes a modified enzyme having protoporphyrogen oxidase (protox) activity, wherein said modified enzyme is resistant to an inhibitor of the naturally occurring form of said enzyme, wherein said modified enzyme comprises at least one of the following amino acid sub-sequences:

(a) KΔ₁₁₁₁, wherein Δ₁₁₁₁ is an amino acid other than alanine,
(b) QΔ₁₉₁₉, wherein Δ₁₉₁₉ is an amino acid other than leucine,
(c) APΔ₁₁, wherein Δ₁ is leucine,
(d) Δ₁₇₁₁G, wherein Δ₁₇ is histidine,

wherein the nucleotide sequence that encodes the naturally occurring form of said enzyme is derived from a plant.

Preferred is a nucleic acid molecule, wherein said modified enzyme comprises the amino acid sub-sequence KΔ₁₁₁₁, wherein Δ₁₁₁₁ is an amino acid other than alanine.

Further preferred is a nucleic acid molecule, wherein Δ₁₁₁₁ is threonine or valine.

Also preferred is a nucleic acid molecule, wherein said modified enzyme comprises the amino acid sub-sequence QΔ₁₉₁₉, wherein Δ₁₉₁₉ is an amino acid other than leucine.

Also preferred is a nucleic acid molecule, wherein Δ₁₉₁₉ is serine.

Also preferred is a nucleic acid molecule, wherein said modified enzyme comprises the amino acid sub-sequence APΔ₁₁, wherein Δ₁₁ is leucine or cysteine.

Also preferred is a nucleic acid molecule, wherein said modified enzyme comprises the amino acid sub-sequence Δ₁₇₁₁G, wherein Δ₁₇ is histidine or alanine, but especially histidine.

Also preferred is a nucleic acid molecule, wherein said modified enzyme comprises the amino acid sub-sequence TΔ₁₈₁₈G, wherein Δ₁₈₁₈ is serine, and the amino acid sub-sequence YVΔ₁₇₁₇G, wherein Δ₁₇₁₇ is threonine.

The invention further provides a nucleic acid molecule, wherein modified enzyme according to the invention further comprises at least one additional amino acid sub-sequence selected from the group consisting of:

(e) QΔ₁₁₁₁, wherein Δ₁₁₁₁ is an amino acid other than proline;
(f) IGGAΔ₁₂, wherein Δ₁₂ is an amino acid other than threonine;
(g) SWXLΔ₁₃, wherein Δ₁₃ is an amino acid other than serine;
(h) \( \Delta_{14} Y \), wherein \( \Delta_{14} \) is an amino acid other than asparagine;

(i) \( \Delta_{15} XGL \), wherein \( \Delta_{15} \) is an amino acid other than tyrosine.

Preferred is a nucleic acid molecule, wherein said additional sub-sequence is \( Q\Delta_{11} S \), wherein \( \Delta_{11} \) is an amino acid other than proline.

Further preferred is a nucleic acid molecule, wherein \( \Delta_{11} \) is leucine.

Also preferred is a nucleic acid molecule, wherein said additional sub-sequence is \( IGGA\Delta_{12} \), wherein \( \Delta_{12} \) is an amino acid other than threonine.

Also preferred is a nucleic acid molecule, wherein \( \Delta_{12} \) is isoleucine or alanine.

Also preferred is a nucleic acid molecule, wherein said additional sub-sequence is \( SWXL\Delta_{13} \), wherein \( \Delta_{13} \) is an amino acid other than serine.

Also preferred is a nucleic acid molecule, wherein \( \Delta_{13} \) is leucine.

Also preferred is a nucleic acid molecule, wherein said additional sub-sequence is \( L\Delta_{14} Y \), wherein \( \Delta_{14} \) is an amino acid other than asparagine.

Also preferred is a nucleic acid molecule, wherein \( \Delta_{14} \) is serine.

Also preferred is a nucleic acid molecule, wherein said additional sub-sequence is \( GA\Delta_{15} XGL \), wherein \( \Delta_{15} \) is an amino acid other than tyrosine.

Also preferred is a nucleic acid molecule, wherein \( \Delta_{15} \) is cysteine.

Especially preferred is a nucleic acid molecule encoding a plant protox enzyme that is capable of being incorporated into a DNA construct used to transform a plant containing wild-type, herbicide-sensitive protox, wherein the nucleic acid molecule has at least one point mutation relative to a wild-type DNA molecule encoding plant protox as mentioned hereinbefore such that upon transformation with the DNA construct the plant contains the DNA molecule, which renders the plant resistant to the application of a herbicide that inhibits naturally occurring plant protox, wherein said nucleic acid molecule can be obtained from a plant selected from the group consisting of \textit{Arabidopsis}, maize, wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, and sugar cane.

Sequences of such DNA molecules are set forth in SEQ ID NOs: 9 (wheat), 11 (soybean), 15 (cotton), 17 (sugar beet), 19 (oilseed rape), 21 (rice), 23 (sorghum), and 36 (sugar cane).

Preferred is a nucleic acid molecule which can be obtained from maize or cotton.

Also comprised by the present invention is a modified enzyme having protoporphyrinogen oxidase (protopx) activity, wherein said modified enzyme is encoded by any of the nucleotide sequences according to the invention mentioned hereinbefore.
The present invention further includes chimeric genes and modified forms of naturally occurring protox genes that can express the inhibitor-resistant plant protox enzymes in plants.

In particular, the present invention relates to a chimeric gene comprising a promoter that is active in a plant operatively linked to a nucleic acid molecule according to the invention. The invention further includes recombinant vector molecules comprising a chimeric gene according to the invention.

Genes encoding inhibitor-resistant plant protox enzymes can be used to confer resistance to protox-inhibitory herbicides in whole plants and as a selectable marker in plant cell transformation methods. Accordingly, the present invention also includes plants, including the descendants thereof, plant tissues and plant seeds containing plant expressible genes encoding these modified protox enzymes. These plants, plant tissues and plant seeds are resistant to protox-inhibitors at levels that normally are inhibitory to the naturally occurring protox activity in the plant.

In particular, the invention relates to a plant cell comprising the nucleic acid molecule according to the invention.

The invention further relates to a plant, plant tissue, plant cell, or plant seed, including the progeny thereof, comprising the nucleic acid molecule according to the invention, wherein said nucleic acid molecule is expressed in said plant, plant tissue, plant cell, or plant seed and confers tolerance thereupon to an inhibitor of naturally occurring protox activity.

Plants encompassed by the invention especially include those that would be potential targets for protox inhibiting herbicides, particularly agronomically important crops such as maize and other cereal crops such as barley, wheat, sorghum, rye, oats, turf and forage grasses, millet and rice. Also comprised are other crop plants such as sugar cane, soybean, cotton, sugar beet, oilseed rape and tobacco.

Preferred is a plant, plant tissue, plant cell, or plant seed, including the progeny thereof, wherein said plant, plant tissue, plant cell, or plant seed is selected from the group consisting of Arabidopsis, soybean, cotton, sugar beet, oilseed rape, maize, wheat, sorghum, rice, and sugar cane comprising the nucleic acid molecule according to the invention.

The present invention also relates to plastid transformation and to the expression of a nucleic acid molecule according to the invention in a plant plastid. In a preferred embodiment, a modified plant protox enzyme as described hereinbefore is expressed in plant plastids to obtain herbicide resistant plants.
In particular, the invention relates to a chimeric gene comprising a promoter functional in a plant plastid, preferably a *clpP* gene promoter, operatively linked to the nucleic acid molecule according to the invention.

In a further embodiment, the present invention is directed to a chimeric gene comprising: (a) a DNA molecule isolated from a plant, which in its native state encodes a polypeptide that comprises a plastid transit peptide, and a mature enzyme that is natively targeted to a plastid of the plant by the plastid transit peptide, wherein the DNA molecule is modified such that it does not encode a functional plastid transit peptide; and (b) a promoter capable of expressing the DNA molecule in a plastid, wherein the promoter is operatively linked to the DNA molecule according to the invention. The DNA molecule may be modified in that at least a portion of the native plastid transit peptide coding sequence is absent from the DNA molecule. Alternatively, the DNA molecule may be modified in that one or more nucleotides of the native plastid transit peptide coding sequence are mutated, thereby rendering an encoded plastid transit peptide nonfunctional. The present invention also relates to plants homoplasmic for chloroplast genomes containing such chimeric genes. In a preferred embodiment, the DNA molecule encodes an enzyme that is naturally inhibited by a herbicidal compound. In this case, such plants are resistant to a herbicide that naturally inhibits the enzyme encoded by a DNA molecule according to the present invention.

A further embodiment of the invention relates to a plastid transformation vector comprising some or all of the elements mentioned hereinbefore including a nucleic acid molecule according to the invention.

Further comprised is a plant plastid comprising a plastid transformation vector according to the invention.

The present invention is also directed to plants made resistant to a herbicide by transforming their plastid genome with a DNA molecule according to the present invention and to methods for obtaining such plants. The invention thus further relates to a plant, plant tissue, plant cell, or plant seed, including the progeny thereof, comprising a plant plastid according to the invention, wherein said modified plant protox enzyme is expressed in said plant, plant tissue, plant cell, or plant seed and confers upon said plant, plant tissue, plant cell, or plant seed tolerance to an inhibitor of naturally occurring protox activity.

The present invention is directed further to methods for the production of plants, including plant material, such as for example plant tissues, protoplasts, cells, calli, organs, plant seeds, embryos, pollen, egg cells, zygotes, together with any other propagating material and plant parts, such as for example flowers, stems, fruits, leaves, roots originating
in transgenic plants or their progeny previously transformed by means of the process of the
invention, which produce an inhibitor-resistant form of the plant protox enzyme provided
herein. Such plants may be stably transformed with a structural gene encoding the
resistant protox, or prepared by direct selection techniques whereby herbicide resistant lines
are isolated, characterized and developed.

The present invention further provides a method for selecting plant cells transformed with a
DNA molecule of the invention that encodes a herbicide-tolerant form of plant protox. The
method comprises introducing the DNA molecule into plant cells whose growth is sensitive
to inhibition by herbicides to which the protox encoded by the DNA molecule is resistant,
thus forming a transformed plant cell. The transformed plant cell whose growth is resistant
to the selected herbicide is identified by selection at a herbicide concentration that inhibits
the growth of untransformed plant cells.

The present invention also provides a novel method for selecting a transplastomic plant cell,
comprising the steps of introducing the above-described chimeric gene into the plastome of
a plant cell; expressing the encoded enzyme in the plastids of the plant cell; and selecting a
cell that is resistant to a herbicidal compound that naturally inhibits the activity of the
enzyme, whereby the resistant cell comprises transformed plastids. In a preferred
embodiment, the enzyme is naturally inhibited by a herbicidal compound and the transgenic
plant is able to grow on an amount of the herbicidal compound that naturally inhibits the
activity of the enzyme.

In another aspect, the present invention is directed to a method for controlling unwanted
vegetation growing at a locus where a herbicide-tolerant, agronomically useful plant, which
is transformed with a DNA molecule according to the present invention that encodes a
herbicide-tolerant form of plant protox, has been cultivated. The method comprises
applying to the locus to be protected an effective amount of herbicide that inhibits naturally
occurring protox activity. In particular, the present invention is directed to a method for
controlling the growth of undesired vegetation, which comprises applying to a population of
the above-described plants an effective amount of an inhibitor of the enzyme.

In preferred embodiment, the invention relates to a method for controlling the growth of
undesired vegetation, wherein said protox-inhibiting herbicide is selected from the group
consisting of an ary luracil, a diphenylether, an oxidiazole, an imide, preferably an imide
having formula V, VI, VII, VIIa, VIII, IX, IXa, or IXb, a phenyl pyrazole, preferably a phenyl
pyrazole having formula XXIV, a pyridyl pyrazole, preferably a pyridyl pyrazole having
formula XXIIIa or XXIIIb, a pyridine derivative, a 3-substituted-2-aryl-4,5,6,7-
tetrahydroindazole, a phenopylate and O-phenylpyrrolidino- and piperidinocarbamate analogs of said phenopylate.

The invention further relates to a method for selectively suppressing the growth of weeds in a field containing a crop of planted crop seeds or plants, comprising the steps of:

(a) planting herbicide tolerant crops or crop seeds, which are plants or plant seeds according to the invention; and

(b) applying to the crops or crop seeds and the weeds in the field a protox-inhibiting herbicide in amounts that inhibit naturally occurring protox activity, wherein the herbicide suppresses the growth of the weeds without significantly suppressing the growth of the crops.

In one aspect, the present invention is directed to an isolated DNA molecule that encodes protoporphyrinogen oxidase (referred to herein as "protox") from sugar cane. The partial DNA coding sequence and corresponding amino acid sequence for a sugar cane protox enzyme are provided as SEQ ID NOs: 36 and 37, respectively.

Further comprised by the present invention are, therefore, isolated DNA molecules encoding the protoporphyrinogen oxidase (protox) enzyme protein from a sugar cane plant, wherein the protein comprises the amino acid sequence given in SEQ ID NO: 37.

In another aspect, the present invention is directed to an isolated DNA molecule that encodes a sugar cane protox enzyme and that comprises a nucleotide sequence that hybridizes to the coding sequence shown in SEQ ID NO: 36 under the following hybridization and wash conditions:

(a) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50 °C; and

(b) wash in 2X SSC, 1% SDS at 50 °C. The present invention is further directed to probes and methods for detecting the presence of genes encoding inhibitor-resistant forms of the plant protox enzyme and quantitating levels of inhibitor-resistant protox transcripts in plant tissue. These methods may be used to identify or screen for plants or plant tissue containing and/or expressing a gene encoding an inhibitor-resistant form of the plant protox enzyme as disclosed herein.

The entire protox sequences according to the invention or portions thereof may be used as probes capable of specifically hybridizing to protox coding sequences and messenger RNA's. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique among protox coding sequences and are preferably at least 10 nucleotides in length, and most preferably at least 20 nucleotides in
length. Such probes may be used to amplify and analyze protox coding sequences from a
calculated melting temperature \( T_m \), which can be easily calculated
The preferred hybridization temperature is in the range of about 25°C below the
calculated melting temperature \( T_m \) and preferably in the range of about 12-15°C below the
calculated melting temperature \( T_m \) and in the case of oligonucleotides in the range of about
5-10°C below the melting temperature \( T_m \).

Comprised by the present invention are DNA molecules that hybridize to a DNA
molecule according to the invention as defined hereinbefore, but preferably to an
oligonucleotide probe obtainable from the DNA molecule comprising a contiguous portion of
the sequence of the protoporphyrinogen oxidase (protox) enzyme at least 10 nucleotides in
length, under moderately stringent conditions.

The invention further embodies the use of a nucleotide probe capable of specifically
hybridizing to a plant protox gene or mRNA of at least 10 nucleotides length in a
polymerase chain reaction (PCR).

In a further embodiment, the present invention provides probes capable of specifically
hybridizing to a eukaryotic DNA sequence encoding a protoporphyrinogen oxidase activity
or to the respective mRNA and methods for detecting the DNA sequences in eukaryotic
organisms using the probes according to the invention.

Protox specific hybridization probes according to the invention may also be used to
map the location of the native eukaryotic protox gene(s) in the genome of a chosen
organism using standard techniques based on the selective hybridization of the probe to
genomic protox sequences. These techniques include, but are not limited to, identification
of DNA polymorphisms identified or contained within the protox probe sequence, and use of
such polymorphisms to follow segregation of the protox gene relative to other markers of
known map position in a mapping population derived from self fertilization of a hybrid of two
polymorphic parental lines (see e.g. Helentjaris et al., Plant Mol. Biol. 5: 109 (1985).
Sommer et al. Biotechniques 12: 82 (1992); D'Ovidio et al., Plant Mol. Biol. 15: 169 (1990)).

While any eukaryotic protox sequence is contemplated to be useful as a probe for mapping protox genes from any eukaryotic organism, preferred probes are those protox sequences from organisms more closely related to the chosen organism, and most preferred probes are those protox sequences from the chosen organism. Mapping of protox genes in this manner is contemplated to be particularly useful in plants for breeding purposes. For instance, by knowing the genetic map position of a mutant protox gene that confers herbicide resistance, flanking DNA markers can be identified from a reference genetic map (see, e.g., Helentjaris, Trends Genet. 3: 217 (1987)). During introgression of the herbicide resistance trait into a new breeding line, these markers can then be used to monitor the extent of protox-linked flanking chromosomal DNA still present in the recurrent parent after each round of back-crossing.

Protox specific hybridization probes according to the invention may also be used to quantitate levels of protox mRNA in an organism using standard techniques such as Northern blot analysis. This technique may be useful as a diagnostic assay to detect altered levels of protox expression that may be associated with particular adverse conditions such as autosomal dominant disorder in humans characterized by both neuropsychiatric symptoms and skin lesions, which are associated with decreased levels of protox activity (Brenner and Bloomer, New Engl. J. Med. 302: 765 (1980)).

A further embodiment of the invention is a method of producing a DNA molecule comprising a DNA portion encoding a protein having protoporphyrinogen oxidase (protox) enzyme activity comprising:

(a) preparing a nucleotide probe capable of specifically hybridizing to a plant protox gene or mRNA, wherein the probe comprises a contiguous portion of the coding sequence for a protox protein according to the present invention of at least 10 nucleotides length;

(b) probing for other protox coding sequences in populations of cloned genomic DNA fragments or cDNA fragments from a chosen organism using the nucleotide probe prepared according to step (a); and

(c) isolating and multiplying a DNA molecule comprising a DNA portion encoding a protein having protoporphyrinogen oxidase (protox) enzyme activity.

A further embodiment of the invention is a method of isolating a DNA molecule from any plant comprising a DNA portion encoding a protein having protoporphyrinogen oxidase (protox) enzyme activity.
(a) preparing a nucleotide probe capable of specifically hybridizing to a plant protox gene or mRNA, wherein the probe comprises a contiguous portion of the coding sequence for a protox protein according to the present invention of at least 10 nucleotides length;

(b) probing for other protox coding sequences in populations of cloned genomic DNA fragments or cDNA fragments from a chosen organism using the nucleotide probe prepared according to step (a); and

(c) isolating a DNA molecule comprising a DNA portion encoding a protein having protoporphyrinogen oxidase (protox) enzyme activity.

The invention further comprises a method of producing an essentially pure DNA sequence coding for a protein exhibiting protoporphyrinogen oxidase (protox) enzyme activity, which method comprises:

(a) preparing a genomic or a cDNA library from a suitable source organism using an appropriate cloning vector;

(b) hybridizing the library with a probe molecule according to the present invention; and

(c) identifying positive hybridizations of the probe to the DNA clones from the library that is clones potentially containing the nucleotide sequence corresponding to the amino acid sequence for protoporphyrinogen oxidase (protox).

The invention further comprises a method of producing an essentially pure DNA sequence coding for a protein exhibiting protoporphyrinogen oxidase (protox) enzyme activity, which method comprises:

(a) preparing total DNA from a genomic or a cDNA library;

(b) using the DNA of step (a) as a template for PCR reaction with primers representing low degeneracy portions of the amino acid sequence of protoporphyrinogen oxidase (protox) according to the present invention.

A further object of the invention is an assay to identify inhibitors of protoporphyrinogen oxidase (protox) enzyme activity that comprises:

(a) incubating a first sample of a protoporphyrinogen oxidase (protox) according to the present invention and its substrate;

(b) measuring an uninhibited reactivity of the protoporphyrinogen oxidase (protox) from step (a);
(c) incubating a first sample of protoporphyrinogen oxidase (protox) and its substrate in the presence of a second sample comprising an inhibitor compound;

(d) measuring an inhibited reactivity of the protoporphyrinogen oxidase (protox) enzyme from step (c); and

(e) comparing the inhibited reactivity to the uninhibited reactivity of protoporphyrinogen oxidase (protox) enzyme.

A further object of the invention is an assay to identify inhibitor-resistant protoporphyrinogen oxidase (protox) mutants that comprises:

(a) incubating a first sample of protoporphyrinogen oxidase (protox) enzyme and its substrate in the presence of a second sample comprising a protoporphyrinogen oxidase (protox) enzyme inhibitor;

(b) measuring an unmutated reactivity of the protoporphyrinogen oxidase (protox) enzyme from step (a);

(c) incubating a first sample of a mutated protoporphyrinogen oxidase (protox) enzyme and its substrate in the presence of a second sample comprising protoporphyrinogen oxidase (protox) enzyme inhibitor;

(d) measuring a mutated reactivity of the mutated protoporphyrinogen oxidase (protox) enzyme from step (c); and

(e) comparing the mutated reactivity to the unmutated reactivity of the protoporphyrinogen oxidase (protox) enzyme.

A further object of the invention is a protox enzyme inhibitor obtained by a method according to the invention.

For recombinant production of the enzyme in a host organism, the protox coding sequence according to the present invention may be inserted into an expression cassette designed for the chosen host and introduced into the host where it is recombinantly produced. The choice of specific regulatory sequences such as promoter, signal sequence, 5' and 3' untranslated sequences, and enhancer, is within the level of skill of the routineer in the art. The resultant molecule, containing the individual elements linked in proper reading frame, may be inserted into a vector capable of being transformed into the host cell.

Suitable expression vectors and methods for recombinant production of proteins are well known for host organisms such as *E. coli* (see, e.g., Studier and Moffatt, *J. Mol. Biol.* 189: 113 (1986); Brosius, *DNA* 8: 759 (1989)), yeast (see, e.g., Schneider and Guarente, *Meth. Enzymol.* 194: 373 (1991)) and insect cells (see, e.g., Luckow and Summers, *Bio/Technol.*
Specific examples include plasmids such as pBluescript (Stratagene, La Jolla, CA), pFLAG (International Biotechnologies, Inc., New Haven, CT), pTrcHis (Invitrogen, La Jolla, CA), and baculovirus expression vectors, e.g., those derived from the genome of Autographica californica nuclear polyhedrosis virus (AcMNPV). A preferred baculovirus/insect system is pVL11392/Sf21 cells (Invitrogen, La Jolla, CA).

Recombinantly produced eukaryotic protox enzyme is useful for a variety of purposes. For example, it may be used to supply protox enzymatic activity in vitro. It may also be used in an in vitro assay to screen known herbicidal chemicals whose target has not been identified to determine if they inhibit protox. Such an in vitro assay may also be used as a more general screen to identify chemicals that inhibit protox activity and that are therefore herbicide candidates. Recombinantly produced eukaryotic protox enzyme may also be used in an assay to identify inhibitor-resistant protox mutants (see International application no. PCT/IB95/00452 filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659, incorporated by reference herein in its entirety). Alternatively, recombinantly produced protox enzyme may be used to further characterize its association with known inhibitors in order to rationally design new inhibitory herbicides as well as herbicide tolerant forms of the enzyme.

In another aspect, the present invention teaches modifications that can be made to the amino acid sequence of any eukaryotic protoporphyrinogen oxidase (referred to herein as "protox") enzyme to yield an inhibitor-resistant form of this enzyme. Preferably, the eukaryotic protox enzyme is a plant protox enzyme. The present invention is directed to inhibitor-resistant protox enzymes having the modifications taught herein, to DNA molecules encoding these modified enzymes, and to chimeric genes capable of expressing these modified enzymes in plants.

The present invention discloses an isolated DNA molecule encoding a modified eukaryotic protoporphyrinogen oxidase (protox) having at least one amino acid modification, wherein the amino acid modification has the property of conferring resistance to a protox inhibitor, that is wherein the modified protox is tolerant to an inhibitor in amounts that inhibit the naturally occurring eukaryotic protox. As used herein "inhibit" refers to a reduction in enzymatic activity observed in the presence of a subject compound compared to the level of activity observed in the absence of the subject compound, wherein the percent level of reduction is preferably at least 10%, more preferably at least 50%, and most preferably at least 90%.
In particular, a DNA molecule is disclosed encoding a modified eukaryotic protoporphyrinogen oxidase (protox) that is a plant protox, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity, particularly a protox selected from the group consisting of an Arabidopsis protox enzyme, a maize protox enzyme, a wheat protox enzyme, a soybean protox enzyme, a cotton protox enzyme, a sugar beet protox enzyme, an oilseed rape protox enzyme, a rice protox enzyme, a sorghum protox enzyme, and a sugar cane protox enzyme having at least one amino acid modification, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity.

As used herein, the expression “substantially conserved amino acid sequences” refers to regions of amino acid homology between polypeptides comprising protox enzymes from different sources. In the present invention, seventeen substantially conserved amino acid sub-sequences, designated 1-19 respectively, are shown in Table 1B. One skilled in the art could align the amino acid sequences of protox enzymes from different sources, as has been done in Table 1A, to identify the sub-sequences therein that make up the substantially conserved amino acid sequences defined herein. Stated another way, a given sub-sequence from one source “corresponds to” a homologous subsequence from a different source. The skilled person could then determine whether the identified sub-sequences have the characteristics disclosed and claimed in the present application.

Therefore, the present invention discloses a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes an enzyme having protoporphyrinogen oxidase (protox) activity, wherein the nucleic acid molecule is capable of being incorporated into a nucleic acid construct used to transform a plant containing wild-type, herbicide-sensitive protox, wherein the nucleotide sequence has at least one point mutation relative to a wild-type nucleotide sequence encoding plant protox, such that upon transformation with the nucleic acid construct the plant is rendered herbicide-tolerant.

A preferred embodiment of the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises at least one of the following amino acid sub-sequences:

(a) APΔ₁F, wherein Δ₁ is an amino acid other than arginine, but especially leucine or cysteine;
(b) FAΔ₂S, wherein Δ₂ is leucine;

- 17 -
(c) $Y_{13}G$, wherein $\Delta_3$ is isoleucine;
(d) $\Delta_1IG$, wherein $\Delta_1$ is histidine or alanine;
(e) $\text{KA}_{18}F$, wherein $\Delta_{18}$ is an amino acid other than alanine;
(f) $\text{Q}_{19}H$, wherein $\Delta_{19}$ is an amino acid other than leucine;
(g) $T_{16}G$, wherein $\Delta_{16}$ is an amino acid other than leucine and $Y_{17}G$, wherein $\Delta_{17}$ is an amino acid other than alanine.

(Table 1B; sub-sequences 1-10 and 18-19, and 16-17).

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $\text{AP}_{1}F$, wherein $\Delta_1$ is leucine or cysteine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $\text{FA}_{2}S$, wherein $\Delta_2$ is leucine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $Y_{3}G$, wherein $\Delta_3$ is an isoleucine.

Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $\text{IG}_{2}$, wherein $\Delta_2$ is histidine or alanine, but especially histidine.

Further preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence $\text{KA}_{18}F$, wherein $\Delta_{18}$ is an amino acid other than alanine. Most preferably, $\Delta_{18}$ is threonine or valine.
Preferred is a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises the amino acid sub-sequence QA_{19}H, wherein A_{19} is an amino acid other than leucine. Most preferably, A_{19} is serine.

Another preferred embodiment of the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises at least one of the following amino acid sub-sequences:

(a) K\Delta_{18}F, wherein \Delta_{18} is an amino acid other than alanine,
(b) Q\Delta_{19}H, wherein \Delta_{19} is an amino acid other than leucine,
(c) A\Delta_{1}F, wherein \Delta_{1} is an amino acid other than arginine, but especially leucine or cysteine,
(d) F\Delta_{2}S, wherein \Delta_{2} is leucine
(e) Y\Delta_{3}G, wherein \Delta_{3} is isoleucine
(f) \Delta_{7}IG, wherein \Delta_{7} is histidine or alanine,
(g) T\Delta_{16}G, wherein \Delta_{16} is an amino acid other than leucine and YV\Delta_{17}G, wherein \Delta_{17} is an amino acid other than alanine,

(Table 1B; sub-sequences 1-10 and 18-19; and 16-17), and wherein the modified enzyme further comprises at least one additional amino acid sub-sequence selected from the group consisting of:

(c) (d)Q\Delta_{11}S, wherein \Delta_{11} is an amino acid other than proline;
(d) (e)IGG\Delta_{12}, wherein \Delta_{12} is an amino acid other than threonine;
(e) S\Delta_{13}X\Delta_{13}, wherein \Delta_{13} is an amino acid other than serine;
(f) L\Delta_{14}Y, wherein \Delta_{14} is an amino acid other than asparagine; and
(g) G\Delta_{15}XG, wherein \Delta_{15} is an amino acid other than tyrosine.

Another preferred embodiment of the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the modified enzyme comprises: the amino acid sub-sequence \Delta_{7}IG, wherein \Delta_{7} is an amino acid other than tyrosine; the amino acid sub-sequences IGG\Delta_{12}, wherein \Delta_{12} is an amino acid other
than threonine; and the amino acid sub-sequence SWXLΔ13, wherein Δ13 is an amino acid other than serine. Most preferably, Δ7 is isoleucine, Δ12 is isoleucine, and Δ13 is leucine.

Yet another preferred embodiment of the present invention is directed to a nucleic acid molecule comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the nucleotide sequence is further characterized in that at least one of the following conditions is met:

(a) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence APΔ1F, wherein Δ1 is leucine;
(b) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence FΔ2S, wherein Δ2 is leucine;
(c) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence YΔ3G, wherein Δ3 is isoleucine;
(d) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence Δ7IG, wherein Δ7 is alanine or histidine;
(e) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence YΔ3G, wherein Δ3 is an amino acid other than alanine, but preferably cysteine or isoleucine, more preferably isoleucine, and the nucleic acid sequence also has a sequence that encodes one of the group consisting of:

(1) sub-sequence QΔ11S, wherein Δ11 is an amino acid other than proline,
(2) sub-sequence IGGAΔ12, wherein Δ12 is an amino acid other than threonine,
(3) sub-sequence SWXLΔ13, wherein Δ13 is an amino acid other than serine,
(4) sub-sequence LΔ14Y, wherein Δ14 is an amino acid other than asparagine, and
(5) sub-sequence GAΔ15XGL, wherein Δ15 is an amino acid other than tyrosine;
(f) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence Δ7IG, wherein Δ7 is an amino acid other than tyrosine, but
preferably threonine, alanine or histidine, more preferably alanine or histidine and most preferably histidine, and the nucleic acid sequence also has a sequence that encodes one of the group consisting of:

(1) sub-sequence QΔ_{11}S, wherein Δ_{11} is an amino acid other than proline,
(2) sub-sequence IGGΔ_{12}, wherein Δ_{12} is an amino acid other than threonine,
(3) sub-sequence SWXLΔ_{13}, wherein Δ_{13} is an amino acid other than serine,
(4) sub-sequence LΔ_{14}Y, wherein Δ_{14} is an amino acid other than asparagine, and
(5) sub-sequence GΔ_{15}XGL, wherein Δ_{15} is an amino acid other than tyrosine; and

(g) the nucleic acid has a sequence that encodes amino acid sub-sequence TΔ_{16}G, wherein Δ_{16} is an amino acid other than leucine, and the nucleic acid sequence also has a sequence that encodes amino acid sub-sequence YVΔ_{17}G, wherein Δ_{17} is an amino acid other than alanine.

(h) KΔ_{18}F, wherein Δ_{18} is an amino acid other than alanine;

(i) QΔ_{19}H, wherein Δ_{19} is an amino acid other than leucine.

Preferably, said nucleic acid sequence has a sequence that encodes amino acid sub-sequence TΔ_{16}G, wherein Δ_{16} is an amino acid other than leucine, and said nucleic acid sequence also has a sequence that encodes amino acid sub-sequence YVΔ_{17}G, wherein Δ_{17} is an amino acid other than alanine. Preferably, the nucleic acid sequence has a sequence that encodes amino acid sub-sequence TΔ_{18}G, wherein Δ_{18} is serine, and said nucleic acid sequence also has a sequence that encodes amino acid sub-sequence YVΔ_{17}G, wherein Δ_{17} is threonine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the arginine occurring at the position corresponding to amino acid 88 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is the DNA molecule wherein the arginine is replaced with a leucine or cysteine.
Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the cysteine occurring at the position corresponding to amino acid 159 of SEQ ID NO:6 is replaced with a leucine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the alanine occurring at the position corresponding to amino acid 175 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the alanine is replaced with a valine or threonine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the leucine occurring at the position corresponding to amino acid 337 of SEQ ID NO:6 is replaced with another amino acid, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity. Particularly preferred is a DNA molecule wherein the leucine is replaced with a serine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the tyrosine occurring at the position corresponding to amino acid 428 of SEQ ID NO:16 is replaced with a histidine or alanine, preferably a histidine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the alanine occurring at the position corresponding to amino acid 220 of SEQ ID NO:2 is replaced with a isoleucine or a tyrosine, preferably a isoleucine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the tyrosine occurring at the position corresponding to amino acid 426 of SEQ ID NO:2 is replaced with a alanine.

Also preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the leucine occurring at the position corresponding to amino acid 347 of SEQ ID NO:6 is replaced with a serine and the alanine occurring at the position corresponding to amino acid 453 of SEQ ID NO:6 is replaced with a threonine.

The present invention further discloses a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox having a first amino acid substitution and a second amino acid substitution; the first amino acid substitution having
the property of conferring resistance to a protox inhibitor; and the second amino acid substitution having the property of enhancing the resistance conferred by the first amino acid substitution but especially a DNA molecule encoding a modified protoporphyrogen oxidase (protox) comprising a plant protox, wherein the plant is selected from the group consisting of maize, wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, sugar cane, and Arabidopsis. Preferred is a DNA molecule wherein the first amino acid substitution occurs at a position selected from the group consisting of:

(a) the position corresponding to the arginine at amino acid 88 of SEQ ID NO:6;
(b) the position corresponding to the alanine at amino acid 175 of SEQ ID NO:6;
(c) the position corresponding to the leucine at amino acid 337 of SEQ ID NO:6.

Particularly preferred is a DNA molecule wherein the first amino acid substitution occurs at a position selected from the group consisting of:

(a) the position corresponding to the arginine at amino acid 88 of SEQ ID NO:6;
(b) the position corresponding to the alanine at amino acid 164 of SEQ ID NO:6;
(c) the position corresponding to the glycine at amino acid 165 of SEQ ID NO:6;
(d) the position corresponding to the tyrosine at amino acid 370 of SEQ ID NO:6;
(e) the position corresponding to the cysteine at amino acid 159 of SEQ ID NO:6;
(f) the position corresponding to the isoleucine at amino acid 419 of SEQ ID NO:6;
(g) the position corresponding to the valine at amino acid 356 of SEQ ID NO:10;
(h) the position corresponding to the serine at amino acid 421 of SEQ ID NO:10;
(i) the position corresponding to the valine at amino acid 502 of SEQ ID NO:10;
(j) the position corresponding to the alanine at amino acid 211 of SEQ ID NO:10;
(k) the position corresponding to the glycine at amino acid 212 of SEQ ID NO:10;
(l) the position corresponding to the isoleucine at amino acid 466 of SEQ ID NO:10;
(m) the position corresponding to the proline at amino acid 369 of SEQ ID NO:12;
(n) the position corresponding to the alanine at amino acid 226 of SEQ ID NO:12;
(o) the position corresponding to the tyrosine at amino acid 432 of SEQ ID NO:12;
(p) the position corresponding to the valine at amino acid 517 of SEQ ID NO:12;
(q) the position corresponding to the tyrosine at amino acid 428 of SEQ ID NO:16;
(r) the position corresponding to the proline at amino acid 365 of SEQ ID NO:16;
(s) the position corresponding to the tyrosine at amino acid 449 of SEQ ID NO:18;
(t) the position corresponding to the alanine at amino acid 175 of SEQ ID NO:6;
and
(u) the position corresponding to the leucine at amino acid 337 of SEQ ID NO:6;

wherein the second amino acid substitution occurs at a position selected from the group consisting of:

(a) the position corresponding to the serine at amino acid 305 of SEQ ID NO:2;
(b) the position corresponding to the threonine at amino acid 249 of SEQ ID NO:2;
(c) the position corresponding to the proline at amino acid 118 of SEQ ID NO:2;
(d) the position corresponding to the asparagine at amino acid 425 of SEQ ID NO:2;
and
(e) the position corresponding to the tyrosine at amino acid 498 of SEQ ID NO:2.

Particularly preferred is a DNA molecule wherein the first amino acid substitution occurs at a position selected from the group consisting of:

(a) the position corresponding to the arginine at amino acid 88 of SEQ ID NO:6;
(b) the position corresponding to the alanine at amino acid 175 of SEQ ID NO:6;
and

(c) the position corresponding to the leucine at amino acid 337 of SEQ ID NO:6.

wherein the second amino acid substitution occurs at a position selected from the group consisting of:

(a) the position corresponding to the serine at amino acid 305 of SEQ ID NO:2;
(b) the position corresponding to the threonine at amino acid 249 of SEQ ID NO:2; 
(c) the position corresponding to the proline at amino acid 118 of SEQ ID NO:2; 
(d) the position corresponding to the asparagine at amino acid 425 of SEQ ID NO:2; and 
(d) the position corresponding to the tyrosine at amino acid 498 of SEQ ID NO:2. Particularly preferred is a DNA molecule wherein the arginine occurring at the position corresponding to residue 88 of SEQ ID NO:6 is replaced with a cysteine or a leucine.

More preferred is a DNA molecule wherein the cysteine occurring at the position corresponding to residue 159 of SEQ ID NO:6 is replaced with an leucine.

Particularly preferred is a DNA molecule wherein the alanine occurring at the position corresponding to residue 175 of SEQ ID NO:6 is replaced with an amino acid selected from the group consisting of threonine and valine.

Particularly preferred is a DNA molecule wherein the leucine occurring at the position corresponding to residue 337 of SEQ ID NO:6 is replaced with a serine.

More preferred is a DNA molecule wherein the tyrosine occurring at the position corresponding to residue 428 of SEQ ID NO:16 is replaced with histidine or alanine.

The present invention is still further directed to a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox having a double amino acid substitution, wherein both amino acid substitutions are required for there to be resistance to a protox inhibitor. Preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox, wherein the plant is selected from the group consisting of maize, wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, sugar cane, and Arabidopsis. More preferred is a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox, wherein the plant is maize.

Preferred is a DNA molecule having a double amino acid substitution, wherein one amino acid substitution occurs at the position corresponding to the leucine at amino acid 347 of SEQ ID NO:6, and wherein the second amino acid substitution occurs at the position corresponding to the alanine at amino acid 453 of SEQ ID NO:6.

Particularly preferred is a DNA molecule having a double amino acid substitution, wherein a leucine occurring at the position corresponding to amino acid 347 of SEQ ID NO:6 is replaced with a serine, and wherein an alanine occurring at the position corresponding to amino acid 453 of SEQ ID NO:6 is replaced with a threonine.
The present invention is further directed to a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the plant protox comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 37.

The present invention is further directed to expression cassettes and recombinant vectors comprising the expression cassettes comprising essentially a promoter, but especially a promoter that is active in a plant, operatively linked to a DNA molecule encoding the protoporphyrinogen oxidase (protox) enzyme from a eukaryotic organism according to the invention. The expression cassette according to the invention may in addition further comprise a signal sequence operatively linked to the DNA molecule, wherein the signal sequence is capable of targeting the protein encoded by the DNA molecule into the chloroplast or the mitochondria.

The invention relates to a chimeric gene, which comprises an expression cassette comprising essentially a promoter, but especially a promoter that is active in a plant, operatively linked to a heterologous DNA molecule encoding a protoporphyrinogen oxidase (protox) enzyme from a eukaryotic organism according to the invention. Preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox) enzyme from sugar cane.

More preferred is a chimeric gene comprising a promoter active in a plant operatively linked to a heterologous DNA molecule encoding a protoporphyrinogen oxidase (protox) from sugar cane comprising the sequence set forth in SEQ ID NO:37.

Particularly preferred is a chimeric gene, wherein the DNA molecule encodes a protox-1 activity, preferably wherein the protein comprises the amino acid sequence set forth in SEQ ID NO:37.

The invention also embodies a chimeric gene, which comprises an expression cassette comprising essentially a promoter, but especially a promoter that is active in a plant, operatively linked to the DNA molecule encoding an protoporphyrinogen oxidase (protox) enzyme from a eukaryotic organism according to the invention, which is resistant to herbicides at levels that inhibit the corresponding unmodified version of the enzyme. Preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox) enzyme from a plant selected from the group consisting of Arabidopsis, sugar cane, soybean, barley, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf and forage grasses, millet, forage and rice. More preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox)
enzyme from a plant selected from the group consisting of soybean, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf grass, and rice. Particularly preferred is a chimeric gene, wherein the DNA molecule encodes an protoporphyrinogen oxidase (protox) enzyme from a plant selected from the group consisting of Arabidopsis, soybean, cotton, sugar beet, oilseed rape, maize, wheat, sorghum, and rice.

Encompassed by the present invention is a chimeric gene comprising a promoter that is active in a plant operatively linked to the DNA molecule according to the invention encoding a modified protoporphyrinogen oxidase (protox) comprising a eukaryotic protox having at least one amino acid modification, wherein the amino acid modification has the property of conferring resistance to a protox inhibitor.

Also encompassed by the present invention is a chimeric gene comprising a promoter that is active in a plant operatively linked to the DNA molecule according to the invention encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox having a first amino acid substitution and a second amino acid substitution; the first amino acid substitution having the property of conferring resistance to a protox inhibitor; and the second amino acid substitution having the property of enhancing the resistance conferred by the first amino acid substitution. Preferred is the chimeric gene additionally comprising a signal sequence operatively linked to the DNA molecule, wherein the signal sequence is capable of targeting the protein encoded by the DNA molecule into the chloroplast or in the mitochondria.

The chimeric gene according to the invention may in addition further comprise a signal sequence operatively linked to the DNA molecule of the invention, wherein the signal sequence is capable of targeting the protein encoded by the DNA molecule into the chloroplast. The chimeric gene according to the invention may in addition further comprise a signal sequence operatively linked to the DNA molecule, wherein the signal sequence is capable of targeting the protein encoded by the DNA molecule into the mitochondria.

Also encompassed by the present invention is any of the DNA sequences mentioned herein before, which is stably integrated into a host genome.

The invention further relates to a recombinant DNA molecule comprising a plant protoporphyrinogen oxidase (protox) according to the invention or a functionally equivalent derivative thereof.

The invention further relates to a recombinant DNA vector comprising the recombinant DNA molecule of the invention.
A further object of the invention is a recombinant vector comprising the chimeric gene according to the invention, wherein the vector is capable of being stably transformed into a host cell.

A further object of the invention is a recombinant vector comprising the chimeric gene according to the invention, wherein the vector is capable of being stably transformed into a plant, plant seeds, plant tissue or plant cell. Preferred is a recombinant vector comprising the chimeric gene according to the invention, wherein the vector is capable of being stably transformed into a plant. The plant, plant seeds, plant tissue or plant cell stably transformed with the vector is capable of expressing the DNA molecule encoding a protoporphyrinogen oxidase (protox). Preferred is a recombinant vector, wherein the plant, plant seeds, plant tissue or plant cell stably transformed with the vector is capable of expressing the DNA molecule encoding a protoporphyrinogen oxidase (protox) from a plant that is resistant to herbicides at levels that inhibit the corresponding unmodified version of the enzyme.

Preferred is a recombinant vector comprising the chimeric gene of the invention comprising a promoter active in a plant operatively linked to a heterologous DNA molecule encoding a protoporphyrinogen oxidase (protox) from sugar cane comprising the sequence set forth in SEQ ID NO:37, wherein the vector is capable of being stably transformed into a host cell.

Also preferred is recombinant vector comprising the chimeric gene of the invention comprising a promoter that is active in a plant operatively linked to the DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox having a first amino acid substitution and a second amino acid substitution; the first amino acid substitution having the property of conferring resistance to a protox inhibitor; and the second amino acid substitution having the property of enhancing the resistance conferred by the first amino acid substitution, wherein the vector is capable of being stably transformed into a plant cell.

Also encompassed by the present invention is a host cell stably transformed with the vector according to the invention, wherein the host cell is capable of expressing the DNA molecule. Preferred is a host cell selected from the group consisting of a plant cell, a bacterial cell, a yeast cell, and an insect cell.

The present invention is further directed to plants and the progeny thereof, plant tissue and plant seeds tolerant to herbicides that inhibit the naturally occurring protox activity in these plants, wherein the tolerance is conferred by a gene expressing a modified inhibitor-resistant protox enzyme as taught herein. Representative plants include any plants
to which these herbicides may be applied for their normally intended purpose. Preferred are 
agronomically important crops, i.e., angiosperms and gymnosperms such as Arabidopsis, 
sugar cane, soybean, barley, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, 
sorghum, rye, oats, tomato, potato, turf and forage grasses, millet, forage, and rice and the 
like. More preferred are agronomically important crops, i.e., angiosperms and gymnosperms 
such as Arabidopsis, cotton, soybean, oilseed rape, sugar beet, maize, rice, wheat, barley, 
oats, rye, sorghum, millet, turf, forage, turf grasses. Particularly preferred are agronomically 
important crops, i.e., angiosperms and gymnosperms such as Arabidopsis, soybean, cotton, 
sugar beet, oilseed rape, maize, wheat, sorghum, and rice.

Preferred is a plant comprising the DNA molecule encoding a modified 
protoporphyrinogen oxidase (protox) comprising a plant protox having a first amino acid 
substitution and a second amino acid substitution; the first amino acid substitution having 
the property of conferring resistance to a protox inhibitor; and the second amino acid 
substitution having the property of enhancing the resistance conferred by the first amino 
acid substitution, wherein the DNA molecule is expressed in the plant and confers upon the 
plant tolerance to a herbicide in amounts that inhibit naturally occurring protox activity. 
Preferred is a plant, wherein the DNA molecule replaces a corresponding naturally 
occurring protox coding sequence. Comprised by the present invention is a plant and the 
progeny thereof comprising the chimeric gene according to the invention, wherein the 
chimeric gene confers upon the plant tolerance to a herbicide in amounts that inhibit 
naturally occurring protox activity.

Encompassed by the present invention are transgenic plant tissue, including plants 
and the progeny thereof, seeds, and cultured tissue, stably transformed with at least one 
chimeric gene according to the invention. Preferred is transgenic plant tissue, including 
plants, seeds, and cultured tissue, stably transformed with at least one chimeric gene that 
comprises an expression cassette comprising essentially a promoter, but especially a 
promoter that is active in a plant, operatively linked to the DNA molecule encoding an 
protoporphyrinogen oxidase (protox) enzyme that is resistant to herbicides at levels that 
inhibit the corresponding unmodified version of the enzyme in the plant tissue.

The present invention is further directed to plants, plant tissue, plant seeds, and plant 
cells tolerant to herbicides that inhibit the naturally occurring protox activity in these plants, 
wherein the tolerance is conferred by increasing expression of wild-type herbicide-sensitive 
protox. This results in a level of a protox enzyme in the plant cell at least sufficient to 
overcome growth inhibition caused by the herbicide. The level of expressed enzyme
generally is at least two times, preferably at least five times, and more preferably at least ten times the natively expressed amount. Increased expression may be due to multiple copies of a wild-type protox gene; multiple occurrences of the coding sequence within the gene (i.e. gene amplification) or a mutation in the non-coding, regulatory sequence of the endogenous gene in the plant cell. Plants having such altered gene activity can be obtained by direct selection in plants by methods known in the art (see, e.g. U.S. Patent No. 5,162,602, and U.S. Patent No. 4,761,373, and references cited therein). These plants also may be obtained by genetic engineering techniques known in the art. Increased expression of a herbicide-sensitive protox gene can also be accomplished by stably transforming a plant cell with a recombinant or chimeric DNA molecule comprising a promoter capable of driving expression of an associated structural gene in a plant cell operatively linked to a homologous or heterologous structural gene encoding the protox enzyme.


Comprised within the scope of the present invention are transgenic plants, in particular transgenic fertile plants transformed by means of the aforesaid processes
and their asexual and/or sexual progeny, which still are resistant or at least tolerant to inhibition by a herbicide at levels that normally are inhibitory to the naturally occurring protox activity in the plant. Progeny plants also include plants with a different genetic background than the parent plant, which plants result from a backcrossing program and still comprise in their genome the herbicide resistance trait according to the invention. Very especially preferred are hybrid plants that are resistant or at least tolerant to inhibition by a herbicide at levels that normally are inhibitory to the naturally occurring protox activity in the plant.

The transgenic plant according to the invention may be a dicotyledonous or a monocotyledonous plant. Preferred are monocotyledonous plants of the Graminaceae family involving Lolium, Zea, Triticum, Triticale, Sorghum, Saccharum, Bromus, Oryzae, Avena, Hordeum, Secale and Setaria plants. More preferred are transgenic maize, wheat, barley, sorghum, rye, oats, sugar cane, turf and forage grasses, millet and rice. Especially preferred are maize, wheat, sorghum, rye, oats, turf grasses and rice.

Among the dicotyledonous plants Arabidopsis, soybean, cotton, sugar beet, oilseed rape, tobacco, tomato, potato, and sunflower are more preferred herein. Especially preferred are soybean, cotton, tobacco, sugar beet, tomato, potato, and oilseed rape.

The expression 'progeny' is understood to embrace both, "asexually" and "sexually" generated progeny of transgenic plants. This definition is also meant to include all mutants and variants obtainable by means of known processes, such as for example cell fusion or mutant selection and that still exhibit the characteristic properties of the initial transformed plant, together with all crossing and fusion products of the transformed plant material. This also includes progeny plants that result from a backcrossing program, as long as the progeny plants still contain the herbicide resistant trait according to the invention.

Another object of the invention concerns the proliferation material of transgenic plants. The proliferation material of transgenic plants is defined relative to the invention as any plant material that may be propagated sexually or asexually in vivo or in vitro. Particularly preferred within the scope of the present invention are protoplasts, cells, calli, tissues, organs, seeds, embryos, pollen, egg cells, zygotes, together with any other propagating material obtained from transgenic plants.

Parts of plants, such as for example flowers, stems, fruits, leaves, roots originating in transgenic plants or their progeny previously transformed by means of the process of the invention and therefore consisting at least in part of transgenic cells, are also an object of the present invention.
A further object of the invention is a method of producing plants, protoplasts, cells, calli, tissues, organs, seeds, embryos, pollen, egg cells, zygotes, together with any other propagating material, parts of plants, such as for example flowers, stems, fruits, leaves, roots originating in transgenic plants or their progeny previously transformed by means of the process of the invention, which therefore produce an inhibitor resistant form of a plant protox enzyme by transforming the plant, plant parts with the DNA according to the invention. Preferred is a method of producing a host cell comprising an isolated DNA molecule encoding a protein from a eukaryote having protoporphyrinogen oxidase (protox) activity comprising transforming the host cell with a recombinant vector molecule according to the invention. Further preferred is a method of producing a plant cell comprising an isolated DNA molecule encoding a protein from a eukaryote having protoporphyrinogen oxidase (protox) activity comprising transforming the plant cell with a recombinant vector molecule according to the invention. Preferred is a method of producing transgenic progeny of a transgenic parent plant comprising an isolated DNA molecule encoding a protein from a eukaryote having protoporphyrinogen oxidase (protox) activity comprising transforming the parent plant with a recombinant vector molecule according to the invention and transferring the herbicide tolerant trait to the progeny of the transgenic parent plant involving known plant breeding techniques.

Preferred is a method for the production of plants, plant tissues, plant seeds and plant parts, which produce an inhibitor-resistant form of the plant protox enzyme, wherein the plants, plant tissues, plant seeds and plant parts have been stably transformed with a structural gene encoding the resistant protox enzyme. Particularly preferred is a method for the production of plants, plant tissues, plant seeds and plant parts, wherein the plants, plant tissues, plant seeds and plant parts have been stably transformed with the DNA according to the invention. Especially preferred is a method for the production of the plants, plant tissues, plant seeds and plant parts, which produce an inhibitor-resistant form of the plant protox enzyme, wherein the plants, plant tissues, plant seeds and plant parts have been prepared by direct selection techniques whereby herbicide resistant lines are isolated, characterized and developed.

The genetic properties engineered into the transgenic seeds and plants described above are passed on by sexual reproduction or vegetative growth and can thus be maintained and propagated in progeny plants. Generally the maintenance and propagation make use of known agricultural methods developed to fit specific purposes such as tilling, sowing or harvesting. Specialized processes such as hydroponics or greenhouse
technologies can also be applied. As the growing crop is vulnerable to attack and damages caused by insects or infections as well as to competition by weed plants, measures are undertaken to control weeds, plant diseases, insects, nematodes, and other adverse conditions to improve yield. These include mechanical measures such as tillage of the soil or removal of weeds and infected plants, as well as the application of agrochemicals such as herbicides, fungicides, gametocides, nematicides, growth regulants, ripening agents and insecticides.

Use of the advantageous genetic properties of the transgenic plants and seeds according to the invention can further be made in plant breeding that aims at the development of plants with improved properties such as tolerance of pests, herbicide tolerance, or stress tolerance, improved nutritional value, increased yield, or improved structure causing less loss from lodging or shattering. The various breeding steps are characterized by well-defined human intervention such as selecting the lines to be crossed, directing pollination of the parental lines, or selecting appropriate progeny plants. Depending on the desired properties different breeding measures are taken. The relevant techniques are well known in the art and include but are not limited to hybridization, inbreeding, backcross breeding, multiline breeding, variety blend, interspecific hybridization, aneuploid techniques, etc. Hybridization techniques also include the sterilization of plants to yield male or female sterile plants by mechanical, chemical or biochemical means. Cross pollination of a male sterile plant with pollen of a different line assures that the genome of the male sterile but female fertile plant will uniformly obtain properties of both parental lines. Thus, the transgenic seeds and plants according to the invention can be used for the breeding of improved plant lines that for example increase the effectiveness of conventional methods such as herbicide or pesticide treatment or allow to dispense with the methods due to their modified genetic properties. Alternatively new crops with improved stress tolerance can be obtained that, due to their optimized genetic “equipment”, yield harvested product of better quality than products that were not able to tolerate comparable adverse developmental conditions.

In seeds production germination quality and uniformity of seeds are essential product characteristics, whereas germination quality and uniformity of seeds harvested and sold by the farmer is not important. As it is difficult to keep a crop free from other crop and weed seeds, to control seedborne diseases, and to produce seed with good germination, fairly extensive and well-defined seed production practices have been developed by seed producers, who are experienced in the art of growing, conditioning and marketing of pure
seed. Thus, it is common practice for the farmer to buy certified seed meeting specific quality standards instead of using seed harvested from his own crop. Propagation material to be used as seeds is customarily treated with a protectant coating comprising herbicides, insecticides, fungicides, bactericides, nematicides, molluscicides or mixtures thereof. Customarily used protectant coatings comprise compounds such as captan, carboxin, thiram (TMTD®), methalaxyl (Apron®), and pirimiphos-methyl (Actellic®). If desired these compounds are formulated together with further carriers, surfactants or application-promoting adjuvants customarily employed in the art of formulation to provide protection against damage caused by bacterial, fungal or animal pests. The protectant coatings may be applied by impregnating propagation material with a liquid formulation or by coating with a combined wet or dry formulation. Other methods of application are also possible such as treatment directed at the buds or the fruit.

It is thus a further object of the present invention to provide plant propagation material for cultivated plants, but especially plant seed that is treated with an seed protectant coating customarily used in seed treatment.

It is a further aspect of the present invention to provide new agricultural methods such as the methods exemplified above, which are characterized by the use of transgenic plants, transgenic plant material, or transgenic seed according to the present invention. Comprised by the present invention is an agricultural method, wherein a transgenic plant or the progeny thereof is used comprising a chimeric gene according to the invention in an amount sufficient to express herbicide resistant forms of herbicide target proteins in a plant to confer tolerance to the herbicide.

To breed progeny from plants transformed according to the method of the present invention, a method such as that which follows may be used: maize plants produced as described in the examples set forth below are grown in pots in a greenhouse or in soil, as is known in the art, and permitted to flower. Pollen is obtained from the mature tassel and used to pollinate the ears of the same plant, sibling plants, or any desirable maize plant. Similarly, the ear developing on the transformed plant may be pollinated by pollen obtained from the same plant, sibling plants, or any desirable maize plant. Transformed progeny obtained by this method may be distinguished from non-transformed progeny by the presence of the introduced gene(s) and/or accompanying DNA (genotype), or the phenotype conferred. The transformed progeny may similarly be selfed or crossed to other plants, as is normally done with any plant carrying a desirable trait. Similarly, tobacco or other transformed plants produced by this method may be selfed or crossed as is known in
the art in order to produce progeny with desired characteristics. Similarly, other transgenic organisms produced by a combination of the methods known in the art and this invention may be bred as is known in the art in order to produce progeny with desired characteristics.

The modified inhibitor-resistant protox enzymes of the invention have at least one amino acid substitution, addition or deletion relative to their naturally occurring counterpart (i.e. inhibitor-sensitive forms that occur naturally in a plant without being manipulated, either directly via recombinant DNA methodology or indirectly via selective breeding, etc., by man). Amino acid positions that may be modified to yield an inhibitor-resistant form of the protox enzyme, or enhance inhibitor resistance, are indicated in bold type in Table 1A in the context of plant protox-1 sequences from Arabidopsis, maize, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, wheat, and sugar cane. The skilled artisan will appreciate that equivalent changes may be made to any plant protox gene having a structure sufficiently similar to the protox enzyme sequences shown herein to allow alignment and identification of those amino acids that are modified according to the invention to generate inhibitor-resistant forms of the enzyme. Such additional plant protox genes may be obtained using standard techniques as described in International application no. PCT/IB95/00452 filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659 whose relevant parts are herein incorporated by reference.

DNA molecules encoding the herbicide resistant protox coding sequences taught herein may be genetically engineered for optimal expression in a crop plant. This may include altering the coding sequence of the resistance allele for optimal expression in the crop species of interest. Methods for modifying coding sequences to achieve optimal expression in a particular crop species are well known (see, e.g. Perlak et al., Proc. Natl. Acad. Sci. USA 88: 3324 (1991); Koziel et al., Bio/technol. 11: 194 (1993)).

Genetically engineering a protox coding sequence for optimal expression may also include operatively linking the appropriate regulatory sequences (i.e. promoter, signal sequence, transcriptional terminators). Examples of promoters capable of functioning in plants or plant cells (i.e., those capable of driving expression of the associated structural genes such as protox in plant cells) include the cauliflower mosaic virus (CaMV) 19S or 35S promoters and CaMV double promoters; nopaline synthase promoters; pathogenesis-related (PR) protein promoters; small subunit of ribulose bisphosphate carboxylase (ssuRUBISCO) promoters, heat shock protein promoter from Brassica with reference to EPA 0 559 603 (hsp80 promoter), Arabidopsis actin promoter and the SuperMas promoter with reference to WO 95/14098 and the like. Preferred promoters will
be those that confer high level constitutive expression or, more preferably, those that confer specific high level expression in the tissues susceptible to damage by the herbicide.

Preferred promoters are the rice actin promoter (McElroy et al., Mol. Gen. Genet. 231: 150 (1991)), maize ubiquitin promoter (EP 0 342 926; Taylor et al., Plant Cell Rep. 12: 491 (1993)), and the PR-1 promoter from tobacco, Arabidopsis, or maize (see U.S. Patent No. 5,614,395 to Ryals et al., incorporated by reference herein in its entirety). The promoters themselves may be modified to manipulate promoter strength to increase protox expression, in accordance with art-recognized procedures.

The inventors have also discovered that another preferred promoter for use with the inhibitor-resistant protox coding sequences is the promoter associated with the native protox gene (i.e. the protox promoter; see copending, co-owned U.S. Patent Application No. 08/808,323, entitled "Promoters from Protoporphyrinogen Oxidase Genes", incorporated by reference herein in its entirety.) The promoter sequence from an Arabidopsis protox-1 gene is set forth in SEQ ID NO:13, the promoter sequence from a maize protox-1 gene is set forth in SEQ ID NO:14, and the promoter sequence from a sugar beet protox-1 gene is set forth in SEQ ID NO:26.

Since the protox promoter itself is suitable for expression of inhibitor-resistant protox coding sequences, the modifications taught herein may be made directly on the native protox gene present in the plant cell genome without the need to construct a chimeric gene with heterologous regulatory sequences. Such modifications can be made via directed mutagenesis techniques such as homologous recombination and selected for based on the resulting herbicide-resistance phenotype (see, e.g. Example 10, Pazkowski et al., EMBO J. 7: 4021-4026 (1988), and U.S. Patent No. 5,487,992, particularly columns 18-19 and Example 8). An added advantage of this approach is that besides containing the native protox promoter, the resulting modified gene will also include any other regulatory elements, such as signal or transit peptide coding sequences, which are part of the native gene.

In the event of transformation of the nuclear genome, signal or transit peptides may be fused to the protox coding sequence in chimeric DNA constructs of the invention to direct transport of the expressed protox enzyme to the desired site of action. Examples of signal peptides include those natively linked to the plant pathogenesis-related proteins, e.g. PR-1, PR-2, and the like. See, e.g., Payne et al., Plant Mol. Biol. 11:89-94 (1988).

Examples of transit peptides include the chloroplast transit peptides such as those described in Von Heijne et al., Plant Mol. Biol. Rep. 9:104-126 (1991); Mazur et al., Plant Physiol. 85: 1110 (1987); Vorst et al., Gene 65: 59 (1988), and mitochondrial transit
peptides such as those described in Boutry et al., Nature 328:340-342 (1987). Chloroplast and mitochondrial transit peptides are contemplated to be particularly useful with the present invention as protox enzymatic activity typically occurs within the mitochondria and chloroplast. Most preferred for use are chloroplast transit peptides, as inhibition of the protox enzymatic activity in the chloroplasts is contemplated to be the primary basis for the action of protox-inhibiting herbicides (Witkowski and Halling, Plant Physiol. 87: 632 (1988); Lehnen et al., Pestic. Biochem. Physiol. 37: 239 (1990); Duke et al., Weed Sci. 39: 465 (1991)). Also included are sequences that result in localization of the encoded protein to various cellular compartments such as the vacuole. See, for example, Neuhaus et al., Proc. Natl. Acad. Sci. USA 88: 10362-10366 (1991) and Chrispeels, Ann. Rev. Plant Physiol. Plant Mol. Biol. 42: 21-53 (1991). The relevant disclosures of these publications are incorporated herein by reference in their entirety.

Chimeric genes of the invention may contain multiple copies of a promoter or multiple copies of the protox structural genes. In addition, the construct(s) may include coding sequences for markers and coding sequences for other peptides such as signal or transit peptides, each in proper reading frame with the other functional elements in the DNA molecule. The preparation of such constructs are within the ordinary level of skill in the art.

Useful markers include peptides providing herbicide, antibiotic or drug resistance, such as, for example, resistance to hygromycin, kanamycin, G418, gentamycin, lincomycin, methotrexate, glyphosate, phosphinothricin, or the like. These markers can be used to select cells transformed with the chimeric DNA constructs of the invention from untransformed cells. Other useful markers are peptidic enzymes that can be easily detected by a visible reaction, for example a color reaction, for example luciferase, β-glucuronidase, or β-galactosidase.

The method of positive selection of genetically transformed cells into which a desired nucleotide sequence can be incorporated by providing the transformed cells with a selective advantage is herein incorporated by reference as WO 94/20627.

Where a herbicide resistant protox allele is obtained via directed mutation of the native gene in a crop plant or plant cell culture from which a crop plant can be regenerated, it may be moved into commercial varieties using traditional breeding techniques to develop a herbicide tolerant crop without the need for genetically engineering the modified coding sequence and transforming it into the plant. Alternatively, the herbicide resistant gene may be isolated, genetically engineered for optimal expression and then transformed into the desired variety.
Genes encoding altered protox resistant to a protox inhibitor can also be used as selectable markers in plant cell transformation methods. For example, plants, plant tissue or plant cells transformed with a transgene can also be transformed with a gene encoding an altered protox capable of being expressed by the plant. The thus-transformed cells are transferred to medium containing the protox inhibitor wherein only the transformed cells will survive. Protox inhibitors contemplated to be particularly useful as selective agents are the diphenylethers (e.g. acifluorfen, 5-[2-chloro-4-(trifluoromethyl)phenoxy]-2-nitrobenzoic acid; its methyl ester; or oxyfluorfen, 2-chloro-1-(3-ethoxy-4-nitrophenoxy)-4-(trifluorobenzene)), oxidiazoles, (e.g. oxidiazon, 3-[2,4-dichloro-5-(1-methylethoxy)phenyl]-5-(1,1-dimethylethyl)-1,3,4-oxadiazol-2-(3H)-one), cyclic imides (e.g. S-23142, N-(4-chloro-2-fluoro-5-propargyloxyphenyl)-3,4,5,6-tetrahydropthalimide; chlorothalim, N-(4-chlorophenyl)-3,4,5,6-tetrahydropthalimide), phenyl pyraZoles (e.g. TNPP-ethyl, ethyl 2-[1-(2,3,4-trichlorophenyl)-4-nitropyrazol-5-ox])propionate; M&B 39279), pyridine derivatives (e.g. LS 82-556), and phenopylate and its O-phenylpyrroldino- and piperidinocarbamate analogs and bicyclic triazolones as disclosed in the International patent application WO 92/04827; EP 532146).

The method is applicable to any plant cell capable of being transformed with an altered protox-encoding gene, and can be used with any transgene of interest. Expression of the transgene and the protox gene can be driven by the same promoter functional on plant cells, or by separate promoters.

Modified inhibitor-resistant protox enzymes of the present invention are resistant to herbicides that inhibit the naturally occurring protox activity. The herbicides that inhibit protox include many different structural classes of molecules (Duke et al., Weed Sci. 39:465 (1991); Nandhali et al., Pesticide Biochem. Physiol. 43:193 (1992); Matrinne et al., FEBS Lett. 245:35 (1989); Yanase and Andoh, Pesticide Biochem. Physiol. 35:70 (1989)), including the diphenylethers (e.g. acifluorifen, 5-[2-chloro-4-(trifluoromethyl)phenoxy]-2-nitrobenzoic acid; its methyl ester; or oxyfluorfen, 2-chloro-1-(3-ethoxy-4-nitrophenoxy)-4-(trifluorobenzene)), oxidiazoles (e.g. oxidiazon, 3-[2,4-dichloro-5-(1-methylethoxy)phenyl]-5-(1,1-dimethylethyl)-1,3,4-oxadiazol-2-(3H)-one), cyclic imides (e.g. S-23142, N-(4-chloro-2-fluoro-5-propargyloxyphenyl)-3,4,5,6-tetrahydropthalimide; chlorothalim, N-(4-chlorophenyl)-3,4,5,6-tetrahydropthalimide), phenyl pyrazoles (e.g. TNPP-ethyl, ethyl 2-[1-(2,3,4-trichlorophenyl)-4-nitropyrazol-5-ox)propionate; M&B 39279), pyridine derivatives (e.g. LS 82-556), and phenopylate and its O-phenylpyrroldino- and piperidinocarbamate analogs.
The diphenylethers of particular significance are those having the general formula

\[
\begin{array}{c}
\text{CF}_3 \quad \text{Cl} \\
\text{O} \\
\text{Cl} \quad \text{R} \\
\text{NO}_2
\end{array}
\]  
(Formula I)

wherein R equals -COONa (Formula II), -CONHSO\(_2\)CH\(_3\) (Formula III) or -COOCH\(_2\)COOCH\(_2\)H\(_5\) (Formula IV; see Maigrot et al., Brighton Crop Protection Conference-Weeds: 47-51 (1989)).

Additional diphenylethers of interest are those where R equals:

\[
\begin{array}{c}
\text{NOCH}_2\text{COOCH}_3 \\
\text{CCH}_2\text{OCH}_3
\end{array}
\]  
(Formula IVa; see Hayashi et al., Brighton Crop Protection Conference-Weeds: 53-58 (1989)).

An additional diphenylether of interest is one having the formula:

\[
\begin{array}{c}
\text{Cl} \\
\text{O} \\
\text{Cl} \quad \text{NO}_2
\end{array}
\]

(Formula IVb; bifenox, see Dest et al., Proc. Northeast Weed Sci. Conf. 27: 31 (1973)).

A further diphenylether of interest is one having the formula:

\[
\begin{array}{c}
\text{CF}_3 \\
\text{O} \\
\text{Cl}
\end{array}
\]

(Formula IVc; oxyfluorfen; see Yih and Switchenbank, J. Agric. Food Chem., 23: 592 (1975))

Yet another diphenylether of interest is one having the formula:

Also of significance are the class of herbicides known as imides, having the general formula

\[
\begin{align*}
    & \quad \begin{array}{c}
    \text{R}_1 \\
    \text{R}_2 \\
    \text{R}_3 \\
    \text{Q}
    \end{array} \\
    \text{wherein } Q \text{ equals}
\end{align*}
\]

(Formula V)

or

(Formula VI)

or

(Formula VII)

or

(Formula VIII)

or

(Formula IX)
Cl

N

CF₃

CH₃

(Formula IXa)

or

Cl

N

OCHF₂

N

CH₃

(Formula IXb)

(see Hemper et al. (1995) in "Proceedings of the Eighth International Congress of Pesticide Chemistry", Ragdale et al., eds., Amer. Chem. Soc, Washington, D.C., pp.42-48 (1994)); and R₁ equals H, Cl or F, R₂ equals Cl and R₃ is an optimally substituted ether, thioether, ester, amino or alkyl group. Alternatively, R₂ and R₃ together may form a 5 or 6 membered heterocyclic ring. Examples of imide herbicides of particular interest are

F

N

N

Cl

SCH₂CO₂CH₃

(Formula VIIa; fluthiacet-methyl, see Miyazawa et al., Brighton Crop Protection Conference-Weeds, pp. 23-28 (1993))

Cl

Cl

N

OCHF₂

N

CH₃

CH₃SO₂NH

(Formula X sulfentrazone, see Van Saun et al., Brighton Crop Protection Conference-Weeds, pp. 77-82 (1991)).
(Formula XI)

(Formula XII)

(see Miura et al., Brighton Crop Protection Conference-Weeds: 35-40 (1993))

(Formula XIII)

(Formula XIV)

(Formula XV)

(Formula XVI)

The most preferred imide herbicides are those classified as aryluracils and having the general formula

![Formula XVII](image)

wherein R signifies the group (C_{2-6}-alkenyloxy)carbonyl-C_{1-4}-alkyl, as disclosed in U.S. Patent No. 5,183,492, herein incorporated by reference.

Also of significance are herbicides having the general formula:

![Formula XVIII](image)

(see Weiler et al., *Brighton Crop Protection Conference-Weeds*, pp. 29-34 (1993));

![Formula XIX](image)

(see Van Saun et al., *Brighton Crop Protection Conference-Weeds*,
Weeds: pp. 19-22 (1993));

N-substituted pyrazoles of the general formula:

\[
\begin{array}{c}
\text{R}_{2} \\
\text{R}_{1}-\text{N} \\
\text{R}_{3} \\
\text{R}_{4} \\
\text{R}_{5} \\
\text{R}_{6}
\end{array}
\]

(Formula XX)

wherein R₁ is C₁-C₄-alkyl, optionally substituted by one or more halogen atoms;
R₂ is hydrogen, or a C₁-C₄-alkoxy, each of which is optionally substituted by one or more halogen atoms, or
R₁ and R₂ together from the group -(CH₂)ₙ-X-, where X is bound at R₂;
R₃ is hydrogen or halogen,
R₄ is hydrogen or C₁-C₄-alkyl,
R₅ is hydrogen, nitro, cyano or the group -COOR₆ or -CONR₇R₈, and
R₆ is hydrogen, C₁-C₆-alkyl, C₂-C₆-alkenyl or C₂-C₆-alkynyl;

(see international patent publications WO 94/08999, WO 93/10100, and U. S. Patent No. 5,405,829 assigned to Schering);

N-phenylpyrazoles, such as:

\[
\begin{array}{c}
\text{NO}_{2} \\
\text{N} \\
\text{NH}_{2} \\
\text{Cl} \\
\text{CF}_{3}
\end{array}
\]

(Formula XXI; nipyraclofen)


and 3-substituted-2-aryl-4,5,6,7-tetrahydroindazoles (Lyga et al. Pesticide Sci. 42:29-36 (1994)).
Also of significance are phenylpyrazoles of the type described in WO 96/01254 and WO 97/00246, both of which are hereby incorporated by reference. (Formula XXII).

Also of significance are pyridyl pyrazoles such as the following:

\[
\text{CGA 374410} \quad \text{(Formula XXIIIa)}
\]

\[
\text{NOA 401954} \quad \text{(Formula XXIIIb)}
\]

Also of significance are phenylpyrazoles having the general formula:

\[
\text{(Formula XXIV: Fluazolate)}
\]
Additional compounds of significance are described in WO98/33927 and US patent 5,856,495, both of which are incorporated herein by reference in their entirety.

Levels of herbicide that normally are inhibitory to the activity of protox include application rates known in the art, and that depend partly on external factors such as environment, time and method of application. For example, in the case of the imide herbicides represented by Formulae V through IX, and more particularly those represented by Formulae X through XVII, the application rates range from 0.0001 to 10 kg/ha, preferably from 0.005 to 2 kg/ha. This dosage rate or concentration of herbicide may be different, depending on the desired action and particular compound used, and can be determined by methods known in the art.

A further object of the invention is a method for controlling the growth of undesired vegetation that comprises applying to a population of the plant selected from a group consisting of Arabidopsis, sugar cane, soybean, barley, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf and forage grasses, millet, forage and rice and the like an effective amount of a protox-inhibiting herbicide. Preferred is a method for controlling the growth of undesired vegetation, which comprises applying to a population of the selected from the group consisting of selected from the group consisting of soybean, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf grasses and rice an effective amount of a protox-inhibiting herbicide. Particularly preferred is a method for controlling the growth of undesired vegetation, which comprises applying to a population of the selected from the group consisting of Arabidopsis, soybean, cotton, sugar beet, oilseed rape, maize, wheat, sorghum, and rice.

The present invention further encompasses a chimeric gene comprising a promoter capable of expression in a plant plastid operatively linked to a DNA molecule of the present invention. A preferred promoter capable of expression in a plant plastid is a promoter isolated from the 5' flanking region upstream of the coding region of a plastid gene, which may come from the same or a different species, and the native product of which is typically found in a majority of plastid types including those present in non-green tissues. Examples of such promoters are promoters of clpP genes, such as the tobacco clpP gene promoter (WO 97/06250, incorporated herein by reference) and the Arabidopsis clpP gene promoter (U.S. Application No. 09/038,878, incorporated herein by reference). Other promoters that
are capable of expressing a DNA molecule in plant plastids are promoters recognized by viral RNA polymerases. Preferred promoters of this type are promoters recognized by a single sub-unit RNA polymerase, such as the T7 gene 10 promoter, which is recognized by the bacteriophage T7 DNA-dependent RNA polymerase. Yet another promoter that is capable of expressing a DNA molecule in plant plastids comes from the regulatory region of the plastid 16S ribosomal RNA operon (Harris et al., Microbiol. Rev. 58:700-754 (1994), Shinozaki et al., EMBO J. 5:2043-2049 (1986), both of which are incorporated herein by reference). The gene encoding the T7 polymerase is preferably transformed into the nuclear genome and the T7 polymerase is targeted to the plastids using a plastid transit peptide. Expression of the DNA molecules in the plastids can be constitutive or can be inducible. These different embodiment are extensively described in WO 98/11235, incorporated herein by reference. The chimeric gene preferably further comprises a 5' untranslated sequence (5' UTR) functional in plant plastids and a plastid gene 3' untranslated sequence (3' UTR) operatively linked to a DNA molecule of the present invention. Preferably, the 3' UTR is a plastid rps16 gene 3' untranslated sequence. In a further embodiment, the chimeric gene comprises a poly-G tract instead of a 3' untranslated sequence.

The present invention also encompasses a plastid transformation vector comprising the chimeric gene described above and flanking regions for integration into the plastid genome by homologous recombination. The plastid transformation vector may optionally comprise at least one chloroplast origin of replication. The present invention also encompasses a plant plastid transformed with such a plastid transformation vector, wherein the DNA molecule is expressible in the plant plastid. The invention also encompasses a plant or plant cell, including the progeny thereof, comprising this plant plastid. In a preferred embodiment, the plant is homoplasmic for transgenic plastids. The plants transformed in the present invention may be monocots or dicots. A preferred monocot is maize and a preferred dicot is tobacco. Other preferred dicots are tomato and potato.

Plastid transformation, in which genes are inserted by homologous recombination into some or all of the several thousand copies of the circular plastid genome present in each plant cell, takes advantage of the enormous copy number advantage over nuclear-expressed genes to permit expression levels that may exceed 10% of the total soluble plant protein. In addition, plastid transformation is desirable because in most plants plastid-encoded traits are not pollen transmissible; hence, potential risks of inadvertent transgene escape to wild relatives of transgenic plants is obviated. Plastid transformation technology is

The basic technique for tobacco chloroplast transformation involves the particle bombardment of leaf tissue or PEG-mediated uptake of plasmid DNA in protoplasts with regions of cloned plastid DNA flanking a selectable antibiotic resistance marker. The 1 to 1.5 kb flanking regions, termed “targeting sequences,” facilitate homologous recombination with the plastid genome and thus allow the replacement or modification of specific regions of the 156 kb tobacco plastid genome. Initially, point mutations in the chloroplast 16S rDNA and rps12 genes conferring resistance to spectinomycin and/or streptomycin were utilized as selectable markers for transformation (Svab, Z., Hajdukiewicz, P., and Maliga, P. (1990) Proc. Natl. Acad. Sci. USA 87, 8526-8530; Staub, J. M., and Maliga, P. (1992) Plant Cell 4, 39-45, incorporated herein by reference). This resulted in stable homoplasmic transformants at a frequency of approximately one per 100 bombardments of target leaves. The presence of cloning sites between these markers allowed creation of a plastid targeting vector for introduction of foreign genes (Staub, J.M., and Maliga, P., EMBO J. 12: 601-606 (1993), incorporated herein by reference). Substantial increases in transformation frequency were obtained by replacement of the recessive rRNA or r-protein antibiotic resistance genes with a dominant selectable marker, the bacterial aadA gene encoding the spectinomycin-detoxifying enzyme aminoglycoside-3'-adenyltransferase (Svab, Z., and Maliga, P. (1993) Proc. Natl. Acad. Sci. USA 90, 913-917, incorporated herein by reference). Previously, this marker had been used successfully for high-frequency transformation of the plastid genome of the green alga Chlamydomonas reinhardtii (Goldschmidt-Clermont, M. (1991) Nucl. Acids Res. 19, 4083-4089, incorporated herein by reference). Recently, plastid transformation of protoplasts from tobacco and the moss Physcomitrella patens has been attained using polyethylene glycol (PEG) mediated DNA uptake (O'Neill et al. (1993) Plant J. 3: 729-738; Koop et al. (1996) Planta 199: 193-201, both of which are incorporated herein by reference).
In a preferred embodiment, the present invention encompasses a chimeric gene comprising a promoter capable of expression in a plant plastid operatively linked to a DNA molecule isolated from a prokaryote or a eukaryote that encodes a native or modified protox enzyme, such as a DNA molecule that encodes a native or modified wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, or sugar cane protox enzyme. Such a DNA molecule is comprised in a plastid transformation vector as described above and plants homoplasmic for transgenic plastid genomes are produced. Expression in plant plastids of a DNA molecule that encodes a modified protox enzyme preferably confers upon the plant tolerance to a herbicide in amounts that inhibit naturally occurring protox activity.

In a further preferred embodiment, the present invention encompasses a chimeric gene comprising (a) a DNA molecule isolated from a plant, which in its native state encodes a polypeptide that comprises a plastid transit peptide, and a mature enzyme that is natively targeted to a plastid of the plant by the plastid transit peptide, wherein the DNA molecule is modified such that it does not encode a functional plastid transit peptide; and (b) a promoter capable of expressing the DNA molecule in a plastid, wherein the promoter is operatively linked to the DNA molecule. In one preferred embodiment, the transit peptide is mutated and thus does not allow the proper transport of the enzyme encoded by the DNA molecule to the desired cell compartment, such as the plastid. In another preferred embodiment, a portion of the transit peptide coding sequence or the entire transit peptide coding sequence is removed from the DNA molecule, preventing the enzyme from being properly targeted to the desired cell compartment.

The chimeric genes described above are inserted in plastid transformation vectors, and the present invention is therefore also directed to plants having their plastid genome transformed with such vectors, whereby the DNA molecule is expressible in plant plastids. Such plants are preferably homoplasmic for transgenic plastids.

In a preferred embodiment, a DNA molecule described immediately above encodes an enzyme that in its wild-type form is inhibited by a herbicide. In a further preferred embodiment, the DNA molecule encodes an enzyme that in its wild-type form is inhibited by a herbicide, but that comprises at least one amino acid change compared to the wild-type enzyme. Such an amino acid change makes the enzyme resistant to compounds that naturally inhibit the wild-type enzyme. In a further preferred embodiment, the DNA molecule encodes an enzyme having protoporphyrinogen oxidase (protox) activity. In a further preferred embodiment, the transit peptide is removed from the DNA molecule as further illustrated in Examples 37-42. Plants homoplasmic for transgenic plastids of the invention
are resistant to high amounts of herbicides such as Formula XVII that inhibit the naturally occurring protox activity (as further illustrated in Example 44).

In another preferred embodiment, the transit peptide of a DNA molecule encoding a 5-enolpyruvyl-3-phosphoshikimate synthase (EPSP synthase) is mutated or removed. The resulting DNA molecule is fused to a promoter capable of expression in plant plastids and homoplasmic plants harboring such constructs in their plastid genomes are obtained. These plants are resistant to herbicidal compounds that naturally inhibit EPSP synthase, in particular glyphosate. In another preferred embodiment, the transit peptide of a DNA molecule encoding an acetoacetyl-CoA thiolase (ALS) is mutated or removed. The resulting DNA molecule is fused to a promoter capable of expression in plant plastids and homoplasmic plants harboring such constructs in their plastid genome are obtained. These plants are resistant to herbicidal compounds that naturally inhibit ALS, in particular sulfonylureas. In another preferred embodiment, the transit peptide of a DNA molecule encoding a acetoxyhydroxyacid synthase (AHAS) is mutated or removed. The resulting DNA molecule is fused to a promoter capable of expression in plant plastids and homoplasmic plants harboring such constructs in their plastid genome are obtained. These plants are resistant to herbicidal compounds that naturally inhibit AHAS, in particular, imidazolinone and sulfonamide herbicides. In another preferred embodiment, the transit peptide of a DNA molecule encoding an acetyl-CoA carboxylase (ACCase) is mutated or removed. The resulting DNA molecule is fused to a promoter capable of expression in plant plastids and homoplasmic plants harboring such constructs in their plastid genome are obtained. These plants are resistant to herbicidal compounds that naturally inhibit ACCase, in particular cyclohexanedione and aryloxypropanoic acid herbicides. In another preferred embodiment, the transit peptide of a DNA molecule encoding a glutamine synthase (GS) is mutated or removed. The resulting DNA molecule is fused to a promoter capable of expression in plant plastids and homoplasmic plants harboring such constructs in their plastid genome are obtained. These plants are resistant to herbicidal compounds that naturally inhibit GS, in particular phosphinothricin and methionine sulfoximine.

The present invention is also further directed to a method of obtaining herbicide-resistant plants by transforming their plastid genome with a chimeric gene comprising (a) a DNA molecule isolated from a plant, which in its native state encodes a polypeptide that comprises a plastid transit peptide, and a mature enzyme that is natively targeted to a
plastid of the plant by the plastid transit peptide, wherein the DNA molecule is modified such that it does not encode a functional plastid transit peptide; and (b) a promoter capable of expressing the DNA molecule in a plastid, wherein the promoter is operatively linked to the DNA molecule. Examples of enzymes that are used in the present invention are cited immediately above, but the applicability of such a method is not limited to the cited examples.

The present invention is still further directed to a novel method for selecting a transplastomic plant cell, comprising the steps of: introducing the above-described chimeric gene into the plastome of a plant cell; expressing the encoded enzyme in the plastids of the plant cell; and selecting a cell that is resistant to a herbicidal compound that naturally inhibits the activity of the enzyme, whereby the resistant cell comprises transformed plastids. In a preferred embodiment, the enzyme is naturally inhibited by a herbicidal compound and the transgenic plant is able to grow on an amount of the herbicidal compound that naturally inhibits the activity of the enzyme. In a further preferred embodiment, the enzyme has protoporphyrinogen oxidase (protox) activity and is modified so that it that confers resistance to protox inhibitors.

A further aspect of the present invention is a novel method for plastid transformation of recalcitrant plants. The methods pioneered for plastid transformation of tobacco and lower plant species rely on non-lethal selection for resistance to antibiotics that preferentially affect the plastid translational apparatus and hence allow photo-heterotrophic transformants to outgrow heterotrophic, non-transformed tissue.

Several factors have likely contributed to the difficulties encountered with plastid transformation of monocots and other dicots. For example, the maize chloroplast 16S ribosomal RNA (rRNA) is naturally resistant to spectinomycin because of the presence of a G at position 1138 in the *Zea mays* 16S rDNA gene (Harris *et al.*, 1994). Thus, utilization of 16s rRNA point mutations that confer spectinomycin and/or streptomycin resistance which have been used successfully as selectable chloroplast markers in *Chlamydomonas* and tobacco (Boynton and Gillham (1993) *in Wu, R. [Ed.] Methods in Enzymology* Vol 217. Academic Press, San Diego, pp. 510-536; Svab *et al.* (1990) Proc. Natl. Acad. Sci. U.S.A. 87: 8526–8530) is not feasible for maize. Natural spectinomycin and streptomycin resistance in maize also obviates the use of the bacterial *aadA* gene encoding aminoglycoside 3”-adenyltransferase, which results in dominant spectinomycin and streptomycin resistance and allows a 100-fold increase in tobacco chloroplast transformation efficiency (Svab and Maliga (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90: 913–
917). Use of kanamycin (the only other antibiotic proven to be useful for chloroplast transformation) is also problematic due to a large excess (ca. 50:1) of nuclear vs. chloroplast-encoded resistance in tobacco following bombardment of the bacterial nptII gene encoding neomycin phosphotransferase (Carrer et al. (1993) Mol. Gen. Genet. 241: 49-56). This has been shown to result from both a high frequency of spontaneous nuclear resistance mutants as well as integration of nptII into the nuclear genome. Since nptII is also a highly effective selectable marker for maize nuclear transformation it is reasonable to expect similar background levels to that observed in tobacco. Spontaneous resistance and a significant excess of selectable marker integration by random, illegitimate recombination into the nuclear genome, rather than homologous integration into the chloroplast genome, would make recovery of bona fide chloroplast transformants difficult if not impossible.

A more fundamental reason for the difficulties encountered with plastid transformation in plant species other than tobacco may have to do with the non-photosynthetic nature of many regenerable cultured plant tissues, especially in maize and Arabidopsis. Tobacco is an exception in that cultured vegetative tissues are regenerable and contain mature differentiated chloroplasts that are photosynthetically competent in the presence of sucrose. Consequently, the current system for selecting tobacco plastid transformants relies on the faster growth rate of transformed cells that can use both reduced and inorganic carbon sources. Moreover, transformed cells do not suffer the chloroplast membrane damage that results from inhibition of plastid protein synthesis in the light. This expression of selectable markers that act preferentially on photosynthetic cells, driven by promoters that have high activity in differentiated chloroplasts, is unlikely to work in non-green tissues containing proplastids (e.g. dark-grown maize Type I callus, somatic embryos) or amyloplasts/leucoplasts (e.g. Arabidopsis root cultures). Plastid transformation in these plants requires a selectable marker that gives strong selection in all plastid types.

A preferred selectable marker for generalized plastid transformation: (1) is active only in the plastid to eliminate nuclear-transformed "escapes"; (2) has a mode of action that does not depend on photosynthetic competence or the presence of fully differentiated chloroplasts; and (3) has a level of resistance that is co-dependent on an adjustable external parameter (e.g. light), rather than being determined solely by the bulk concentration of a selective agent, so that selection pressure can vary during selection to facilitate segregation of the many-thousand plastid genome copies.

In a preferred embodiment, such a selectable marker gene involves the use of a chimeric gene comprising an isolated DNA molecule encoding a plastid-targeted enzyme
having in its natural state a plastid transit peptide, wherein the DNA molecule is modified such that the transit peptide either is absent or does not function to target the enzyme to the plastid, wherein the DNA molecule is operatively linked to a promoter capable of expression in plant plastids. In a preferred embodiment, a DNA molecule of the present invention encodes an enzyme that is naturally inhibited by a herbicide. In another preferred embodiment, the DNA molecule encodes a protoporphyrinogen IX oxidase ("protox"). In a preferred embodiment, the protoporphyrinogen IX oxidase gene is from Arabidopsis thaliana and in a more preferred embodiment, the protoporphyrinogen IX oxidase gene is from Arabidopsis thaliana and comprises at least one amino acid substitution. Preferably, an amino acid substitution results in tolerance of the enzyme against inhibition by an herbicide which naturally inhibits the activity of the enzyme. Low concentrations of herbicide are thought to kill wildtype plants due to light-sensitive intermediates which build up when the plastid-localized protox enzyme is inhibited. Production of these photosensitizing compounds does not require differentiated chloroplasts or active photosynthesis, which is a key factor for successful plastid transformation of plants whose regenerable cultured tissues are of non-photosynthetic nature.

Another key feature is to have expression of the selectable marker gene in non-green plastids. In a preferred embodiment, the invention encompasses the use of promoters that are capable of expression of operatively linked DNA molecules in plastids of both green and non-green tissue. In particular, one such promoter comes from the regulatory region of the plastid 16S ribosomal RNA operon. Another candidate is the promoter and 5' UTR from the plastid clpP gene. The clpP gene product is expressed constitutively in plastids from all plant tissues, including those that do not contain chloroplasts (Shanklin (1995) Plant Cell 7: 1713-22).

Other DNA molecules may be co-introduced in plant plastids using the method described above. In a preferred embodiment, a plastid transformation vector of the present invention contains a chimeric gene allowing for selection of transformants as described above and at least one other gene fused to a promoter capable of expression in plant plastids. The other such gene may, for example, confer resistance to insect pests, or to fungal or bacterial pathogens, or may encode one or more value-added traits.

BRIEF DESCRIPTION OF THE SEQUENCES IN THE SEQUENCE LISTING
SEQ ID NO:1: DNA coding sequence for an *Arabidopsis thaliana* protox-1 protein.
SEQ ID NO:2: *Arabidopsis* protox-1 amino acid sequence encoded by SEQ ID NO:1.
SEQ ID NO:3: DNA coding sequence for an *Arabidopsis thaliana* protox-2 protein.
SEQ ID NO:4: *Arabidopsis* protox-2 amino acid sequence encoded by SEQ ID NO:3.
SEQ ID NO:5: DNA coding sequence for a maize protox-1 protein.
SEQ ID NO:6: Maize protox-1 amino acid sequence encoded by SEQ ID NO:5.
SEQ ID NO:7: DNA coding sequence for a maize protox-2 protein.
SEQ ID NO:8: Maize protox-2 amino acid sequence encoded by SEQ ID NO:7.
SEQ ID NO:9: Partial DNA coding sequence for a wheat protox-1 protein.
SEQ ID NO:10: Partial wheat protox-1 amino acid sequence encoded by SEQ ID NO:9.
SEQ ID NO:11: DNA coding sequence for a soybean protox-1 protein.
SEQ ID NO:12: Soybean protox-1 protein encoded by SEQ ID NO:11.
SEQ ID NO:13: Promoter sequence from *Arabidopsis thaliana* protox-1 gene.
SEQ ID NO:14: Promoter sequence from maize protox-1 gene.
SEQ ID NO:15: DNA coding sequence for a cotton protox-1 protein.
SEQ ID NO:16: Cotton protox-1 amino acid sequence encoded by SEQ ID NO:15.
SEQ ID NO:17: DNA coding sequence for a sugar beet protox-1 protein.
SEQ ID NO:18: Sugar beet protox-1 amino acid sequence encoded by SEQ ID NO:17.
SEQ ID NO:19: DNA coding sequence for an oilseed rape protox-1 protein.
SEQ ID NO:20: Oilseed rape protox-1 amino acid sequence encoded by SEQ ID NO:19.
SEQ ID NO:21: Partial DNA coding sequence for a rice protox-1 protein.
SEQ ID NO:22: Partial rice protox-1 amino acid sequence encoded by SEQ ID NO:21.
SEQ ID NO:23: Partial DNA coding sequence for a sorghum protox-1 protein.
SEQ ID NO:24: Partial sorghum protox-1 amino acid sequence encoded by SEQ ID NO:23.
SEQ ID NO:25: Maize protox-1 intron sequence.
SEQ ID NO:26: Promoter sequence from sugar beet protox-1 gene.
SEQ ID NO:27: Pclp_P1a - plastid *clpP* gene promoter top strand PCR primer.
SEQ ID NO:28: Pclp_P1b - plastid *clpP* gene promoter bottom strand PCR primer.
SEQ ID NO:29: Pclp_P2b - plastid *clpP* gene promoter bottom strand PCR primer.
SEQ ID NO:30: Trps16_P1a - plastid *rps16* gene top strand PCR primer.
SEQ ID NO:31: Trps16_p1b - plastid *rps16* gene bottom strand PCR primer.
SEQ ID NO:34: APRTXP1a - top strand PCR primer.
SEQ ID NO:35: APRTXP1b - bottom strand PCR primer.
SEQ ID NO:36: Partial DNA coding sequence for a sugar cane protox-1 protein.
SEQ ID NO:37: Partial sugar cane protox-1 amino acid sequence encoded by SEQ ID NO:36.

1. DEPOSITS

The following vector molecules have been deposited with Agricultural Research Service, Patent Culture Collection (NRRL), Northern Regional Research Center, 1815 North University Street, Peoria, Illinois 61604, U.S.A on the dates indicated below:

Wheat protox-1a, in the pBluescript SK vector, was deposited March 19, 1996, as pWDC-13 (NRRL #B21545).

Soybean protox-1, in the pBluescript SK vector, was deposited December 15, 1995 as pWDC-12 (NRRL #B-21516).

Cotton protox-1, in the pBluescript SK vector, was deposited July 1, 1996 as pWDC-15 (NRRL #B-21594).

Sugar beet protox-1, in the pBluescript SK vector, was deposited July 29, 1996, as pWDC-16 (NRRL #B-21595N).

Oilseed rape protox-1, in the pBluescript SK vector, was deposited August 23, 1996, as pWDC-17 (NRRL #B-21615).

Rice protox-1, in the pBluescript SK vector, was deposited December 6, 1996, as pWDC-18 (NRRL #B-21648).

Sorghum protox-1, in the pBluescript SK vector, was deposited December 6, 1996, as pWDC-19 (NRRL #B-21649).

Resistant mutant pAraC-2Cys, in the pMut-1 plasmid, was deposited on November 14, 1994 under the designation pWDC-7 with the Agricultural Research Culture Collection and given the deposit designation NRRL #21339N.

AraPT1Pro containing the Arabidopsis protox-1 promoter was deposited December 15, 1995, as pWDC-11 (NRRL #B-21515).

A plasmid containing the maize protox-1 promoter fused to the remainder of the maize protox-1 coding sequence was deposited March 19, 1996 as pWDC-14 (NRRL #B-21546).

A plasmid containing the sugar beet protox-1 promoter was deposited December 6, 1996, as pWDC-20 (NRRL #B-21650).
EXAMPLES

The invention will be further described by reference to the following detailed examples. These examples are provided for purposes of illustration only, and are not intended to be limiting unless otherwise specified. Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by Ausubel (ed.), *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc. (1994); T. Maniatis, E. F. Fritsch and J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor laboratory, Cold Spring Harbor, NY (1989); and by T.J. Silhavy, M.L. Berman, and L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984).

Section A. Isolation And Characterization Of Plant
Protoporphyrinogen Oxidase (Protox) Genes

Example 1: Isolation of a Wheat Protox-1 cDNA Based on Sequence Homology to a Maize Protox-1 Coding Sequence

Total RNA prepared from *Triticum aestivum* (cv Kanzler) was submitted to Clontech for custom cDNA library construction in the Lambda Uni-Zap vector. Approximately 50,000 pfu of the cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5; see Example 2 of International application no. PCT/IB95/00452, filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659) labeled with 32P-dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50°C. Wash conditions were 2X SSC, 1% SDS at 50°C. (Church and Gilbert, *Proc. Natl. Acad. Sci. USA* 81: 1991-1995 (1984), hereby incorporated by reference in its entirety.) Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest wheat protox-1 cDNA obtained from initial screening efforts, designated "wheat protox-1", was 1489-bp in length. Wheat protox-1 lacks coding
sequence for the transit peptide plus approximately 126 amino acids of the mature coding sequence based on comparison with the other known plant protox peptide sequences.

A second screen was performed to obtain a longer wheat protox cDNA. For this screen, a *Triticum aestivum* (cv Kanzler) cDNA library was prepared internally using the lambda Uni-Zap vector. Approximately 200,000 pfu of the cDNA library was screened as indicated above, except that the wheat protox-1 cDNA was used as a probe and hybridization and wash conditions were at 65 °C instead of 50 °C. The longest wheat cDNA obtained from this screening effort, designated "wheat protox-1a", was 1811-bp in length. The nucleotide sequence of this cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs:9 and 10, respectively. Based on comparison with the other known plant protox peptide sequences and with corresponding genomic sequence, this cDNA is either full-length or missing only a few transit peptide codons (Table 1A). This wheat protein sequence is 91% identical (95% similar) to the maize protox-1 protein sequence set forth in SEQ ID NO:6.

Wheat protox-1a, in the pBluescript SK vector, was deposited March 19, 1996, as pWDC-13 (NRRL #B21545).

Example 2: Isolation of a Soybean Protox-1 cDNA Based on Sequence Homology to an *Arabidopsis* Protox-1 Coding Sequence

A Lambda Uni-Zap cDNA library prepared from soybean (v Williams 82, epicotyls) was purchased from Stratagene. Approximately 50,000 pfu of the library was plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto Colony/Plaque Screen membranes (NEN Dupont). The plaque lifts were probed with the *Arabidopsis* protox-1 cDNA (SEQ ID NO:1; see Example 1 of International application no. PCT/IB95/00452, filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659)) labeled with 32P-dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50 °C. Wash conditions were 2X SSC, 1% SDS at 50 °C. (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequence of the cDNA inserts was determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest soybean cDNA obtained, designated "soybean protox-1", is full-length based on comparison with the other known plant protox peptide sequences (Table 1A). Soybean protox-1 is
1847-bp in length and encodes a protein of 58.8 kDa. The nucleotide sequence of this cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs:11 and 12, respectively. The soybean protein is 78% identical (87% similar) to the *Arabidopsis* protox-1 protein.

Soybean protox-1, in the pBluescript SK vector, was deposited December 15, 1995 as pWDC-12 (NRRL #B-21516).

Example 3: Isolation of a Cotton Protox-1 cDNA Based on Sequence Homology to a Maize Protox-1 Coding Sequence

A Lambda Uni-Zap cDNA library prepared from *Gossypium hirsutum* L. (72 hr. dark grown cotyledons) was obtained from Dr. Dick Trelease, Dept. of Botany, Arizona State University (Ni W. and Trelease R.N., *Arch. Biochem. Biophys.* 289: 237-243 (1991)). Approximately 50,000 pfu of the library was plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto Colony/Plaque Screen membranes (NEN Dupont). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5) labeled with 32P-dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50 °C. Wash conditions were 2X SSC, 1% SDS at 50 °C. (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequence of the cDNA inserts was determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest cotton cDNA obtained, designated "cotton protox-1", appears to be full-length based on comparison with the other known plant protox peptide sequences (Table 1A).

Cotton protox-1 is 1826-bp in length and encodes a protein of 58.2 kDa. The nucleotide sequence of this cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs:13 and 14, respectively. The cotton protein is 77% identical (86% similar) to the maize protox-1 protein.

Cotton protox-1, in the pBluescript SK vector, was deposited July 1, 1996 as pWDC-15 (NRRL #B-21594).
Example 4: Isolation of a Sugar Beet Protox-1 cDNA Based on Sequence Homology to an Arabidopsis Protox-1 Coding Sequence

A Lambda-Zap cDNA library prepared from Beta vulgaris was obtained from Dr. Philip Rea, Dept. of Botany, Plant Science Institute, Philadelphia, PA (Yongcheol Kim, Eugene J. Kim, and Philip A. Rea, Plant Physiol. 106: 375-382 (1994)). Approximately 50,000 pfu of the cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the Arabidopsis protox-1 cDNA (SEQ ID NO:1) labeled with 32P-dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50°C. Wash conditions were 2X SSC, 1% SDS at 50°C. (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest sugar beet protox-1 cDNA obtained, designated "sugar beet protox-1", is full-length based on comparison with the other known plant protox peptide sequences (Table 1A). Sugar beet protox-1 is 1910-bp in length and encodes a protein of 60 kDa. The nucleotide sequence of this cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOS:15 and 16, respectively. The sugar beet protein is 73% identical (82% similar) to the Arabidopsis protox-1 protein.

Sugar beet protox-1, in the pBluescript SK vector, was deposited July 29, 1996, as pWDC-16 (NRRL #B-21595N).

Example 5: Isolation of an Oilseed Rape Protox-1 cDNA Based on Sequence Homology to an Arabidopsis Protox-1 Coding Sequence

A Lambda Uni-Zap II cDNA library prepared from Brassica napus (3-4 wk. mature green leaves) was obtained from Dr. Guenther Ochs, Institut Fuer Allgemeine Botanik, Johannes Gutenberg-Universitaet Mainz, Germany (Guenther Ochs, Gerald Schock, and Aloysius Wild, Plant Physiol. 103: 303-304 (1993)). Approximately 50,000 pfu of the cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the Arabidopsis protox-1 cDNA (SEQ ID NO:1) labeled with
32P-dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50°C. Wash conditions were 2X SSC, 1% SDS at 50°C. (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest oilseed rape protox-1 cDNA obtained, designated "rape protox-1", is full-length based on comparison with the other known plant protox peptide sequences (Table 1A). Rape protox-1 is 1784-bp in length and encodes a protein of 57.3kD. The nucleotide sequence of this cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs: 17 and 18, respectively. The oilseed rape protein is 87% identical (92% similar) to the Arabidopsis protox-1 protein.

Rape protox-1, in the pBluescript SK vector, was deposited August 23, 1996, as pWDC-17 (NRRL #B-21615).

Example 6: Isolation of a Rice Protox-1 cDNA Based on Sequence Homology to a Maize Protox-1 Coding Sequence

A Lambda gt11 cDNA library prepared from Oryza sativa (5 day etiolated shoots) was purchased from Clontech. Approximately 50,000 pfu of the cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5) labeled with 32P-dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50°C. Wash conditions were 2X SSC, 1% SDS at 50°C. (Church and Gilbert (1984)). Positively hybridizing plaques were purified, and lambda DNA was prepared using the Wizard Lambda-Prep kit (Promega). The cDNA inserts were subcloned as EcoRI fragments into the pBluescript SK vector using standard techniques. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest rice protox-1 cDNA obtained, designated "rice protox-1", was 1224-bp in length. Rice protox-1 lacks coding sequence for the transit peptide plus approximately 172 amino acids of the mature coding sequence based on comparison with the other known plant...
protox peptide sequences (Table 1A). The nucleotide sequence of this partial cDNA and
the amino acid sequence it encodes are set forth in SEQ ID NOs:19 and 20, respectively.

Rice protox-1, in the pBluescript SK vector, was deposited December 6, 1996, as
pWDC-18 (NRRL #B-21648).

Example 7: Isolation of a Sorghum Protox-1 cDNA Based on Sequence Homology to a
Maize Protox-1 Coding Sequence

A Lambda-Zap II cDNA library prepared from Sorghum bicolor (3-6 day green
seedlings) was obtained from Dr. Klaus Pfizenmaier, Institute of Cell Biology and
Immunology, University of Stuttgart, Germany (Harald Wajant, Karl-Wolfgang Mundry, and
Klaus Pfizenmaier, Plant Mol. Biol. 26: 735-746 (1994)). Approximately 50,000 pfu of the
cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and
duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The
plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5) labeled with 32P-
dCTP by the random priming method (Life Technologies). Hybridization conditions were 7%
sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50 °C. Wash
conditions were 2X SSC, 1% SDS at 50 °C. (Church and Gilbert (1984)). Positively
hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The
sequences of the cDNA inserts were determined by the chain termination method using
dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest
sorghum protox-1 cDNA obtained, designated "sorghum protox-1", was 1590-bp in length.
Sorghum protox-1 lacks coding sequence for the transit peptide plus approximately 44
amino acids of the mature coding sequence based on comparison with the other known
plant protox peptide sequences (Table 1A). The nucleotide sequence of this partial cDNA
and the amino acid sequence it encodes are set forth in SEQ ID NOs:21 and 22,
respectively.

Sorghum protox-1, in the pBluescript SK vector, was deposited December 6, 1996,
as pWDC-19 (NRRL #B-21649).
Example 8: Isolation of a Sugar Cane Protox-1 cDNA Based on Sequence Homology to a Maize Protox-1 Coding Sequence

A Lambda-Zap II cDNA library prepared from sugar cane was obtained from Henrik Albert of USDA/ARS at the Hawaii Agricultural Research Center. Approximately 50,000 pfu of the cDNA library were plated at a density of approximately 5,000 pfu per 10 cm Petri dish and duplicate filter lifts were made onto nitrocellulose membranes (Schleicher and Schuell). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5) labeled with 32P-dCTP by the random priming method (Life Technologies). Hybridization conditions were 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50°C. Wash conditions were 2X SSC, 1% SDS at 50°C. (Church and Gilbert (1984)). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. The sequences of the cDNA inserts were determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). The longest sugar cane protox-1 cDNA obtained, designated "sugar cane protox-1", was 633-bp in length. Sugar cane protox-1 lacks coding sequence for the transit peptide plus approximately 382 amino acids of the mature coding sequence based on comparison with the other known plant protox peptide sequences (Table 1A). The nucleotide sequence of this partial cDNA and the amino acid sequence it encodes are set forth in SEQ ID NOs:36 and 37, respectively.

Example 9: Demonstration of Plant Protox Clone Sensitivity to Protox Inhibitory Herbicides in a Bacterial System

Liquid cultures of protox-1/SASX38, protox-2/SASX38 and pBluescript/XL1-Blue were grown in L amp100. One hundred microliter aliquots of each culture were plated on L amp100 media containing various concentrations (1.0nM-10mM) of a protox inhibitory aryluracil herbicide of formula XVII. Duplicate sets of plates were incubated for 18 hours at 37°C.

The protox+ E. coli strain XL1-Blue showed no sensitivity to the herbicide at any concentration, consistent with reported resistance of the native bacterial enzyme to similar herbicides. The protox-1/SASX38 was clearly sensitive, with the lawn of bacteria almost entirely eliminated by inhibitor concentrations as low as 10nM. The protox-2/SASX38 was also sensitive, but only at a higher concentration (10μM) of the herbicide. The herbicide was
effective even on plates maintained almost entirely in the dark. The toxicity of the herbicide was entirely eliminated by the addition of 20 μg/ml hematin to the plates.

The different herbicide tolerance between the two plant protox strains is likely the result of differential expression from these two plasmids, rather than any inherent difference in enzyme sensitivity. Protox-1/SASX38 grows much more slowly than protox-2/SASX38 in any heme-deficient media. In addition, the Mtprotox-2/SASX38 strain, with a growth rate comparable to Arab protox-1/SASX38, is also very sensitive to herbicide at the lower (10-100nM) concentrations.

Section B: Identification and Characterization of Plant Protox Genes Resistant to Protox-Inhibitory Herbicides

Example 10: Selecting for Plant Protox Genes Resistant to Protox-Inhibitory Herbicides in the E. coli Expression System

An Arabidopsis thaliana (Landsberg) cDNA library in the plasmid vector pFL61 (Minet et al., Plant J. 2:417-422 (1992) was obtained and amplified. The E. coli hemG mutant SASX38 (Sasarman et al., J. Gen. Microbiol. 113:297(1979)) was obtained and maintained on L media containing 20μg/ml hematin (United States Biochemicals). The plasmid library was transformed into SASX38 by electroporation using the Bio-Rad Gene Pulser and the manufacturer's conditions. The electroporated cells were plated on L agar containing 100μg/ml ampicillin at a density of approximately 500,000 transformants/10cm plate. The cells were then incubated at 37°C for 40 hours in low light and selected for the ability to grow without the addition of exogenous heme. Heme prototrophs were recovered at a frequency of 400/10^7 from the pFL61 library. Sequence analysis of twenty-two complementing clones showed that nine were of the type designated "protox-1," the protox gene expected to express a chloroplastic protox enzyme.

The pFL61 library is a yeast expression library, with the Arabidopsis cDNAs inserted bidirectionally. These cDNAs can also be expressed in bacteria. The protox cDNAs apparently initiate at an in-frame ATG in the yeast PGK 3' sequence approximately 10 amino acids 5' to the NotI cloning site in the vector and are expressed either from the lacZ promoter 300bp further upstream or from an undefined cryptic bacterial promoter. Because protox-1 cDNAs that included significant portions of a chloroplast transit sequence inhibited
the growth of the E. coli SASX38 strain, the clone with the least amount of chloroplast transit sequence attached was chosen for mutagenesis/herbicide selection experiments. This clone, pSLV19, contains only 17 amino acids of the putative chloroplast transit peptide, with the DNA sequence beginning at bp-151 of the Arabidopsis protox-1 cDNA (SEQ ID NO:1).

The plasmid pSLV19 was transformed into the random mutagenesis strain XL1-Red (Stratagene, La Jolla, CA). The transformation was plated on L media containing 50ug/ml ampicillin and incubated for 48 hours at 37°C. Lawns of transformed cells were scraped from the plates and plasmid DNA prepared using the Wizard Megaprep kit (Promega, Madison, WI). Plasmid DNA isolated from this mutator strain is predicted to contain approximately one random base change per 2000 nucleotides (see Greener et al., Strategies 7(2):32-34 (1994).

The mutated plasmid DNA was transformed into the hemG mutant SASX38 (Sasarman et al., J. Gen. Microbiol. 113:297 (1979) and plated on L media containing various concentrations of protox-inhibiting herbicide (formula XVII). The plates were incubated for 2 days at 37°C. Plasmid DNA was isolated from all colonies that grew in the presence of herbicide concentrations that effectively killed the wild type strain. The isolated DNA was then transformed into SASX38 and plated again on herbicide to ensure that the resistance observed was plasmid-borne. The protox coding sequence from plasmids passing this screen was excised by NotI digestion, recloned into an unmutagenized vector, and tested again for the ability to confer herbicide tolerance. The DNA sequence of protox cDNAs that conferred herbicide resistance was then determined and mutations identified by comparison with the wild type Arabidopsis protox-1 sequence (SEQ ID NO:1).

A single coding sequence mutant was recovered from the first mutagenesis experiment. This mutant leads to enhanced herbicide "resistance" only by increasing growth rate. It contains a C to A mutation at nucleotide 197 in SEQ ID NO:1 in the truncated chloroplast transit sequence of pSLV19, converting an ACG codon for threonine to an AAG codon for lysine at amino acid 56 of SEQ ID NO:2, and resulting in better complementation of the bacterial mutant. This plasmid also contains a silent coding sequence mutation at nucleotide 1059, with AGT (Ser) changing to AGC (Ser). This plasmid was designated pMut-1.

The pMut-1 plasmid was then transformed into the mutator XL1-Red strain as described above and the mutated DNA was isolated and plated on an herbicide concentration that is lethal to the unmutagenized pMut-1 protox gene. Herbicide tolerant
colonies were isolated after two days at 37°C and analyzed as described above. Multiple plasmids were shown to contain herbicide resistant protox coding sequences. Sequence analysis indicated that the resistant genes fell into two classes. One resistance mutation identified was a C to T change at nucleotide 689 in the *Arabidopsis* protox-1 sequence set forth in SEQ ID NO:1. This change converts a GCT codon for alanine at amino acid 220 of SEQ ID NO:2 to a GTT codon for valine, and was designated pAraC-1Val (see, Table 1B; sub-sequence 3).

A second class of herbicide resistant mutant contains an A to G change at nucleotide 1307 in the *Arabidopsis* protox-1 sequence. This change converts a TAC codon for tyrosine at amino acid 426 to a TGC codon for cysteine, and was designated pAraC-2Cys (see, Table 1B; sub-sequence 7).

A third resistant mutant has a G to A change at nucleotide 691 in the *Arabidopsis* protox-1 sequence. This mutation converts a GGT codon for glycine at amino acid 221 to an AGT codon for serine at the codon position adjacent to the mutation in pAraC-1. This plasmid was designated pAraC-3Ser (see, Table 1B; sub-sequence 4).

Resistant mutant pAraC-2Cys, in the pMut-1 plasmid, was deposited on November 14, 1994 under the designation pWDC-7 with the Agricultural Research Culture Collection and given the deposit designation NRRL #21339N.

Example 11: Additional Herbicide-Resistant Codon Substitutions at Positions Identified in the Random Screen

The amino acids identified as herbicide resistance sites in the random screen are replaced by other amino acids and tested for function and for herbicide tolerance in the bacterial system. Oligonucleotide-directed mutagenesis of the *Arabidopsis* protox-1 sequence is performed using the Transformer Site-Directed Mutagenesis Kit (Clontech, Palo Alto, CA). After amino acid changes are confirmed by sequence analysis, the mutated plasmids are transformed into SASX38 and plated on L-amp<sup>100</sup> media to test for function and on various concentrations of protox-inhibiting herbicide to test for tolerance.

This procedure is applied to the alanine codon at nucleotides 688-690 and to the tyrosine codon at nucleotides 1306-1308 of the *Arabidopsis* protox-1 sequence (SEQ ID NO:1). The results demonstrate that the alanine codon at nucleotides 688-690 can be changed to a codon for valine (pAraC-1Val), threonine (pAraC-1Thr), leucine (pAraC-1Leu), cysteine (pAraC-1Cys), or isoleucine (pAraC-1Ile) to yield an herbicide-resistant protox
enzyme that retains function (see, Table 1B; sub-sequence 3). The results further demonstrate that the tyrosine codon at nucleotides 1306-1308 can be changed to a codon for cysteine (pAraC-2Cys), isoleucine (pAraC-2Ile), leucine (pAraC-2Leu), threonine (pAraC-2Thr), methionine (pAraC-2Met), valine (pAraC-2Val), or alanine (pAraC-2Ala) to yield an herbicide-resistant protox enzyme that retains function (see, Table 1B; sub-sequence 7).

Example 12: Isolation of Additional Mutations that Increase Enzyme Function and/or Herbicide Tolerance of Previously Identified Resistant Mutants

Plasmids containing herbicide resistant protox genes are transformed into the mutator strain XL1-Red and mutated DNA is isolated as described above. The mutated plasmids are transformed into SASX38 and the transformants are screened on herbicide concentrations (formula XVII) sufficient to inhibit growth of the original "resistant" mutant. Tolerant colonies are isolated and the higher tolerance phenotype is verified as being coding sequence dependent as described above. The sequence of these mutants is determined and mutations identified by comparison to the progenitor sequence.

This procedure was applied to the pAraC-1Val mutant described above. The results demonstrate that the serine codon at amino acid 305 (SEQ ID NO:2) can be changed to a codon for leucine to yield an enzyme with higher tolerance to protox-inhibiting herbicides than the pAraC-1Val mutant alone. This second site mutation is designated AraC305Leu (see, Table 1B; sub-sequence 13). The same results are demonstrated for the threonine codon at amino acid 249, where a change to either isoleucine or to alanine leads to a more tolerant enzyme (see, Table 1B; sub-sequence 12). These changes are designated AraC249Ile and AraC249Ala, respectively.

The procedure was also applied to the pAraC-2Cys mutant described above. The results demonstrate that the proline codon at amino acid 118 (SEQ ID NO:2) can be changed to a codon for leucine to yield an enzyme with higher tolerance to protox-inhibiting herbicides than the pAraC-1Cys mutant alone. This mutation is designated AraC118Leu (see, Table 1B; sub-sequence 11). The same results are demonstrated for the serine codon at amino acid 305, where a change to leucine leads to a more tolerant pAraC-2Cys enzyme (see, Table 1B; sub-sequence 13). This change was also isolated with the pAraC-1Val mutant as described above and is designated AraC305Leu. Additional mutations that enhance the herbicide resistance of the pAraC-2Cys mutant include an asparagine to serine change at amino acid 425, designated AraC425Ser (Table 1B; sub-sequence 14),
and a tyrosine to cysteine at amino acid 498, designated AraC498Cys (Table 1B; sub-sequence 15).

These changes (Table 1B; sub-sequences 11-15) are referred to as "second site" mutations, because they are not sufficient to confer herbicide tolerance alone, but rather enhance the function and/or the herbicide tolerance of an already mutant enzyme. This does not preclude the possibility that other amino acid substitutions at these sites could suffice to produce an herbicide tolerant enzyme since exhaustive testing of all possible replacements has not been performed.

Example 13: Combining Identified Resistance Mutations with Identified Second Site Mutations to Create Highly Functional/Highly Tolerant Protox Enzymes

The AraC305Leu mutation described above was found to enhance the function/herbicide resistance of both the AraC-1Val and the AraC-2Cys mutant plasmids. In an effort to test the general usefulness of this second site mutation, it was combined with the AraC-2Leu, AraC-2Val, and AraC-2Ile mutations and tested for herbicide tolerance. In each case, the AraC305Leu change significantly increased the growth rate of the resistant protox mutant on protox-inhibiting herbicide. Combinations of the AraC-2Ile resistant mutant with either the second site mutant AraC249Ile or AraC118Leu also produced more highly tolerant mutant protox enzymes. The AraC249Ile mutation demonstrates that a second site mutation identified as enhancing an AraC-1 (sub-sequence 3) mutant may also increase the resistance of an AraC-2 (sub-sequence 7) mutant. A three mutation plasmid containing AraC-2Ile, AraC305Leu, and AraC249Ile (Table 1B; sub-sequences 7, 13, and 12) has also been shown to produce a highly functional, highly herbicide tolerant protox-1 enzyme.

Example 14: Identification of Sites in the Maize Protox-1 Gene that Can Be Mutated to Give Herbicide Tolerance

The pMut-1 Arabidopsis protox -1 plasmid described above is very effective when used in mutagenesis/screening experiments in that it gives a high frequency of genuine coding sequence mutants, as opposed to the frequent up-promoter mutants that are isolated when other plasmids are used. In an effort to create an efficient plasmid screening system for maize protox-1, the maize cDNA was engineered into the pMut-1 vector in
approximately the same sequence context as the *Arabidopsis* cDNA. Using standard methods of overlapping PCR fusion, the 5' end of the pMut-1 *Arabidopsis* clone (including 17 amino acids of chloroplast transit peptide with one mis-sense mutation as described above) was fused to the maize protox-1 cDNA sequence starting at amino acid number 14 of the maize sequence (SEQ ID NO:6). The 3' end of the maize cDNA was unchanged. A *NdeI* restriction site was placed on both ends of this fusion, and the chimeric gene was cloned into the pFL61 plasmid backbone from pMut-1. Sequence analysis revealed a single nucleotide PCR-derived silent mutation that converts the ACG codon at nucleotides 745-747 (SEQ ID NO:5) to an ACT codon, both of which encode threonine. This chimeric Arab-maize protox-1 plasmid was designated pMut-3.

The pMut-3 plasmid was transformed into the mutator XL1-Red strain as described above and the mutated DNA was isolated and plated on a herbicide concentration (formula XVII) that was lethal to the unmutagenized pMut-3 maize protox gene. Herbicide tolerant colonies were isolated after two days at 37°C and analyzed as described above. This analysis revealed multiple plasmids containing herbicide resistant protox coding sequences. Sequence analysis showed 5 single base changes that individually result in an herbicide tolerant maize protox-1 enzyme. Three of these mutations correspond to amino acid changes previously shown to confer tolerance at the homologous position in the *Arabidopsis* protox-1 gene. Two of the three are pMzC-1Val and pMzC-1Thr, converting the alanine (GCT) at amino acid 164 (SEQ ID NO:6) to either valine (GAT) or to threonine (ACT). This position corresponds to the pAraC-1 mutations described above (see, Table 1B; sub-sequence 3). The third analogous change, pMzC-3Ser, converts the glycine (GGT) at amino acid 165 to Serine (AGT), corresponding to the AraC-3Ser mutation described above (see, Table 1B; sub-sequence 4). These results serve to validate the expectation that herbicide-tolerant mutations identified in one plant protox gene may also confer herbicide tolerance in an equivalent plant protox gene from another species.

Two of the mutations isolated from the maize protox-1 screen result in amino acid changes at residues not previously identified as herbicide resistance sites. One change (Mz159Phe) converts cysteine (TGC) to phenylalanine (TTC) at amino acid 159 of the maize protox-1 sequence (SEQ ID NO:6) (see, Table 1B; sub-sequence 2). The second (Mz419Thr) converts isoleucine (ATA) to threonine (ACA) at amino acid 419 (see, Table 1B; sub-sequence 9).

Additional amino acid substitutions were made and tested at three of the maize mutant sites. Tolerance was demonstrated when glycine 165 was changed to leucine
(pMzc-3Leu) or when cysteine 159 was changed to either leucine (Mz159Leu) or to lysine (Mz159Lys) (see, Table 1B; sub-sequences 4 and 2). Tolerant enzymes were also created by changing isoleucine 419 to histidine (Mz419His), glycine (Mz419Gly), or asparagine (Mz419Asn) (see, Table 1B; sub-sequence 9).

Individual amino acid changes that produced highly herbicide tolerant *Arabidopsis* protox-1 enzymes were engineered into the maize protox-1 gene by site-directed mutagenesis as described above. Bacterial testing demonstrated that changing the alanine (GCT) at amino acid 164 (SEQ ID NO:6) to leucine (CTT) produced a highly tolerant maize enzyme (pMzc-1Leu) (see, Table 1B; sub-sequence 3). No mutation analogous to the AraC-2 site (Table 1B; sub-sequence 7) in *Arabidopsis* was isolated in the maize random screen. However, changing this site, tyrosine 370 in the maize enzyme (SEQ ID NO:8), to either isoleucine (pMzc-2Ile) or methionine (pMzc-2Met) did produce herbicide tolerant enzymes (see, Table 1B; sub-sequence 7).

Additional mutant screens performed as described earlier in this example, except using formulas XXIIIa and XXIIIb instead of XVII, identified three additional amino acid changes that confer tolerant protox enzymes. One, using formula XXIIIb, demonstrated that changing the arginine (CGT) at amino acid 88 (SEQ ID NO:6) to cysteine (TGT) produced a highly tolerant maize enzyme (Mz88Cys) (see, Table 1B; sub-sequence 1). Another, using formula IVc, demonstrated that changing the arginine (CGT) at amino acid 88 (SEQ ID NO:6) to leucine (CTT) produced a highly tolerant maize enzyme (Mz88Cys) (see, Table 1B; sub-sequence 1). Another, using formula XXIIIa, demonstrated that changing both the leucine (TTA) at amino acid 347 (SEQ ID NO:6) to serine (TCA) and the alanine (GCA) at amino acid 453 (SEQ ID NO:6) to threonine (ACA) produced a highly tolerant maize enzyme (Mz347Ser453Thr) (see, Table 1B; sub-sequences 16 and 17). Unlike the second site mutations described above, which increase enzyme function and/or herbicide tolerance of previously identified resistant mutants, Mz347Ser453Thr is a “double mutant” that requires that both mutations be present for herbicide tolerance.

Additional mutant screens performed as described earlier in this example identified two additional amino acid changes that confer tolerant protox enzymes. One, using formula XXIV, demonstrated that changing the alanine (GCA) at amino acid 175 (SEQ ID NO:6) to valine (GTA) or to threonine (ACA) produced a highly tolerant maize enzyme (Mz175Val and Mz175Thr, respectively) (see, Table 1B; sub-sequence 18). Another, using formula IVc, demonstrated that changing the leucine (TTG) at amino acid 337 (SEQ ID NO:6) to serine
(TCG) produced a highly tolerant maize enzyme (Mz337Ser) (see, Table 1B; sub-sequence 19).

Example 15: Identification of Sites in the Wheat Protox-1 Gene that can be Mutated to Give Herbicide Tolerance

To create an efficient plasmid screening system for wheat protox-1, the wheat cDNA was engineered into the pMut-1 vector as described above for the maize cDNA. This chimeric Arab-wheat protox-1 plasmid is designated pMut-4. The pMut-4 DNA was mutated and screened for herbicide tolerance as described above. This analysis revealed multiple plasmids containing herbicide resistant protox coding sequences. Sequence analysis showed 7 single base changes that individually result in an herbicide tolerant wheat protox-1 enzyme. Four of these mutations correspond to amino acid changes previously shown to confer tolerance at the homologous position in the Arabidopsis and/or in the maize protox-1 gene. Two, pWhtC-1Val and pWhtC-1Thr, convert the alanine (GCT) at amino acid 211 (SEQ ID NO:10) to valine (GAT) and to threonine (ACT), respectively. This position corresponds to the pAraC-1 mutations described above (see, Table 1B; sub-sequence 3). The third analogous change, pWhtC-3Ser, converts the glycine (GGT) at amino acid 212 to serine (AGT), corresponding to the AraC-3Ser mutation described above (see, Table 1B; sub-sequence 4). The fourth, Wht466Thr, converts isoleucine (ATA) to threonine (ACA) at amino acid 466, corresponding to the Mz419Thr mutant from maize (see, Table 1B; sub-sequence 9).

Three of the mutations isolated from the wheat protox-1 screen result in amino acid changes at residues not previously identified as herbicide resistance sites. One change (Wht356Leu) converts valine (GTT) to leucine (CTT) at amino acid 356 of the wheat protox-1 sequence (SEQ ID NO:10) (see, Table 1B; sub-sequence 6). A second (Wht421Pro) converts serine (TCT) to proline (CCT) at amino acid 421 (see, Table 1B; sub-sequence 8). The third (Wht502Ala) converts valine (GTT) to alanine (GCT) at amino acid 502 (see, Table 1B; sub-sequence 10).
Example 16: Identification of Sites in the Soybean Protox-1 Gene that can be Mutated to Give Herbicide Tolerance

To create an efficient plasmid screening system for soybean protox-1, the soybean cDNA was engineered into the pMut-1 vector as described above for the maize cDNA. This chimeric Arab-soybean protox-1 plasmid is designated pMut-5. The pMut-5 DNA was mutated and screened for herbicide tolerance as described above. This analysis revealed multiple plasmids containing herbicide resistant protox coding sequences. Sequence analysis showed 4 single base changes that individually result in an herbicide tolerant soybean protox-1 enzyme. Two of these mutations correspond to amino acid changes previously shown to confer tolerance at the homologous position in the Arabidopsis and/or in the wheat protox-1 gene. One, pSoyC-1Thr, converts the alanine (GCA) at amino acid 226 (SEQ ID NO:12) to threonine (ACA). This position corresponds to the pAraC-1Thr mutation described above (see, Table 1B; sub-sequence 3). The second analogous change, Soy517Ala, converts the valine (GTT) at amino acid 517 to alanine (GCT), corresponding to the Wht502Ala mutation from wheat (see, Table 1B; sub-sequence 10).

Two of the mutations isolated from the soybean protox-1 screen result in amino acid changes at a residue not previously identified as an herbicide resistance site. One change (Soy369Ser) converts proline (CCT) to serine (TCT) at amino acid 369 of the soybean protox-1 sequence (SEQ ID NO:12) (see, Table 1B; sub-sequence 5). A second (Soy369His) converts this same proline369 to histidine (CAT) (see, Table 1B; sub-sequence 5).

Individual amino acid changes that produced highly herbicide tolerant Arabidopsis protox-1 enzymes were engineered into the soybean protox-1 gene by site directed mutagenesis as described above. Bacterial testing demonstrated that changing the alanine (GCA) at amino acid 226 (SEQ ID NO:12) to leucine (pSoyC-1Leu) produced a tolerant soybean enzyme (see, Table 1B; sub-sequence 3). Changing the tyrosine (TAC) at amino acid 432 (SEQ ID NO:12) to either leucine (pSoyC-2Leu) or isoleucine (pSoyC-2Ile) also produced herbicide tolerant enzymes (see, Table 1B; sub-sequence 7).
Example 17: Identification of Sites in the Sugar Beet Protox-1 Gene that can be Mutated to Give Herbicide Tolerance

To create an efficient plasmid screening system for sugar beet protox-1, the sugar beet cDNA was engineered into the pMut-1 vector as described above for the maize cDNA. This chimeric Arab-sugar beet protox-1 plasmid is designated pMut-6. The pMut-6 DNA was mutated and screened for herbicide tolerance as described above. This analysis revealed multiple plasmids containing herbicide resistant protox coding sequences. Sequence analysis showed a single base change that results in an herbicide tolerant sugar beet protox-1 enzyme. This change (pSugC-2Cys) converts tyrosine (TAC) at amino acid 449 to cysteine (TGC) and is analogous to the AraC-2 mutations in Arabidopsis (see, Table 1B; sub-sequence 7).

Individual amino acid changes that produced highly herbicide tolerant Arabidopsis protox-1 enzymes were engineered into the sugar beet protox-1 gene by site directed mutagenesis as described above. Bacterial testing demonstrated that changing the tyrosine (TAC) at amino acid 449 to leucine (pSugC-2Leu), isoleucine (pSugC-2Ile), valine (pSugC-2Val), or methionine (pSugC-2Met) produced herbicide tolerant sugar beet enzymes (see, Table 1B; sub-sequence 7).

Example 18: Identification of Sites in the Cotton Protox-1 Gene that can be Mutated to Give Herbicide Tolerance

In an effort to create an efficient plasmid screening system for cotton protox-1, the cotton cDNA was engineered into the pMut-1 vector as described above for the maize cDNA. This chimeric Arab-cotton protox-1 plasmid is designated pMut-7. The pMut-7 DNA was mutated and screened for herbicide tolerance as described above. This analysis revealed multiple plasmids containing herbicide resistant protox coding sequences. Sequence analysis showed 3 single base changes that individually result in an herbicide tolerant cotton protox-1 enzyme. Three mutants, pCotC-2Cys, pCotC-2His and pCotC-2Arg, change tyrosine (TAC) at amino acid 428 (SEQ ID NO:16) to cysteine (TGC), histidine (CAC) and to arginine (CGC), respectively (see, Table 1B; sub-sequence 7). Arginine is a novel substitution giving tolerance at this previously identified AraC-2 (sub-sequence 7) site. The third mutation (Cot365Ser) converts proline (CCT) to serine (TCT) at amino acid 365. 

- 72 -
This change corresponds to the soybean mutant Soy369Ser (see, Table 1B; sub-sequence 5).

Example 19: Demonstration of Resistant Mutations' Cross-Tolerance to Various Protox-Inhibiting Compounds

Resistant mutant plasmids, originally identified based on resistance against a single protox inhibitory herbicide, were tested against a spectrum of other protox inhibiting compounds. For this test, the SASX38 strain containing the wild-type plasmid is plated on a range of concentrations of each compound to determine the lethal concentration for each one. Resistant mutant plasmids in SASX38 are plated and scored for the ability to survive on a concentration of each compound at least 10 fold higher than the concentration that is lethal to the SASX38 strain containing the wild-type plasmid.

Results from bacterial cross-tolerance testing, illustrated in Tables 3A and 3B, show that each of the mutations identified confers tolerance to a variety of protox inhibiting compounds.

2. Section C: Expression of Herbicide-Resistant Protox Genes in Transgenic Plants

Example 20: Engineering of Plants Tolerant to Protox-Inhibiting Herbicides by Homologous Recombination or Gene Conversion

Because the described mutant coding sequences effectively confer herbicide tolerance when expressed under the control of the native protox promoter, targeted changes to the protox coding sequence in its native chromosomal location represent an alternative means for generating herbicide tolerant plants and plant cells. A fragment of protox DNA containing the desired mutations, but lacking its own expression signals (either promoter or 3' untranslated region) can be introduced by any of several art-recognized methods (for instance, Agrobacterium transformation, direct gene transfer to protoplasts, microprojectile bombardment), and herbicide-tolerant transformants selected. The introduced DNA fragment also contains a diagnostic restriction enzyme site or other sequence polymorphism that is introduced by site-directed mutagenesis in vitro without
changing the encoded amino acid sequence (i.e. a silent mutation). As has been previously reported for various selectable marker and herbicide tolerance genes (see, e.g., Paszkowski et al., EMBO J. 7: 4021-4026 (1988); Lee et al., Plant Cell 2: 415-425 (1990); Risseeuw et al., Plant J. 7: 109-119 (1995)). some transformants are found to result from homologous integration of the mutant DNA into the protox chromosomal locus, or from conversion of the native protox chromosomal sequence to the introduced mutant sequence. These transformants are recognized by the combination of their herbicide-tolerant phenotype, and the presence of the diagnostic restriction enzyme site in their protox chromosomal locus.

3. Example 21: Construction of Plant Transformation Vectors

Numerous transformation vectors are available for plant transformation, and the genes of this invention can be used in conjunction with any such vectors. The selection of vector for use will depend upon the preferred transformation technique and the target species for transformation. For certain target species, different antibiotic or herbicide selection markers may be preferred. Selection markers used routinely in transformation include the nptII gene, which confers resistance to kanamycin and related antibiotics (Messing & Vierra, Gene 19: 259-268 (1982); Bevan et al., Nature 304:184-187 (1983)), the bar gene, which confers resistance to the herbicide phosphinothricin (White et al., Nucl Acids Res 18: 1062 (1990), Spencer et al. Theor Appl Genet 79: 625-631(1990)), the hph gene, which confers resistance to the antibiotic hygromycin (Blochinger & Diggelmann, Mol Cell Biol 4: 2929-2931), and the dhfr gene, which confers resistance to methotrexate (Bourouis et al., EMBO J. 2(7): 1099-1104 (1983)).

I. Construction of Vectors Suitable for Agrobacterium Transformation

Many vectors are available for transformation using Agrobacterium tumefaciens. These typically carry at least one T-DNA border sequence and include vectors such as pBIN19 (Bevan, Nucl. Acids Res. (1984)) and pXYZ. Below the construction of two typical vectors is described.

Construction of pCIB200 and pCIB2001: The binary vectors pCIB200 and pCIB2001 are used for the construction of recombinant vectors for use with Agrobacterium and was constructed in the following manner. pTJS75kan was created by Nael digestion of pTJS75
(Schmidhauser & Helinski, *J Bacteriol.* 164: 446-455 (1985)) allowing excision of the tetracycline-resistance gene, followed by insertion of an *Acl* fragment from pUC4K carrying an NPTII (Messing & Vierra, *Gene* 19: 259-268 (1982); Bevan *et al.*, *Nature* 304: 184-187 (1983); McBride *et al.*, *Plant Molecular Biology* 14: 266-276 (1990)). *Xhol* linkers were ligated to the *EcoRV* fragment of pCIB7, which contains the left and right T-DNA borders, a plant selectable *nos/nptII* chimeric gene and the pUC polylinker (Rothstein *et al.*, *Gene* 53: 153-161 (1987)), and the *Xhol*-digested fragment was cloned into *SalI*-digested pTJS75kan to create pCIB200 (see also EP 0 332 104, example 19). pCIB200 contains the following unique polylinker restriction sites: *EcoRI*, *SstI*, *KpnI*, *BglII*, *XbaI*, and *SalI*. pCIB2001 is a derivative of pCIB200, which is created by the insertion into the polylinker of additional restriction sites. Unique restriction sites in the polylinker of pCIB2001 are *EcoRI*, *SstI*, *KpnI*, *BglII*, *XbaI*, *SalI*, *MluI*, *BclI*, *AvrII*, *ApaI*, *HpaI*, and *Stul*. pCIB2001, in addition to containing these unique restriction sites also has plant and bacterial kanamycin selection, left and right T-DNA borders for *Agrobacterium*-mediated transformation, the RK2-derived *tra* function for mobilization between *E. coli* and other hosts, and the *OriT* and *OriV* functions also from RK2. The pCIB2001 polylinker is suitable for the cloning of plant expression cassettes containing their own regulatory signals.

Construction of pCIB10 and Hygromycin Selection Derivatives Thereof: The binary vector pCIB10 contains a gene encoding kanamycin resistance for selection in plants, T-DNA right and left border sequences and incorporates sequences from the wide host-range plasmid pRK252 allowing it to replicate in both *E. coli* and *Agrobacterium*. Its construction is described by Rothstein *et al.*, *Gene* 53: 153-161 (1987). Various derivatives of pCIB10 have been constructed that incorporate the gene for hygromycin B phosphotransferase described by Gritz *et al.*, *Gene* 25: 179-188 (1983)). These derivatives enable selection of transgenic plant cells on hygromycin only (pCIB743), or hygromycin and kanamycin (pCIB715, pCIB717).

II. Construction of Vectors Suitable for non-*Agrobacterium* Transformation.

Transformation without the use of *Agrobacterium tumefaciens* circumvents the requirement for T-DNA sequences in the chosen transformation vector and consequently vectors lacking these sequences can be utilized in addition to vectors such as the ones described above that contain T-DNA sequences. Transformation techniques that do not rely on *Agrobacterium* include transformation via particle bombardment, protoplast uptake (*e.g.* PEG and electroporation) and microinjection. The choice of vector depends largely on
the preferred selection for the species being transformed. Below, the construction of some

typical vectors is described.

Construction of pCIB3064: pCIB3064 is a pUC-derived vector suitable for direct gene

transfer techniques in combination with selection by the herbicide basta (or

phosphinothricin). The plasmid pCIB246 comprises the CaMV 35S promoter in operational

fusion to the E. coli GUS gene and the CaMV 35S transcriptional terminator and is

described in the PCT published application WO 93/07278. The 35S promoter of this vector

contains two ATG sequences 5' of the start site. These sites were mutated using standard

PCR techniques in such a way as to remove the ATG's and generate the restriction sites

SspI and PvuII. The new restriction sites were 96 and 37-bp away from the unique Sall site

and 101 and 42-bp away from the actual start site. The resultant derivative of pCIB246 was

designated pCIB3025. The GUS gene was then excised from pCIB3025 by digestion with

Sall and SacI, the termini rendered blunt and religated to generate plasmid pCIB3060. The

plasmid pJIT82 was obtained from the John Innes Centre, Norwich and the a 400-bp Smal

fragment containing the bar gene from Streptomyces viridochromogenes was excised and

inserted into the Hpal site of pCIB3060 (Thompson et al. EMBO J 6: 2519-2523 (1987)).

This generated pCIB3064, which comprises the bar gene under the control of the CaMV

35S promoter and terminator for herbicide selection, a gene for ampicillin resistance (for

selection in E. coli) and a poly linker with the unique sites Sphi, PstI, HindIII, and BamHI.

This vector is suitable for the cloning of plant expression cassettes containing their own

regulatory signals.

Construction of pSOG19 and pSOG35: pSOG35 is a transformation vector that

utilizes the E. coli gene dihydrofolate reductase (DHFR) as a selectable marker conferring

resistance to methotrexate. PCR was used to amplify the 35S promoter (−800-bp), intron 6

from the maize Adh1 gene (−550-bp) and 18-bp of the GUS untranslated leader sequence

from pSOG10. A 250-bp fragment encoding the E. coli dihydrofolate reductase type II gene

was also amplified by PCR and these two PCR fragments were assembled with a Sacl-PstI

fragment from pBl221 (Clontech), which comprised the pUC19 vector backbone and the

nopaline synthase terminator. Assembly of these fragments generated pSOG19, which

contains the 35S promoter in fusion with the intron 6 sequence, the GUS leader, the DHFR

gene and the nopaline synthase terminator. Replacement of the GUS leader in pSOG19

with the leader sequence from Maize Chlorotic Mottle Virus (MCMV) generated the vector

pSOG35. pSOG19 and pSOG35 carry the pUC gene for ampicillin resistance and have

HindIII, Sphi, PstI and EcoRI sites available for the cloning of foreign sequences.
4. Example 22: Construction of Plant Expression Cassettes

Gene sequences intended for expression in transgenic plants are firstly assembled in expression cassettes behind a suitable promoter and upstream of a suitable transcription terminator. These expression cassettes can then be easily transferred to the plant transformation vectors described above in Example 21.

I. Promoter Selection

The selection of a promoter used in expression cassettes will determine the spatial and temporal expression pattern of the transgene in the transgenic plant. Selected promoters will express transgenes in specific cell types (such as leaf epidermal cells, mesophyll cells, root cortex cells) or in specific tissues or organs (roots, leaves or flowers, for example) and this selection will reflect the desired location of expression of the transgene. Alternatively, the selected promoter may drive expression of the gene under a light-induced or other temporally regulated promoter. A further alternative is that the selected promoter be chemically regulated. This would provide the possibility of inducing expression of the transgene only when desired and caused by treatment with a chemical inducer.

II. Transcriptional Terminators

A variety of transcriptional terminators are available for use in expression cassettes. These are responsible for the termination of transcription beyond the transgene and its correct polyadenylation. Appropriate transcriptional terminators are those that are known to function in plants and include the CaMV 3SS terminator, the \textit{tml} terminator, the nopaline synthase terminator, the \textit{pea rbcS} E9 terminator, as well as terminators naturally associated with the plant protox gene (i.e. "protox terminators"). These can be used in both monocotyledons and dicotyledons.

III. Sequences for the Enhancement or Regulation of Expression

Numerous sequences have been found to enhance gene expression from within the transcriptional unit and these sequences can be used in conjunction with the genes of this invention to increase their expression in transgenic plants.
Various intron sequences have been shown to enhance expression, particularly in monocotyledonous cells. For example, the introns of the maize *Adh1* gene have been found to significantly enhance the expression of the wild-type gene under its cognate promoter when introduced into maize cells. Intron 1 was found to be particularly effective and enhanced expression in fusion constructs with the chloramphenicol acetyltransferase gene (Callis *et al.*, *Genes Develop.* 1: 1183-1200 (1987)). In the same experimental system, the intron from the maize *bronzel* gene had a similar effect in enhancing expression (Callis *et al.*, *supra*). Intron sequences have been routinely incorporated into plant transformation vectors, typically within the non-translated leader.

A number of non-translated leader sequences derived from viruses are also known to enhance expression, and these are particularly effective in dicotyledonous cells. Specifically, leader sequences from Tobacco Mosaic Virus (TMV, the "W-sequence"), Maize Chlorotic Mottle Virus (MCMV), and Alfalfa Mosaic Virus (AMV) have been shown to be effective in enhancing expression (*e.g.* Gallie *et al.* *Nucl. Acids Res.* 15: 8693-8711 (1987); Skuzeski *et al.* *Plant Molec. Biol.* 15: 65-79 (1990)).

IV. Targeting of the Gene Product Within the Cell

Various mechanisms for targeting gene products are known to exist in plants and the sequences controlling the functioning of these mechanisms have been characterized in some detail. For example, the targeting of gene products to the chloroplast is controlled by a signal sequence that is found at the amino terminal end of various proteins and that is cleaved during chloroplast import yielding the mature protein (*e.g.* Comai *et al.* *J. Biol. Chem.* 263: 15104-15109 (1988)). These signal sequences can be fused to heterologous gene products to effect the import of heterologous products into the chloroplast (van den Broeck *et al.* *Nature* 313: 358-363 (1985)). DNA encoding for appropriate signal sequences can be isolated from the 5' end of the cDNAs encoding the RUBISCO protein, the CAB protein, the EPSP synthase enzyme, the GS2 protein and many other proteins that are known to be chloroplast localized.

Other gene products are localized to other organelles such as the mitochondrion and the peroxisome (*e.g.* Unger *et al.* *Plant Molec. Biol.* 13: 411-418 (1989)). The cDNAs encoding these products can also be manipulated to effect the targeting of heterologous gene products to these organelles. Examples of such sequences are the nuclear-encoded ATPases and specific aspartate amino transferase isoforms for mitochondria. Targeting to
cellular protein bodies has been described by Rogers et al., *Proc. Natl. Acad. Sci. USA* 82: 6512-6516 (1985)).

In addition, sequences have been characterized that cause the targeting of gene products to other cell compartments. Amino terminal sequences are responsible for targeting to the ER, the apoplast, and extracellular secretion from aleurone cells (Koehler & Ho, *Plant Cell* 2: 769-783 (1990)). Additionally, amino terminal sequences in conjunction with carboxy terminal sequences are responsible for vacuolar targeting of gene products (Shinshi et al., *Plant Molec. Biol. 14*: 357-368 (1990)).

By the fusion of the appropriate targeting sequences described above to transgene sequences of interest it is possible to direct the transgene product to any organelle or cell compartment. For chloroplast targeting, for example, the chloroplast signal sequence from the RUBISCO gene, the CAB gene, the EPSP synthase gene, or the GS2 gene is fused in frame to the amino terminal ATG of the transgene. The signal sequence selected should include the known cleavage site and the fusion constructed should take into account any amino acids after the cleavage site that are required for cleavage. In some cases this requirement may be fulfilled by the addition of a small number of amino acids between the cleavage site and the transgene ATG or alternatively replacement of some amino acids within the transgene sequence. Fusions constructed for chloroplast import can be tested for efficacy of chloroplast uptake by *in vitro* translation of *in vitro* transcribed constructions followed by *in vitro* chloroplast uptake using techniques described by (Bartlett et al. In: Edelmann et al. (Eds.) *Methods in Chloroplast Molecular Biology*, Elsevier. pp. 1081-1091 (1982); Wasmann et al. *Mol. Gen. Genet. 205*: 446-453 (1986)). These construction techniques are well known in the art and are equally applicable to mitochondria and peroxisomes. The choice of targeting that may be required for expression of the transgenes will depend on the cellular localization of the precursor required as the starting point for a given pathway. This will usually be cytosolic or chloroplastic, although it may is some cases be mitochondrial or peroxisomal. The products of transgene expression will not normally require targeting to the ER, the apoplast or the vacuole.

The above described mechanisms for cellular targeting can be utilized not only in conjunction with their cognate promoters, but also in conjunction with heterologous promoters so as to effect a specific cell targeting goal under the transcriptional regulation of a promoter that has an expression pattern different to that of the promoter from which the targeting signal derives.
5. Example 23: Transformation of Dicotyledons

Transformation techniques for dicotyledons are well known in the art and include \textit{Agrobacterium}-based techniques and techniques that do not require \textit{Agrobacterium}. Non-\textit{Agrobacterium} techniques involve the uptake of exogenous genetic material directly by protoplasts or cells. This can be accomplished by PEG or electroporation mediated uptake, particle bombardment-mediated delivery, or microinjection. Examples of these techniques are described by Paszkowski \textit{et al.}, \textit{EMBO J} 3: 2717-2722 (1984), Potrykus \textit{et al.}, \textit{Mol. Gen. Genet.} 199: 169-177 (1985), Reich \textit{et al.}, \textit{Biotechnology} 4: 1001-1004 (1986), and Klein \textit{et al.}, \textit{Nature} 327: 70-73 (1987). In each case the transformed cells are regenerated to whole plants using standard techniques known in the art.

\textit{Agrobacterium}-mediated transformation is a preferred technique for transformation of dicotyledons because of its high efficiency of transformation and its broad utility with many different species. The many crop species that are routinely transformable by \textit{Agrobacterium} include tobacco, tomato, sunflower, cotton, oilseed rape, potato, soybean, alfalfa and poplar (EP 0 317 511 (cotton), EP 0 249 432 (tomato, to Calgene), WO 87/07299 (\textit{Brassica}, to Calgene), US 4,795,855 (poplar)).

Transformation of the target plant species by recombinant \textit{Agrobacterium} usually involves co-cultivation of the \textit{Agrobacterium} with explants from the plant and follows protocols well known in the art. Transformed tissue is regenerated on selectable medium carrying the antibiotic or herbicide resistance marker present between the binary plasmid T-DNA borders.

6. Example 24: Transformation of Monocotyledons

Transformation of most monocotyledon species has now also become routine. Preferred techniques include direct gene transfer into protoplasts using PEG or electroporation techniques, and particle bombardment into callus tissue. Transformations can be undertaken with a single DNA species or multiple DNA species (\textit{i.e.} co-transformation) and both these techniques are suitable for use with this invention. Co-transformation may have the advantage of avoiding complex vector construction and of generating transgenic plants with unlinked loci for the gene of interest and the selectable marker, enabling the removal of the selectable marker in subsequent generations, should
this be regarded desirable. However, a disadvantage of the use of co-transformation is the
less than 100% frequency with which separate DNA species are integrated into the genome
(Schocher et al., Biotechnology 4: 1093-1096 (1986)).

Patent Applications EP 0 292 435 (to Ciba-Geigy), EP 0 392 225 (to Ciba-Geigy) and
WO 93/07278 (to Ciba-Geigy) describe techniques for the preparation of callus and
protoplasts from an elite inbred line of maize, transformation of protoplasts using PEG or
electroporation, and the regeneration of maize plants from transformed protoplasts.
Gordon-Kamm et al., Plant Cell 2: 603-618 (1990)) and Fromm et al., Biotechnology 8: 833-
839 (1990)) have published techniques for transformation of A188-derived maize line using
particle bombardment. Furthermore, application WO 93/07278 (to Ciba-Geigy) and Koziel
et al., Biotechnology 11: 194-200 (1993)) describe techniques for the transformation of élite
inbred lines of maize by particle bombardment. This technique utilizes immature maize
embryos of 1.5-2.5 mm length excised from a maize ear 14-15 days after pollination and a
PDS-1000He Biolistics device for bombardment.

Transformation of rice can also be undertaken by direct gene transfer techniques
utilizing protoplasts or particle bombardment. Protoplast-mediated transformation has been
described for Japonica-types and Indica-types (Zhang et al., Plant Cell Rep 7: 379-384
(1990)). Both types are also routinely transformable using particle bombardment (Christou

Patent Application EP 0 332 581 (to Ciba-Geigy) describes techniques for the
generation, transformation and regeneration of Pooidae protoplasts. These techniques
allow the transformation of Dactylis and wheat. Furthermore, wheat transformation was
been described by Vasil et al., Biotechnology 10: 667-674 (1992)) using particle
bombardment into cells of type C long-term regenerable callus, and also by Vasil et al.,
(1993) using particle bombardment of immature embryos and immature embryo-derived
callus. A preferred technique for wheat transformation, however, involves the
transformation of wheat by particle bombardment of immature embryos and includes either
a high sucrose or a high maltose step prior to gene delivery. Prior to bombardment, any
number of embryos (0.75-1 mm in length) are plated onto MS medium with 3% sucrose
(Murashige & Skoog, Physiologia Plantarum 15: 473-497 (1962)) and 3 mg/l 2,4-D for
induction of somatic embryos, which is allowed to proceed in the dark. On the chosen day
of bombardment, embryos are removed from the induction medium and placed onto the
osmoticum (i.e. induction medium with sucrose or maltose added at the desired concentration, typically 15%). The embryos are allowed to plasmolyze for 2-3 h and are then bombarded. Twenty embryos per target plate is typical, although not critical. An appropriate gene-carrying plasmid (such as pCIB3064 or pSG35) is precipitated onto micrometer size gold particles using standard procedures. Each plate of embryos is shot with the DuPont Biolistics helium device using a burst pressure of ~1000 psi using a standard 80 mesh screen. After bombardment, the embryos are placed back into the dark to recover for about 24 h (still on osmoticum). After 24 hrs, the embryos are removed from the osmoticum and placed back onto induction medium where they stay for about a month before regeneration. Approximately one month later the embryo explants with developing embryogenic callus are transferred to regeneration medium (MS + 1 mg/liter NAA, 5 mg/liter GA), further containing the appropriate selection agent (10 mg/l basta in the case of pCIB3064 and 2 mg/l methotrexate in the case of pSOG35). After approximately one month, developed shoots are transferred to larger sterile containers known as "GA7s" that contained half-strength MS, 2% sucrose, and the same concentration of selection agent.

Patent application WO 94/13822 describes methods for wheat transformation and is hereby incorporated by reference.

Example 25: Isolation of the Arabidopsis thaliana Protox-1 Promoter Sequence

A Lambda Zap II genomic DNA library prepared from Arabidopsis thaliana (Columbia, whole plant) was purchased from Stratagene. Approximately 125,000 phage were plated at a density of 25,000 pfu per 15 cm Petri dish and duplicate lifts were made onto Colony/Plaque Screen membranes (NEN Dupont). The plaque lifts were probed with the Arabidopsis protox-1 cDNA (SEQ ID NO:1 labeled with 32P-dCTP by the random priming method (Life Technologies). Hybridization and wash conditions were at 65°C as described in Church and Gilbert, Proc. Natl. Acad. Sci. USA 81: 1991-1995 (1984). Positively hybridizing plaques were purified and in vivo excised into pBluescript plasmids. Sequence from the genomic DNA inserts was determined by the chain termination method using dideoxy terminators labeled with fluorescent dyes (Applied Biosystems, Inc.). One clone, AraPT1Pro, was determined to contain 580-bp of Arabidopsis sequence upstream from the initiating methionine (ATG) of the protox-1 protein coding sequence. This clone also contains coding sequence and introns that extend to-bp 1241 of the protox-1 cDNA
sequence. The 580-bp 5' noncoding fragment is the putative *Arabidopsis* protox-1 promoter, and the sequence is set forth in SEQ ID NO:13.

AraPT1Pro was deposited December 15, 1995, as pWDC-11 (NRRL #B-21515)

Example 26: Construction of Plant Transformation Vectors Expressing Altered Protox-1 Genes Behind the Native *Arabidopsis* Protox-1 Promoter

A full-length cDNA of the appropriate altered *Arabidopsis* protox-1 cDNA was isolated as an *EcoRl*-XhoI partial digest fragment and cloned into the plant expression vector pCGN1761ENX (see Example 9 of International application no. PCT/IB95/00452 filed June 8, 1995, published Dec. 21, 1995 as WO 95/34659). This plasmid was digested with *Ncol* and *BamHl* to produce a fragment comprised of the complete protox-1 cDNA plus a transcription terminator from the 3' untranslated sequence of the *tml* gene of *Agrobacterium tumefaciens*. The AraPT1Pro plasmid described above was digested with *Ncol* and *BamHl* to produce a fragment comprised of pBluescript and the 580-bp putative *Arabidopsis* protox-1 promoter. Ligation of these two fragments produced a fusion of the altered protox cDNA to the native protox promoter. The expression cassette containing the protox-1 promoter/protox-1 cDNA/*tml* terminator fusion was excised by digestion with KpnI and cloned into the binary vector pClIB200. The binary plasmid was transformed by electroporation into *Agrobacterium* and then into *Arabidopsis* using the vacuum infiltration method (Bechtold *et al.*, *C.R. Acad. Sci. Paris* 316: 1194-1199 (1993). Transformants expressing altered protox genes were selected on kanamycin or on various concentrations of protox inhibiting herbicide.

Example 27: Production of Herbicide Tolerant Plants by Expression of a Native Protox-1 Promoter/Altered Protox-1 Fusion

Using the procedure described above, an *Arabidopsis* protox-1 cDNA containing a TAC to ATG (Tyrosine to Methionine) change at nucleotides 1306-1308 in the protox-1 sequence (SEQ ID NO:1) was fused to the native protox-1 promoter fragment and transformed into *Arabidopsis thaliana*. This altered protox-1 enzyme (AraC-2Met) has been shown to be >10-fold more tolerant to various protox-inhibiting herbicides than the naturally occurring enzyme when tested in the previously described bacterial expression system. Seed from the vacuum infiltrated plants was collected and plated on a range (10.0nM-
1.0uM) of a protox inhibitory aryluracil herbicide of formula XVII. Multiple experiments with wild type Arabidopsis have shown that a 10.0nM concentration of this compound is sufficient to prevent normal seedling germination. Transgenic seeds expressing the AraC-2Met altered enzyme fused to the native protox-1 promoter produced normal Arabidopsis seedlings at herbicide concentrations up to 500nM, indicating at least 50-fold higher herbicide tolerance when compared to wild-type Arabidopsis. This promoter/altered protox enzyme fusion therefore functions as an effective selectable marker for plant transformation. Several of the plants that germinated on 100.0nM of protox-inhibiting herbicide were transplanted to soil, grown 2-3 weeks, and tested in a spray assay with various concentrations of the protox-inhibiting herbicide. When compared to empty vector control transformants, the AraPT1Pro/AraC-2Met transgenics were >10-fold more tolerant to the herbicide spray.

EXAMPLE 28: Demonstration of resistant mutations' cross-tolerance to various protox-inhibiting compounds in an Arabidopsis germination assay.

Using the procedure described above, an Arabidopsis protox-1 cDNA containing both a TAC to ATC (tyrosine to isoleucine) change at nucleotides 1306-1308 and a TCA to TTA (serine to leucine) change at nucleotides 945-947 in the protox-1 sequence (SEQ ID NO:1) was fused to the native protox-1 promoter fragment and transformed into Arabidopsis thaliana. This altered protox-1 enzyme (AraC-2Ile + AraC305Leu) has been shown to be >10-fold more tolerant to a protox inhibitory aryluracil herbicide of formula XVII than the naturally occurring enzyme when tested in a bacterial system (see Examples 9-13). Homozygous Arabidopsis lines containing this fusion were generated from transformants that showed high tolerance to a protox inhibiting herbicide in a seedling germination assay as described above. The seed from one line was tested for cross-tolerance to various protox-inhibitory compounds by repeating the germination assay on concentrations of the compounds that had been shown to inhibit germination of wild-type Arabidopsis. The results from these experiments are shown in Table 4.
7. Example 29: Isolation of a Maize Protox-1 Promoter Sequence

A Zea Mays (Missouri 17 inbred, etiolated seedlings) genomic DNA library in the Lambda FIX II vector was purchased from Stratagene. Approximately 250,000 pfu of the library was plated at a density of 50,000 phage per 15 cm plate and duplicate lifts were made onto Colony/Plaque screen membranes (NEN Dupont). The plaque lifts were probed with the maize protox-1 cDNA (SEQ ID NO:5) labeled with 32P-dCTP by the random priming method (Life Technologies). Hybridization and wash conditions were at 65°C as described in Church and Gilbert, *Proc. Natl. Acad. Sci. USA* 81: 1991-1995 (1984). Lambda phage DNA was isolated from three positively hybridizing phage using the Wizard Lambda Preps DNA Purification System (Promega). Analysis by restriction digest, hybridization patterns, and DNA sequence analysis identified a lambda clone containing approximately 3.5 kb of maize genomic DNA located 5’ to the maize protox-1 coding sequence previously isolated as a cDNA clone. This fragment includes the maize protox-1 promoter. The sequence of this fragment is set forth in SEQ ID NO:14. From nucleotide 1 to 3532, this sequence is comprised of 5’ noncoding sequence. From nucleotide 3533 to 3848, this sequence encodes the 5’ end of the maize protox-1 protein.

A plasmid containing the sequence of SEQ ID NO:14 fused to the remainder of the maize protox-1 coding sequence was deposited March 19, 1996 as pWDC-14 (NRRL #B-21546).

Example 30: Construction of Plant Transformation Vectors Expressing Altered Protox-1 Genes Behind the Native Maize Protox-1 Promoter

The 3848-bp maize genomic fragment (SEQ ID NO:14) was excised from the isolated lambda phage clone as a *SalI-KpnI* partial digest product and ligated to a *KpnI-NotI* fragment derived from an altered maize protox-1 cDNA that contained an alanine to leucine change at amino acid 164 (SEQ ID NO:6). This created a fusion of the native maize protox-1 promoter to a full length cDNA that had been shown to confer herbicide tolerance in a bacterial system (Examples 9-14). This fusion was cloned into a pUC18 derived vector containing the CaMV 35S terminator sequence to create a protox promoter/altered protox cDNA/terminator cassette. The plasmid containing this cassette was designated pWC0-1.

A second construct for maize transformation was created by engineering the first intron found in the coding sequence from the maize genomic clone back into the maize
cDNA. The insertion was made using standard overlapping PCR fusion techniques. The intron (SEQ ID NO:25) was 93-bp long and was inserted between nucleotides 203 and 204 of SEQ ID NO:6, exactly as it appeared in natural context in the lambda clone described in Example 29. This intron-containing version of the expression cassette was designated pWC0-2.

8. Example 31: Demonstration of Maize Protox-1 Promoter Activity in Transgenic Maize Plants

Maize plants transformed with maize protox promoter/ altered protox fusions were identified using PCR analysis with primers specific for the transgene. Total RNA was prepared from the PCR positive plants and reverse-transcribed using Superscript M-MLV (Life Technologies) under recommended conditions. Two microliters of the reverse transcription reaction was used in a PCR reaction designed to be specific for the altered protox sequence. While untransformed controls give no product in this reaction, approximately 85% of plants transformed with pWC0-1 gave a positive result, indicating the presence of mRNA derived from the transgene. This demonstrates some level of activity for the maize protox promoter. The RNA’s from the transgenic maize plants were also subjected to standard northern blot analysis using the radiolabeled maize protox cDNA fragment from SEQ ID NO:6 as a probe. Protox-1 mRNA levels significantly above those of untransformed controls were detected in some of the transgenic maize plants. This elevated mRNA level is presumed to be due to expression of altered protox-1 mRNA from the cloned maize protox promoter.

9. Example 32: Isolation of a Sugar Beet Protox-1 Promoter Sequence

A genomic sugar beet library was prepared by Stratagene in the Lambda Fix II vector. Approximately 300,000 pfu of the library was plated and probed with the sugar beet protox-1 cDNA sequence (SEQ ID NO:17) as described for maize in Example 29. Analysis by restriction digest, hybridization patterns and DNA sequence analysis identified a lambda clone containing approximately 7 kb of sugar beet genomic DNA located 5’ to the sugar beet coding sequence previously isolated as a cDNA clone. A PstI-Sall fragment of 2606-bp was subcloned from the lambda clone into a pBluescript vector. This fragment contains 2068-bp of 5’ noncoding sequence and includes the sugar beet protox-1 promoter.
sequence. It also includes the first 453-bp of the protox-1 coding sequence and the 85-bp first intron contained in the coding sequence. The sequence of this fragment is set forth in SEQ ID NO:26.

A plasmid containing the sequence of SEQ ID NO:26 was deposited December 6, 1996 as pWDC-20 (NRRL #B-21650).

Example 33: Construction of Plant Transformation Vectors Expressing Altered Sugar Beet Protox-1 Genes Behind the Native Sugar Beet Protox-1 Promoter

The sugar beet genomic fragment (SEQ ID NO:26) was excised from the genomic subclone described in Example 32 as a SacI-BsrGI fragment that includes 2068-bp of 5' noncoding sequence and the first 300-bp of the sugar beet protox-1 coding sequence. This fragment was ligated to a BsrGI-NotI fragment derived from an altered sugar beet protox-1 cDNA that contained a tyrosine to methionine change at amino acid 449 (SEQ ID NO:18). This created a fusion of the native sugar beet protox-1 promoter to a full length cDNA that had been shown to confer herbicide tolerance in a bacterial system (Examples 9-14). This fusion was cloned into a pUC18 derived vector containing the CaMV 35S terminator sequence to create a protox promoter/altered protox cDNA/terminator cassette. The plasmid containing this cassette was designated pWC0-3.

Example 34: Production of Herbicide Tolerant Plants by Expression of a Native Sugar Beet Protox-1 Promoter/Altered Sugar Beet Protox-1 Fusion

The expression cassette from pWC0-3 is transformed into sugar beet using any of the transformation methods applicable to dicot plants, including Agrobacterium, protoplast, and biolistic transformation techniques. Transgenic sugar beets expressing the altered protox-1 enzyme are identified by RNA-PCR and tested for tolerance to protox-inhibiting herbicides at concentrations that are lethal to untransformed sugar beets.

Section D: Expression of Protox Genes in Plant Plastids

Example 35: Preparation of a Chimeric Gene Containing the Tobacco Plastid clpP Gene Promoter and Native clpP 5' Untranslated Sequence Fused to a GUS Reporter Gene and Plastid rps16 Gene 3' Untranslated Sequence in a Plastid Transformation Vector
I. Amplification of the Tobacco Plastid clpP Gene Promoter and Complete 5' Untranslated RNA (5' UTR).

Total DNA from \textit{N. tabacum} c.v. "Xanthi NC" was used as the template for PCR with a left-to-right "top strand" primer comprising an introduced \textit{EcoRI} restriction site at position -197 relative to the ATG start codon of the constitutively expressed plastid \textit{clpP} gene (primer Pclp_P1a: 5'\text{-}GCGGAATTCATCTTTATTTATCATGAAAG-3' (SEQ ID NO:27); \textit{EcoRI} restriction site underlined) and a right-to-left "bottom strand" primer homologous to the region from -21 to -1 relative to the ATG start codon of the \textit{clpP} promoter that incorporates an introduced \textit{Ncol} restriction site at the start of translation (primer Pclp_P2b: 5'\text{-}GCGGCCATGGTAAATGAAAGAAGAAGACTAAA-3' (SEQ ID NO:28); \textit{Ncol} restriction site underlined). This PCR reaction was undertaken with \textit{Pfu} thermostable DNA polymerase (Stratagene, La Jolla CA) in a Perkin Elmer Thermal Cycler 480 according to the manufacturer's recommendations (Perkin Elmer/Roche, Branchburg, NJ) as follows: 7 min 95°C, followed by 4 cycles of 1 min 95°C / 2 min 43°C / 1 min 72°C, then 25 cycles of 1 min 95°C / 2 min 55°C / 1 min 72°C. The 213-bp amplification product comprising the promoter and 5' untranslated region of the \textit{clpP} gene containing an \textit{EcoRI} site at its left end and an \textit{Ncol} site at its right end and corresponding to nucleotides 74700 to 74505 of the \textit{N. tabacum} plastid DNA sequence (Shinozaki \textit{et al.}, \textit{EMBO J}. 5: 2043–2049 (1986)) was gel purified using standard procedures and digested with \textit{EcoRI} and \textit{Ncol} (all restriction enzymes were purchased from New England Biolabs, Beverly, MA).

II. Amplification of the Tobacco Plastid rps16 Gene 3' Untranslated RNA Sequence (3' UTR).

Total DNA from \textit{N. tabacum} c.v. "Xanthi NC" was used as the template for PCR as described above with a left-to-right "top strand" primer comprising an introduced \textit{XbaI} restriction site immediately following the TAA stop codon of the plastid \textit{rps16} gene encoding ribosomal protein S16 (primer rps16P_1a (5'\text{-}GCGTCTAGATCAACCGAAATTCAATGAAAG-3' (SEQ ID NO:30); \textit{XbaI} restriction site underlined) and a right-to-left "bottom strand" primer homologous to the region from +134 to +151 relative to the TAA stop codon of \textit{rps16} that incorporates an introduced \textit{HindIII} restriction site at the 3' end of the \textit{rps16} 3' UTR (primer rps16P_1b (5'\text{-}.
CGCAAGCTTCAATGGGAAGCAATGATA-3' (SEQ ID NO:31); HindIII restriction site underlined). The 169-bp amplification product comprising the 3’ untranslated region of the rps16 gene containing an XbaI site at its left end and a HindIII site at its right end and containing the region corresponding to nucleotides 4943 to 5093 of the N. tabacum plastid DNA sequence (Shinozaki et al., 1986) was gel purified and digested with XbaI and HindIII.

III. Ligation of a GUS Reporter Gene Fragment to the clpP Gene Promoter and 5’ and 3’ UTR’s.

An 1864-bp β-glucuronidase (GUS) reporter gene fragment derived from plasmid pRAJ275 (Clontech) containing an Ncol restriction site at the ATG start codon and an XbaI site following the native 3' UTR was produced by digestion with Ncol and XbaI. This fragment was ligated in a four-way reaction to the 201-bp EcoRI/Ncol clpP promoter fragment, the 157-bp XbaI/HindIII rps16 3’UTR fragment, and a 3148-bp EcoRI/HindIII fragment from cloning vector pGEM3Zf(-) (Promega, Madison WI) to construct plasmid pPH138. Plastid transformation vector pPH140 was constructed by digesting plasmid pPRV111a (Zoubenko et al. 1994) with EcoRI and HindIII and ligating the resulting 7287-bp fragment to a 2222-bp EcoRI/HindIII fragment of pPH138.

Example 36: Preparation of a Chimeric Gene Containing the Tobacco Plastid clpP Gene Promoter Plus Tobacco Plastid psbA Gene Minimal 5’ Untranslated Sequence Fused to a GUS Reporter Gene and Plastid rps16 Gene 3’ Untranslated Sequence in a Plastid Transformation Vector

Amplification of the tobacco plastid clpP gene promoter and truncated 5’ untranslated RNA (5’ UTR): Total DNA from N. tabacum c.v. “Xanthi NC” was used as the template for PCR as described above with the left-to-right “top strand” primer Pclp_P1a (SEQ ID NO:27) and a right-to-left “bottom strand” primer homologous to the region from -34 to -11 relative to the ATG start codon of the clpP promoter that incorporates an introduced XbaI restriction site in the clpP 5’ UTR at position -11 (primer Pclp_P1b: 5’-GCGTCTAGAAGAAGAACTAATATATATTCAC-3’ (SEQ ID NO:29); XbaI restriction site underlined). The 202-bp amplification product comprising the promoter and truncated 5’ UTR of the clpP gene containing an EcoRI site at its left end and an XbaI site at its right
end was gel purified and digested with XbaI. The XbaI site was subsequently filled in with Klenow DNA polymerase (New England Biolabs) and the fragment digested with EcoRI. This was ligated in a five-way reaction to a double stranded DNA fragment corresponding to the final 38 nucleotides and ATG start codon of the tobacco plastid psbA gene 5' UTR (with an NcoI restriction site overhang introduced into the ATG start codon) that was created by annealing the synthetic oligonucleotides minpsb_U (top strand: 5’-GGGAGTCCCTGATGATTAAATAAAACCAAGATTATTAC-3’ (SEQ ID NO:32)) and minpsb_L (bottom strand: 5’-CATGGTAAATCTTGTATTTAATCATCAGGGACTCCC-3’ (SEQ ID NO:33); NcoI restriction site 5’ overhang underlined), the NcoI/XbaI GUS reporter gene fragment described above, the XbaI/HindIII rps16 3’UTR fragment described above, and the EcoRI/HindIII pGEM3Zf(-) fragment described above to construct plasmid pH139. Plastid transformation vector pPH144 was constructed by digesting plasmid pPRV111a (Zoubenko, et al., Nucleic Acids Res 22: 3819-3824 (1994)) with EcoRI and HindIII and ligating the resulting 7287-bp fragment to a 2251-bp EcoRI/HindIII fragment of pH139.

Example 37: Preparation of a Chimeric Gene Containing the Tobacco Plastid clpP Gene Promoter and Complete 5' Untranslated Sequence Fused to the Arabidopsis thaliana Protox-1 Coding Sequence and Plastid rps16 Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

Miniprep DNA from plasmid AraC-2Met carrying an Arabidopsis thaliana NotI insert that includes cDNA sequences from the Protoporphyrinogen IX Oxidase ("protox") gene encoding a portion of the amino terminal plastid transit peptide, the full-length cDNA and a portion of the 3' untranslated region was used as the template for PCR as described above using a left-to-right "top strand" primer (with homology to nucleotides +172 to +194 relative to the ATG start codon of the full length precursor protein) comprising an introduced NcoI restriction site and new ATG start codon at the deduced start of the mature protox protein coding sequence (primer APRTXP1a: 5’-GGGACCATTGGATTGTGATTGTCGCGGAGG-3’ (SEQ ID NO:34); NcoI restriction site underlined) and a right-to-left "bottom strand" primer homologous to nucleotides +917 to +940 relative to the native ATG start codon of the protox precursor protein (primer APRTXP1b: 5’-CTCCGCTCTCCAGCTTATGTGATA-3’ (SEQ ID NO:35)). The 778-bp product was digested with NcoI and SfI and the resulting 682-bp fragment ligated to an 844-bp SfuI/NotI DNA fragment of AraC-2Met comprising the 3' portion of the protox coding
sequence and a 2978-bp \textit{Ncol/NotI} fragment of the cloning vector pGEMSZf(+) (Promega, Madison WI) to construct plasmid pH141. Plastid transformation vector pH143 containing the \textit{clipP} promoter driving the Formula XVII-resistant AraC-2Met protox gene with the \textit{rps16} 3' UTR was constructed by digesting pH141 with \textit{Ncol} and \textit{SspI} and isolating the 1491-bp fragment containing the complete protox coding sequence, digesting the \textit{rps16P_1a} and \textit{rps16P_1b} PCR product described above with \textit{HindIII}, and ligating these to a 7436-bp \textit{Ncol/HindIII} fragment of pH140.

Example 38: Preparation of a Chimeric Gene Containing the Tobacco Plastid \textit{clipP} Gene Promoter Plus Tobacco Plastid \textit{psbA} Gene Minimal 5' Untranslated Sequence Fused to the \textit{Arabidopsis thaliana} Protox-1 Coding Sequence and Plastid \textit{rps16} Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

Plastid transformation vector pH145 containing the \textit{clipP} promoter/\textit{psbA} 5' UTR fusion driving the Formula XVII-resistant AraC-2Met protox gene with the \textit{rps16} 3' UTR was constructed by digesting pH141 with \textit{Ncol} and \textit{SspI} and isolating the 1491-bp fragment containing the complete protox coding sequence, digesting the \textit{rps16P_1a} and \textit{rps16P_1b} PCR product described above with \textit{HindIII}, and ligating these to a 7465-bp \textit{Ncol/HindIII} fragment of pH144.
Example 39: Preparation of a Chimeric Gene Containing the Tobacco Plastid clpP Gene Promoter and 5' Untranslated Sequence Fused to the EPSP Synthase Coding Sequence and Plastid rps16 Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

A cDNA library is screened for the 5-enolpyruvyl-3-phosphoshikimate synthase (EPSP synthase) gene (U.S. Patent Nos. 5,310,667, 5,312,910, and 5,633,435, all incorporated herein by reference). A plasmid clone containing the full length EPSP synthase gene cDNA is isolated by standard techniques of molecular cloning. PCR primers are designed for amplification of the mature-size EPSP synthase coding sequence from this plasmid using a top strand primer having a 5' extension containing an Ncol restriction site inserted at amino acid −1 from the deduced mature protein start, thus creating an ATG start codon at this position, and a bottom strand primer having a 5' extension containing an Xbal restriction site downstream of the stop codon of the EPSP mature coding sequence in the amplified PCR product. The PCR amplification is performed using the designated primers and plasmid DNA template according to standard protocols. Amplified products are cloned and sequenced and a Ncol-Xbal DNA fragment containing the complete mature EPSP synthase coding sequence is isolated by restriction digest with Ncol and Xbal, electrophoresis on a 0.8% TAE agarose gel, and phenol extraction of the excised band. A plastid transformation vector containing the clpP promoter directing transcription of the mature-sized EPSP synthase gene with the rps16 3' UTR is constructed by digesting pPH140 with Ncol and Xbal and purifying the fragment containing the vector backbone, 5' and 3' plastid integration targeting sequences, aadA selectable marker cassette, and clpP promoter / rps16 3' UTR expression sequences. This product is ligated in a two-way reaction with the Ncol-Xbal DNA fragment containing the mature-sized EPSP synthase coding sequence isolated as described above.

Example 40: Preparation of a Chimeric Gene Containing the Tobacco Plastid clpP Gene Promoter and 5' Untranslated Sequence Fused to the ALS Coding Sequence and Plastid rps16 Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

A cDNA library is screened for the acetolactate synthase (ALS) gene (U.S. Patent No. 5,013,659). A plasmid clone containing the full length ALS gene cDNA is isolated by standard techniques of molecular cloning. PCR primers are designed for amplification of the
mature-size ALS coding sequence from this plasmid using a top strand primer having a 5' extension containing an Ncol restriction site inserted at amino acid –1 from the deduced mature protein start, thus creating an ATG start codon at this position, and a bottom strand primer having a 5’ extension containing an XbaI restriction site downstream of the stop codon of the ALS mature coding sequence in the amplified PCR product. The PCR amplification is performed using the designated primers and plasmid DNA template according to standard protocols. Amplified products are cloned and sequenced and a Ncol-XbaI DNA fragment containing the complete mature ALS coding sequence is isolated by restriction digest with Ncol and XbaI, electrophoresis on a 0.8% TAE agarose gel, and phenol extraction of the excised band.

A plastid transformation vector containing the clpP promoter driving the mature-sized ALS gene with the rps16 3' UTR is constructed by digesting pH140 with Ncol and XbaI and purifying the fragment containing the vector backbone, 5' and 3' plastid integration targeting sequences, aadA selectable marker cassette, and clpP promoter / rps16 3' UTR expression sequences. This product is ligated in a two-way reaction with the Ncol-XbaI DNA fragment containing the mature-sized ALS coding sequence isolated as described above.

Example 41: Preparation of a Chimeric Gene Containing the Tobacco Plastid clpP Gene Promoter and 5' Untranslated Sequence Fused to the AHAS Coding Sequence and Plastid rps16 Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

A cDNA library is screened for the acetohydroxyacid synthase (AHAS) gene (U.S. Patent No. 4,761,373). A plasmid clone containing the full length AHAS gene cDNA is isolated by standard techniques of molecular cloning. PCR primers are designed for amplification of the mature-size AHAS coding sequence from this plasmid using a top strand primer having a 5' extension containing an Ncol restriction site inserted at amino acid –1 from the deduced mature protein start, thus creating an ATG start codon at this position, and a bottom strand primer having a 5’ extension containing an XbaI restriction site downstream of the stop codon of the AHAS mature coding sequence in the amplified PCR product. The PCR amplification is performed using the designated primers and plasmid DNA template according to standard protocols. Amplified products are cloned and sequenced and a Ncol-XbaI DNA fragment containing the complete mature AHAS coding
sequence is isolated by restriction digest with *Ncol* and *Xbal*, electrophoresis on a 0.8% TAE agarose gel, and phenol extraction of the excised band.

A plastid transformation vector containing the *clpP* promoter driving the mature-sized AHAS gene with the *rps16* 3' UTR is constructed by digesting pH140 with *Ncol* and *Xbal* and purifying the fragment containing the vector backbone, 5' and 3' plastid integration targeting sequences, *aadA* selectable marker cassette, and *clpP* promoter / *rps16* 3' UTR expression sequences. This product is ligated in a two-way reaction with the *Ncol*-*Xbal* DNA fragment containing the mature-sized AHAS coding sequence isolated as described above.

Example 42: Preparation of a Chimeric Gene Containing the Tobacco Plastid *clpP* Gene Promoter and 5' Untranslated Sequence Fused to the ACCase Coding Sequence and Plastid *rps16* Gene 3' Untranslated Sequence in a Vector for Tobacco Plastid Transformation

A cDNA library is screened for the acetylcoenzyme A carboxylase (ACCase) gene (U.S. Patent No. 5,162,602). A plasmid clone containing the full length ACCase gene cDNA is isolated by standard techniques of molecular cloning. PCR primers are designed for amplification of the mature-size ACCase coding sequence from this plasmid using a top strand primer having a 5' extension containing an *Ncol* restriction site inserted at amino acid −1 from the deduced mature protein start, thus creating an ATG start codon at this position, and a bottom strand primer having a 5' extension containing an *Xbal* restriction site downstream of the stop codon of the ACCase mature coding sequence in the amplified PCR product. The PCR amplification is performed using the designated primers and plasmid DNA template according to standard protocols. Amplified products are cloned and sequenced and a *Ncol*-*Xbal* DNA fragment containing the complete mature ACCase coding sequence is isolated by restriction digest with *Ncol* and *Xbal*, electrophoresis on a 0.8% TAE agarose gel, and phenol extraction of the excised band.

A plastid transformation vector containing the *clpP* promoter driving the mature-sized ACCase gene with the *rps16* 3' UTR is constructed by digesting pH140 with *Ncol* and *Xbal* and purifying the fragment containing the vector backbone, 5' and 3' plastid integration targeting sequences, *aadA* selectable marker cassette, and *clpP* promoter / *rps16* 3' UTR expression sequences. This product is ligated in a two-way reaction with the *Ncol*-*Xbal*
DNA fragment containing the mature-sized ACCase coding sequence isolated as described above.

10. Example 43: Biolistic Transformation of the Tobacco Plastid Genome

Seeds of Nicotiana tabacum c.v. 'Xanthi nc' were germinated seven per plate in a 1" circular array on T agar medium and bombarded 12-14 days after sowing with 1 μm tungsten particles (M10, Biorad, Hercules, CA) coated with DNA from plasmids pH143 and pH145 essentially as described in Svab, Z. and Maliga, P. (1993) PNAS 90, 913–917. Bombarded seedlings were incubated on T medium for two days after which leaves were excised and placed abaxial side up in bright light (350-500 μmol photons/m²/s) on plates of RMOP medium (Svab, Z., Hajdukiewicz, P. and Maliga, P. (1990) PNAS 87, 8526–8530) containing 500 μg/ml spectinomycin dihydrochloride (Sigma, St. Louis, MO). Resistant shoots appearing underneath the bleached leaves three to eight weeks after bombardment were subcloned onto the same selective medium, allowed to form callus, and secondary shoots isolated and subcloned. Complete segregation of transformed plastid genome copies (homoplasmicity) in independent subclones was assessed by standard techniques of Southern blotting (Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor). BamH1/EcoRI-digested total cellular DNA (Mettler, I. J. (1987) Plant Mol Biol Reporter 5, 346–349) was separated on 1% Tris-borate (TBE) agarose gels, transferred to nylon membranes (Amersham) and probed with 32P-labeled random primed DNA sequences corresponding to a 0.7 kb BamH1/HindIII DNA fragment from pC8 containing a portion of the rps7/12 plastid targeting sequence. Homoplasmic shoots are rooted aseptically on spectinomycin-containing MS/IBA medium (McBride, K. E. et al. (1994) PNAS 91, 7301-7305) and transferred to the greenhouse.

Example 44: Assessment of Herbicide Tolerance in Nt_pPH143 and Nt_pPH145 Plastid Transformant Lines

Primary homoplasmic transformant lines transformed with pH143 (line Nt_pPH143) or with pH145 (line Nt_pPH145), which were obtained as described in Example 43, were grown to maturity in the greenhouse. Flowers were either: (a) self-pollinated, (b) pollinated with wildtype tobacco (c.v. Xanthi nc), or (c) used as pollen donors to fertilize emasculated
flowers of wildtype Xanthi plants. Plastid segregation of the linked spectinomycin resistance marker was verified by uniparental female inheritance of the spectinomycin-resistance phenotype in each transformant line using a minimum of 50 seeds per selection pool derived from either selfed or backcross capsules. Additional self or wildtype backcross (Xanthi pollen parent) seeds were germinated in soil. 36 plants of each line (143 1B-1, 143 1B-4, 143 4A-2, 143 4A-5, 145 7A-5, 145 7A-6, 145 8A-3) plus 36 wildtype Xanthi plants as isogenic controls were grown in separate 6" clay pots in a controlled environment cubicle. In order to assess tolerance to the protox inhibitor Formula XVII, plants of Xanthi and the seven transformant lines were distributed into eight identical 16-pot flats (2 plants of each type per flat). The flats were sprayed with Formula XVII until runoff at concentrations of either 0, 0.5, 2.5, 5, 10, 25, 50, or 100 mg Formula XVII per liter. Solutions were made up in water using 4 g/liter or 40 g/liter stock solutions of Formula XVII dissolved in dimethylsulfoxide (DMSO) and used immediately after preparation. Twenty microliters of the wetting agent Silwett was added to each 200 ml volume of herbicide solution for a final concentration of 0.01%. Flats were sprayed in the late afternoon and allowed to dry overnight before transfer to the growth cubicle. Tolerance was assessed by comparing leaf damage and wilting to the untransformed Xanthi controls at 0, 18 hrs, 48 hrs, and 6 days post-application. Severe damage was apparent on the Xanthi plants at all concentrations above 0.5 mg/l, and complete wilting/burn down occurred above 2.5 mg/l. Only slight damage occurred on the Nt_pPH143 plants even at the highest concentration (100 mg/liter), and the plants soon outgrew the bleached spots (the appearance of Xanthi at 0.5 mg/liter was approximately equivalent to Nt_pPH143 1B-1 at 100 mg/liter, giving a tolerance of ca. 200-fold).

Example 45: Plastid Transformation of Maize

Type I embryogenic callus cultures (Green et al. (1983) in A. Fazelahmad, K. Downey, J. Schultz, R.W. Voellmy, eds. Advances in Gene Technology: Molecular Genetics of Plants and Animals. Miami Winter Symposium Series, Vol. 20. Academic Press, N.Y.) of the proprietary genotypes CG00526 and CG00714 are initiated from immature embryos, 1.5 - 2.5 mm in length, from greenhouse grown material. Embryos are aseptically excised from surface-sterilized ears approximately 14 days after pollination. Embryos of CG00526 are placed on D callus initiation media with 2% sucrose and 5mg/L chloramben
(Duncan et al. (1985) Planta 165: 322-332) while those of CG00714 are placed onto KM callus initiation media with 3% sucrose and 0.75mg/L 2,4-d (Kao and Michayluk (1975) Planta 126, 105-110). Embryos and embryogenic cultures are subsequently cultured in the dark. Embryogenic responses are removed from the explants after ~14 days. CG00526 responses are placed onto D callus maintenance media with 2% sucrose and 0.5mg/L 2,4-d while those of CG00714 are placed onto KM callus maintenance media with 2% sucrose and 5mg/L Dicamba. After 3 to 8 weeks of weekly selective subculture to fresh maintenance media, high quality compact embryogenic cultures are established. Actively growing embryogenic callus pieces are selected as target tissue for gene delivery. The callus pieces are plated onto target plates containing maintenance medium with 12% sucrose approximately 4 hours prior to gene delivery. The callus pieces are arranged in circles, with radii of 8 and 10mm from the center of the target plate. Plasmid DNA is precipitated onto gold microcarriers as described in the DuPont Biolistics manual. Two to three μg of each plasmid is used in each 6 shot microcarrier preparation. Genes are delivered to the target tissue cells using the PDS-1000He Biolistics device. The settings on the Biolistics device are as follows: 8 mm between the rupture disc and the macrocarrier, 10 mm between the macrocarrier and the stopping screen and 7 cm between the stopping screen and the target. Each target plate is shot twice using 650psi rupture discs. A 200 X 200 stainless steel mesh (McMaster-Carr, New Brunswick, NJ) is placed between the stopping screen and the target tissue.

Five days later, the bombed callus pieces are transferred to maintenance medium with 2% sucrose and 0.5mg/L 2,4-d, but without amino acids, and containing 750 or 1000 nM Formula XVII. The callus pieces are placed for 1 hour on the light shelf 4-5 hours after transfer or on the next day, and stored in the dark at 27°C for 5-6 weeks. Following the 5-6 week primary selection stage, yellow to white tissue is transferred to fresh plates containing the same medium supplemented with 500 or 750 nM Formula XVII. 4-5 hours after transfer or on the next day, the tissues are placed for 1 hour on the light shelf and stored in the dark at 27°C for 3-4 weeks. Following the 3-4 week secondary selection stage, the tissues are transferred to plates containing the same medium supplemented with 500 nM Formula XVII. Healthy growing tissue is placed for 1 hour on the light shelf and stored in the dark at 27°C. It is subcultured every two weeks until the colonies are large enough for regeneration. At that point, colonies are transferred to a modified MS medium (Murashige and Skoog (1962) Physiol. Plant 15: 473-497) containing 3% sucrose (MS3S) with no selection agent and placed in the light. For CG00526, 0.25mg/L ancymidol and 0.5mg/L kinetin are added.
to this medium to induce embryo germination, while for CG00714, 2mg/L benzyl adenine is added. Regenerating colonies are transferred to MS3S media without ancyclidol and kinetin, or benzyl adenine, for CG00526 or CG00714, respectively, after 2 weeks. Regenerating shoots with or without roots are transferred to boxes containing MS3S medium and small plants with roots are eventually recovered and transferred to soil in the greenhouse.

Table 1A
Alignment of the full-length and partial protox-1 amino acid sequences from Arabidopsis ("Arabpt-1"; SEQ ID NO:2), maize ("Mzpt-1"; SEQ ID NO:6), wheat ("Wtpt-1"; SEQ ID NO:10), soybean ("Soybeanpt-1"; SEQ ID NO:12), cotton ("Cottonpt-1"; SEQ ID NO:16), sugar beet ("Sugpt-1"; SEQ ID NO:18), oilseed rape ("Rapept-1"; SEQ ID NO:20), rice ("Ricept-1"; SEQ ID NO:22), sorghum ("Sorghumpt-1"; SEQ ID NO:24), and sugar cane ("Scpt-1"; SEQ ID NO:37). Alignment was performed using the PileUp program (GCG package, University of Wisconsin, Madison, WI). Positions that may be modified according to the teachings herein to confer or enhance inhibitor resistance are shown in bold type.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapept-1</td>
<td>MDLSLRRP.. QPFLSPFSNP FPRSFRKYKPL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabpt-1</td>
<td>MEKSLLRPPT QSSLPSFSKP NLRGVYKPL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sorghumpt-1</td>
<td>.........</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mzpt-1</td>
<td>.........</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wtpt-1</td>
<td>.........</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ricept-1</td>
<td>.........</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cottonpt-1</td>
<td>.........</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soybeanpt-1</td>
<td>MV SVFNEILPPP NQTLRPSLH SPTSFTSPT RKFRPSRN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sugpt-1</td>
<td>MKSMALSNCI PQTQCMPLRS SGHYRGNCFI LSIPCSLIGR RGYSHHKRR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scpt-1</td>
<td>.........</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>51</th>
<th>52</th>
<th>53</th>
<th>54</th>
<th>55</th>
<th>56</th>
<th>57</th>
<th>58</th>
<th>59</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapept-1</td>
<td>NLRCSVSAGS VVGSSTIEGG GGGKTVDADC VIVGFIISGL CIAQALVTKH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabpt-1</td>
<td>RLRCSVAGGP TVGSSKIEGG GGT.TITTDC VIVGFIISGL CIAQALATKH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sorghumpt-1</td>
<td>.........</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 98 -
Mzpt-1 ............... ...........ADC VVVGGGISGL CTAQALATRH
Wtpt-1 RVRPRCATAS SATETPAAPG VRL..SAEC VIVGAGISGL CTAQALATRY
Ricept-1 ............... ............... ............... ...........
Cottonpt-1 KLRCSLAEGP TISSSKIDGG ESS..IADC VIVGGGISGL CIAQALATKHH
Soybeanpt-1 ILRCSIAEES TASSPKTR.. DSA..PVDC VVVGGGSVGSL CIAQALATKHH
Sugpt-1 MMSMCSTSSG SKSAVEKAGS GSGAGGLLDC VIVGGGISGL CIAQALCTKHH
Scpt-1 ............... ............... ............... ............... ...........

101
Rapept-1 PDA..AKNVM VTEAKDRVGG NIIT..REEQ GFLWEEGPN PSQPSPDMLTM
Arabpt-1 PDA..APNLI VTEAKDRVGG NIIT..REEN GFLWEEGPN PSQPSPDMLTM
Sorghumpt-1 ............... ............... STVERPEE GYLWEEGPN PSQPSPDVLTM
Mzpt-1 ...G..VGDVL VTEARARPGG NITTVERPEE GYLWEEGPN PSQPSPDVLTM
Wtpt-1 ...G..VSDLL VTEARDRPGG NITTVERPDE GYLWEEGPN PSQPSPDVLTM
Ricept-1 ............... ............... ............... ............... ...........
Cottonpt-1 RDV..ASNVV VTEARDRVGG NITTVER..D GYLWEEGPN PSQPSPDILTM
Soybeanpt-1 ...A..ANNVV VTEARDRVGG NITTMER..D GYLWEEGPN PSQPSPDMLTM
Sugpt-1 SSSSLSPNFL VTEAKDRVGG NIVTVE..AD GYIWEESPNS PSQPSPDAVLTM
Scpt-1 ............... ............... ............... ............... ...........

151
Rapept-1 VVDSLKDDL VLGDPTAPRF VLWNGKLRPV PSKTLTDLPPF DLMSIGGKIR
Arabpt-1 VVDSLKDDL VLGDPTAPRF VLWNGKLRPV PSKTLTDLPPF DLMSIGGKIR
Sorghumpt-1 AVDSLKDDL VFGDPNAPRF VLWEGKLRPV PSKPADLPFF DLMSIPGKLR
Mzpt-1 AVDSLKDDL VFGDPNAPRF VLWEGKLRPV PSKPADLPFF DLMSIPGKLR
Wtpt-1 AVDSLKDDL VFGDPNAPRF VLWEGKLRPV PSKPADLPFF SLMSIPGKLR
Ricept-1 ............... ............... ............... ............... ...........
Cottonpt-1 AVDSLKDDL VLGDPTAPRF VLWEGKLRPV PSKPDTDLPPF DLMSIAGKLR
Soybeanpt-1 VVDSLKDEL VLGDPTAPRF VLWNRKLRPV PGKLTDLPPF DLMSIGGKIR
Sugpt-1 AVDSLKDEL VLGDPTAPRF VLWNDKLRPV PSSSLTLDPFF DLMTIPGKIR
Scpt-1 ............... ............... ............... ............... ...........

201
Rapept-1 AGFGAIGIRP SPPG'REEVE EFVRRNLGDE VFERLIEPFCC SGVYAGDPASK
Arabpt-1 AGFGALGIRP SPPG'REEVE EFVRRNLGDE VFERLIEPFCC SGVYAGDPASK
Sorghumpt-1 AGLGALGIRP PAPG'REEVE EFVRRNLGAE VFERLIEPFCC SGVYAGDPASK
Mzpt-1 AGLGALGIRP PAPG'REEVE EFVRRNLGAE VFERLIEPFCC SGVYAGDPASK

250

-99-
Wtpt-1  AGLGALGIRP PPPGRESVE EFVRRNLGAE VFERLIEPPC SGVYAGDPSK
Ricept-1 .................................................................
Cottonpt-1 AGFGAIGIRP PPPGYESVE EFVRRNLGAE VFERFIEPPC SGVYAGDPSK
Soybeanpt-1 AGFGAIGIRP PPPGHEESVE EFVRRNLGDE VFERLIEPPC SGVYAGDPSK
Sugpt-1  AALGALGFRP SPPHESVE HFVRRNLGDE VFERLIEPPC SGVYAGDPAK
Scpt-1 .................................................................

251
Rapept-1  LSKMAAFGKV WKLEENGSI IGGAFKAIQA KKNAPKTTRD PRLPKPKGQT
Arabpt-1  LSKMAAFGKV WKLEQNGSGI IGGTFKAIQE RKNAPKAERD PRLPKPKGQT
Sorghumpt-1 LSKMAAFGKV WRLEAGGSI IGGTIKTTQIE RGKPNQPRFD PRLPKPKGQT
Mzt-1  LSKMAAFGKV WRLEETGGSI IGGTIKTQIE RSKNPQPRFD ARLPKPKGQT
Wtpt-1  LSKMAAFGKV WRLEEIGGSI IGGTIKAIQD KGKPNKRPFD PRLPAFKGQT
Ricept-1  RAFKMAAFGKV WRLEDGSGSI IGGTIKTIQIE RGKPNKPRFD PRLPTPKGQT
Cottonpt-1 LSKMAAFGKV WKLEEGGSI IGGTFKTIQIE RKNTPKPRFD PRLPKPKGQT
Soybeanpt-1 LSKMAAFGKV WKLEKNGGSI IGGTFKAIQIE RGNASKKPRFD PRLPKPKGQT
Sugpt-1  LSKMAAFGKV WKLEQKGGSI IGGTTLKAIQIE RSNPKKPRFD QRLPKPKGQT
Scpt-1 .................................................................

301
Rapept-1  VGSRKGLTM LPEAISARLG DKVKVSWKLS SITKLASGEG SLTYETPEGI
Arabpt-1  VGSRKGLRM LPEAISARLG SKVKLSWKLS GITKLESGGY NLTYETPDGL
Sorghumpt-1 VASFRKGLAM LPNATSSLG SKVKLSWKLT SMKDSKGKGY VLEYETPEGV
Mzt-1  VASFRKGLAM LPNATSSLG SKVKLSWKLT SITKSDDKGY VLEYETPEGV
Wtpt-1  VASFRKGLAM LPNATSSLG SKVKLSWKLT SITKADNNQGY VGLETPEGV
Ricept-1  VASFRKGLTM LPDAITSLG SKVKLSWKLT SITKSDNQGY ALVETYPEGV
Cottonpt-1 VGSFRKGLTM LPEAISARLG SNVKLSWKLS SITKLDGGY NLTFETPEGM
Soybeanpt-1 VGSFRKGLTM LPDAISARLG NKVKLSWKLS SISKLDGESY SLTYETPEGV
Sugpt-1  VGSFRKGLVM LTPAISARLG SRVKLSWKLS SIVKSLMGEY SLTYETPGDL
Scpt-1 .................................................................
Rapept-1  VTVQSKSVVM TVP.SH.VASSL LRPLS.DSAAE ALSKLYYPFV AAVSISYAKE
Arabpt-1  VSVQSKSVVM TVP.SH.VASGL LRPLSE.SAAE ALSKLYYPFV AAVSISYPKE
Sorghumpt-1  VLQAKSVIM TIPSYVASDI LRPLS.GDAAD VLSRFYYPPV AAUTVSYPKE
Mzpt-1  VSVQAKSVIM TIPSYVASNI LRPLS.SDAAD ALSRFYYPFV AAVTVSYPKE
Wtpt-1  VSVQAKSVIM TIPSYVASDI LRPLS.IDAAD ALSKFYYPPV AAUTVSYPKE
Ricept-1  VSVQAKTYVM TIPSYVASDI LRPLS.SDAAD ALSIFYYPPV AAUTVSYPKE
Cottonpt-1  VSLQRSVVM TIPSYVASNL LHPLS.AADA ALSQFYYPFV AVTTVSYPKE
Soybeanpt-1  VSLQCTVVL TIPSYVASTL LRPLS.AADA ALSKFYYPPV AAVSISYPKE
Sugpt-1  VSVRTKSVVM TVP.SYVASRL LRPLS.DSAAD SLSKFYYPFV AAVSLSYYPE
Scpt-1

Rapept-1  AIRSECLIDG ELKGFGQQLHP RTQKVETLGT IYSSSLFPNR APPGRVLLLN
Arabpt-1  AIRTECLIDG ELKGFGQQLHP RTQGVETLGT IYSSSLFPNR APPGRILLLLN
Sorghumpt-1  AIRKECLIDG ELQGFQQLHP RSQGVTQLGT IYSSSLFPNR APPGRVLLLN
Mzpt-1  AIRKECLIDG ELQGFQQLHP RSQGVTQLGT IYSSSLFPNR APPGRVLLLN
Wtpt-1  AIRKECLIDG ELQGFQQLHP RSQGVTQLGT IYSSSLFPNR APPGRVLLLN
Ricept-1  AIRKECLIDG ELQGFQQLHP RSQGVTQLGT IYSSSLFPNR APPGRVLLLN
Cottonpt-1  AIRKECLIDG ELKGFGQQLHP RSQGIELTQI IYSSSLFPNR APGRVLLLN
Soybeanpt-1  AIRSECLIDG ELKGFGQQLHP RSQGVTQILG IYSSSLFPNR APPGRVLLLN
Sugpt-1  AIRSECLIDG ELQGFQQLHP RSQGVTQLGT IYSSSLFPGR APPGRLLLLS
Scpt-1

Rapept-1  YIGGATNTGI LSKSE.GELVE AVDRDLRKML IKPSSTDPLV LGVKLWPQAI
Arabpt-1  YIGGSTNTGI LSKSE.GELVE AVDRDLRKML IKPNSTDKL LGVRWQPQAI
Sorghumpt-1  YIGGATNTGI VSKTESELVE AVDRDLRKML INPTAVDPLV LGVRWQPQAI
Mzpt-1  YIGGATNTGI VSKTESELVE AVDRDLRKML INSTAVDPLV LGVRWQPQAI
Wtpt-1  YIGGATNTGI VSKTESELVE AVDRDLRKML INPRAADPLA LGVRWQPQAI
Ricept-1  YIGGSTNTGI VSKTESELVE AVDRDLRKML INPRAV.DPLV LGVRWQPQAI
Cottonpt-1  YIGGATNTGI LSKTE.GELVE AVDRDLRKML INPN.AKDPLV LGVRWQPQAI
Soybeanpt-1  YIGGATNTGI LSKTESELVE TVDRDLRKML INPNAQDPVF VGRVLWPQAI
Sugpt-1  YIGGAKNP.GI LNKSKDNLAK TVDKDL.RML INPDALPRV LGVRWQPQAI
Scpt-1

- 101 -
Rapept-1   PQFLIGHIDL VDAAKASLSS SGHEGLFLGG NYVAGVALGR CVEGAYETAT
Arabpt-1   PQFLVGHPDI LDTAKSLTS SGYEGLFLGG NYVAGVALGR CVEGAYETAI
Sorghumpt-1 PQFLVGHLDL LEAAKSLDQ GGYNGLFLGG NYVAGVALGR CIEGAYESAA
Mzpt-1     PQFLVGHLDDL LEAAKALDR GGYDGLFLGG NYVAGVALGR CVEGAYESAS
Wtpt-1     PQFLNGHLDR LAAKSLAQ GGYDGLFLGG KYVAGVALGR CIEGAYESAS
Ricept-1   PQFLIGHLDH LEAAKSLGK GGYDGLFLGG NYVAGVALGR CVEGAYESAS
Cottonpt-1 PQFLVGHLDL LDSAKMALRD SGFHGLFLGG NYVSGVALGR CVEGAYEVA
Soybeanpt1 PQFLVGHLDL LVAKASINR TGFEGLFLGG NYVSGVALGR CVEGAYEVA
Sugpt-1    PQFSIGHFDL LDAAKALTD TGVGLFLGG NYVSGVALGR CIEGAYESAA
Scpt-1     PQFLVGHLDL LEAAKSLDR GGYDGLFLGG NYVAGVALGR CVEGAYESAS

Rapept-1   QVNDFMSRHY YK*
Arabpt-1   EVNNFMSRHY YK*
Sorghumpt-1 QIYDFTKYA YK*
Mzpt-1     QISDFLTKYA YK*
Wtpt-1     QVDFTKYA YK*
Ricept-1   QISDLYTKYA YK*
Cottonpt-1 EVKEFLSQYA YK*
Soybeanpt1 EVNDFLTNRV YK*
Sugpt-1    EVVDFLSQYS DK*
Scpt-1     QIYDFTKYA YK*

<table>
<thead>
<tr>
<th>Sub-sequence #</th>
<th>Sub-sequence AA sequence</th>
<th>Δn AA wild-type</th>
<th>Δn AA substitutions</th>
<th>Corresponding AA position of Δn in Table 1A</th>
<th>Exemplary mutants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>APΔ1F</td>
<td>R</td>
<td>C, L</td>
<td>169</td>
<td>Mz88Cys, Mz88Leu</td>
</tr>
<tr>
<td>2</td>
<td>FΔ2S</td>
<td>C</td>
<td>F, L, K</td>
<td>240</td>
<td>Mz159Phe, Mz159Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mz159Lys</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---------</td>
<td></td>
</tr>
</tbody>
</table>
| 3 | YΔ₃G | A | V, T, L, C, I | pAraC-1Val  
|   |   |   |   | pAraC-1Thr  
|   |   |   |   | pAraC-1Leu  
|   |   |   |   | pAraC-1Cys  
|   |   |   |   | pAraC-1Ile  
|   |   |   |   | pMzC-1Val  
|   |   |   |   | pMzC-1Thr  
|   |   |   |   | pMzC-1Leu  
|   |   |   |   | pWhtC-1Val  
|   |   |   |   | pWhtC-1Thr  
|   |   |   |   | pSoyC-1Thr  
|   |   |   |   | pSoyC-1Leu |
| 4 | AΔ₄D | G | S, L | 246  
|   |   |   |   | pAraC-3Ser  
|   |   |   |   | pMzC-3Ser  
|   |   |   |   | pMzC-3Leu  
|   |   |   |   | pWhtC-3Ser |
| 5 | YΔ₅P | P | S, H | 388  
|   |   |   |   | Soy369Ser  
|   |   |   |   | Soy369His  
|   |   |   |   | Cot365Ser |
| 6 | PΔ₅A | V | L | 390  
|   |   |   |   | Wht356Leu |
| 7 | Δ₇IG | Y | C, I, L, T, M, V, A, H, R | 451  
|   |   |   |   | pAraC-2Cys  
|   |   |   |   | pAraC-2Ile  
|   |   |   |   | pAraC-2Leu  
|   |   |   |   | pAraC-2Thr  
|   |   |   |   | pAraC-2Met  
|   |   |   |   | pAraC-2Val  
|   |   |   |   | pAraC-2Ala  
|   |   |   |   | pMzC-2Ile  
|   |   |   |   | pMzC-2Met  
|   |   |   |   | pSoyC-2Leu  
|   |   |   |   | pSoyC-2Ile  
<p>|   |   |   |   | pSugC-2Cys |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th>pSugC-2Leu</th>
<th>pSugC-2Ile</th>
<th>pSugC-2Val</th>
<th>pSugC-2Met</th>
<th>pCotC-2Cys</th>
<th>pCotC-2His</th>
<th>pCotC-2Arg</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>YIGGΔ₈</td>
<td>A, S</td>
<td>P</td>
<td>455</td>
<td>Wht421Pro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>AΔ₉P</td>
<td>I</td>
<td>T, H, G, N</td>
<td>500</td>
<td>Mz419Thr</td>
<td>Mz419His</td>
<td>Mz419Gly</td>
<td>Mz419Asn</td>
<td>Wht466Thr</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>GΔ₁₀A</td>
<td>V</td>
<td>A</td>
<td>536</td>
<td>Wht502Ala</td>
<td>Soy517Ala</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>KAΔ₁₈F</td>
<td>A</td>
<td>T, V</td>
<td>256</td>
<td>Mz175Thr</td>
<td>Mz175Val</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>QΔ₁₉H</td>
<td>L</td>
<td>S</td>
<td>418</td>
<td>Mz337Ser</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Second-site mutations**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th>AraC1 18Leu</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>QΔ₁₁S</td>
<td>P</td>
<td>L</td>
<td>143</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AraC2 49Ala</td>
</tr>
<tr>
<td>12</td>
<td>IGGΔ₁₂</td>
<td>T</td>
<td>I, A</td>
<td>274</td>
</tr>
<tr>
<td>13</td>
<td>SWXLΔ₁₃</td>
<td>S, T</td>
<td>L</td>
<td>330</td>
</tr>
<tr>
<td>14</td>
<td>LΔ₁₄Y</td>
<td>N</td>
<td>S</td>
<td>450</td>
</tr>
<tr>
<td>15</td>
<td>GΔ₁₅XGL</td>
<td>Y, H, F, V</td>
<td>C</td>
<td>523</td>
</tr>
</tbody>
</table>
Table 2

Comparison of the *Arabidopsis* (SEQ ID NO:4) and maize (SEQ ID NO:8) protox-2 amino acid sequences. Identical residues are denoted by the vertical bars between the two sequences. Alignment was performed using the GAP program described in Deveraux *et al.*, *Nucleic Acids Res.* 12:387-395 (1984). Percent similarity: 75.889 / percent identity: 57.905.

<table>
<thead>
<tr>
<th>16</th>
<th>TΔ19G</th>
<th>L</th>
<th>S</th>
<th>428</th>
<th>Mz34</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>YVΔ17G</td>
<td>A, (S)</td>
<td>T</td>
<td>534</td>
<td>7Ser-453Thr</td>
</tr>
</tbody>
</table>

| 1 | .................. | MASQAVAD.HQIEAVSGKRVAV | 21 |
| 1 | MLALTASASSASSHPYRHASAHTRRPRLRRAVLAMAGSDDPRAAPARSVA | 50 |
| 22 | VGAVGSCLASSAYKLKSRGLTVTVFEDGRVGGKLRRMIONGLIWDEGANT | 71 |
| 51 | VGAVGSCLASSAYLRQSQGVTVFEDRAGKIRTNSEGGFWDEGANT | 100 |
| 72 | MTEEAPEVGSLLDDLGLREKQQPISQKRYTRNGVPMLTPNPIELVT | 121 |
| 101 | MTEGWEASRLDDGLQDKQQPYNSQHKRYIVKDGAPALIPSDPISLMK | 150 |
| 122 | SSVLSTQSFQILLEPLWKK...KSSKVSDASAEVSEFQRHFGQE | 167 |
| 151 | SSVLSTSKIALFFEPFLYYKANKTRNSGKVSEHLSVSFGFCERHFGE | 200 |
| 168 | VVDYLIDPFPVGGTSAADPDSLMSMKHDWPNVEKLFSFGSIIVGAIERTKFA | 217 |
| 201 | VVDYFVDPPVAGTSAGDPLSIRHAFLPNLIRKGYGSGVGVAILSKLA | 250 |
218 AKGGKSRDTKSSPGTKGSRGSFSFKGGMQILPDYLCKSLSHDEINLDSK 267
251 AKGDPVKTRHDSSGKRNRVRFSFHGMQSLINALHENVGDDNVKLTGE
268 VLSLS..YNSSGRQENWSLSCVSHNETQPQ...NPHYDAVIMTAPLCNVK 312
301 VLSLACTFGVPALGRWSISVDSDSDDKDLSNQTFDAVIMTAPLSNVR 350
313 EMKVNKGQPFQNLNPEINYMPSVLITFTKEKVRPLEGFGVLPK 362
351 RMKFTKGGAPVLDFLPMDYLPLSMLMTAKFKDDVKKPLGFGVLPYK 400
363 E.QKHGFKTLGTLFSSMFPDRSPDVHLVYTTTFIGGSRNQELAKASTDEL 411
401 EQQKHGLKTLGTLFSSMFPDRAPDDQYLYTTTFVGGSNRDLAGAPTSIL 450
412 KQVVTSDLQRLLGVEQEPVSNHYYWKRKAFLDYDSSYDSVMEAIDKMEND 461
451 KQLVTSDELKLLGVEQPTFVKHVVWGNAPFLYGHDDYSSVLEAEKMEKN 500
462 LPGFFYAGNHHRGGLSVGKSIAGCAADLVISYLESCSNDKPKNDL* 509
462 LPGFFYAGNRSKGDGLAVGSVIASGSKAADLAISYLESHTKHNNSH*... 545

Table 3A
Cross tolerance of plant protox mutants to various protox inhibitors.

<table>
<thead>
<tr>
<th>Formula</th>
<th>AraC-1Val</th>
<th>AraC-2Cys</th>
<th>AraC-1Thr</th>
<th>AraC-3Thr</th>
<th>Mzc-1Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>XVII</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Vlla</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>IV</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>XV</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>XI</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>XVI</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>XII</td>
<td>+</td>
<td>-</td>
<td>++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>XIV</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>*X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

+ = 10X or more tolerant than WT  
++ = 100X or more tolerant than WT  
- = no cross tolerance  
* = this compound was tested but provided no information

**Table 3B**  
Cross tolerance of plant protox mutants to various protox inhibitors.
### Table 4

Cross tolerance to various protox inhibitors in a seed germination assay.

<table>
<thead>
<tr>
<th>Formula</th>
<th>Common name</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>acifluorofen</td>
<td>+</td>
</tr>
<tr>
<td>III</td>
<td>fomasafen</td>
<td>+</td>
</tr>
<tr>
<td>IV</td>
<td>fluoroglycofen</td>
<td>±</td>
</tr>
<tr>
<td>IVb</td>
<td>bifenox</td>
<td>+</td>
</tr>
<tr>
<td>IVc</td>
<td>oxyfluorofen</td>
<td>+</td>
</tr>
<tr>
<td>IVd</td>
<td>lactofen</td>
<td>±</td>
</tr>
<tr>
<td>VIIa</td>
<td>fluthiacet-methyl</td>
<td>++</td>
</tr>
<tr>
<td>X</td>
<td>sulfentrazone</td>
<td>+</td>
</tr>
<tr>
<td>XI</td>
<td>flupropazil</td>
<td>++</td>
</tr>
<tr>
<td>XIV</td>
<td>flumiclorac</td>
<td>+</td>
</tr>
<tr>
<td>XVI</td>
<td>flumioxazin</td>
<td>+++</td>
</tr>
<tr>
<td>XVII</td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>XXIa</td>
<td>BAY 11340</td>
<td>+</td>
</tr>
<tr>
<td>XXII</td>
<td></td>
<td>++</td>
</tr>
</tbody>
</table>

± ≤ 10X more tolerant than wt
+ ≥ 10X more tolerant than wt
++ ≥ 100X more tolerant than wt
+++ ≥ 1000X more tolerant than wt

Various modifications of the invention described herein will become apparent to those skilled in the art. Such modifications are intended to fall within the scope of the appended claims.
INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 55, line 6-33

B. IDENTIFICATION OF DEPOSIT

Name of depository institution
Agricultural Research Service, Patent Culture Collection (NRRL)
Northern Regional Research Center
International Depository Authority

Address of depository institution (including postal code and country)
1815 North University Street
Peoria, Illinois 61604
United States of America (U.S.A)

Date of deposit 19 March 1996 (19.03.96)
Accession Number NRRL B-21545

C. ADDITIONAL INDICATIONS (leave blank if not applicable)
This information is continued on an additional sheet

We request the Expert Solution where available.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. 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")

For receiving Office use only
This sheet was received with the international application 30 JUN 2000
Authorized officer Mrs. H. Fransz

For International Bureau use only
This sheet was received by the International Bureau on:
Authorized officer

Form PCT/RO/134 (July 1998)
INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description
on page 55, line 6-33

B. IDENTIFICATION OF DEPOSIT

<table>
<thead>
<tr>
<th>Name of depositary institution</th>
<th>Agricultural Research Service, Patent Culture Collection (NRRL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Northern Regional Research Center</td>
</tr>
<tr>
<td></td>
<td>International Depository Authority</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Address of depository institution (including postal code and country)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1815 North University Street</td>
</tr>
<tr>
<td>Peoria, Illinois 61604</td>
</tr>
<tr>
<td>United States of America (U.S.A)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date of deposit</th>
<th>15 December 1995 (15.12.95)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accession Number</td>
<td>NRRL B-21516</td>
</tr>
</tbody>
</table>

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet

We request the Expert Solution where available.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (If the indications are not for all designated States)

E. 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")

For receiving Office use only

☐ This sheet was received with the international application

3 0 JUN 2000

Authorized officer

Mrs. H. Fransz

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

Form PCT/RQ/134 (July 1998)
INDICATIONS RELATING TO DEPOSITED MICROORGANISM OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 55, line 6-33

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution

Agricultural Research Service, Patent Culture Collection (NRRL)
Northern Regional Research Center
International Depositary Authority

Address of depositary institution (including postal code and country)

1815 North University Street
Peoria, Illinois 61604
United States of America (U.S.A.)

Date of deposit

01 July 1996 (01.07.96)

Accession Number

NRRL B-21594

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet

We request the Expert Solution where available.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. 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”)

For receiving Office use only

☑ This sheet was received with the international application 30 JUN 2000

Authorized officer

Mrs. H. Fransz

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer
INDICATIONS RELATING TO DEPOSITED MICROORGANISM 
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description
on page 55, line 6-33

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution
Agricultural Research Service, Patent Culture Collection (NRRL)
Northern Regional Research Center
International Depository Authority

Address of depositary institution (including postal code and country)

1815 North University Street
Peoria, Illinois 61604
United States of America (U.S.A)

Date of deposit 29 July 1996 (29.07.96) Accession Number NRRL B-21595N

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet

We request the Expert Solution where available.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. 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")

For receiving Office use only

This sheet was received with the international application
30 JUN 2000

Authorized officer: Mrs. H. Fransz

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer
INDICATIONS RELATING TO DEPOSITED MICROORGANISM OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 55, line 6-33.

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depository institution
Agricultural Research Service, Patent Culture Collection (NRRL)
Northern Regional Research Center
International Depository Authority

Address of depository institution (including postal code and country)
1815 North University Street
Peoria, Illinois 61604
United States of America (U.S.A.)

Date of deposit
23 August 1996 (23.08.96)

Accession Number
NRRL B-21615

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet

We request the Expert Solution where available.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. 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")

For receiving Office use only

This sheet was received with the international application
3 0 JUN 2000

Authorized officer
Mrs. H. Fransz

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer
INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 55, line 6-33

B. IDENTIFICATION OF DEPOSIT

Name of depositary institution
Agricultural Research Service, Patent Culture Collection (NRRL)
Northern Regional Research Center
International Depository Authority

Address of depositary institution (including postal code and country)
1815 North University Street
Peoria, Illinois 61604
United States of America (U.S.A)

Date of deposit
06 December 1996 (06.12.96)

Accession Number
NRRL B-21648

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

We request the Expert Solution where available.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. 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")

For receiving Office use only
This sheet was received with the international application
3 0 JUN 2000
Authorized officer
Mrs. H. Fransz

For International Bureau use only
This sheet was received by the International Bureau on:

Authorized officer
INDICATIONS RELATING TO DEPOSITED MICROORGANISM OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 55, line 6-33

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution
Agricultural Research Service, Patent Culture Collection (NRRL)
Northern Regional Research Center
International Depository Authority

Address of depositary institution (including postal code and country)
1815 North University Street
Peoria, Illinois 61604
United States of America (U.S.A)

Date of deposit
06 December 1996 (06.12.96)
Accession Number
NRRL B-21649

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet

We request the Expert Solution where available.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. 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")

For receiving Office use only

This sheet was received with the international application
30 JUN 2000

Authorized officer
Mrs. H. Franz

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer
INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 55, line 6-33

B. IDENTIFICATION OF DEPOSIT

Name of depositary institution
Agricultural Research Service, Patent Culture Collection (NRRL)
Northern Regional Research Center
International Depositary Authority

Address of depositary institution (including postal code and country)
1815 North University Street
Peoria, Illinois 61604
United States of America (U.S.A)

Date of deposit
14 November 1994 (14.11.94)
Accession Number
NRRL B-21339N

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

We request the Expert Solution where available.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. 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")
INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description
on page 55, line 6-33

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution
Agricultural Research Service, Patent Culture Collection (NRRL)
Northern Regional Research Center
International Depository Authority

Address of depositary institution (including postal code and country)
1815 North University Street
Peoria, Illinois 61604
United States of America (U.S.A)

Date of deposit
15 December 1995 (15.12.95)

Accession Number
NRRL B-21515

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet

We request the Expert Solution where available...

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. 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")

Form PCT/RO/134 (July 1998)

- 117 -
INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 55, line 6-33

B. IDENTIFICATION OF DEPOSIT

<table>
<thead>
<tr>
<th>Name of depositary institution</th>
<th>Agricultural Research Service, Patent Culture Collection (NRRL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Northern Regional Research Center</td>
</tr>
<tr>
<td></td>
<td>International Depositary Authority</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Address of depositary institution (including postal code and country)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1815 North University Street</td>
</tr>
<tr>
<td>Peoria, Illinois 61604</td>
</tr>
<tr>
<td>United States of America (U.S.A)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date of deposit</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>19 March 1996</td>
<td>NRRL B-21546</td>
</tr>
</tbody>
</table>

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

We request the Expert Solution where available.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE

(if the indications are not for all designated States)

E. 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")

For receiving Office use only

☐ This sheet was received with the international application

30 JUN 2000

Authorized officer

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

Form PCT/R0/134 (July 1998)
INDICATIONS RELATING TO DEPOSITED MICROORGANISM OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 55, line 6-33.

<table>
<thead>
<tr>
<th>B. IDENTIFICATION OF DEPOSIT</th>
<th>Further deposits are identified on an additional sheet</th>
</tr>
</thead>
</table>
| Name of depository institution | Agricultural Research Service, Patent Culture Collection (NRRL)  
|                              | Northern Regional Research Center  
|                              | International Depository Authority  
| Address of depository institution (including postal code and country) | 1815 North University Street  
|                              | Peoria, Illinois 61604  
|                              | United States of America (U.S.A)  
| Date of deposit | 06 December 1996 (06.12.96)  
| Accession Number | NRRL B-21650  

C. ADDITIONAL INDICATIONS (leave blank if not applicable)  
This information is continued on an additional sheet

We request the Expert Solution where available.

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. 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")
What is Claimed is:

1. A nucleic acid molecule comprising a nucleotide sequence that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein said modified enzyme is resistant to an inhibitor of the naturally occurring form of said enzyme, wherein said modified enzyme comprises at least one of the following amino acid sub-sequences:
   (a) $\text{KA}\Delta_{18}\text{F}$, wherein $\Delta_{18}$ is an amino acid other than alanine,
   (b) $\text{Q}\Delta_{19}\text{H}$, wherein $\Delta_{19}$ is an amino acid other than leucine,
   (c) $\text{AP}\Delta_{1}\text{F}$, wherein $\Delta_{1}$ is an amino acid other than arginine,
   (d) $\text{F}\Delta_{2}\text{S}$, wherein $\Delta_{2}$ is leucine,
   (e) $\text{Y}\Delta_{3}\text{G}$, wherein $\Delta_{3}$ is isoleucine,
   (f) $\Delta_{7}\text{IG}$, wherein $\Delta_{7}$ is histidine or alanine,
   (g) $\text{T}\Delta_{18}\text{G}$, wherein $\Delta_{18}$ is an amino acid other than leucine and $\text{YV}\Delta_{17}\text{G}$, wherein $\Delta_{17}$ is an amino acid other than alanine,
wherein the nucleotide sequence that encodes the naturally occurring form of said enzyme is derived from a plant.

2. A nucleotide acid molecule according to claim 1 comprising a nucleotide sequence that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein said modified enzyme is resistant to an inhibitor of the naturally occurring form of said enzyme, wherein said modified enzyme comprises at least one amino acid sub-sequence selected from the group consisting of:
   (a) $\text{KA}\Delta_{18}\text{F}$, wherein $\Delta_{18}$ is an amino acid other than alanine,
   (b) $\text{Q}\Delta_{19}\text{H}$, wherein $\Delta_{19}$ is an amino acid other than leucine,
   (c) $\text{AP}\Delta_{1}\text{F}$, wherein $\Delta_{1}$ is an amino acid other than arginine,
   (d) $\text{F}\Delta_{2}\text{S}$, wherein $\Delta_{2}$ is leucine,
   (e) $\text{Y}\Delta_{3}\text{G}$, wherein $\Delta_{3}$ is isoleucine,
   (f) $\Delta_{7}\text{IG}$, wherein $\Delta_{7}$ is histidine or alanine,
   (g) $\text{T}\Delta_{18}\text{G}$, wherein $\Delta_{18}$ is an amino acid other than leucine and $\text{YV}\Delta_{17}\text{G}$, wherein $\Delta_{17}$ is an amino acid other than alanine,
wherein said nucleotide sequence that encodes said modified enzyme hybridizes to any one of the nucleotide sequences set forth in SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:9,
3. A nucleic acid molecule according to any one of claims 1 or 2 comprising a nucleotide sequence that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein said modified enzyme is resistant to an inhibitor of the naturally occurring form of said enzyme, wherein said modified enzyme comprises at least one of the following amino acid sub-sequences:

(a) \( \text{KAD}_{18}\text{F} \), wherein \( \Delta_{18} \) is an amino acid other than alanine,
(b) \( \text{QAD}_{19}\text{H} \), wherein \( \Delta_{19} \) is an amino acid other than leucine,
(c) \( \text{APD}_{1}\text{F} \), wherein \( \Delta_{1} \) is an amino acid other than arginine;
(d) \( \Delta_{7}\text{IG} \), wherein \( \Delta_{7} \) is histidine,

wherein the nucleotide sequence that encodes the naturally occurring form of said enzyme is derived from a plant.

4. The nucleic acid molecule according to claim 1, wherein said modified enzyme comprises the amino acid sub-sequence \( \text{KAD}_{18}\text{F} \), wherein \( \Delta_{18} \) is an amino acid other than alanine, preferably threonine or valine, but especially valine.

5. The nucleic acid molecule according to claim 1, wherein said modified enzyme comprises the amino acid sub-sequence \( \text{QAD}_{19}\text{H} \), wherein \( \Delta_{19} \) is an amino acid other than leucine, preferably serine.

6. The nucleic acid molecule according to claim 1, wherein said modified enzyme comprises the amino acid sub-sequence \( \text{APD}_{1}\text{F} \), wherein \( \Delta_{1} \) is leucine or cysteine.

7. The nucleic acid molecule according to claim 1, wherein said modified enzyme comprises the amino acid sub-sequence \( \Delta_{7}\text{IG} \), wherein \( \Delta_{7} \) is histidine or alanine, preferably histidine.
8. A nucleic acid molecule comprising a nucleotide sequence that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein said modified enzyme is resistant to an inhibitor of the naturally occurring form of said enzyme, and wherein said modified enzyme comprises a first amino acid sub-sequence and further at least one additional amino acid sub-sequence selected from the group consisting of:

(a) QΔ₁₁S, wherein Δ₁₁ is an amino acid other than proline;
(b) IGGΔ₁₂, wherein Δ₁₂ is an amino acid other than threonine;
(c) SWXLΔ₁₃, wherein Δ₁₃ is an amino acid other than serine;
(d) LΔ₁₄Y, wherein Δ₁₄ is an amino acid other than asparagine;
(e) GΔ₁₅XGL, wherein Δ₁₅ is an amino acid other than tyrosine.

9. A nucleic acid molecule according to claim 8, wherein said first sub-sequence is a sequence according to any one of claims 1 to 7.

10. The nucleic acid molecule according to any one of claims 8 and 9, wherein said additional sub-sequence is QΔ₁₁S, wherein Δ₁₁ is an amino acid other than proline, preferably leucine.

11. The nucleic acid molecule according to any one of claims 8 and 9, wherein said additional sub-sequence is IGGΔ₁₂, wherein Δ₁₂ is an amino acid other than threonine, preferably isoleucine or alanine.

12. The nucleic acid molecule according to any one of claims 8 and 9, wherein said additional sub-sequence is SWXLΔ₁₃, wherein Δ₁₃ is an amino acid other than serine, preferably leucine.

13. The nucleic acid molecule according to any one of claims 8 and 9, wherein said additional sub-sequence is LΔ₁₄Y, wherein Δ₁₄ is an amino acid other than asparagine, preferably serine.
14. The nucleic acid molecule according to any one of claims 8 and 9, wherein said additional sub-sequence is $G\Delta_{12}XGL$, wherein $\Delta_{15}$ is an amino acid other than tyrosine, preferably cysteine.

15. The nucleic acid molecule according to any one of claims 1 to 9, wherein wherein the modified enzyme comprises: the amino acid sub-sequence $\Delta_7IG$, wherein $\Delta_7$ is an amino acid other than tyrosine; preferably isoleucine, the amino acid sub-sequences $IGG\Delta_{12}$, wherein $\Delta_{12}$ is an amino acid other than threonine; preferably isoleucine, and the amino acid sub-sequence $SWXL\Delta_{13}$, wherein $\Delta_{13}$ is an amino acid other than serine, preferably leucine.

16. A nucleic acid molecule according to any one of claims 1 and 8 to 9 comprising a nucleotide sequence isolated from a plant that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein the modified enzyme is resistant to an inhibitor of a naturally occurring protox enzyme, wherein the nucleotide sequence is further characterized in that at least one of the following conditions is met:

(a) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $AP\Delta_1F$, wherein $\Delta_1$ is leucine;
(b) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $F\Delta_2S$, wherein $\Delta_2$ is leucine;
(c) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $Y\Delta_3G$, wherein $\Delta_3$ is isoleucine;
(d) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $\Delta_7IG$, wherein $\Delta_7$ is alanine or histidine;
(e) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence $Y\Delta_3G$, wherein $\Delta_3$ is an amino acid other than alanine, but preferably cysteine or isoleucine, more preferably isoleucine, and the nucleic acid sequence also has a sequence that encodes one of the group consisting of:

(1) sub-sequence $Q\Delta_{11}S$, wherein $\Delta_{11}$ is an amino acid other than proline,
(2) sub-sequence $IGG\Delta_{12}$, wherein $\Delta_{12}$ is an amino acid other than threonine,
(3) sub-sequence SWXLA_{13}, wherein \( \Delta_{13} \) is an amino acid other than serine,

(4) sub-sequence L\( \Delta_{14} \)Y, wherein \( \Delta_{14} \) is an amino acid other than asparagine, and

(5) sub-sequence G\( \Delta_{15} \)XGL, wherein \( \Delta_{15} \) is an amino acid other than tyrosine;

(f) the nucleic acid sequence has a sequence that encodes amino acid sub-sequence \( \Delta_{7} \)IG, wherein \( \Delta_{7} \) is an amino acid other than tyrosine, but preferably threonine, alanine or histidine, more preferably alanine or histidine and most preferably histidine, and the nucleic acid sequence also has a sequence that encodes one of the group consisting of:

(1) sub-sequence Q\( \Delta_{11} \)S, wherein \( \Delta_{11} \) is an amino acid other than proline,

(2) sub-sequence IGG\( \Delta_{12} \), wherein \( \Delta_{12} \) is an amino acid other than threonine,

(3) sub-sequence SWXLA_{13}, wherein \( \Delta_{13} \) is an amino acid other than serine,

(4) sub-sequence L\( \Delta_{14} \)Y, wherein \( \Delta_{14} \) is an amino acid other than asparagine, and

(5) sub-sequence G\( \Delta_{15} \)XGL, wherein \( \Delta_{15} \) is an amino acid other than tyrosine; and

(g) the nucleic acid has a sequence that encodes amino acid sub-sequence T\( \Delta_{16} \)G, wherein \( \Delta_{16} \) is an amino acid other than leucine, and the nucleic acid sequence also has a sequence that encodes amino acid sub-sequence YVA_{17}G, wherein \( \Delta_{17} \) is an amino acid other than alanine.

(h) K\( \Delta_{18} \)F, wherein \( \Delta_{18} \) is an amino acid other than alanine;

(i) Q\( \Delta_{19} \)H, wherein \( \Delta_{19} \) is an amino acid other than leucine.

17. A nucleic acid molecule according to claim 16, wherein said sub-sequence encodes amino acid sub-sequence T\( \Delta_{16} \)G, wherein \( \Delta_{16} \) is an amino acid other than leucine, preferably serine, and said nucleic acid sequence also has a sequence that encodes amino acid sub-sequence YVA_{17}G, wherein \( \Delta_{17} \) is an amino acid other than alanine, preferably threonine.
18. A nucleic acid molecule according to claim 1 encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the arginine occurring at the position corresponding to amino acid 88 of SEQ ID NO:6 is replaced with another amino acid, preferably with a leucine, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity.

19. A nucleic acid molecule according to claim 1 encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the cysteine occurring at the position corresponding to amino acid 159 of SEQ ID NO:6 is replaced with a leucine.

20. A nucleic acid molecule according to claim 1 encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the alanine occurring at the position corresponding to amino acid 175 of SEQ ID NO:6 is replaced with another amino acid, preferably with a valine or threonine, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity.

21. A nucleic acid molecule according to claim 1 encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the leucine occurring at the position corresponding to amino acid 337 of SEQ ID NO:6 is replaced with another amino acid, preferably with a serine, wherein the modified protox is tolerant to a herbicide in amounts that inhibit the naturally occurring protox activity.

22. A nucleic acid molecule according to claim 1 encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the tyrosine occurring at the position corresponding to amino acid 428 of SEQ ID NO:16 is replaced with a histidine or alanine, preferably with a histidine.

23. A nucleic acid molecule according to claim 1 encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the alanine occurring at the position corresponding to amino acid 220 of SEQ ID NO:2 is replaced with an isoleucine or a tyrosine.
24. A nucleic acid molecule according to claim 1 encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox wherein the tyrosine occurring at the position corresponding to amino acid 426 of SEQ ID NO:2 is replaced with a alanine.

25. A nucleic acid molecule according to any one of claims 1 and 8 to 9 encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox having a first amino acid substitution and a second amino acid substitution; the first amino acid substitution having the property of conferring resistance to a protox inhibitor; and the second amino acid substitution having the property of enhancing the resistance conferred by the first amino acid substitution but especially a DNA molecule encoding a modified protoporphyrinogen oxidase (protox) comprising a plant protox, wherein the plant is selected from the group consisting of maize, wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, sugar cane, and Arabidopsis.

26. A nucleic acid molecule according to claim 25 wherein the first amino acid substitution occurs at a position selected from the group consisting of:

(a) the position corresponding to the arginine at amino acid 88 of SEQ ID NO:6;
(b) the position corresponding to the alanine at amino acid 175 of SEQ ID NO:6;
and
(c) the position corresponding to the leucine at amino acid 337 of SEQ ID NO:6.

27. Particularly preferred is a DNA molecule wherein the first amino acid substitution occurs at a position selected from the group consisting of:

(a) the position corresponding to the arginine at amino acid 88 of SEQ ID NO:6;
(b) the position corresponding to the alanine at amino acid 164 of SEQ ID NO:6;
(c) the position corresponding to the glycine at amino acid 165 of SEQ ID NO:6;
(d) the position corresponding to the tyrosine at amino acid 370 of SEQ ID NO:6;
(e) the position corresponding to the cysteine at amino acid 159 of SEQ ID NO:6;
(f) the position corresponding to the isoleucine at amino acid 419 of SEQ ID NO:6;
(g) the position corresponding to the valine at amino acid 356 of SEQ ID NO:10;
(h) the position corresponding to the serine at amino acid 421 of SEQ ID NO:10;
(i) the position corresponding to the valine at amino acid 502 of SEQ ID NO:10;
(j) the position corresponding to the alanine at amino acid 211 of SEQ ID NO:10;
(k) the position corresponding to the glycine at amino acid 212 of SEQ ID NO:10;
(l) the position corresponding to the isoleucine at amino acid 466 of SEQ ID NO:10;
(m) the position corresponding to the proline at amino acid 369 of SEQ ID NO:12;
(n) the position corresponding to the alanine at amino acid 226 of SEQ ID NO:12;
(o) the position corresponding to the tyrosine at amino acid 432 of SEQ ID NO:12;
(p) the position corresponding to the valine at amino acid 517 of SEQ ID NO:12;
(q) the position corresponding to the tyrosine at amino acid 428 of SEQ ID NO:16;
(r) the position corresponding to the proline at amino acid 365 of SEQ ID NO:16;
(s) the position corresponding to the tyrosine at amino acid 449 of SEQ ID NO:18;
(t) the position corresponding to the alanine at amino acid 175 of SEQ ID NO:6;
and
(u) the position corresponding to the leucine at amino acid 337 of SEQ ID NO:6;

wherein the second amino acid substitution occurs at a position selected from the group consisting of:

(a) the position corresponding to the serine at amino acid 305 of SEQ ID NO:2;
(b) the position corresponding to the threonine at amino acid 249 of SEQ ID NO:2;
(c) the position corresponding to the proline at amino acid 118 of SEQ ID NO:2;
(d) the position corresponding to the asparagine at amino acid 425 of SEQ ID NO:2; and
(e) the position corresponding to the tyrosine at amino acid 498 of SEQ ID NO:2.

28. A nucleic acid molecule according to claim 27, wherein the first amino acid substitution occurs at a position selected from the group consisting of:

(a) the position corresponding to the arginine at amino acid 88 of SEQ ID NO:8;
(b) the position corresponding to the alanine at amino acid 175 of SEQ ID NO:6;

and

(c) the position corresponding to the leucine at amino acid 337 of SEQ ID NO:6.

wherein the second amino acid substitution occurs at a position selected from the group consisting of:

(a) the position corresponding to the serine at amino acid 305 of SEQ ID NO:2;
(b) the position corresponding to the threonine at amino acid 249 of SEQ ID NO:2;
(c) the position corresponding to the proline at amino acid 118 of SEQ ID NO:2;
(d) the position corresponding to the asparagine at amino acid 425 of SEQ ID NO:2; and
(e) the position corresponding to the tyrosine at amino acid 498 of SEQ ID NO:2.

29. A nucleic acid molecule according to claim 1 having a double amino acid substitution, wherein one amino acid substitution occurs at the position corresponding to the leucine at amino acid 347 of SEQ ID NO:6, and wherein the second amino acid substitution occurs at the position corresponding to the alanine at amino acid 453 of SEQ ID NO:6.

30. A nucleic acid molecule according to claim 29 having a double amino acid substitution, wherein a leucine occurring at the position corresponding to amino acid 347 of SEQ ID NO:6 is replaced with a serine, and wherein an alanine occurring at the position corresponding to amino acid 453 of SEQ ID NO:6 is replaced with a threonine.

31. The nucleic acid molecule according to any one of claims 1 to 30, wherein said plant is selected from the group consisting of *Arabidopsis*, maize, wheat, soybean, cotton, sugar beet, oilseed rape, rice, sorghum, and sugar cane, preferably maize or cotton.

32. A modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein said modified enzyme is encoded by the nucleotide sequence of any one of claims 1 to 30.

33. A chimeric gene comprising a promoter that is active in a plant operatively linked to the nucleic acid molecule any one of claims 1 to 30.

34. A recombinant vector comprising the chimeric gene according to claim 33.

35. A plant cell comprising the nucleic acid molecule according to any one of claims 1 to 30.

36. A plant, plant tissue, plant cell, or plant seed, including the progeny thereof, comprising the nucleic acid molecule according to any one of claims 1 to 30, wherein said nucleic acid molecule is expressed in said plant, plant tissue, plant cell, or plant seed or the
progeny thereof confers tolerance thereupon to an inhibitor of naturally occurring protox activity.

37. A plant, plant tissue, plant cell, or plant seed, including the progeny thereof, according to claim 36, wherein said plant, plant tissue, plant cell, or plant seed or the progeny thereof is selected from the group consisting of Arabidopsis, sugar cane, soybean, barley, cotton, tobacco, sugar beet, oilseed rape, maize, wheat, sorghum, rye, oats, turf and forage grasses, millet, forage and rice.

38. A plant, plant tissue, plant cell, or plant seed, including the progeny thereof, according to claim 36, wherein said plant, plant tissue, plant cell, or plant seed or the progeny thereof is selected from the group consisting of Arabidopsis, soybean, cotton, sugar beet, oilseed rape, maize, wheat, sorghum, rice, and sugar cane.

39. A plant, plant tissue, plant cell, or plant seed, including the progeny thereof, selected from the group consisting of Arabidopsis, soybean, cotton, sugar beet, oilseed rape, maize, wheat, sorghum, rice, and sugar cane, comprising the nucleic acid molecule according to any one of claims 1 to 30.

40. A chimeric gene according to claim 33 comprising a promoter functional in a plant plastid operatively linked to the nucleic acid molecule according to any one of claims 1 to 30.

41. A chimeric gene according to claim 40, wherein said promoter functional in a plant plastid is a clpP gene promoter.

42. A plastid transformation vector comprising the nucleic acid molecule according to any one of claims 1 to 30.

43. A plant plastid comprising the plastid transformation vector according to claim 42.

44. A plant, plant tissue, plant cell, or plant seed, including the progeny thereof, comprising a plant plastid according to claim 43, wherein said modified plant protox enzyme is expressed in said plant, plant tissue, plant cell, or plant seed and confers upon said plant,
plant tissue, plant cell, or plant seed tolerance to an inhibitor of naturally occurring protox activity.

45. A method for controlling the growth of undesired vegetation, which comprises applying to a population of a plant comprising a plant cell according to any one of claims 35 to 39 and 44 an effective amount of a protox-inhibiting herbicide.

46. A method for controlling the growth of undesired vegetation, which comprises applying to a population of a plant comprising a plant plastid according to claim 43 an effective amount of a protox-inhibiting herbicide.

47. A method according to any one of claims 45 and 46, wherein said protox-inhibiting herbicide is selected from the group consisting of an ary luracil, a diphenylether, an oxidiazole, an imide, a phenyl pyrazole, a pyridine derivative, a 3-substituted-2-aryl-4,5,6,7-tetrahydroindazole, a phenopylate and O-phenylpyrrolidino- and piperidinocarbamate analogs of said phenopylate.

48. A method according to claim 47, wherein said protox-inhibiting herbicide is an imide having formula V, VI, VII, VIIa, VIII, IX, IXa, or IXb.

49. A method according to claim 47, wherein said protox-inhibiting herbicide is a pyridyl pyrazole having formula XXIIa or XXIIb.

50. A method according to claim 47, wherein said protox-inhibiting herbicide is a phenylpyrazole having formula XXIV.

51. A method for selectively suppressing the growth of weeds in a field containing a crop of planted crop seeds or plants, comprising the steps of:
   (a) planting herbicide tolerant crops or crop seeds, which comprise a plant cell according to any one of claims 35 to 39 and 44; or a plant plastid according to claim 43; and
   (b) applying to the crops or crop seeds and the weeds in the field a protox-inhibiting herbicide in amounts that inhibit naturally occurring protox activity,
wherein the herbicide suppresses the growth of the weeds without significantly suppressing the growth of the crops.

52. A nucleic acid molecule comprising a nucleotide sequence that encodes a modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein said modified enzyme is resistant to an inhibitor of the naturally occurring form of said enzyme, wherein said modified enzyme comprises at least one amino acid sub-sequence selected from the group consisting of:

(a) KA$_{18}$F, wherein A$_{18}$ is an amino acid other than alanine,
(b) QA$_{19}$H, wherein A$_{19}$ is an amino acid other than leucine,
(c) AP$_{1}$F, wherein A$_{1}$ is leucine,
(d) A$_{7}$IG, wherein A$_{7}$ is histidine,

wherein said nucleotide sequence that encodes said modified enzyme hybridizes to any one of the nucleotide sequences set forth in SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23 or SEQ ID NO:36 under the following conditions:

(i) hybridization in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO4 pH 7.0, 1 mM EDTA at 50$^\circ$ C; and
(ii) wash in 2X SSC, 1% SDS at 50$^\circ$ C.

53. The nucleic acid molecule according to claim 52, wherein the nucleotide sequence that encodes the naturally occurring form of said enzyme is derived from a plant.

54. A modified enzyme having protoporphyrinogen oxidase (protox) activity, wherein said modified enzyme is encoded by the nucleotide sequence of claim 52.
SEQUENCE LISTING

110> Novartis

120> Herbicide-Tolerant Protoporphyrinogen Oxidase

130> PB/5-30911A

140>

141>

150> US 09/373691
151> 1999-08-13
160> 44

170> PatentIn Ver. 2.2

210> 1
211> 1719
212> DNA
213> Arabidopsis thaliana

220>
221> CDS
222> (31)...(1644)
223> Arabidopsis protox-1

400> 1
 tgacaaatt ccgaattctc tgcgattccc atg gag tta tct ctt ctc cgk cgk 54
          Met Glu Leu Ser Leu Leu Arg Pro

                     1       5
   aca gct caa tcg ctt ctt ccg tcg ttt tcg aag ccc aat ctc cga tta 102
   Thr Thr Gln Ser Leu Leu Pro Ser Phe Ser Lys Pro Asn Leu Arg Leu
       10      15      20
   aat gtt tat aag cct ctt aga ctc cgk tgt tca tgt gcc ggt gga cca 150
   Asn Val Tyr Lys Pro Leu Arg Leu Arg Cys Ser Val Ala Gly Gly Pro
       25      30      35      40

- 1 -
<table>
<thead>
<tr>
<th>Codon</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Glu</td>
</tr>
<tr>
<td>GTC</td>
<td>Gly</td>
</tr>
<tr>
<td>GGA</td>
<td>Gly</td>
</tr>
<tr>
<td>TTC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>AAT</td>
<td>Thr</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>AGT</td>
<td>Ser</td>
</tr>
<tr>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>ACT</td>
<td>Ala</td>
</tr>
<tr>
<td>AGT</td>
<td>Ser</td>
</tr>
<tr>
<td>GTC</td>
<td>Val</td>
</tr>
<tr>
<td>TTA</td>
<td>Ala</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td>ACT</td>
<td>Ala</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAT</td>
<td>Asp</td>
</tr>
<tr>
<td>GTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAA</td>
<td>Lys</td>
</tr>
<tr>
<td>GCG</td>
<td>Gly</td>
</tr>
<tr>
<td>CAA</td>
<td>Ala</td>
</tr>
<tr>
<td>CGG</td>
<td>Arg</td>
</tr>
<tr>
<td>TGG</td>
<td>Leu</td>
</tr>
<tr>
<td>CAC</td>
<td>Arg</td>
</tr>
<tr>
<td>GAG</td>
<td>Asp</td>
</tr>
<tr>
<td>GCC</td>
<td>Gly</td>
</tr>
<tr>
<td>TCC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCC</td>
<td>Gly</td>
</tr>
<tr>
<td>TCA</td>
<td>Ala</td>
</tr>
<tr>
<td>ACG</td>
<td>Thr</td>
</tr>
<tr>
<td>TTC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>AAT</td>
<td>Thr</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>AGT</td>
<td>Ser</td>
</tr>
<tr>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td>ACT</td>
<td>Ala</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>TTA</td>
<td>Ala</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td>ACT</td>
<td>Ala</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAT</td>
<td>Asp</td>
</tr>
<tr>
<td>GTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAA</td>
<td>Lys</td>
</tr>
<tr>
<td>GCG</td>
<td>Gly</td>
</tr>
<tr>
<td>CAA</td>
<td>Ala</td>
</tr>
<tr>
<td>CGG</td>
<td>Arg</td>
</tr>
<tr>
<td>TGG</td>
<td>Leu</td>
</tr>
<tr>
<td>CAC</td>
<td>Arg</td>
</tr>
<tr>
<td>GAG</td>
<td>Asp</td>
</tr>
<tr>
<td>GCC</td>
<td>Gly</td>
</tr>
<tr>
<td>TCC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCC</td>
<td>Gly</td>
</tr>
<tr>
<td>TCA</td>
<td>Ala</td>
</tr>
<tr>
<td>ACG</td>
<td>Thr</td>
</tr>
<tr>
<td>TTC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>AAT</td>
<td>Thr</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>AGT</td>
<td>Ser</td>
</tr>
<tr>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td>ACT</td>
<td>Ala</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>TTA</td>
<td>Ala</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td>ACT</td>
<td>Ala</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAT</td>
<td>Asp</td>
</tr>
<tr>
<td>GTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAA</td>
<td>Lys</td>
</tr>
<tr>
<td>GCG</td>
<td>Gly</td>
</tr>
<tr>
<td>CAA</td>
<td>Ala</td>
</tr>
<tr>
<td>CGG</td>
<td>Arg</td>
</tr>
<tr>
<td>TGG</td>
<td>Leu</td>
</tr>
<tr>
<td>CAC</td>
<td>Arg</td>
</tr>
<tr>
<td>GAG</td>
<td>Asp</td>
</tr>
<tr>
<td>GCC</td>
<td>Gly</td>
</tr>
<tr>
<td>TCC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCC</td>
<td>Gly</td>
</tr>
<tr>
<td>TCA</td>
<td>Ala</td>
</tr>
<tr>
<td>ACG</td>
<td>Thr</td>
</tr>
<tr>
<td>TTC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>AAT</td>
<td>Thr</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>AGT</td>
<td>Ser</td>
</tr>
<tr>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td>ACT</td>
<td>Ala</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>TTA</td>
<td>Ala</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td>ACT</td>
<td>Ala</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAT</td>
<td>Asp</td>
</tr>
<tr>
<td>GTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAA</td>
<td>Lys</td>
</tr>
<tr>
<td>GCG</td>
<td>Gly</td>
</tr>
<tr>
<td>CAA</td>
<td>Ala</td>
</tr>
<tr>
<td>CGG</td>
<td>Arg</td>
</tr>
<tr>
<td>TGG</td>
<td>Leu</td>
</tr>
<tr>
<td>CAC</td>
<td>Arg</td>
</tr>
<tr>
<td>GAG</td>
<td>Asp</td>
</tr>
<tr>
<td>GCC</td>
<td>Gly</td>
</tr>
<tr>
<td>TCC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCC</td>
<td>Gly</td>
</tr>
<tr>
<td>TCA</td>
<td>Ala</td>
</tr>
<tr>
<td>ACG</td>
<td>Thr</td>
</tr>
<tr>
<td>TTC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>AAT</td>
<td>Thr</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>AGT</td>
<td>Ser</td>
</tr>
<tr>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td>ACT</td>
<td>Ala</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>TTA</td>
<td>Ala</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>GGT</td>
<td>Gly</td>
</tr>
<tr>
<td>ACT</td>
<td>Ala</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAT</td>
<td>Asp</td>
</tr>
<tr>
<td>GTA</td>
<td>Val</td>
</tr>
<tr>
<td>GAA</td>
<td>Lys</td>
</tr>
<tr>
<td>GCG</td>
<td>Gly</td>
</tr>
<tr>
<td>CAA</td>
<td>Ala</td>
</tr>
<tr>
<td>CGG</td>
<td>Arg</td>
</tr>
<tr>
<td>TGG</td>
<td>Leu</td>
</tr>
<tr>
<td>CAC</td>
<td>Arg</td>
</tr>
<tr>
<td>GAG</td>
<td>Asp</td>
</tr>
<tr>
<td>GCC</td>
<td>Gly</td>
</tr>
<tr>
<td>TCC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCC</td>
<td>Gly</td>
</tr>
<tr>
<td>TCA</td>
<td>Ala</td>
</tr>
<tr>
<td>ACG</td>
<td>Thr</td>
</tr>
<tr>
<td>TTC</td>
<td>Thr</td>
</tr>
<tr>
<td>GCT</td>
<td>Val</td>
</tr>
<tr>
<td>CTA</td>
<td>Val</td>
</tr>
<tr>
<td>AAT</td>
<td>Thr</td>
</tr>
<tr>
<td>ATT</td>
<td>Thr</td>
</tr>
<tr>
<td>AGT</td>
<td>Ser</td>
</tr>
</tbody>
</table>
ccg tca cct cca ggt cgt gaa gaa tct gtg gag gag ttt gta cgg cgt
Pro Ser Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg
185 190 195 200

aac ctc ggt gat gag gtt ttt gag cgc ctg att gaa ccg ttt tgt tca
Asn Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser
205 210 215

ggt gtt tat gct ggt gat cct tca aaa ctg agc atg aaa gca gcg ttt
Gly Val Tyr Ala Gly Asp Pro Ser Leu Ser Met Lys Ala Ala Phe
220 225 230

ggg aag gtt tgg aaa cta gag cca aat ggt gga agc ata ata ggt ggt
Gly Lys Val Trp Iys Leu Glu Glu Asn Gly Gly Ser Ile Ile Gly Gly
235 240 245

act ttt aag gca att cag gag agg aaa aac gct ccc aag gca gaa cga
Thr Phe Lys Ala Ile Gln Arg Lys Asn Ala Pro Lys Ala Glu Arg
250 255 260

gac cgc ctc cca aaa cca cag ggc cca aca gtt ggt tct ttc agg
Asp Pro Arg Leu Pro Iys Pro Gln Gly Gln Thr Val Gly Ser Phe Arg
265 270 275 280

aag gga ctt cga atg ttt cca gaa gca ata tct gca aga tta gtt agc
Lys Gly Leu Arg Met Leu Pro Glu Ala Ile Ser Ala Arg Leu Gly Ser
285 290 295

aaa gtt aag ttg tct tgg aag ctc tca ggt atc act aag ctc cag gag agg
Lys Val Lys Leu Ser Trp Lys Leu Ser Gly Ile Thr Lys Leu Glu Ser
300 305 310

gga gga tac aac tta aca tat gag act cca gat ggt tta gtt tcc gtt
Gly Gly Tyr Asn Leu Thr Tyr Glu Thr Pro Asp Gly Leu Val Ser Val
315 320 325

cag agc aaa aat gtt gta atg acc atg cca tct cat gtt gca aat ggt
1062
Gln Ser Lys Ser Val Val Met Thr Val Pro Ser His Val Ala Ser Gly
330 335 340

cct ttg cgc ccc ctu tct gaa tct gct gca aat gca ctc tca aac cta
Leu Leu Arg Pro Leu Ser Glu Ser Ala Ala Asn Ala Leu Ser Lys Leu
345 350 355 360

tat tac cca cca gtt gca gca gta tct atc tct tac ccg aag gaa gca
Tyr Tyr Pro Val Ala Ala Val Ser Tyr Pro Lys Glu Ala
365 370 375

atc cga aca gaa tgt ttg ata gat ggt gaa cta aag gtt ccc ccc caa
Ile Arg Thr Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln
380 385 390

TTG CAT CCA CGC AGG CAA GGA GTT GAA ACA TTG ACT ATC TAC AGG
Leu His Pro Arg Thr Glu Gly Val Val Leu Gly Thr Ile Tyr Ser
395 400 405

tcc tca ctc ttt cca aat cgc gca ccg ccc gga aga att tgg ctg ttc
Ser Ser Leu Phe Pro Asn Arg Ala Pro Pro Gly Arg Ile Leu Leu Leu
410 415 420

AAC TAC ATT GGC GGG TCT ACA AAC ACC GGA ATT CTG TCC AAG TCT GAA
Asn Tyr Ile Gly Gly Ser Thr Asn Thr Gly Ile Leu Ser Lys Ser Glu
425 430 435 440

ggt gag tta gtt gaa gca gtt gac aga gat ttc agg aaa atg cta att
Gly Glu Leu Val Glu Ala Val Asp Arg Leu Arg Lys Met Leu Ile
445 450 455

AAG CCT AAT TCG ACC GAT CCA CTT AAA TTA GGA GTT AGG GTA TGG CTT
Lys Pro Asn Ser Thr Asp Pro Leu Lys Leu Gly Val Arg Val Trp Pro
460 465 470

caa gcc att cct cag ttt cta gtt ggt cac ttt gat atc ctt gac acg
Gln Ala Ile Pro Gln Phe Leu Val Gly His Phe Asp Ile Leu Asp Thr
494

- 4 -
gct aaa tca tct cta acg tct tcg ggc tac gaa ggg cta ttt ttg ggt 1542
Ala Lys Ser Ser Leu Thr Ser Ser Gly Tyr Glu Gly Leu Phe Leu Gly
490 495 500

ggc aat tac gtc gct ggt gta gcc tta ggc cgg tgt gta gaa ggc gca 1590
Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala
505 510 515 520

tat gaa acc ggc att gag gtc aac aac ttc atg tca cgg tac gct tac 1638
Tyr Glu Thr Ala Ile Glu Val Asn Asn Phe Met Ser Arg Tyr Ala Tyr
525 530 535

aag taa atgtaaaaa ttaatatcctc cagcttgccgt gagttttatt aaatatatttg 1694
Lys

agatatccaa aaaaaa aaaa aaaaa aaaaa
1719

<210> 2
<211> 537
<212> PRT
<213> Arabidopsis thaliana

<400> 2
Met Glu Leu Ser Leu Leu Arg Pro Thr Thr Gln Ser Leu Leu Pro Ser
1 5 10 15

Phe Ser Lys Pro Asn Leu Arg Leu Asn Val Tyr Lys Pro Leu Arg Leu
20 25 30

Arg Cys Ser Val Ala Gly Gly Pro Thr Val Gly Ser Ser Lys Ile Glu
35 40 45

Gly Gly Gly Thr Thr Ile Thr Thr Asp Cys Val Ile Val Gly Gly
50 55 60
Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala Thr Lys His Pro
65 70 75 80

Asp Ala Ala Pro Asn Leu Ile Val Thr Glu Ala Lys Asp Arg Val Gly
85 90 95

Gly Asn Ile Ile Thr Arg Glu Glu Asn Gly Phe Leu Trp Glu Glu Gly
100 105 110

Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr Met Val Val Asp
115 120 125

Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Thr Ala Pro Arg
130 135 140

Phe Val Leu Trp Asn Gly Lys Leu Arg Pro Val Pro Ser Lys Leu Thr
145 150 155 160

Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly Gly Lys Ile Arg Ala
165 170 175

Gly Phe Gly Ala Leu Gly Ile Arg Pro Ser Pro Pro Gly Arg Glu Glu
180 185 190

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

Arg Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser
210 215 220

Lys Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Lys Leu Glu Gln
225 230 235 240

Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys Ala Ile Gln Glu Arg
245 250 255
Lys Asn Ala Pro Lys Ala Glu Arg Asp Pro Arg Leu Pro Lys Pro Gln
260 265 270

Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Arg Met Leu Pro Glu
275 280 285

Ala Ile Ser Ala Arg Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu
290 295 300

Ser Gly Ile Thr Lys Leu Glu Ser Gly Gly Tyr Asn Leu Thr Tyr Glu
305 310 315 320

Thr Pro Asp Gly Leu Val Ser Val Gln Ser Lys Ser Val Val Met Thr
325 330 335

Val Pro Ser His Val Ala Ser Gly Leu Leu Arg Pro Leu Ser Glu Ser
340 345 350

Ala Ala Asn Ala Leu Ser Lys Leu Tyr Tyr Pro Pro Val Ala Ala Val
355 360 365

Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg Thr Glu Cys Leu Ile Asp
370 375 380

Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Thr Gln Gly Val
385 390 395 400

Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala
405 410 415

Pro Pro Gly Arg Ile Leu Leu Leu Asn Tyr Ile Gly Gly Ser Thr Asn
420 425 430

Thr Gly Ile Leu Ser Lys Ser Glu Gly Glu Leu Val Glu Ala Val Asp
435 440 445

Arg Asp Leu Arg Lys Met Leu Ile Lys Pro Asn Ser Thr Asp Pro Leu

- 7 -
Lys Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val
450 455 460
Gly His Phe Asp Ile Leu Asp Thr Ala Lys Ser Ser Leu Thr Ser Ser
465 470 475 480 485 490 495
Gly Tyr Glu Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala
500 505 510
Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala Ile Glu Val Asn
515 520 525
Asn Phe Met Ser Arg Tyr Ala Tyr Lys
530 535

<210> 3
<211> 1738
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (70) .. (1596)
<223> Arabidopsis protox-2

<400> 3
tttttttactt attttcgctca ctgcttttoga ctggtcagag attttgactc tgaattgttg 60
cagataagca atg ggc ttc gga gca gta gca gat cat caa att gaa ggc gtt 111
Met Ala Ser Gly Ala Val Val Asp His Gln Ile Glu Ala Val
1 5 10
tca gga aag aag gtc gca gtc gta ggt gca ggt gta agt gga ctt ggc 159
Ser Gly Lys Arg Val Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala
gct  tcc  tac  aag  ttg  aaa  tca  agg  ggt  ttg  aat  gtg  act  gtc  ttt  gaa
Ala  Ala  Tyr  Lys  Leu  Lys  Ser  Arg  Gly  Leu  Asn  Val  Thr  Val  Phe  Glu
35 40 45

gct  gat  gga  aga  gta  ggt  ggg  aag  ttg  aga  agt  gtt  aga  gaa  aat  ggt
Ala  Asp  Gly  Arg  Val  Gly  Gly  Lys  Leu  Arg  Ser  Val  Met  Gln  Asn  Gly
50 55 60

att  tgg  gat  gaa  gga  gca  aac  acc  atg  act  gag  gct  gag  cca  gaa
Leu  Ile  Trp  Asp  Glu  Gly  Ala  Asn  Thr  Met  Thr  Glu  Ala  Glu  Pro  Glu
65 70 75

gtt  ggg  agt  tta  ctt  gat  gat  ctt  ggg  ctt  cgt  gag  aaa  caa  caa  ttt
Val  Gly  Ser  Leu  Leu  Asp  Leu  Gly  Leu  Arg  Glu  Lys  Gln  Gln  Phe
80 85 90

cca  att  tca  cag  aaa  aag  cgg  tat  att  gtt  cgg  aat  ggt  gta  cct  gtt
Pro  Ile  Ser  Gln  Lys  Iys  Arg  Tyr  Ile  Val  Arg  Asn  Gly  Val  Pro  Val
95 100 105 110

atg  cta  cct  acc  aat  ccc  ata  gag  ctc  gtc  aca  agt  agt  gtt  ctc  tct
Met  Leu  Pro  Thr  Asn  Pro  Ile  Glu  Leu  Val  Thr  Ser  Ser  Val  Leu  Ser
115 120 125

acc  caa  tct  aag  ttt  caa  atc  ttg  ttg  gaa  cca  ttt  tta  tgg  aag  aaa
Thr  Gln  Ser  Lys  Phe  Gln  Ile  Leu  Leu  Glu  Pro  Phe  Leu  Trp  Lys  Lys
130 135 140

aag  tcc  tca  aaa  gtc  tca  gat  gca  tct  gct  gaa  gaa  agt  gta  agc  gag
Lys  Ser  Ser  Lys  Val  Ser  Asp  Ala  Ser  Ala  Glu  Glu  Ser  Val  Ser  Glu
145 150 155

ttc  ttt  caa  cgc  cat  ttt  gga  caa  gag  gtt  gtt  gac  tat  ctc  atc  gac
Phe  Phe  Gln  Arg  His  Phe  Gly  Gln  Glu  Val  Val  Asp  Tyr  Leu  Ile  Asp
160 165 170
cct ttt gtt ggt gga aca agt gct gcg gac cct gat tcc ctt tca atg
Pro Phe Val Gly Gly Thr Ser Ala Ala Asp Pro Asp Ser Leu Ser Met
175 180 185 190

aag cat tct ttc cca gat ctc tgg aat gta gag aaa agt ttt gcc tct
Lys His Ser Phe Pro Asp Leu Trp Asn Val Glu Lys Ser Phe Gly Ser
195 200 205

att ata gtc ggt gca atc aga aca aag ttt gct gct aaa gtt ggt aaa
Ile Ile Val Gly Ala Ile Arg Thr Lys Phe Ala Ala Lys Gly Gly Lys
210 215 220

agt aga gac aca aag agt tct cct gcg aca aaa aag gtt tcg cgt ggg
Ser Arg Asp Thr Lys Ser Ser Pro Gly Thr Lys Lys Gly Ser Arg Gly
225 230 235

tca ttc tct ttt aag ggg gga atg cag att ctt cct gat aoc ttg tgc
Ser Phe Ser Phe Lys Gly Gly Met Gln Ile Leu Pro Asp Thr Leu Cys
240 245 250

aaa agt ctc tca cat gat gag atc aat tta gac tcc aag gta ctc tct
Lys Ser Leu Ser His Asp Glu Ile Asn Leu Asp Ser Lys Val Leu Ser
255 260 265 270

ttg tct tac aat tct gga tca aga cag gag aac tgg tca tta tct tgt
Leu Ser Tyr Asn Ser Gly Ser Arg Gln Glu Asn Trp Ser Leu Ser Cys
275 280 285

gtt tcg cat aat gaa acg cag aga caa aac ccc cat tat gat gct gta
Val Ser His Glu Thr Gln Arg Gly Glu Asn Pro His Tyr Asp Ala Val
290 295 300

att atg acg gct cct ctg tgc aat gtg aag gag atg aag gtt atg aaa
Ile Met Thr Ala Pro Leu Cys Asn Val Lys Glu Met Lys Val Met Lys
305 310 315
Gly Gly Gln Pro Phe Gln Leu Asn Phe Leu Pro Glu Ile Asn Tyr Met

320 325 330

Pro Leu Ser Val Leu Ile Thr Thr Phe Thr Lys Glu Lys Val Lys Arg
335 340 345 350

Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Ser Lys Glu Gln Lys His
355 360 365

Gly Phe Lys Thr Leu Gly Thr Leu Phe Ser Ser Met Met Phe Pro Asp
370 375 380

Arg Ser Pro Ser Asp Val His Leu Tyr Thr Thr Phe Ile Gly Gly Ser
385 390 395

Arg Asn Gln Glu Leu Ala Iys Ala Ser Thr Asp Glu Leu Lys Gln Val
400 405 410

Val Thr Ser Asp Leu Gln Arg Leu Leu Gly Val Glu Gly Glu Pro Val
415 420 425 430

Ser Val Asn His Tyr Tyr Trp Arg Lys Ala Phe Pro Leu Tyr Asp Ser
435 440 445

Ser Tyr Asp Ser Val Met Glu Ala Ile Asp Lys Met Glu Asn Asp Leu
450 455 460

Cct ggg ttc ttc tat gca ggt aat cat cga ggg ggg ctc tct gtt ggg 1503
Pro Gly Phe Phe Tyr Ala Gly Asn His Arg Gly Gly Leu Ser Val Gly
465 470 475

aaa tca ata gca tca ggt tgc aaa gca gct gac ctt gtg atc tca tac 1551
Lys Ser Ile Ala Ser Gly Cys Lys Ala Ala Asp Leu Val Ile Ser Tyr
480 485 490
ctg gag tct tgc tca aat gac aag aaa cca aat gac agc tta taa 1596
Leu Glu Ser Cys Ser Asn Asp Lys Lys Pro Asn Asp Ser Leu
495 500 505
cattgtcaag gtgcgtctct ttttatcact tacctttaaa actttgtaaa tgcaacaacg 1656
cgcccgtgca ga tagc caca ac gca gga gcc cactaaggtct cactaattcc 1716
agaataaact atttatgtaa aa 1738

<210> 4
<211> 508
<212> PRT
<213> Arabidopsis thaliana

<400> 4
Met Ala Ser Gly Ala Val Ala Asp His Gln Ile Glu Ala Val Ser Gly
1 5 10 15

Lys Arg Val Ala Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala
20 25 30

Tyr Lys Leu Lys Ser Arg Gly Leu Asn Val Thr Val Phe Glu Ala Asp
35 40 45

Gly Arg Val Gly Gly Lys Leu Arg Ser Val Met Gln Asn Gly Leu Ile
50 55 60

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

Ser Gln Lys Lys Arg Tyr Ile Val Arg Asn Gly Val Pro Val Met Leu

Pro Thr Asn Pro Ile Glu Leu Val Thr Ser Ser Val Leu Ser Thr Gln

Ser Lys Phe Gln Ile Leu Leu Glu Pro Phe Leu Trp Lys Lys Lys Ser

Ser Lys Val Ser Asp Ala Ser Ala Glu Glu Ser Val Ser Glu Phe Phe

Gln Arg His Phe Gly Gln Glu Val Val Asp Tyr Leu Ile Asp Pro Phe

Val Gly Gly Thr Ser Ala Ala Asp Pro Asp Ser Leu Ser Met Lys His

Ser Phe Pro Asp Leu Trp Asn Val Glu Lys Ser Phe Gly Ser Ile Ile

Val Gly Ala Ile Arg Thr Lys Phe Ala Ala Lys Gly Gly Lys Ser Arg

Asp Thr Lys Ser Ser Pro Gly Thr Lys Lys Gly Ser Arg Gly Ser Phe

Ser Phe Lys Gly Gly Met Gln Ile Leu Pro Asp Thr Leu Cys Lys Ser

Leu Ser His Asp Glu Ile Asn Leu Asp Ser Lys Val Leu Ser Leu Ser
Tyr Asn Ser Gly Ser Arg Gln Glu Asn Trp Ser Leu Ser Cys Val Ser
275 280 285
His Asn Glu Thr Gln Arg Gln Asn Pro His Tyr Asp Ala Val Ile Met
290 295 300
Thr Ala Pro Leu Cys Asn Val Lys Glu Met Lys Val Met Lys Gly Gly
305 310 315 320
Gln Pro Phe Gln Leu Asn Phe Leu Pro Glu Ile Asn Tyr Met Pro Leu
325 330 335
Ser Val Leu Ile Thr Thr Phe Thr Lys Glu Lys Val Lys Arg Pro Leu
340 345 350
Glu Gly Phe Gly Val Leu Ile Pro Ser Lys Glu Gln Lys His Gly Phe
355 360 365
Lys Thr Leu Gly Thr Leu Phe Ser Ser Met Met Phe Pro Asp Arg Ser
370 375 380
Pro Ser Asp Val His Leu Tyr Thr Thr Phe Ile Gly Gly Ser Arg Asn
385 390 395 400
Gln Glu Leu Ala Lys Ala Ser Thr Asp Glu Leu Lys Gln Val Val Thr
405 410 415
Ser Asp Leu Gln Arg Leu Leu Gly Val Glu Gly Glu Pro Val Ser Val
420 425 430
Asn His Tyr Tyr Trp Arg Lys Ala Phe Pro Leu Tyr Asp Ser Ser Tyr
435 440 445
Asp Ser Val Met Glu Ala Ile Asp Lys Met Glu Asn Asp Leu Pro Gly
450 455 460
Phe Hty Ala Gly Asn His Arg Gly Gly Leu Ser Val Gly Lys Ser
465          470          475          480
Ile Ala Ser Gly Cys Lys Ala Ala Asp Leu Val Ile Ser Tyr Leu Glu
485          490          495
Ser Cys Ser Asn Asp Lys Lys Pro Asn Asp Ser Leu
500          505

<210> 5
<211> 1691
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (1)...(1443)
<223> Maize protox-1 c-DNA (not full-length)

<400> 5
gga gac tgc gtc gtg gtc ggg gga ggtat ggc ctc tgt gc g
Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu Cys Thr Ala
1      5     10     15

cag ggc ctg gcc acg ccg cac ggc gtc ggg gac ggt ctt gtc acg gag
Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu Val Thr Glu
20      25     30

gcc ggc gcc ggc ccc gcc ggc aac att acc acc gtc gag cgc ccc gag
Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Glu
35      40     45

gaa ggg tac ctc tgg gag gag ggt ccc aac agc ttc cag ccc tcc gag
Glu Gly Tyr Leu Trp Glu Gly Gy Asn Ser Phe Gln Pro Ser Asp
50      55     60
ccc gtt ctc acc atg gcc gtg gac agc gga ctg aag gat gac ttg gtt 240
Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val
65  70  75  80

ttt ggg gac cca aac gcg cgg ctg ttc gtg tgg gag ggg aag ctg 288
Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu
85  90  95

agg ccc gtg cca tcc aag ccc gcc gac ctc ccc gtt ttg gat ccc atg 336
Arg Pro Val Pro Ser Iys Pro Ala Asp Leu Pro Phe Asp Leu Met
100 105 110

agc atc cca ggg aag ctc agy gcc ggt cta ggc ggc ggg ctt ggc atc cgc 384
Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg
115 120 125

cog cct cct cca ggc cgc gaa gag tca gtt ggg gag tgg gtt ggc ggc 432
Pro Pro Pro Gly Arg Glu Glu Val Glu Glu Phe Val Arg Arg
130 135 140

aac ctc ggt gct gag gtc ttt gag cgc ctc att gag cct ttc tgc tca 480
Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser
145 150 155 160

ggt gtc tat gct ggt gat cct tct aag ctc agc atg aag gct gca ttt 528
Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe
165 170 175

ggg aag gtt tgg cgg ttg gaa gaa act gga ggt agt att att ggt gga 576
Gly Lys Val Trp Arg Leu Glu Glu Thr Gly Gly Ser Ile Ile Gly Gly
180 185 190

acc atc aac aca att cag gag agg agc aag aat cca aaa cca cca cog agg 624
Thr Ile Lys Thr Ile Lsn Glu Arg Ser Lys Asn Pro Lys Pro Pro Arg
195 200 205

gat gcc cgc ctt cgg aag cca aaa ggg cag aca gtt gca tct ttc cgg 672
Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Ala Ser Phe Arg  
210 215 220

aag ggt ctt gcc atg ctt cca aat gcc att aca tcc agc ttg ggt agt 720
Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Thr Ser Ser Leu Gly Ser  
225 230 235 240

aaa gtc aaa cta tca tgg aaa ctc agc agc att aca aaa tca gat gac 768
Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ser Asp Asp  
245 250 255

aag gga tat gtt ttg gag tat gaa acg cca gaa ggg gtt gtt tcg gtt 816
Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro Glu Gly Val Val Ser Val  
260 265 270

cag gct aaa agt gtt atc atg act att cca tca tat gtt gct agc aac 864
Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asn  
275 280 285

att ttg cgt cca ctt cca agc gat gct gct gca gat gct cta tca aga ttc 912
Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala Asp Ala Leu Ser Arg Phe  
290 295 300

tat tat cca ccc gtt gct gct gta act gtt tcg tat cca aag gaa gca 960
Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala  
305 310 315 320

att aga aaa gaa tgc tta att gat ggg gaa ctc cag ggc ttt ggc cag 1008
Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln  
325 330 335

ttg cat cca cgt agt cca gga ttg gag aca tta gga aca ata tac agt 1056
Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser  
340 345 350

tcc tca ctc ttt cca aat cgt gct cct gac ggt ggg gtt gta cct tta 1104
Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val Leu Leu Leu
aac tac ata gga ggt gct aca aac aca gga att gtt tcc aag act gaa
Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser Lys Thr Glu
370 375 380

agt gag ctc gaa gca gtt gac gct gac gtt gaa atg ctt ata
Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile
385 390 395 400

aat tct aca gca gtg gac cct tta gtc ctt ggt gtt cga gtt tgg cca
Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg Val Trp Pro
405 410 415

ca a gcc ata cct cag ttc ctg gta gga cat ctt gat ctt ctg gaa gcc
Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Glu Ala
420 425 430

gca aac gct gcc ctg gac cpa ggt ggc tac gat ggg ctg ttc cta gga
Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu Phe Leu Gly
435 440 445

ggg aac tat gtt gca gga gtt gcc ctg ggc aga tgc gtt gag ggc gcy
Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala
450 455 460

tat gaa agt gcc tcg caa ata tct gac ttc ttg acc aag tat gcc tac
Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys Tyr Ala Tyr
465 470 475 480

aag tga tgaagaga agttgagcgc taactgtttaa togtttatatgt tgcaatatag
Lys

aggtgccctcc gggaaaaaa aagcttgaat agtttttttt attccttattt tggtaattgc
1553

atctttctc tttttttctat cagtaattag ttatatattta gttctgtagg agattgttct
1613
gttcactgcc cttcaaaaga aattttat ttcatttttt tatgagagct gtgctactta 1673

aaaaaa aaaaaa 1691

<210> 6
<211> 481
<212> PRT
<213> Zea mays

<400> 6
Ala Asp Cys Val Val Val Gly Gly Gly Ile Ser Gly Leu Cys Thr Ala

1  5  10  15

Gln Ala Leu Ala Thr Arg His Gly Val Gly Asp Val Leu Val Thr Glu

20  25  30

Ala Arg Ala Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Glu

35  40  45

Glu Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp

50  55  60

Pro Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val

65  70  75  80

Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu

85  90  95

Arg Pro Val Pro Ser Lys Pro Ala Asp Leu Pro Phe Phe Asp Leu Met

100  105  110

Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg

115  120  125

Pro Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg

130  135  140
Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser
145 150 155 160

Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe
165 170 175

Gly Lys Val Trp Arg Leu Glu Glu Thr Gly Gly Ser Ile Ile Gly Gly
180 185 190

Thr Ile Lys Thr Ile Glu Gly Arg Ser Lys Asn Pro Lys Pro Pro Arg
195 200 205

Asp Ala Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Ala Ser Phe Arg
210 215 220

Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Thr Ser Ser Leu Gly Ser
225 230 235 240

Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ser Asp Asp
245 250 255

Lys Gly Tyr Val Leu Glu Tyr Glu Thr Pro Glu Gly Val Val Ser Val
260 265 270

Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asn
275 280 285

Ile Leu Arg Pro Leu Ser Ser Asp Ala Ala Asp Ala Leu Ser Arg Phe
290 295 300

Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala
305 310 315 320

Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln
325 330 335
Leu His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser
340     345     350

Ser Ser Leu Phe Pro Asn Arg Ala Pro Asp Gly Arg Val Leu Leu Leu
355     360     365

Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Val Ser Lys Thr Glu
370     375     380

Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile
385     390     395     400

Asn Ser Thr Ala Val Asp Pro Leu Val Leu Gly Val Arg Val Trp Pro
405     410     415

Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Glu Ala
420     425     430

Ala Lys Ala Ala Leu Asp Arg Gly Gly Tyr Asp Gly Leu Phe Leu Gly
435     440     445

Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Gly Ala
450     455     460

Tyr Glu Ser Ala Ser Gln Ile Ser Asp Phe Leu Thr Lys Tyr Ala Tyr
465     470     475     480

Lys

<210> 7
<211> 2061
<212> DNA
<213> Zea mays

<220>
<221> CDS
<222> (64)・・・(1698)
<223> Maize protox-2

<400> 7
ctctctctcc ctccactctcc aagcaaaaga caaatccccca tccagttcga aacccctaact 60

caa atg ctc gct ttg act gcc tca gcc tca tcc gct tcc cat cct 108

Met Leu Ala Leu Thr Ala Ser Ala Ser Ala Ser Ser His Pro
1 10 15

tat cgc cac gcc tcc ggc ggc cac act cgt cgc ccc cgc cta cgt gog gtc 156

Tyr Arg His Ala Ser Ala His Thr Arg Arg Pro Arg Leu Arg Ala Val
20 25 30

cgc atg gog gcc tcc gac gac ccc cgt gca gog ccc gcc aga tgc 204

Leu Ala Met Ala Gly Ser Asp Asp Pro Arg Ala Ala Pro Ala Arg Ser
35 40 45

gtc ggc gtc gtc ggc ggc ggg gtc agc ggg gtc gog gog gog tac agg 252

Val Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala Tyr Arg
50 55 60

cgc aga cag agc ggc gtg aac gta acg gtg ttc gaa gog gcc gac agg 300

Leu Arg Gln Ser Gly Val Asn Val Thr Val Phe Glu Ala Ala Asp Arg
65 70 75

ggc gga gga aag ata cgg acc aat tcc gag gcc ggg ttt gtc tgt gat 348

Ala Gly Gly Lys Ile Arg Thr Asn Ser Glu Gly Gly Phe Val Trp Asp
80 85 90 95

gaa gga gct aac acc atg aca gaa ggt gaa tgg gag gcc agt aga cttg 396

Glu Gly Ala Asn Thr Met Thr Glu Gly Glu Trp Glu Ala Ser Arg Leu
100 105 110

att gat gat ctt ggt cta caa gac aaa cag cag tat cct aac tcc caa 444

Ile Asp Asp Leu Gly Leu Gln Asp Lys Gln Gln Tyr Pro Asn Ser Gln
115 120 125

cac aag cgt tac att gtc aaa gat gga gca cca gca ctg att cct tcg 492
His Lys Arg Tyr Ile Val Lys Asp Gly Ala Pro Ala Leu Ile Pro Ser
130 135 140

gat ccc att tcg cta atg aaa agc agt gtt ctt tcg aca aaa tca aag 540
Asp Pro Ile Ser Leu Met Lys Ser Ser Val Leu Ser Thr Lys Ser Lys
145 150 155

att gcc tta ttt ttt gaa cca ttt ctc tac aag aaa gct aac aca aga 588
Ile Ala Leu Phe Phe Glu Pro Phe Leu Tyr Lys Ala Asn Thr Arg
160 165 170 175

aac tct gga aaa gtt tct gag gag cac ttg agt gag agt gtt ggg agc 636
Asn Ser Gly Lys Val Ser Glu Glu His Leu Ser Glu Ser Val Gly Ser
180 185 190

ttc tgt gaa cgc cac ttt gga aga gaa gtt gtt gac tat ttt gtt gat 684
Phe Cys Glu Arg His Phe Gly Arg Glu Val Val Asp Tyr Phe Val Asp
195 200 205

cca ttt gta gct gga aca agt gca gga gat cca gag tca cta tct att 732
Pro Phe Val Ala Gly Thr Ser Ala Gly Asp Pro Glu Ser Leu Ser Ile
210 215 220

cgt cat gca ttc cca gca ttg tgg aat ttg gaa aga aag tat gtt tca 780
Arg His Ala Phe Pro Ala Leu Trp Asn Leu Glu Arg Lys Tyr Gly Ser
225 230 235

gtt att gtt tgt gcc atc ttg tgt aat tta gaa gaa aag ttc gga cca 828
Val Ile Val Gly Ala Ile Leu Ser Lys Leu Ala Ala Lys Gly Asp Pro
240 245 250 255

gta aag aca aga cat gat tca tca ggg aaa aag agg aat aca cga gtt 876
Val Lys Thr Arg His Asp Ser Ser Gly Lys Arg Arg Asn Arg Arg Val
260 265 270

- 23 -
tcg ttt tca ttt cat ggt gga atg cag tca cta ata aat gca ctt cac
Ser Phe Ser Phe His Gly Gly Met Gln Ser Leu Ile Asn Ala Leu His
275 280 285

aat gaa gtt gga gat gat aat gtg aag ctt ggt aca gaa gtt ttg tca
Asn Glu Val Gly Asp Asp Asn Val Lys Leu Gly Thr Glu Val Leu Ser
290 295 300

ttg gca tgt aca ttt gat gga gtt cct gca cta ggc agg tgg tca att
Leu Ala Cys Thr Phe Asp Gly Val Pro Ala Leu Gly Arg Trp Ser Ile
305 310 315

tct gtt gat tgg aag gtt gat gac aag gag gac ctt gct agt aac caa
Ser Val Asp Ser Lys Asp Ser Gly Asp Lys Asp Leu Ala Ser Asn Gln
320 325 330 335

acc ttt gat gct gtt ata atg aca gct cca ttg tca aat gtc cgg agg
Thr Phe Asp Ala Val Ile Met Thr Ala Pro Leu Ser Asn Val Arg Arg
340 345 350

atg aag ttc acc aaa ggt gga gct ccg gtt gtt ctt gac ttt ctt cct
Met Lys Phe Thr Lys Gly Ala Pro Val Val Leu Asp Phe Leu Pro
355 360 365

aag atg gat tat cta cca cta tct ctc atg gtg act gct ttt aag aag
Lys Met Asp Tyr Leu Pro Leu Ser Leu Met Val Thr Ala Phe Lys Lys
370 375 380

gat gat gtc aag aaa cct ctg gaa gga ttt ggg gtc tta ata cct tac
Asp Asp Val Lys Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Tyr
385 390 395

aag gaa cag caa aaa cat ggt ctg aaa acc ctt ggg act ctc ttt tcc
Lys Glu Gln Gln Lys His Gly Leu Lys Thr Leu Gly Thr Leu Phe Ser
400 405 410 415
tca atg atg ttc cca gat cga gct cct gat gac caa tat tta tat aca 1356
Ser Met Met Phe Pro Asp Arg Ala Pro Asp Asp Gln Tyr Leu Tyr Thr
   420   425   430
aca ttt gtt ggg ggt agc cac aat aga gat ctt gct gga gct cca aag 1404
Thr Phe Val Gly Gly Ser His Asn Arg Asp Leu Ala Gly Ala Pro Thr
   435   440   445
tct att ctg aaa caa ctt gtt acc tct gac ctt aaa aaa ctc ttg ggc 1452
Ser Ile Leu Lys Gln Leu Val Thr Ser Asp Leu Lys Lys Leu Leu Gly
   450   455   460
gta gag ggg caa cca act ttt gtc aag cat gta tac tgg gga aat gct 1500
Val Glu Gly Glu Pro Thr Phe Val Lys His Val Tyr Trp Gly Asn Ala
   465   470   475
ttt cct ttg tat ggc cat gat tat agt tct gta ttg gaa gct ata gaa 1548
Phe Pro Leu Tyr Gly His Asp Tyr Ser Ser Val Leu Glu Ala Ile Glu
   480   485   490   495
aag atg gag aaa aac ctt cca ggg ttc ttc tac gca gga aat agc aag 1596
Lys Met Glu Lys Asp Leu Pro Gly Phe Phe Tyr Ala Gly Asn Ser Lys
   500   505   510
gat ggg ctt gct gtt gga agt gtt ata gct tca gga agc aag gct gct 1644
Asp Gly Leu Ala Val Gly Ser Val Ile Ala Ser Gly Ser Lys Ala Ala
   515   520   525
gac ctt gca atc tca tat ctt gaa tct cac acc aag cat aat gat tca 1692
Asp Leu Ala Ile Ser Tyr Leu Glu Ser His Thr Lys His Asn Asn Ser
   530   535   540
cat tga aagagctctga cctatctct agcagttgtc gacaaatctt tccagttcct 1748
His
     545

gtacagtga aaccgatgctg ttgctgctctt ccagcgattg gacccatttcc tccagttccat

-25-
```
ctctgtttgaa catccaccag aaggtagtc acatgtgtaa gtgggaaaat gaggttaaa 1868
actattatgg cggcgcgaaat gtctctttttt gtttttcctca caagtggcct acgacacttg 1928
atgttggaata tacattttaaa tttgttgaat tgggtggaga cacatgcgtg acgtgtaata 1988
tttgcctatt gtgttttttag cagtagttt ggccagattg tgcctttacgc ctttaaaaaa 2048

<210> 8
<211> 544
<212> PRT
<213> Zea mays

<400> 8
Met Leu Ala Leu Thr Ala Ser Ala Ser Ala Ser His Pro Tyr
   1   5   10   15

Arg His Ala Ser Ala His Thr Arg Arg Pro Arg Leu Arg Ala Val Leu
   20  25   30

Ala Met Ala Gly Ser Asp Asp Pro Arg Ala Ala Pro Ala Arg Ser Val
   35  40   45

Ala Val Val Gly Ala Gly Val Ser Gly Leu Ala Ala Ala Tyr Arg Leu
   50  55   60

Arg Gln Ser Gly Val Asm Val Thr Val Phe Glu Ala Ala Asp Arg Ala
   65  70   75   80

Gly Gly Lys Ile Arg Thr Asm Ser Glu Gly Phe Val Trp Asp Glu
   85  90   95

Gly Ala Asm Thr Met Thr Glu Gly Glu Trp Glu Ala Ser Arg Leu Ile
```
Asp Asp Leu Gly Leu Gln Asp Lys Gln Gln Tyr Pro Asn Ser Gln His
100 105 110
Lys Arg Tyr Ile Val Lys Asp Gly Ala Pro Ala Leu Ile Pro Ser Asp
115 120 125
Pro Ile Ser Leu Met Lys Ser Ser Val Leu Ser Thr Lys Ser Lys Ile
130 135 140
Ala Leu Phe Phe Glu Pro Phe Leu Tyr Lys Lys Ala Asn Thr Arg Asn
145 150 155 160
Ser Gly Lys Val Ser Glu Glu His Leu Ser Glu Ser Val Gly Ser Phe
165 170 175
Cys Glu Arg His Phe Gly Arg Glu Val Val Asp Tyr Phe Val Asp Pro
180 185 190
Phe Val Ala Gly Thr Ser Ala Gly Asp Pro Glu Ser Leu Ser Ile Arg
195 200 205
His Ala Phe Pro Ala Leu Trp Asn Leu Glu Arg Lys Tyr Gly Ser Val
210 215 220
Ile Val Gly Ala Ile Leu Ser Lys Leu Ala Ala Lys Gly Asp Pro Val
225 230 235 240
Lys Thr Arg His Asp Ser Ser Gly Lys Arg Arg Asn Arg Arg Val Ser
245 250 255
Phe Ser Phe His Gly Gly Met Gln Ser Leu Ile Asn Ala Leu His Asn
255 260 265 270
Glu Val Gly Asp Asp Asn Val Lys Leu Gly Thr Glu Val Leu Ser Leu
275 280 285
290 295 300
| Ala Cys Thr Phe Asp Gly Val Pro Ala Leu Gly Arg Trp Ser Ile Ser |
|---|---|---|---|---|
| 305 | 310 | 315 | 320 |

| Val Asp Ser Lys Asp Ser Gly Asp Lys Asp Leu Ala Ser Asn Gln Thr |
|---|---|---|---|---|
| 325 | 330 | 335 |

| Phe Asp Ala Val Ile Met Thr Ala Pro Leu Ser Asn Val Arg Arg Met |
|---|---|---|---|---|
| 340 | 345 | 350 |

| Lys Phe Thr Lys Gly Gly Ala Pro Val Val Leu Asp Phe Leu Pro Lys |
|---|---|---|---|---|
| 355 | 360 | 365 |

| Met Asp Tyr Leu Pro Leu Ser Leu Met Val Thr Ala Phe Lys Asp |
|---|---|---|---|---|
| 370 | 375 | 380 |

| Asp Val Lys Lys Pro Leu Glu Gly Phe Gly Val Leu Ile Pro Tyr Lys |
|---|---|---|---|---|
| 385 | 390 | 395 | 400 |

| Glu Gln Gln Lys His Gly Leu Lys Thr Leu Gly Thr Leu Phe Ser Ser |
|---|---|---|---|---|
| 405 | 410 | 415 |

| Met Met Phe Pro Asp Arg Ala Pro Asp Asp Gln Tyr Leu Tyr Thr Thr |
|---|---|---|---|---|---|
| 420 | 425 | 430 |

| Phe Val Gly Gly Ser His Asn Arg Asp Leu Ala Gly Ala Pro Thr Ser |
|---|---|---|---|---|---|
| 435 | 440 | 445 |

| Ile Leu Lys Gln Leu Val Thr Ser Asp Leu Lys Leu Leu Gly Val |
|---|---|---|---|---|
| 450 | 455 | 460 |

| Glu Gly Gln Pro Thr Phe Val Lys His Val Tyr Trp Gly Asn Ala Phe |
|---|---|---|---|---|
| 465 | 470 | 475 | 480 |

| Pro Leu Tyr Gly His Asp Tyr Ser Ser Val Leu Glu Ala Ile Glu Lys |
|---|---|---|---|
| 485 | 490 | 495 |
Met Glu Lys Asn Leu Pro Gly Phe Phe Tyr Ala Gly Asn Ser Lys Asp
500 505 510

Gly Leu Ala Val Gly Ser Val Ile Ala Ser Gly Ser Lys Ala Ala Asp
515 520 525

Leu Ala Ile Ser Tyr Leu Glu Ser His Thr Lys His Asn Asn Ser His
530 535 540

<210> 9
<211> 1811
<212> DNA
<213> Triticum aestivum

<220>
<221> CDS
<222> (3) ...(1589)
<223> wheat protox-1

<400> 9
gc gca aca atg gcc acc gcc acc gtc gcg gcc ggc gcc gcc gcc gcg tog cog ctc cgc
47
    Ala Thr Met Ala Thr Ala Thr Val Ala Ala Ala Ser Pro Leu Arg
1  5 10 15

ggc agg gtc acc ggg cgc cca cac cgc gtc gcc cgc cog cgt tgc gct acc
95
    Gly Arg Val Thr Gly Arg Pro His Arg Val Arg Pro Arg Cys Ala Thr
20 25 30

gcg agc agc gcc gag act cog ggc gcc gcc gcc gcc gcc gcc gcc gtc ggg ctc tcc
143
    Ala Ser Ser Ala Thr Glu Thr Pro Ala Ala Pro Gly Val Arg Leu Ser
35 40 45

gcg gaa tgc gtc att gtg gcc gcc gcc atc agc ggc ctc tgc acc gcg
191
    Ala Glu Cys Val Ile Val Gly Ala Gly Ile Ser Gly Leu Cys Thr Ala
50 55 60

-29-
cag ggc ctg gcc acc cga tac gcc gtc agc gac ctg ctc gtc aog gag 239
Gln Ala Leu Ala Thr Arg Tyr Gly Val Ser Asp Leu Leu Val Thr Glu
65 70 75
gcc cgc gac gcc ctc gcc ggc gcc aac ctc acc ctc gtc gag cgt ccc gag 287
Ala Arg Asp Arg Pro Gly Gly Asn Ile Thr Thr Val Glu Arg Pro Asp
80 85 90 95
gag ggg tac ctg tgg gag gag gga ccc aac agc ttc cag ccc tcc gag 335
Glu Gly Tyr Leu Trp Glu Glu Gly Pro Ser Ser Phe Gln Pro Ser Asp
100 105 110
ccg gtc ctc acc atg gcc gtc gac agc ggg ctc aag gat gac ttg gtg 383
Pro Val Leu Thr Met Ala Val Asp Gly Leu Lys Asp Asp Leu Val
115 120 125
ttc ggg gac ccc aac gcg ccc cgg ttc gtt gtc ctg tgg gag ggg aag ctt 431
Phe Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu
130 135 140
agg ccc gtt ccg ccg tgg aag cca ggc gac ctg cct ttc ttc agc ctc atg 479
Arg Pro Val Pro Ser Lys Pro Gly Asp Leu Pro Phe Phe Ser Leu Met
145 150 155
agt atc cct ggg aag ctc agg gcc ggc ctt gcc gcg ctc ggc att cgc 527
Ser Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg
160 165 170 175
cca cct ccc cca ggg cgc gag gag tgg gtg gag gag ttt gtg cgc cgc 575
Pro Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg
180 185 190
aac ctc ggt gcc gag gtc ttt gag cgc ctc atc gag cct ttc tgc tca 623
Asn Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser
195 200 205
ggt gta tat gct ggt gat cct tog aag ctt agt atg aag gct gca ttt 671
Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe
210 215 220

Gly Lys Val Trp Arg Leu Glu Glu Ile Gly Gly Ser Ile Ile Gly Gly
225 230 235

Thr Ile Lys Ala Ile Gln Asp Gly Lys Asn Pro Lys Pro Pro Arg
240 245 250 255

Asp Pro Arg Leu Pro Ala Pro Lys Gly Gln Thr Val Ala Ser Phe Arg
260 265 270

Lys Gly Leu Ala Met Leu Pro Asn Ala Ile Ala Ser Arg Leu Gly Ser
275 280 285

Lys Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ala Asp Asn
290 295 300

Gln Gly Tyr Val Leu Gly Tyr Glu Thr Pro Glu Gly Leu Val Ser Val
305 310 315

Gln Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asp
320 325 330 335

Ile Leu Arg Pro Leu Ser Ile Asp Ala Ala Asp Ala Leu Ser Lys Phe
340 345 350

Tyr Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala
355 360 365
- 32 -
<table>
<thead>
<tr>
<th></th>
<th>Ala</th>
<th>Thr</th>
<th>Met</th>
<th>Ala</th>
<th>Thr</th>
<th>Val</th>
<th>Ala</th>
<th>Ala</th>
<th>Ala</th>
<th>Ala</th>
<th>Ala</th>
<th>Ser</th>
<th>Pro</th>
<th>Leu</th>
<th>Arg</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>5</td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Val</td>
<td>Thr</td>
<td>Gly</td>
<td>Arg</td>
<td>Pro</td>
<td>His</td>
<td>Arg</td>
<td>Val</td>
<td>Arg</td>
<td>Pro</td>
<td>Arg</td>
<td>Cys</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>25</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Ser</td>
<td>Ala</td>
<td>Thr</td>
<td>Glu</td>
<td>Thr</td>
<td>Pro</td>
<td>Ala</td>
<td>Ala</td>
<td>Pro</td>
<td>Gly</td>
<td>Val</td>
<td>Arg</td>
<td>Leu</td>
<td>Ser</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td></td>
<td>40</td>
<td></td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Cys</td>
<td>Val</td>
<td>Ile</td>
<td>Val</td>
<td>Gly</td>
<td>Ala</td>
<td>Gly</td>
<td>Ile</td>
<td>Ser</td>
<td>Gly</td>
<td>Leu</td>
<td>Cys</td>
<td>Thr</td>
<td>Ala</td>
<td>Gln</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>55</td>
<td></td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Leu</td>
<td>Ala</td>
<td>Thr</td>
<td>Arg</td>
<td>Tyr</td>
<td>Gly</td>
<td>Val</td>
<td>Ser</td>
<td>Asp</td>
<td>Leu</td>
<td>Leu</td>
<td>Val</td>
<td>Thr</td>
<td>Glu</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td></td>
<td>70</td>
<td></td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Asp</td>
<td>Arg</td>
<td>Pro</td>
<td>Gly</td>
<td>Gly</td>
<td>Asn</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Val</td>
<td>Glu</td>
<td>Arg</td>
<td>Pro</td>
<td>Asp</td>
<td>Glu</td>
<td></td>
</tr>
</tbody>
</table>

- 33 -
Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro
100 105 110
Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Asp Leu Val Phe
115 120 125
Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly Lys Leu Arg
130 135 140
Pro Val Pro Ser Lys Pro Gly Asp Leu Pro Phe Phe Ser Leu Met Ser
145 150 155 160
Ile Pro Gly Lys Leu Arg Ala Gly Leu Gly Ala Leu Gly Ile Arg Pro
165 170 175
Pro Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn
180 185 190
Leu Gly Ala Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly
195 200 205
Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly
210 215 220
Lys Val Trp Arg Leu Glu Glu Ile Gly Gly Ser Ile Ile Gly Gly Thr
225 230 235 240
Ile Lys Ala Ile Gln Asp Lys Gly Lys Asn Pro Lys Pro Pro Arg Asp
245 250 255
Pro Arg Leu Pro Ala Pro Lys Gly Gln Thr Val Ala Ser Phe Arg Lys
260 265 270
Gly Leu Ala Met Leu Pro Asn Ala Ile Ala Ser Arg Leu Gly Ser Lys
275 280 285
Val Lys Leu Ser Trp Lys Leu Thr Ser Ile Thr Lys Ala Asp Asn Gln
290    295    300
Gly Tyr Val Leu Gly Tyr Glu Thr Pro Glu Gly Leu Val Ser Val Gln
305    310    315    320
Ala Lys Ser Val Ile Met Thr Ile Pro Ser Tyr Val Ala Ser Asp Ile
325    330    335
Leu Arg Pro Leu Ser Ile Asp Ala Ala Asp Ala Leu Ser Lys Phe Tyr
340    345    350
Tyr Pro Pro Val Ala Ala Val Thr Val Ser Tyr Pro Lys Glu Ala Ile
355    360    365
Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Gln Gly Phe Gly Gln Leu
370    375    380
His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser
385    390    395    400
Ser Leu Phe Pro Asn Arg Ala Pro Ala Gly Arg Val Leu Leu Leu Asn
405    410    415
Tyr Ile Gly Gly Ser Thr Asn Thr Gly Ile Val Ser Lys Thr Glu Ser
420    425    430
Asp Leu Val Gly Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile Asn
435    440    445
Pro Arg Ala Ala Asp Pro Leu Ala Leu Gly Val Arg Val Trp Pro Gln
450    455    460
Ala Ile Pro Gln Phe Leu Ile Gly His Leu Asp Arg Leu Ala Ala Ala
465    470    475    480
Lys Ser Ala Leu Gly Gln Gly Gly Tyr Asp Gly Leu Phe Leu Gly Gly
485 490 495

Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala Tyr
500 505 510

Glu Ser Ala Ser Gln Val Ser Asp Phe Leu Thr Lys Tyr Ala Tyr Lys
515 520 525

<210> 11
<211> 1847
<212> DNA
<213> Glycine max

<220>
<221> CDS
<222> (55) . . (1683)
<223> soybean protox-1

<400> 11
ctttagcaca gtgttgaaga taacgaacga atagtgcacct tactgtaacc aacc atg
   57
   Met
   1

gtt tcc gtc ttc aac gag atc cta ttc cgc cgc aac caa acc ctt ctt
Val Ser Val Phe Asn Glu Ile Leu Phe Pro Pro Asn Gln Thr Leu Leu
   5 10 15

cgc ccc tcc tct cat tcc cca acc tct ttc acc tct ccc act cga
Arg Pro Ser Leu His Ser Pro Thr Ser Phe Phe Thr Ser Pro Thr Arg
   20 25 30

aaa ttc cct cgc tct cgc cct aac cct att cta cgc tgc tcc att gcg
Lys Phe Pro Arg Ser Arg Pro Arg Pro Phe Pro Ile Leu Arg Ser Ile Ala
   35 40 45

- 36 -
gag gaa tcc acc gcg tct cgc ccc aaa acc aga gac tcc gcc ccc gtc 249
Glu Glu Ser Thr Ala Ser Pro Pro Lys Thr Arg Asp Ser Ala Pro Val
50 55 60 65

gac tgc gtc gtc ggc gga ggc gtc agc ggc ctc tgc atc gcc cag 297
Asp Cys Val Val Val Gly Gly Gly Val Ser Gly Leu Cys Ile Ala Gln
70 75 80

gcc ctc gcc acc aaa cac gcc aat gcc aac gtc gtc gtc acg gag gcc 345
Ala Leu Ala Thr Lys His Ala Asn Ala Asn Val Val Val Thr Glu Ala
85 90 95

cga gac cgc gtc ggc ggc ggc aac atc acc acg atg gag agg gac gga tac 393
Arg Asp Arg Val Gly Gly Asn Ile Thr Thr Met Glu Arg Asp Gly Tyr
100 105 110

tct tgg gaa gga gcc ccc aac agc ttc cag cct tct gat cca atg ctc 441
Leu Trp Glu Glu Pro Ser Asn Ser Phe Gln Pro Ser Asp Pro Met Leu
115 120 125

acc atg gtg gtg gac agt ggt tta aag gat gag ctt gtt ttg ggg gat 489
Thr Met Val Val Asp Gly Leu Lys Asp Glu Leu Val Leu Gly Asp
130 135 140 145

cct gat gca cct cgg ttt gtg ttg tgg aac agg aag ttg agg ccc gtc 537
Pro Asp Ala Pro Arg Phe Val Leu Trp Asn Arg Lys Leu Arg Pro Val
150 155 160

ccc ggg aag ctg act gat ttg cct ttc ttt gcc tgg ggc atg agc att ggt 585
Pro Gly Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Gly
165 170 175

ggc aaa atc agg gct ggc ttt ggt ggc ctt gga att cgg cct cct cct 633
Gly Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg Pro Pro Pro
180 185 190

cca ggt cat gag gaa tcc gtt gaa gag ttt gtt cgt cgg aac ctt ggt 681

- 37 -
Pro Gly His Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly 
195  200  205

gat gag gtt ttt gaa cgg ttt ata gag cct ttt tgt tca ggg gtc tat.  729
Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr 
210  215  220  225

gca ggc gat cct tca aaa tta agt atg aaa gca gca ttc ggg aaa gtt 
777
Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Lys Val 
230  235  240

tgg aag ctg gaa aaa aat ggt ggt agc att att ggt gga act ttc aaa  825
Trp Lys Leu Glu Lys Asn Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys 
245  250  255

gca ata cca gag aga aat gga gct tca aaa cca cct cga gat cgc cgt 
873
Ala Ile Gln Glu Arg Asn Gly Ala Ser Lys Pro Pro Arg Asp Pro Arg 
260  265  270

c tg cca aaa cca aaa ggt cag act gtt gga tct ttc cgg aag gga ctt  921
Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu 
275  280  285

acc atg tgt cct gat gca att tct ggc aga cta ggc aac aaa gta aag  969
Thr Met Leu Pro Asp Ala Ile Ser Ala Arg Leu Gly Asn Lys Val Lys 
290  295  300  305

tta tgt cgg aag ctt tca agt att agt aaa ctg gat agt gga gag tac  1017
Leu Ser Trp Lys Leu Ser Ser Ile Ser Lys Leu Asp Ser Gly Glu Tyr 
310  315  320

agt tgt aca tat gaa cca gaa gga gtt tct tgt cag tgc aaa  1065
Ser Leu Thr Tyr Glu Thr Pro Glu Gly Val Val Ser Leu Gln Cys Lys 
325  330  335

act gtt gtc ctg acc att cct tcc tat gtt gct agt aca ttg ctg cgt  1113
Thr Val Val Leu Thr Ile Pro Ser Tyr Val Ala Ser Thr Leu Leu Arg
cct ctg tct gct gct gca gat gca ctt tca aag ttt tat tac cct
Pro Leu Ser Ala Ala Ala Ala Asp Ala Leu Ser Lys Phe Tyr Tyr Pro
355 360 365

cca gtt gct gca gtt tcc ata tcc tat cca aaa gaa gct att aga tca
Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg Ser
370 375 380 385

gaa tgc ttg ata gat ggt gag ttg aag ggg ttt ggt caa ttg cat cca
Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Glu Leu His Pro
390 395 400

cgt agc caa gga gtg gaa aca tta gga act ata tac agc tca tca cta
Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu
405 410 415

 TTC CCC AAC CGA CGA CCA CCT GGA AGG GTT CTA CTC TTG AAT TAC ATT
Phe Pro Asn Arg Ala Pro Pro Gly Arg Val Leu Leu Leu Asn Tyr Ile
420 425 430

gga gga gca act aat act gga att tta tcg aag acg gac agt gaa ctt
Gly Gly Ala Thr Asn Thr Gly Ile Leu Ser Lys Thr Asp Ser Glu Leu
435 440 445

gtg gaa aca gtt gat cga gat ttg agg aaa atc ctt ata aac cca aat
Val Glu Thr Val Asp Arg Asp Leu Arg Lys Ile Leu Ile Asn Pro Asn
450 455 460 465

gcc cag gat cca ttt gta gtg ggg gtt aga ctt tgt ctt cca gct att
Ala Gln Asp Pro Phe Val Val Gly Val Arg Leu Trp Pro Gln Ala Ile
470 475 480

cca cag ttc tta gtt ggc cat ctt gat ctt cta gat gtt gct aag gct
Pro Gln Phe Leu Val Gly His Leu Asp Leu Leu Asp Val Ala Lys Ala
485 490 495

- 39 -
tct atc aga aat act ggg ttt gaa ggg ctc ttc ctt ggg ggt aat tat 1593
Ser Ile Arg Asn Thr Gly Phe Glu Gly Leu Phe Leu Gly Gly Asn Tyr
500 505 510

gtg tct ggt gtt gcc ttg gga cga tgc gtt gag gga gcc tat gag gta 1641
Val Ser Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Val
515 520 525

gca gct gaa gta aac gat ttt ctc aca aat aga gtg tac aaa 1683
Ala Ala Glu Val Asn Asp Phe Leu Thr Asn Arg Val Tyr Lys
530 535 540

tagtagcagt tttttttttt gtggtggaat gggtgatggg actctcgtgt tccattgaat 1743
tataataatg tgaagtttcc tcaaatgtct tcagatagtt tttggcgggt tctatgcgtg 1803
ataatgtaaa atcctcttta gttttgaaaa aaaaaaaaaa aaaa 1847

<210> 12
<211> 543
<212> PRT
<213> Glycine max

<400> 12
Met Val Ser Val Phe Asn Glu Ile Leu Phe Pro Pro Asn Gln Thr Leu
1  5  10  15

Leu Arg Pro Ser Leu His Ser Pro Thr Ser Phe Phe Thr Ser Pro Thr
20  25  30

Arg Iys Phe Pro Arg Ser Arg Pro Asn Pro Ile Leu Arg Cys Ser Ile
35  40  45

Ala Glu Glu Ser Thr Ala Ser Pro Pro Lys Thr Arg Asp Ser Ala Pro
50  55  60

-40-
<table>
<thead>
<tr>
<th>Residue</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val</td>
<td>Asp</td>
</tr>
<tr>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Ala</td>
</tr>
<tr>
<td>85</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Arg</td>
</tr>
<tr>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Leu</td>
</tr>
<tr>
<td>115</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Thr</td>
</tr>
<tr>
<td>130</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Pro</td>
</tr>
<tr>
<td>145</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Pro</td>
</tr>
<tr>
<td>165</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
</tr>
<tr>
<td>180</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
</tr>
<tr>
<td>195</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Asp</td>
</tr>
<tr>
<td>210</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Ala</td>
</tr>
<tr>
<td>225</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Trp</td>
</tr>
<tr>
<td>245</td>
<td></td>
</tr>
</tbody>
</table>
Lys Ala Ile Gln Glu Arg Asn Gly Ala Ser Lys Pro Pro Arg Asp Pro
  260         265         270

Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly
  275         280         285

Leu Thr Met Leu Pro Asp Ala Ile Ser Ala Arg Leu Gly Asn Lys Val
  290         295         300

Lys Leu Ser Trp Lys Leu Ser Ser Ile Ser Lys Leu Asp Ser Gly Glu
  305         310         315         320

Tyr Ser Leu Thr Tyr Glu Thr Pro Glu Gly Val Val Ser Leu Gln Cys
  325         330         335

Lys Thr Val Val Leu Thr Ile Pro Ser Tyr Val Ala Ser Thr Leu Leu
  340         345         350

Arg Pro Leu Ser Ala Ala Ala Ala Asp Ala Leu Ser Lys Phe Tyr Tyr
  355         360         365

Pro Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala Ile Arg
  370         375         380

Ser Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His
  385         390         395         400

Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser
  405         410         415

Leu Phe Pro Asm Arg Ala Pro Pro Gly Arg Val Leu Leu Leu Asn Tyr
  420         425         430

Ile Gly Gly Ala Thr Asn Thr Gly Ile Leu Ser Lys Thr Asp Ser Glu
  435         440         445

Leu Val Glu Thr Val Asp Arg Asp Leu Arg Lys Ile Leu Ile Asn Pro
450 455 460
Asn Ala Gln Asp Pro Phe Val Val Gly Val Arg Leu Trp Pro Gln Ala
465 470 475 480
Ile Pro Gln Phe Leu Val Gly His Leu Asp Leu Asp Val Ala Lys
485 490 495
Ala Ser Ile Arg Asn Thr Gly Phe Glu Gly Leu Phe Leu Gly Gly Gly Asn
500 505 510
Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu
515 520 525
Val Ala Ala Glu Val Asn Asp Phe Leu Thr Asn Arg Val Tyr Lys
530 535 540

<210> 13
<211> 583
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> misc_feature
<222> (1) .. (583)
<223> Arabidopsis protox-1 promoter

<400> 13
gaattccogat cgaattatat aattatcata aatttgataa agcatgttgc cttttatta 60
agagttttaa taaagtttgg taataatgga ctttgacttc aaaaactgatt ctcatgtaat 120
taattaatat ttacatcaca aattggtcac taatattacc aaattaatat actaaaatgt 180
taattgcaa ataaaacact aatttcaaat aaggggtcat tattgataac acgtattgaa 240
cttgataaag caaagcaaaa ataatgggtt tcaaggttgg gtttatatat gacaaaaaaaa 300

aaaaaaagtt tggtttatata tctattgggc ctataaccat gtataacaaaa tttagggctta 360

actaaaataaa taaaaataaac gtaatggtcc ttttttatatt tgggtcaaac ccaacctcta 420

acccaaacca aaaaagggat atacgggtacg gtacacagacr ttaatgggttg tggatgctca 480

gggtgatatt ttcgctgatgc ttcctctccc ttctcgagaa gattacccca tctgaaaaaa 540

accaagaagc tgacaaattt ccaaaatttct gttcgatttccc atg 583

<210> 14
<211> 3848
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)..(3848)
<223> maize protox-1 promoter

<400> 14
tcgatcttttc taggcgtgatc cccaaatctt cctcogaagc cccctggtgcc tctgccccctt 60
ggagctgggtg gctggaaga gctttgctgt tgcccgaagc atttgaggt attatgtgac 120
tctgtagacg gacttcccttc gttcgtcactt tgagtgggat ttagttggga cctgacgctgc 180
tcacagtgga ttcctctccc gagaaggcccttg gtcattttggg aagatctgta atctttttccc 240
tcctttttgggc gaaaaatctgt caggctggat gtaacctccgg atctttcttgaa gcagcttttcc 300
caggttttttg gaggcctgctcc tggcgaataa ttggtgctgta gttcctggagc gaagaccccc 360
gatcagttggcc tcaagcaca tttcatttggg caccgttaggc gtttgtggcc tcaatcgc 420
gaaccttccgt accatgacct aaggtacct ttcggtgatct tgttgccatt ggaacagagc 480
ctgagccttg aacgacttccg tttgaaggcc ttggaagcta gtaaccaaca tgtgcttaag 540
cctctgccc acacgctatag tccccctgccc gaaagaagaa tacacagt tgtgactacat gcc 600
cccgactgcc atgacggaagg acttgcgcatc gactacagtg ttgaccccata agaagagatat 660
agttgccttcg tagcctcatca gaaactgtct tggatatgag tgcoccatcat acatgggggag 720
cggagtggcc ttgtatgag ggccctatccg ggtacgcctgc agttctgtgc cccagggagc 780
agcactcatca aaagtaaagg cattcatgatt aatcatcatca tacattccca cctggtgaa 840
taacctctccg tgcgaagacct cccctgctttg ggcctctccg ttcctgttct cttgaccaag 900
atgaagcact ccctcagttgg ctctgctgat cttcccttgg agatcagccca gtgcaccatc 960
cctctccttc ttctttttgta ctcttggtatg gatgatetcc atgtcatctca tctcttgttc 1020
caactctcct tctggaggtg tcgactggtt ggccctcttc ttcggtcttc gcagcctctcg 1080
aagagaaaga gtgtctttgt gttgggtcag cggctgcagt gcagttggtcc cttggtctga 1140
agctttcttc ggtggcctgta ccaggggccg ttctggtgccg agttgggtgca aaagggttca 1200
ctagaggtgg gagcccaatgt tggtggacttc tcaagttgtca tgtgtaaga acaagggcaac 1260
aacaatgttt aaatatataa agctttctcag ttccgagaac tattttctct tgtgggtataa 1320
tgatctttcag aacgaaagagt cccttcctat gtcgatatat gttaatagaa ggagggactt 1380
atgaaatgtga agagacaaaca tagaacaatcc tgtagcattg ttaattcata acatatttat 1440
tttactggaa aaatagaaac aatattgaaat taaatatgta cctttggcct gacagaagat 1500
aaaagtacaa gcttgacgca cagaccaagtca cagtcagtg tgaacagttc gggtgtactg 1560
ttcattatt tatagccaca ggacacagcc ttggagaaat tacagtcatt cccttttacat 1620
ttaacattga cttatatgaa aactctatgag gactggatag cctttttccc ttttaagtcgg 1680
tgccttttttc cggagattaag ccgaatctcc ctgctgcata gctctgagac atcgagcaaacc 1740
ttcgtcagga tcatgcctct ctcattgtgt atgctttttaa tcctgaattc gaaggtaacct 1800
gtcataaatc cataacttggaa agacattttg aatattatgt ttggaggaaacc ttgagaggac 1860
gaagccccc aacagtcgtg tttttgagga ccttggagaag atgaaagcctt ccacaacaagac 1920
catatcataa acacacacta tccacaacac cgcaccccaat caccctctcat tggcctccac 1980
aacaacacca attaggttgt tggttttaaat ttttttaggtt caatgggttc atccaccatcc 2040
actg tcactc cacaacacta atatcaataa acagactcaaa tccaccaaac tgcacccatac 2100
cataaaccog cccccacccttt ctacgcgcttc gccagaaacc agaaaccctgg atccagagtt 2160
cacaacttaa aacagccataa ctcccacctct ggcaatcgcga tcaaggctcct ttttttcca 2220
atcaacacaa attaatttcc gcatcogata atcaagccat ctctctacta tggtttttaag 2280
tggtgctcct actaggttat ttatggacta atcaacctggct tatctcatac aataacatat 2340
cagtcactt caagtgttac tcaattacca aaacagaatt atacctttag aaagagttta 2400
tcgcactagtc actcaatttcc caaaaactaaaa cttttagactt tcattgtatga catcccaacat 2460
gacactgtac tggactaaac cacttttttaa gctacaaag gagaaaaaat aactaattttt 2520
cgtagttgta ggaagctttaag tatagcttca cacaatagtt taagggaagcc ccccaaggagg 2580
TTTTAAAGTCGCTTTCCAC TTTAAACTTC TTCATCCTT CTTTAACCTC 2640
aaaatttgac attttatcac cccttaactc tttaaacat tttaattaca ttcttactag 2700
attataagatg atttgtttgtg gaaaaagtttt taagacatgt ttacacattg attaataatca 2760
tttggtcataat ttgcctagagt taatatctaat cttatataaa ctattagaga taccttcacg 2820
agctctaaat atttttatttt tttcattattg gaattttgttt aagaattctta tagaccccttt 2880
ttgggtggttt aaaaagcttttg ccagttttttt aacaagttttt tttttctattt tttgaaaattt 2940
tcttggaaacct cactctcaac ccggtagagct atttatattttctacaattt atctacacaa 3000
aatacacttt atggaaattttc ttcggaaacct acctctaaccc cggtagaatg aatttgaattg 3060
aaaaattaaac caaccttacgg aatocgcacc aatatgctga tttaagtggta catggataac 3120
tataaggac cccctagagat atacttaaatt gtttcagaaact ggagggttat tttttgaagt 3180
ttgatgggga aaataagccac taacgtagttg tccagagagat aaaaaagggtta ttttttttcag 3240
aaatatatttg gctgcaatttg atcctgtgccc ttaaattcag cctggcaacca aggccaggttt 3300
cctagagcaga caagggccac cgcacccttg gccccgtcagg cgaagcaggtt tttggtcaga 3360
ccttggaggg gattggatat caaoggaacc aacctagcacc ggcgaatgca gttccagccc 3420
acctgtaacc ttcccggtgct cctcctttaaa ctccaaagccc aaoggcctca ccccatctcg 3480
tcggtctcatc cactcagcgc cacaaggcgct cagctccgca acgagcccgag aaatgttcgc 3540
ccgccagcgg acggcgatgg gacacccttgc aatcgccgcra atccaaogggc cccgaattaac 3600
tcgcgagcctc cgcctacagtg gacctcagcgt gcgcgcagctc gcggtggcgcgc gcggccgcgcgc 3660
cgagggccagc gcatccccgg gcggccgcgt gcgcgcggcc gcgcgcggagc tgggcggaggg 3720
catacagtgcc ctctgcaccc gcgcagcgct ggccacgcgg cacggcgtcg gggacggtct 3780
tgtcaacgag gcgcgcgcgc gcgcgcgcgc caacattacc accgtogacgc gcgcgcgagga 3840
agggtaacc 3848

<210> 15
<211> 1826
<212> DNA
<213> Gossypium hirsutum

<220>
<221> CDS
<222> (31) .. (1647)
<223> cotton protox-1 coding sequence

<400> 15
cctctgcgtc gcctggcccc accaccaata atg ago gct cta atc gac ctt tact 54
         Met Thr Ala Leu Ile Asp Leu Ser
         1         5

         ctt ctc cgt tcc tgg ccc tcc gtt tcc ctt tcc tcc ata ccc cac cac
         Leu Leu Arg Ser Ser Pro Ser Val Ser Pro Phe Ser Ile Pro His His
         10        15        20

cag cat ccg ccc cgc ttt cgt aaa cct ttg aag ctc cga tgc tcc ctc
         Gln His Pro Pro Arg Phe Arg Lys Pro Phe Lys Leu Arg Cys Ser Leu
         25        30        35        40

gcc gag ggt ccc aag att tcc tca tct aaa atc gac ggg gga gaa tca
         Ala Glu Gly Pro Thr Ile Ser Ser Ser Lys Ile Asp Gly Gly Glu Ser
         45        50        55

tcc atc gcg gat tgc gtc atc gtt gga ggt atc agt gga ctt tgc
         Ser Ile Ala Asp Cys Val Ile Val Gly Gly Ile Ser Gly Leu Cys
         60        65        70
ATT GCT CAA GCT CTC GCC ACC AAC AAG CAC CGT GAC GTC GCT TCC AAT GTG
Ile Ala Glu Ala Leu Ala Thr Lys His Arg Asp Val Ala Ser Asn Val
75 80 85

ATT GTG ACG GAG GCC AGA GAC CGT GTG GTG GCC AAC AGT TTT CAG CCC
Ile Val Thr Glu Ala Arg Asp Arg Val Gly Gly Asn Ile Thr Thr Val
90 95 100

GAG AGA GAT GGA TAT CTG TGG GAA GAA GGC CCC AAC AGT TTT CAG CCC
Glu Arg Asp Gly Tyr Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro
105 110 115 120

TCC GAT CCT ATT CTA ACC ATG GCC GTG GAT AGT GGA TTG AAG GAC GAT
Ser Asp Pro Ile Leu Thr Met Ala Val Asp Ser Gly Leu Lys Arg Asp
125 130 135

TTG GTC TTA GGT GAC CCA AAC CCG CTC AAT GCA CGA TTA CTA TGG GAG GGA
Leu Val Leu Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Glu Gly
140 145 150

AAA CTA AGG CCT GTG CCC TCC AAC ACC GAC TTG CCC TTT TTT GAT
Lys Leu Arg Pro Val Pro Ser Lys Pro Thr Asp Leu Pro Phe Phe Asp
155 160 165

TTG ATG AGC ATT GCT GGA AAA CTT AGG GCT GGG TTC GGG GCT ATT GGC
Leu Met Ser Ile Ala Gly Lys Leu Arg Ala Gly Phe Gly Ala Ile Gly
170 175 180

ATT CGG CCT CCC CCT CGT TAT GAA GAA TCG GTG GAG GAG TTT GTG
Ile Arg Pro Pro Pro Gly Tyr Glu Ser Val Glu Glu Phe Val
185 190 195 200

CGG CGT AAT CTT GGT GCT GAG GTT TTT GAA CGG TTT ATT GAA CCA TTT
Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg Phe Ile Glu Pro Phe
205 210 215

- 49 -
tgt tca ggt gtt tat gca ggg gat cct tca aaa tta agc atg aaa gca
Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala
220 225 230

gca ttt gga aga gta tgg aag cta gaa gag att ggt ggc agc atc att
Ala Phe Gly Arg Val Trp Lys Leu Glu Glu Ile Gly Ser Ile Ile
235 240 245

ggt ggc act ttc aag cca atc cag gag aga aat aag cca cct aag cca
Gly Gly Thr Phe Lys Thr Ile Gln Glu Arg Asn Lys Thr Pro Lys Pro
250 255 260

ccc aga gac ccg cgt ctt cca aac ccg aag ggc caa aca gtt gga tct
Pro Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly Glu Thr Val Gly Ser
265 270 275 280

aaa gga ctt acc atg ctt gag gcg att gct aac agt ttg
Phe Arg Lys Gly Thr Met Leu Pro Ala Ile Ala Asn Ser Leu
285 290 295

agc aat gta aca tta tct tgg aag ctt tgc agt att acc aaa ttg
Gly Ser Asn Val Lys Leu Ser Trp Leu Ser Ser Ile Thr Lys Leu
300 305 310

agc aat gga ggg tat aac ttg aca ttt gaa aca cct gaa gga atg gta
Gly Asn Gly Tyr Asn Leu Thr Phe Glu Thr Pro Glu Gly Met Val
315 320 325

tct ctt cag agt aga agt gtt gta atg acc att cca tcc cat gtt gcc
Ser Leu Gln Ser Arg Ser Val Val Met Thr Ile Pro Ser His Val Ala
330 335 340

agt aac ttg ttg cat ctc tct tcg gct gct gct gca gat gcc tta tcc
Ser Asn Leu Leu His Pro Leu Ser Ala Ala Ala Asp Ala Leu Ser
345 350 355 360

caa ttt tat tat cct cca gtt gca tca gtc aca gtc tcc tat cca aaa
1158
Gln Phe Tyr Tyr Pro Pro Val Ala Ser Val Thr Val Ser Tyr Pro Lys  

<table>
<thead>
<tr>
<th></th>
<th>365</th>
<th>370</th>
<th>375</th>
</tr>
</thead>
</table>

gaa gcc att cga aaa gaa tgt ttg att gat ggt gaa ctt aag ggg ttt  
Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe  

<table>
<thead>
<tr>
<th></th>
<th>380</th>
<th>385</th>
<th>390</th>
</tr>
</thead>
</table>

ggc cag ttg cac cca gcgc agc caa gga att gaa act tta ggg acg ata  
Gly Gln Leu His Pro Arg Ser Gln Gly Ile Glu Thr Leu Gly Thr Ile  

<table>
<thead>
<tr>
<th></th>
<th>395</th>
<th>400</th>
<th>405</th>
</tr>
</thead>
</table>

tac aat tca cta ctt ttcc ccc aat cga gct cca tct ggc agg gtc ttc  
Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro Ser Gly Arg Val Leu  

<table>
<thead>
<tr>
<th></th>
<th>410</th>
<th>415</th>
<th>420</th>
</tr>
</thead>
</table>

cct ttc aac tac ata gga gga gct acc aac act gga aat ttg tcc aag  
Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Leu Ser Lys  

<table>
<thead>
<tr>
<th></th>
<th>425</th>
<th>430</th>
<th>435</th>
<th>440</th>
</tr>
</thead>
</table>

act gaa ggg gaa ctt gta gaa gca gtt gat cgt gat ttg aga aaa atg  
Thr Glu Gly Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met  

<table>
<thead>
<tr>
<th></th>
<th>445</th>
<th>450</th>
<th>455</th>
</tr>
</thead>
</table>

cct ata aat cct aat gca aag gat cct ctt gtt ggt gta gta aga gta  
Leu Ile Asn Pro Asn Ala Lys Asp Pro Leu Val Leu Gly Val Arg Val  

<table>
<thead>
<tr>
<th></th>
<th>460</th>
<th>465</th>
<th>470</th>
</tr>
</thead>
</table>

ctg cca aca gcc att cca cag ttc ttg ggt cat ttg gat ctc ctt  
Trp Pro Lys Ala Ile Pro Glu Phe Leu Val Gly His Leu Asp Leu Leu  

<table>
<thead>
<tr>
<th></th>
<th>475</th>
<th>480</th>
<th>485</th>
</tr>
</thead>
</table>

gat aat gca aca agt gct ctc agg gat tct ggg ttt cat gga ctc tgt ttt  
Asp Ser Ala Lys Met Ala Leu Arg Asp Ser Gly Phe His Gly Leu Phe  

<table>
<thead>
<tr>
<th></th>
<th>490</th>
<th>495</th>
<th>500</th>
</tr>
</thead>
</table>

ctt ggg gcc aac tat gta tct ggt gtt gca tta gga cgg tgt gtc gaa  
Leu Gly Gly Asn Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Val Glu  

<table>
<thead>
<tr>
<th></th>
<th>500</th>
<th>505</th>
<th>510</th>
</tr>
</thead>
</table>
505 510 515 520

Gly Ala Tyr Glu Val Ala Ala Glu Val Lys Glu Phe Leu Ser Gln Tyr

525 530 535

gca tac aaa taatatgtgaa attctttgta ggctgcacaat gtagaagtc

1687

Ala Tyr Lys

1638

gttatgat agtactcttt tagctaaaaa atgggtacgg gtttttttttg ttatgttccctt 1747

gaccactttt tggggttttct attagaacctt cattattgta ttcattgttg caatatcaaa 1807

aaaaaaaaa aaaaaaaa

1826

<210> 16
<211> 539
<212> PRT
<213> Gossypium hirsutum

<400> 16

Met Thr Ala Leu Ile Asp Leu Ser Leu Leu Arg Ser Ser Pro Ser Val

1 5 10 15

Ser Pro Phe Ser Ile Pro His His Gln His Pro Pro Arg Phe Arg Lys

20 25 30

Pro Phe Lys Leu Arg Cys Ser Leu Ala Glu Gly Pro Thr Ile Ser Ser

35 40 45

Ser Lys Ile Asp Gly Gly Glu Ser Ser Ile Ala Asp Cys Val Ile Val

50 55 60

Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala Thr Lys

65 70 75 80
His Arg Asp Val Ala Ser Asn Val Ile Val Thr Glu Ala Arg Asp Arg
 85                     90                      95
Val Gly Gly Asn Ile Thr Thr Val Glu Arg Asp Gly Tyr Leu Trp Glu
 100                    105                     110
Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Pro Ile Leu Thr Met Ala
 115                    120                     125
Val Asp Ser Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Asn Ala
 130                    135                    140
Pro Arg Phe Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys
 145                    150                    155                    160
Pro Thr Asp Leu Pro Phe Phe Asp Leu Met Ser Ile Ala Gly Lys Leu
 165                    170                    175
Arg Ala Gly Phe Gly Ala Ile Gly Ile Gly Arg Pro Pro Pro Gly Tyr
 180                    185                    190
Glu Glu Ser Val Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val
 195                    200                    205
Phe Glu Arg Phe Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp
 210                    215                    220
Pro Ser Lys Leu Ser Met Lys Ala Ala Phe Gly Arg Val Trp Lys Leu
 225                    230                    235                    240
Glu Glu Ile Gly Gly Ser Ile Ile Gly Gly Thr Phe Lys Thr Ile Gln
 245                    250                    255
Glu Arg Asn Lys Thr Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Lys
 260                    265                    270
Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Thr Met Leu

- 53 -
<table>
<thead>
<tr>
<th>275</th>
<th>280</th>
<th>285</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro Glu Ala Ile Ala Asn Ser Leu Gly Ser Asn Val Lys Leu Ser Trp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>290</td>
<td>295</td>
<td>300</td>
</tr>
<tr>
<td>Lys Leu Ser Ser Ile Thr Lys Leu Gly Asn Gly Gly Tyr Asn Leu Thr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>305</td>
<td>310</td>
<td>315</td>
</tr>
<tr>
<td>Phe Glu Thr Pro Glu Gly Met Val Ser Leu Gln Ser Arg Ser Val Val</td>
<td></td>
<td></td>
</tr>
<tr>
<td>325</td>
<td>330</td>
<td>335</td>
</tr>
<tr>
<td>Met Thr Ile Pro Ser His Val Ala Ser Asn Leu Leu His Pro Leu Ser</td>
<td></td>
<td></td>
</tr>
<tr>
<td>340</td>
<td>345</td>
<td>350</td>
</tr>
<tr>
<td>Ala Ala Ala Ala Asp Ala Leu Ser Gln Phe Tyr Tyr Pro Pro Val Ala</td>
<td></td>
<td></td>
</tr>
<tr>
<td>355</td>
<td>360</td>
<td>365</td>
</tr>
<tr>
<td>Ser Val Thr Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>370</td>
<td>375</td>
<td>380</td>
</tr>
<tr>
<td>Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Ser Gln</td>
<td></td>
<td></td>
</tr>
<tr>
<td>385</td>
<td>390</td>
<td>395</td>
</tr>
<tr>
<td>Gly Ile Glu Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>405</td>
<td>410</td>
<td>415</td>
</tr>
<tr>
<td>Arg Ala Pro Ser Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala</td>
<td></td>
<td></td>
</tr>
<tr>
<td>420</td>
<td>425</td>
<td>430</td>
</tr>
<tr>
<td>Thr Asn Thr Gly Ile Leu Ser Lys Thr Glu Gly Glu Leu Val Glu Ala</td>
<td></td>
<td></td>
</tr>
<tr>
<td>435</td>
<td>440</td>
<td>445</td>
</tr>
<tr>
<td>Val Asp Arg Asp Leu Arg Lys Met Leu Ile Asn Pro Asn Ala Lys Asp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>450</td>
<td>455</td>
<td>460</td>
</tr>
<tr>
<td>Pro Leu Val Leu Gly Val Arg Val Trp Pro Lys Ala Ile Pro Gln Phe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>465</td>
<td>470</td>
<td>475</td>
</tr>
</tbody>
</table>
Leu Val Gly His Leu Asp Leu Leu Asp Ser Ala Lys Met Ala Leu Arg
485 490 495
Asp Ser Gly Phe His Gly Leu Phe Leu Gly Gly Asn Tyr Val Ser Gly
500 505 510
Val Ala Leu Gly Arg Cys Val Glu Gly Ala Tyr Glu Val Ala Ala Glu
515 520 525
Val Lys Glu Phe Leu Ser Gln Tyr Ala Tyr Lys
530 535

<210> 17
<211> 1910
<212> DNA
<213> Beta vulgaris

<220>
<221> CDS
<222> (1)...(1680)
<223> sugar beet protox-1 coding sequence

<400> 17
atg aaa tca atg ggc tta tca aac tgc att cca cag aca cag tgc atg 48
Met Lys Ser Met Ala Leu Ser Asn Cys Ile Pro Gln Thr Gln Cys Met
1 5 10 15
cca ttg cgc agc agc ggg cat tac agg ggt aat tgt atc atg ttg tca 96
Pro Leu Arg Ser Ser Gly His Tyr Arg Gly Asn Cys Ile Met Leu Ser
20 25 30
att cca tgt agt tta att gga aga cga ggt tat tat tca cat aag aag 144
Ile Pro Cys Ser Leu Ile Gly Arg Arg Gly Tyr Tyr Ser His Lys Lys
35 40 45

-55-
agg agg atg agc atg agt tgc agc agc tca ggc tca aag tca gcg 192
Arg Arg Met Ser Met Ser Cys Ser Thr Ser Ser Gly Ser Lys Ser Ala
50 55 60

gtt aaa gaa gca gga tca gga tca ggt gca gga gga ttg cta gac tgc 240
Val Lys Glu Ala Gly Ser Gly Ser Gly Ala Gly Gly Leu Leu Asp Cys
65 70 75 80

gta atc gtt gga ggt gga att agc ggg ctt tgc atc gcg cag gct ctt 288
Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu
85 90 95

tgt aca aaa cac tcc tct tcc tct tta tcc cca aat ttt ata gtt aca 336
Cys Thr Lys His Ser Ser Ser Ser Leu Ser Leu Pro Asn Phe Ile Val Thr
100 105 110

gag gcc aaa gac aga gtt ggc ggc aac atc gtc act gtg gag gcc gat 384
Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Val Thr Val Glu Ala Asp
115 120 125

ggc tat atc tgg gag gag gga ccc aat agc ttc cag cct tcc gac gcg 432
Gly Tyr Ile Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Ala
130 135 140

gtg ctc acc atg gcg gtc gac agt ggc ttg aaa gat gag ttg gtg ctc 480
Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu
145 150 155 160

gga gat ccc aat gct cct cgc ttt gta cta tgg aat gac aaa tta agg 528
Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Asn Asp Lys Leu Arg
165 170 175

ccc gta cct tcc agt ctc acc gac ctc cct ttc ttc gac ctc atg acc 576
Pro Val Pro Ser Ser Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Thr
180 185 190

att cgc ggc aag att agg gct gct ctt ggt gct ctc gga ttt cgc cct 624
Ile Pro Gly Lys Ile Arg Ala Ala Leu Gly Ala Leu Gly Phe Arg Pro
195  200  205

tct cct cca cct cat gag gaa tct gtt gaa cac ttt gtt cgt cgt aat  672
Ser Pro Pro Pro His Glu Glu Ser Val Glu His Phe Val Arg Arg Asn
210  215  220

ctc gga gat gag gtc ttt gaa cgc ttg att gaa ccc ttt tgg tca ggt  720
Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly
225  230  235  240

gtg tat gcc ggt gat cct gcc aag ctg agt atg aaa gct gct ttt ggg  768
Val Tyr Ala Gly Asp Pro Ala Lys Leu Ser Met Lys Ala Ala Phe Gly
245  250  255

aag gtc tgg aag ttg gag caa aag ggt ggc agc ata att ggt gcc act  816
Lys Val Trp Lys Leu Glu Gln Gly Gly Gly Ser Ile Ile Gly Gly Thr
260  265  270

ctc aaa gct ata cag gaa aga ggg agt aat cct aag ccc cct gct gac  864
Leu Lys Ala Ile Gln Glu Arg Gly Ser Asn Pro Lys Pro Pro Arg Asp
275  280  285

cag cgc ctc cct aaa cca aag ggt cag act gtt gga tcc ttt aga aag  912
Gln Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys
290  295  300

gga ctc gtt atg ttg cct acc gcc att tct gct cga ctt ggc agt aca  960
Gly Leu Val Met Leu Pro Thr Ala Ile Ser Ala Arg Leu Gly Ser Arg
305  310  315  320

gtg aaa cta tct tgg acc ctt tct agt atg gta aag tca ctc aat gga 1008
Val Iys Leu Ser Trp Thr Leu Ser Ser Ile Val Lys Ser Leu Asn Gly
325  330  335

gaa tat agt ctg act tat gat acc cca gat ggc ttg gtt tct gta aag 1056
Glu Tyr Ser Leu Thr Tyr Asp Thr Pro Asp Gly Leu Val Ser Val Arg

- 57 -
acc aaa agt gtt gtg atg act gtt cca tca tat gtt gca agt agg ctt
Thr Lys Ser Val Val Met Thr Val Pro Ser Tyr Val Ala Ser Arg Leu
355 360 365

ctt cgt cca ctt tca gac tct gct gca gat tct ctt tca aaa ttt tac
Leu Arg Pro Leu Ser Asp Ser Ala Ala Asp Ser Leu Ser Lys Phe Tyr
370 375 380

tat cca cca gtt gca gca gtg tca ctt tcc tat cct aaa gaa gcg atc
Tyr Pro Pro Val Ala Ala Val Ser Leu Ser Tyr Pro Lys Glu Ala Ile
385 390 395 400

aga tca gaa tgc ttg att aat ggt gaa ctt cca ggt ttc ggg cca cta
Arg Ser Glu Cys Leu Ile Asn Gly Glu Leu Gln Gly Phe Gly Gln Leu
405 410 415

cat ccc cgc agt cag ggt gtg gaa acc ttg gga aca att tat agt tcg
His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser
420 425 430

tct ctt ttc cct ggt cga gca cca cct ggt agg atc ttg atc ttg agc
Ser Leu Phe Pro Gly Arg Ala Pro Pro Gly Arg Ile Leu Ile Leu Ser
435 440 445

tac atc gga gtt gct aaa aat cct ggc ata tta aac aag tcg aaa gat
Tyr Ile Gly Gly Ala Lys Asn Pro Gly Ile Leu Asn Lys Ser Lys Asp
450 455 460

gaa ctt gcc aag cca gtt gac aag gac ctc aga aga atg ctt ata aat
Glu Leu Ala Lys Thr Val Asp Lys Asp Leu Arg Arg Met Leu Ile Asn
465 470 475 480

cct gat gca aaa ctt cct gct gta ctg ggt gtt gga gta tgg cct cca
Pro Asp Ala Lys Leu Pro Arg Val Leu Gly Val Arg Val Trp Pro Gln
485 490 495
gca ata ccc cag ttt tct att ggg cac ttt gat ctc gat gct gca 1536
Ala Ile Pro Gln Phe Ser Ile Gly His Phe Asp Leu Leu Asp Ala Ala
500 505 510

aaa gct gct ctg aca gat aca ggg gtc aaa gga ctg ttt ctt ggt ggc 1584
Lys Ala Ala Leu Thr Asp Thr Gly Val Lys Gly Leu Phe Leu Gly Gly
515 520 525

aac tat gtt tca ggt gtt gcc ttg ggg cgg tgt ata gag ggt gct tat 1632
Asn Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala Tyr
530 535 540

gag tct gca gct gag gta gta gat ttc ctc tca cag tac tca gac aaa 1680
Glu Ser Ala Ala Glu Val Val Asp Phe Leu Ser Gln Tyr Ser Asp Lys
545 550 555 560

tagagcttca gcatacgttg taattcaaca cagccttttt tgtatcttgtt gtgogogcat 1740
gtagtctgtt cgtggtgtctgaa gattgatta gttgctctgct tgtgtgatcc acaagaatttt 1800
tgatggaatt ttccagatgt gggcattat atgttgctgt cttataatcc cttaatgtgt 1860
acgttttagtg aattacacgg catttgatga ctaaaaaaaaa aaaaaa
1910

<210> 18
<211> 560
<212> PRT
<213> Beta vulgaris

<400> 18
Met Lys Ser Met Ala Leu Ser Asn Cys Ile Pro Gln Thr Gln Cys Met
1 5 10 15

Pro Leu Arg Ser Ser Gly His Tyr Arg Gly Asn Cys Ile Met Leu Ser
20 25 30

- 59 -
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ile Pro Cys Ser Leu Ile Gly Arg Arg Gly Tyr Tyr Ser His Lys Lys</td>
<td>35</td>
</tr>
<tr>
<td>Arg Arg Met Ser Met Ser Cys Ser Thr Ser Ser Gly Ser Lys Ser Ala</td>
<td>50</td>
</tr>
<tr>
<td>Val Lys Glu Ala Gly Ser Gly Ser Gly Ala Gly Gly Leu Leu Asp Cys</td>
<td>65</td>
</tr>
<tr>
<td>Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu</td>
<td>85</td>
</tr>
<tr>
<td>Cys Thr Lys His Ser Ser Ser Ser Leu Ser Ser Pro Asn Phe Ile Val Thr</td>
<td>100</td>
</tr>
<tr>
<td>Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Val Thr Val Glu Ala Asp</td>
<td>115</td>
</tr>
<tr>
<td>Gly Tyr Ile Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp Ala</td>
<td>130</td>
</tr>
<tr>
<td>Val Leu Thr Met Ala Val Asp Ser Gly Leu Lys Asp Glu Leu Val Leu</td>
<td>145</td>
</tr>
<tr>
<td>Gly Asp Pro Asn Ala Pro Arg Phe Val Leu Trp Asn Asp Lys Leu Arg</td>
<td>165</td>
</tr>
<tr>
<td>Pro Val Pro Ser Ser Leu Thr Asp Leu Pro Phe Phe Asp Leu Met Thr</td>
<td>180</td>
</tr>
<tr>
<td>Ile Pro Gly Lys Ile Arg Ala Ala Leu Gly Ala Gly Phe Arg Pro</td>
<td>195</td>
</tr>
<tr>
<td>Ser Pro Pro Pro His Glu Glu Ser Val Glu His Phe Val Arg Arg Asn</td>
<td>210</td>
</tr>
</tbody>
</table>

-60-
Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser Gly  
225 230 235 240

Val Tyr Ala Gly Asp Pro Ala Lys Leu Ser Met Lys Ala Ala Phe Gly  
245 250 255

Lys Val Trp Lys Leu Glu Gln Lys Gly Gly Ser Ile Ile Gly Gly Thr  
260 265 270

Leu Lys Ala Ile Gln Glu Arg Gly Ser Asn Pro Lys Pro Pro Arg Asp  
275 280 285

Gln Arg Leu Pro Lys Pro Lys Gly Gln Thr Val Gly Ser Phe Arg Lys  
290 295 300

Gly Leu Val Met Leu Pro Thr Ala Ile Ser Ala Arg Leu Gly Ser Arg  
305 310 315 320

Val Lys Leu Ser Trp Thr Leu Ser Ser Ile Val Lys Ser Leu Asn Gly  
325 330 335

Glu Tyr Ser Leu Thr Tyr Asp Thr Pro Asp Gly Leu Val Ser Val Arg  
340 345 350

Thr Lys Ser Val Val Met Thr Val Pro Ser Tyr Val Ala Ser Arg Leu  
355 360 365

Leu Arg Pro Leu Ser Asp Ser Ala Ala Asp Ser Leu Ser Lys Phe Tyr  
370 375 380

Tyr Pro Pro Val Ala Ala Val Ser Leu Ser Tyr Pro Lys Glu Ala Ile  
385 390 395 400

Arg Ser Glu Cys Leu Ile Asn Gly Glu Leu Gln Gly Phe Gly Gln Leu  
405 410 415

His Pro Arg Ser Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser Ser
Ser Leu Phe Pro Gly Arg Ala Pro Pro Gly Arg Ile Leu Ile Leu Ser  
        435  440  445  
Tyr Ile Gly Gly Ala Lys Asn Pro Gly Ile Leu Asn Lys Ser Lys Asp  
        450  455  460  
Glu Leu Ala Lys Thr Val Asp Lys Asp Leu Arg Arg Met Leu Ile Asn  
        465  470  475  480  
Pro Asp Ala Lys Pro Arg Val Leu Gly Val Arg Val Trp Pro Gln  
        485  490  495  
Ala Ile Pro Gln Phe Ser Ile Gly His Phe Asp Leu Leu Asp Ala Ala  
        500  505  510  
Lys Ala Ala Leu Thr Asp Thr Gly Val Lys Gly Leu Phe Leu Gly Gly  
        515  520  525  
Asn Tyr Val Ser Gly Val Ala Leu Gly Arg Cys Ile Glu Gly Ala Tyr  
        530  535  540  
Glu Ser Ala Ala Glu Val Val Asp Phe Leu Ser Gln Tyr Ser Asp Lys  
        545  550  555  560  

<210> 19  
<211> 1784  
<212> DNA  
<213> Brassica napus  

<220>  
<221> CDS  
<222> (47) .. (1654)  
<223> Oilseed rape protox-1 coding sequence
<table>
<thead>
<tr>
<th>Codon</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>gggccc</td>
<td>19 Met</td>
</tr>
<tr>
<td>ccaattgag</td>
<td>20 Leu</td>
</tr>
<tr>
<td>gattctcctt</td>
<td>25 Arg</td>
</tr>
<tr>
<td>ttcgccggcg</td>
<td>30 Pro</td>
</tr>
<tr>
<td>atg</td>
<td>35 Leu</td>
</tr>
<tr>
<td>gat</td>
<td>40 Arg</td>
</tr>
<tr>
<td>tta</td>
<td>45 Pro</td>
</tr>
<tr>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>103</td>
<td>151</td>
</tr>
<tr>
<td>199</td>
<td>247</td>
</tr>
<tr>
<td>295</td>
<td>343</td>
</tr>
<tr>
<td>391</td>
<td>439</td>
</tr>
<tr>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>
gat gat cta gtc ttg gga gat cct act gct cog agg ttt gtg ttg tgg
Asp Asp Leu Val Leu Gly Asp Pro Thr Ala Pro Arg Phe Val Leu Trp
135 140 145

aat ggg aag ctg agg cog tgt cog tgt gag gaa cta act gac ttg oct ttc
Asn Gly Lys Leu Arg Pro Val Pro Ser Lys Leu Thr Asp Leu Pro Phe
150 155 160

ttt gac ttg atg att gga ggg aag att aga gct ggg ttt ggt gcc
Phe Asp Leu Met Ser Ile Gly Gly Lys Ile Arg Ala Gly Phe Gly Ala
165 170 175

att ggt att cga cct tcg gtt gag gta cca gtt gaa gag
Ile Gly Ile Arg Pro Ser Pro Pro Gly Arg Glu Ser Val Glu Glu
180 185 190 195

ttt gta agg cgt aat ctt ggt gat gag gtt ttt gag cgc ttg att gaa
Phe Val Arg Arg Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu
200 205 210

ccc ttt tgc tca ggt gtt tat gco gga gat cct gco aaa ctc agt atg
Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ala Lys Leu Ser Met
215 220 225

aaa gca gct ttt ggg aag gtt tgg aag cta gag gag aat ggt ggg agc
Lys Ala Ala Phe Gly Lys Val Trp Lys Leu Glu Glu Asn Gly Gly Ser
230 235 240

atc att ggt ggt gct ttt aag gca att caa ggc aag aat aaa gct ccc
Ile Ile Gly Gly Ala Phe Lys Ala Ile Gln Ala Lys Asn Lys Ala Pro
245 250 255

aag aca acc cga gat cog cgt ctg cca aag cca aag ggc caa act gtt
Lys Thr Thr Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly Gln Thr Val
260 265 270 275

ggt tct ttc agg aaa gga ctc aca atg ctg cca gag gca atc tcc gca
919
Gly Ser Phe Arg Lys Gly Leu Thr Met Leu Pro Glu Ala Ile Ser Ala
280      285      290

agg ttg ggt gac aag gtg aaa gtt tct tgg aag ctc tca agt atc act 967
Arg Leu Gly Asp Lys Val Lys Val Ser Trp Lys Leu Ser Ser Ile Thr 295     300     305

aag ctc ggc agc gga gaa tat agc tta act tac gaa act ccg gag ggt 1015
Lys Leu Ala Ser Gly Glu Tyr Ser Leu Thr Tyr Glu Thr Pro Glu Gly 310     315     320

ata gtc act gta cag agc aaa agt gta gtg atg act gta cca tct cat 1063
Ile Val Thr Val Gln Ser Lys Ser Val Val Met Thr Val Pro Ser His 325     330     335

gtt gct agt agt ctc ttg cgc cct ctc tct gat tct gca gct gaa ggc 1111
Val Ala Ser Ser Leu Leu Arg Pro Leu Ser Asp Ser Ala Ala Glu Ala 340     345     350     355

ctc tca aaa ctc tac tat ccc gca gtt gca gcc gta tcc atc tca tac 1159
Leu Ser Lys Leu Tyr Tyr Pro Pro Val Ala Ala Val Ser Ile Ser Tyr 360     365     370

gcg aaa gaa gca atc cga agc gaa tgc tta ata gat ggt gaa cta aaa 1207
Ala Lys Glu Ala Ile Arg Ser Glu Cys Leu Ile Asp Gly Glu Leu Lys 375     380     385

ggg ttc ggc cag ttg cat cca cgc acg caa aaa gtt gaa act ctt gga 1255
Gly Phe Gly Gln Leu His Pro Arg Thr Gln Lys Val Glu Thr Leu Gly 390     395     400

aca ata tac agt tca tct ttt ccc aac cga gca ccc cct gga aga 1303
Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro Pro Gly Arg 405     410     415

gta ttg cta ttg aac tac atc ggt gga gct acc aac act ggg atc tta 1351
Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr Gly Ile Leu

- 65 -
tca aag tcg gaa ggt gag tta gtg gaa gca gta gat aga gac ttg agg 1399
Ser Lys Ser Glu Gly Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg
440 445 450
aag atg ctg ata aag cca agc tcg acc gat cca ctt gta ctt gga gta 1447
Lys Met Leu Ile Lys Pro Ser Ser Thr Asp Pro Leu Val Leu Gly Val
455 460 465
aaa tta tgg cct caa gcc att cct cag ttt ctg ata ggt cac att gat 1495
Lys Leu Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly His Ile Asp
470 475 480
ttg gta gac gca gcg aaa gca tcg ctc tcg tca tct ggt cat gag ggc 1543
Leu Val Asp Ala Ala Lys Ala Ser Leu Ser Ser Ser Gly His Glu Gly
485 490 495
tta ttc ttg ggt gga aat tac gtt ggc ggt gta gca ttg ggt cgg tgt 1591
Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys
500 505 510 515
gtg gaa ggt gct tat gaa act gca acc caa gtg aat gat ttc atg tca 1639
Val Glu Gly Ala Tyr Glu Thr Ala Thr Glu Val Asn Asp Phe Met Ser
520 525 530
agg tat gct tac aag taatgttaacg cagcaacgat ttgatactaa gtagtagatt 1694
Arg Tyr Ala Tyr Lys
535
ttgctgttt gactttaaga acactctgtt ttggaatat tcaagtctgt gattgagtaa 1754
attatgtat tattactaaa aaaaaaaaa 1784

<210> 20
<211> 536
<212> PRT
<213> Brassica napus

<400> 20
Met Asp Leu Ser Leu Leu Arg Pro Gln Pro Phe Leu Ser Pro Phe Ser
   1    5   10   15
Asn Pro Phe Pro Arg Ser Arg Pro Tyr Lys Pro Leu Asn Leu Arg Cys
   20   25  30
Ser Val Ser Gly Gly Ser Val Val Gly Ser Ser Thr Ile Glu Gly Gly
   35   40  45
Gly Gly Gly Lys Thr Val Thr Ala Asp Cys Val Ile Val Gly Gly Gly
   50   55  60
Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Val Thr Lys His Pro Asp
   65   70  75  80
Ala Ala Lys Asn Val Met Val Thr Glu Ala Lys Asp Arg Val Gly Gly
   85   90  95
Asn Ile Ile Thr Arg Glu Glu Gln Gly Phe Leu Trp Glu Glu Gly Pro
  100  105 110
Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr Met Val Val Asp Ser
 115  120 125
Gly Leu Lys Asp Asp Leu Val Leu Gly Asp Pro Thr Ala Pro Arg Phe
 130  135 140
Val Leu Trp Asn Gly Lys Leu Arg Pro Val Pro Ser Lys Leu Thr Asp
 145  150 155 160
Leu Pro Phe Phe Asp Leu Met Ser Ile Gly Gly Lys Ile Arg Ala Gly
 165  170 175
Phe Gly Ala Ile Gly Ile Arg Pro Ser Pro Pro Gly Arg Glu Glu Ser
180 185 190
Val Glu Glu Phe Val Arg Arg Asn Leu Gly Asp Glu Val Phe Glu Arg
195 200 205
Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ala Lys
210 215 220
Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Lys Leu Glu Glu Asn
225 230 235 240
Gly Gly Ser Ile Ile Gly Gly Ala Phe Lys Ala Ile Gln Ala Lys Asn
245 250 255
Lys Ala Pro Lys Thr Thr Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly
260 265 270
Gln Thr Val Gly Ser Phe Arg Lys Gly Leu Thr Met Leu Pro Glu Ala
275 280 285
Ile Ser Ala Arg Leu Gly Asp Lys Val Lys Val Ser Trp Lys Leu Ser
290 295 300
Ser Ile Thr Lys Leu Ala Ser Gly Glu Tyr Ser Leu Thr Tyr Glu Thr
305 310 315 320
Pro Glu Gly Ile Val Thr Val Gln Ser Lys Ser Val Val Met Thr Val
325 330 335
Pro Ser His Val Ala Ser Ser Leu Leu Arg Pro Leu Ser Asp Ser Ala
340 345 350
Ala Glu Ala Leu Ser Lys Leu Tyr Tyr Pro Pro Val Ala Ala Val Ser
355 360 365
Ile Ser Tyr Ala Lys Glu Ala Ile Arg Ser Glu Cys Leu Ile Asp Gly
Glu Leu Lys Gly Phe Gly Gln Leu His Pro Arg Thr Gln Lys Val Glu
Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
Pro Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr
Gly Ile Leu Ser Lys Ser Glu Gly Leu Val Glu Ala Val Asp Arg
Asp Leu Arg Lys Met Leu Ile Lys Pro Ser Ser Thr Asp Pro Leu Val
Leu Gly Val Lys Leu Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly
His Ile Asp Leu Val Asp Ala Ala Lys Ala Ser Leu Ser Ser Ser Gly
His Glu Gly Leu Phe Leu Gly Gln Tyr Val Ala Gly Val Ala Leu
Gly Arg Cys Val Glu Gly Ala Tyr Glu Thr Ala Thr Gln Val Asn Asp
Phe Met Ser Arg Tyr Ala Tyr Lys

<210> 21
<211> 1224
<212> DNA
<213> Oryza sativa
<220>
<221> CDS
<222> (1) .. (936)
<223> rice protox-1 partial coding sequence

<400> 21

```
<table>
<thead>
<tr>
<th>Arg</th>
<th>Ala</th>
<th>Leu</th>
<th>Lys</th>
<th>Ala</th>
<th>Ala</th>
<th>Phe</th>
<th>Gly</th>
<th>Lys</th>
<th>Val</th>
<th>Trp</th>
<th>Arg</th>
<th>Leu</th>
<th>Glu</th>
<th>Asp</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

gga ggt agg att att ggt gga acc atc aag aca atc cag gag agg ggg 96

```
<table>
<thead>
<tr>
<th>Gly</th>
<th>Gly</th>
<th>Ser</th>
<th>Ile</th>
<th>Ile</th>
<th>Gly</th>
<th>Gly</th>
<th>Thr</th>
<th>Ile</th>
<th>Lys</th>
<th>Thr</th>
<th>Ile</th>
<th>Gln</th>
<th>Glu</th>
<th>Arg</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>25</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```
<table>
<thead>
<tr>
<th>Lys</th>
<th>Asn</th>
<th>Pro</th>
<th>Lys</th>
<th>Pro</th>
<th>Pro</th>
<th>Arg</th>
<th>Asp</th>
<th>Pro</th>
<th>Arg</th>
<th>Leu</th>
<th>Pro</th>
<th>Thr</th>
<th>Pro</th>
<th>Lys</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```
cag aca gtt gca tct ttc agg aag ggt ctc act atg ctc cag gat gct 192
```

```
<table>
<thead>
<tr>
<th>Gln</th>
<th>Thr</th>
<th>Val</th>
<th>Ala</th>
<th>Ser</th>
<th>Phe</th>
<th>Arg</th>
<th>Lys</th>
<th>Gly</th>
<th>Leu</th>
<th>Thr</th>
<th>Met</th>
<th>Leu</th>
<th>Pro</th>
<th>Asp</th>
<th>Ala</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```
att aca tct agg ttg ggt agc aaa gtc aaa ctt tca tgg aag ttg aca 240
```

```
<table>
<thead>
<tr>
<th>Ile</th>
<th>Thr</th>
<th>Ser</th>
<th>Arg</th>
<th>Leu</th>
<th>Gly</th>
<th>Ser</th>
<th>Iys</th>
<th>Val</th>
<th>Lys</th>
<th>Leu</th>
<th>Ser</th>
<th>Trp</th>
<th>Lys</th>
<th>Leu</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```
agc att aca aag tca gac aac aaa gga tat gca tta gtg tat gaa aca 288
```

```
<table>
<thead>
<tr>
<th>Ser</th>
<th>Ile</th>
<th>Thr</th>
<th>Iys</th>
<th>Ser</th>
<th>Asp</th>
<th>Asn</th>
<th>Lys</th>
<th>Gly</th>
<th>Tyr</th>
<th>Ala</th>
<th>Leu</th>
<th>Val</th>
<th>Tyr</th>
<th>Glu</th>
<th>Thr</th>
</tr>
</thead>
<tbody>
<tr>
<td>85</td>
<td>90</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```
cca gaa ggg gtt gtc tgg gtt gta gtt gca gct aaa act gtt gtc atg acc atc 336
```

```
<table>
<thead>
<tr>
<th>Pro</th>
<th>Glu</th>
<th>Gly</th>
<th>Val</th>
<th>Val</th>
<th>Ser</th>
<th>Val</th>
<th>Gln</th>
<th>Ala</th>
<th>Lys</th>
<th>Thr</th>
<th>Val</th>
<th>Val</th>
<th>Met</th>
<th>Thr</th>
<th>Ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>105</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```
cca tca tat gtt gct agt gat atc ttg cgg cca ctt tca agt gat gca 384
```

```
| Pro | Ser | Tyr | Val | Ala | Ser | Asp | Ile | Leu | Arg | Pro | Leu | Ser | Ser | Asp | Ala |
```

- 70 -
gca gag gct tgct ca ata ttg tat tat cca cca gtt gct gct gta act
 Ala Asp Ala Leu Ser Ile Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr
 130  135  140

gtt tca tat cca aaa gaa gca att aga aaa gaa tgg tta att gag gga
 Val Ser Tyr Pro Lys Glu Ala Ile Arg Ile Glu Cys Leu Ile Asp Gly
 145  150  155  160

gag ctc cag ggt ttc ggc cag ctc cat ccc gct agt cag gga gtt gag
 Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu
 165  170  175

act tta gga aca ata tat agc tca tca ctc ttt cca aat cgt gct cca
 Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
 180  185  190

gct gga agg gtt ttc ctt ctg aac tac ata gga gtt tct aca aat aca
 Ala Gly Arg Val Leu Leu Leu Leu Asn Tyr Ile Gly Gly Ser Thr Asn Thr
 195  200  205

ggg att gtt tcc aag act gaa agt gag ctg gta gaa gca gtt gag cgt
 Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg
 210  215  220

gac ctc agg aag atg ctg ata aat cct aga gca gtt gct gac cct ttg gtc
 Asp Leu Arg Lys Met Leu Ile Asn Pro Arg Ala Val Asp Pro Leu Val
 225  230  235  240

ctt ggc gtc cgg gta cgg cca cca gcc ata cca cag ttc ccc att ggc
 Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly
 245  250  255

cat ctt gat cat ctt gac ggt gca aac tct gcc ctg gcc aac ggt ggg
 His Leu Asp His Leu Glu Ala Ala Lys Ser Ala Leu Gly Lys Gly
 260  265  270
tat gat gga ttg ttc ctc gga ggg aac tat gtt gca gga gtt gcc ctg 864
Tyr Asp Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
  275  280  285

ggc cga tgc gtt gaa ggt gca tat gag agt gcc tca caa ata tct gac 912
Gly Arg Cys Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp
  290  295  300

tac ttg acc aag tac gcc tac aag tgatcaaggt tgtgctgctc ctttggcac 966
Tyr Leu Thr Lys Tyr Ala Tyr Lys
  305  310

atagatgtga ggtttctgc agcaaaaatt tcatggtgcat ctttttatcc tgattctaat 1026
tagttgaat ttagaattgt agaggaatgt tccatttgca gttcataata gtgtgttcaga 1086
tttcagccat tcaatttgtg cacccattta ctatatgtag tatgatcttg taagtactac 1146
taagaacaaa tcaattatat tttcctgcaaatgtgacatctt aatogtcaagc aatccagtt 1206
actagtaaaa aaaaaaaaa 1224

<210> 22
<211> 312
<212> PRT
<213> Oryza sativa

<400> 22
Arg Ala Leu Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Asp Thr
  1   5  10  15
Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly
  20  25  30
Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Thr Pro Lys Gly
<table>
<thead>
<tr>
<th></th>
<th>35</th>
<th>40</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Thr Met Leu Pro Asp Ala</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ile Thr Ser Arg Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr</td>
<td>65</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Ser Ile Thr Lys Ser Asp Asn Lys Gly Tyr Ala Leu Val Tyr Glu Thr</td>
<td>85</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Pro Glu Gly Val Val Ser Val Gln Ala Lys Thr Val Val Met Thr Ile</td>
<td>100</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Ser Asp Ala</td>
<td>115</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>Ala Asp Ala Leu Ser Ile Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr</td>
<td>130</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly</td>
<td>145</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu</td>
<td>165</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro</td>
<td>180</td>
<td>185</td>
</tr>
<tr>
<td></td>
<td>Ala Gly Arg Val Leu Leu Leu Tyr Ile Gly Gly Ser Thr Asn Thr</td>
<td>195</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg</td>
<td>210</td>
<td>215</td>
</tr>
<tr>
<td></td>
<td>Asp Leu Arg Lys Met Leu Ile Asn Pro Arg Ala Val Asp Pro Leu Val</td>
<td>225</td>
<td>230</td>
</tr>
</tbody>
</table>

- 73 -
Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Ile Gly
245  250  255
His Leu Asp His Leu Glu Ala Ala Lys Ser Ala Leu Gly Lys Gly Gly
260  265  270
Tyr Asp Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
275  280  285
Gly Arg Cys Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Ser Asp
290  295  300
Tyr Leu Thr Iys Tyr Ala Tyr Lys
305  310

<210> 23
<211> 1590
<212> DNA
<213> Sorghum bicolor

<220>
<221> CDS
<222> (1)..(1320)
<223> sorghum protox-1 partial coding sequence

<400> 23
  tcc acc gtc gag cgc ccc gag gaa ggg tac ctc tgg gag gag ggt ccc 48
  Ser Thr Val Glu Arg Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro
  1  5  10  15

  aac agc ttc cag cca tcc gac ccc gtt ctc tcc atg gcc gtt gac agc 96
  Asn Ser Phe Glu Pro Ser Asp Pro Val Leu Ser Met Ala Val Asp Ser
  20  25  30

  ggg ctg aag gat gac ctg gtt ttt ggg gac ccc aac gog cca ogg ttc 144

  -74 -
Gly Leu Lys Asp Asp Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe
35 40 45

gtg ctg tgg gag ggg aag ctg agg ccc gtg cca tcc aag ccc gcc gac 192
Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp
50 55 60

ctc ccg ttc ttc gat ctc atg agc atc cct ggc aag ctc agg gcc ggt 240
Leu Pro Phe Phe Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly
65 70 75 80

c tc gc g g c ct g c gc a t c c gc g c gc g gc g g g g g t c a 288
Leu Gly Ala Leu Gly Ile Arg Pro Pro Ala Pro Gly Arg Glu Glu Ser
85 90 95

gtg gag gag ttt gtg cgc cgc aac ctc ggt gct gag gtc ttt gag ggc 336
Val Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg
100 105 110

ctc att gag cct ttc tgc tca ggt gtc tat gct ggc gat cct tcc aag 384
Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys
115 120 125

ctc agt atg aag gct gca ttt ggg aag gtg tgg cgg tta gaa gaa gct 432
Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Ala
130 135 140

gga ggt agt att att ggt gga acc atc aag a cg att cag gag agg ggc 480
Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly
145 150 155 160

aag aat cca aaa cca ccg agg gat ccc cgc ctt ccg aag cca aaa ggg 528
Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly
165 170 175

cag aca grr gca tct ttc agg aag ggt ctt gcc atg ctt cca aat gcc 576
Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Ala Met Leu Pro Asn Ala
atc aca tcc agc ttg ggt agt aaa gtc aca cta tca tgg aat ctc acg
Ile Thr Ser Ser Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr
180 185 190

Ile Thr Ser Ser Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr
195 200 205

agc atg aca aca tca gat ggc aag ggg tat gtt ttg gag tat gaa aca
Ser Met Thr Lys Ser Asp Gly Lys Gly Tyr Val Leu Glu Tyr Glu Thr
210 215 220

Pro Glu Gly Val Val Leu Val Gln Ala Lys Ser Val Ile Met Thr Ile
225 230 235 240

cca gaa ggg gtt ttg ttg cag gct aca aat gtt atc atg acc att
Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Gly Asp Ala
245 250 255

gca gat gtt cta tca aga ttc tat tat cca cca gtt gct gct gta acg
Ala Asp Val Leu Ser Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr
260 265 270

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

gtt tgc tat cca aag gaa gca att aga aca gaa tgc tta att gat ggg
Val Ser Tyr Pro Lys Glu Ala Ile Arg Lys Glu Cys Leu Ile Asp Gly
290 295 300

gaa ctc cag ggt ttt ggc cag ttg cat cca cgt aat cca aat gta gtt gag
Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu
290 295 300

aca tta gga aca ata tac agc tca tca ctc ttt cca aat cgt gct cct
Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
305 310 315 320

gct ggt agg gtt tta ctt cta aac tac ata gga ggt gct aca aac aca
Ala Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr
325 330 335

- 76 -
Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg
340 345 350

Asp Leu Arg Lys Met Leu Ile Asn Pro Thr Ala Val Asp Pro Leu Val
355 360 365

Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly
370 375 380

His Leu Asp Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Gln Gly Gly
385 390 395 400

Tyr Asn Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu
405 410 415

Gly Arg Cys Ile Glu Gly Ala Tyr Glu Ser Ala Ala Gln Ile Tyr Asp
420 425 430

Phe Leu Thr Lys Tyr Ala Tyr Lys
435 440

tttatgagag ctgtgcagac ttatgtcaag ttttactgta aagtcacac ggtgttagttg 1410

catcttatt ttgttatttc acctcttctgt ttccccccct ctaatatggagttttagaatttttttttt 1470

tagttatagtt ggttgtcactgt cctactaaaaa aatgttatgtttaa ttttctttttgctttgccttttg 1530

ttttatgagag ctgtgcagac ttatgtcaag ttttactgta aagtcacac ggtgttagttg 1590
<210> 24
<211> 440
<212> PRT
<213> Sorghum bicolor

<400> 24
Ser Thr Val Glu Arg Pro Glu Glu Gly Tyr Leu Trp Glu Glu Gly Pro
1      5     10      15
Asn Ser Phe Gln Pro Ser Asp Pro Val Leu Ser Met Ala Val Asp Ser
20     25     30
Gly Leu Lys Asp Leu Val Phe Gly Asp Pro Asn Ala Pro Arg Phe
35     40     45
Val Leu Trp Glu Gly Lys Leu Arg Pro Val Pro Ser Lys Pro Ala Asp
50     55     60
Leu Pro Phe Phe Asp Leu Met Ser Ile Pro Gly Lys Leu Arg Ala Gly
65     70     75     80
Leu Gly Ala Leu Gly Ile Arg Pro Pro Ala Pro Gly Arg Glu Glu Ser
85     90     95
Val Glu Glu Phe Val Arg Arg Asn Leu Gly Ala Glu Val Phe Glu Arg
100    105    110
Leu Ile Glu Pro Phe Cys Ser Gly Val Tyr Ala Gly Asp Pro Ser Lys
115    120    125
Leu Ser Met Lys Ala Ala Phe Gly Lys Val Trp Arg Leu Glu Glu Ala
130    135    140
Gly Gly Ser Ile Ile Gly Gly Thr Ile Lys Thr Ile Gln Glu Arg Gly
145    150    155    160
Lys Asn Pro Lys Pro Pro Arg Asp Pro Arg Leu Pro Lys Pro Lys Gly
165 170 175

Gln Thr Val Ala Ser Phe Arg Lys Gly Leu Ala Met Leu Pro Asn Ala
180 185 190

Ile Thr Ser Ser Leu Gly Ser Lys Val Lys Leu Ser Trp Lys Leu Thr
195 200 205

Ser Met Thr Iys Ser Asp Gly Lys Gly Tyr Val Leu Glu Tyr Glu Thr
210 215 220

Pro Glu Gly Val Val Leu Val Gln Ala Lys Ser Val Ile Met Thr Ile
225 230 235 240

Pro Ser Tyr Val Ala Ser Asp Ile Leu Arg Pro Leu Ser Gly Asp Ala
245 250 255

Ala Asp Val Leu Ser Arg Phe Tyr Tyr Pro Pro Val Ala Ala Val Thr
260 265 270

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

Glu Leu Gln Gly Phe Gly Gln Leu His Pro Arg Ser Gln Gly Val Glu
290 295 300

Thr Leu Gly Thr Ile Tyr Ser Ser Ser Leu Phe Pro Asn Arg Ala Pro
305 310 315 320

Ala Gly Arg Val Leu Leu Leu Asn Tyr Ile Gly Gly Ala Thr Asn Thr
325 330 335

Gly Ile Val Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg
340 345 350

Asp Leu Arg Lys Met Leu Ile Asn Pro Thr Ala Val Asp Pro Leu Val

- 79 -
Leu Gly Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly

355 360 365

His Leu Asp Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Gln Gly Gly

385 390 395 400

Tyr Asn Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu

405 410 415

Gly Arg Cys Ile Glu Gly Ala Tyr Glu Ser Ala Ala Gln Ile Tyr Asp

420 425 430

Phe Leu Thr Lys Tyr Ala Tyr Lys

435 440

<210> 25
<211> 93
<212> DNA
<213> Zea mays

<220>
<221> misc_feature
<222> (1)...(93)
<223> maize protox-1 intron sequence

<400> 25
gtaogctcct cgctgggcc gcagcgtcct cttctcagac tcatgocgcag ccatggaatt 60
gagatgctga atgatattta taogcgcgocg cag

93
Beta vulgaris

misc_feature
(2601)..<(2606)
SalI site

misc_feature
(1)..<(538)
partial cDNA of sugar beet protox-1

sugar beet protox-1 promoter region (partial
sequence of the ~3 kb PstI-SalI fragment subcloned
from pWDC-20)

ctgcaggggg agggaagag agacgcccgc ggtgagggag gggagacccg gacggtgagg 60

gaggggagaa cgcccgggtg aaggagggaa gaacgcgatg gtgagggagg ggagaacgccg 120
acgcgcaggg gaggggggata actcgagcct gcggaggagt gaggagggagc acggggagcgc 180
gcaggggaggg ggggaacccgt gcggggagag ggagacccgg ggggcccgcaca aggtgggttt 240
actgggggtat ggagggccgg cgtagagfaat agtaacagag gaggagttg gggtgtcagg 300
gtggagaag ggtaagaaag gggaagaag agaatttaca ttatcttaac caaaccaccac 360
tctaatctca aggggttttc tttcccttccc tctcctctccc ctttctttgt tctattcctt 420
ttaccctgttt gcacaccac gcccccttat tattgagccgg aggaagtattg tagagatgtg 480
cacaaaaacta cttaaggtgga taactttatat atatactgtgg gattaattgga attaagtgcc 540
cacaataatga ctataaatta cttcgttaatc ttttaggaact atggttgtgca cyaaaataaca 600
taaaactgtg tatatattgg ctttatttagt gtaattgacct cataaatata ttcttaacat 660
aatcgtggtta tgtatgtggttt ttataaaca agaggataaggt ttacaccaaat gttatcttta 720
ttagatgtga gtttagaatca cttcgggact tttgaagatgt atgacacatt ttattatagc 780
ttttatatga tgtctttgtg gtttttatga tagtatttagg ttttaaggccaa gttggaagta 840
tatgatggag aagtaacagta tataggagac aatgtgtttg cttgccccata tggatgaaa 900
gataatagtt acaacgcacact gagcaatgac ctctctcttag ttgtaatattttttctcga 960
cgtgatgaaa gtacacaacaa gatattaggt cttcaagcttc caagataaag agattgtatg 1020
aattttggtg ttgtttccac atcatttttt tacaoctggag acacaaactaa accaatgatg 1080
agttttgtgg aatcgagattt gccoccaagt ctatattttcag ctagccaaagc agtcgtaaacc 1140
atgttagtca aacaccacata gctacaatgt ttagggatgt tgaacaaaaa atttggttat 1200
tctttgggta ccatatagttg agttgttgac tatggattaa acaaaatcac tatataaagt 1260
c tgtgatgag aagcatcgc gcaccagaca ccatgttact ttgtgtgatt tacaagtttt 1320
attagatgtg tttgtagggaaaa tttgtggag ggtgagctgt atgatttttaa ttccttttgtg 1380
ggtgttcac atgcacttctg ttatgtgagac atcttcagct tatatttttaa ggccggttagt 1440
gagatgattttttttaa acaacttctcg atttccatgt aatattaaaaa ggtgttttgtat 1500
aaatacatgt taagataccc aagaaagagc aacctctaaaaa caaataaaaa aaattaagct 1560
gcttaaatcat ttttcacagtt actttttact ttaacacca cttattactgt aatctatatc 1620
ctgtaaggaat gcattttcac gtcatacat gcgaatcaag acctctctca ttgaagggaga 1680
taatattgct ctctaaaccc ccgtaaaga catttttagc atccagagaa atttcgatttc 1740
agttaaatt gcataataaa ccagagaaac aaatccagat gttagtcagt ccagctcatc 1800
aggtcaatgc ctgagagtctt aaaaagatcc gtatcctttaa gcataagtag gtatggaggt 1860
gagttacaaa ggtaaggtac cggtacgca ccaacctccac caaacaagta tgggttagaag 1920
atatcatgtta tcgtttatctt agagttactt ttataaaaaa ctttttaact agaacaagtt 1980
gtttcatcttt gataaaggt taattagaat tccccagcaa ccaagaaggg gatatagagg 2040
ataagggagg cgagagagcg agagagagat gaaatacaag gcggttatcag actgcattcc 2100
acagacacag tgcagcctcat tgcagcagcag cgggcattac aggggcaatt gtatcatgttt 2160
gtcaatccta tgcgattttta tttggaagagct aggttatatt tcacataaga agagggagat 2220
gagcagtcatg tgcagcagcag gtcaggtctc aaagtcagcag gttaaagaag cagagcagg 2280
atcaggatat gcgacgggag gattgctaga ctgcgtaatc gttggaggtg gaattacgag 2340
gctttgcctcg cgcgagctct tttgtacaa gcgcagctcat ttatccccca attttatagt 2400
gacagaggcc aagacagagc tgccgggcaaa cactgcctctg tgggagggcg atggctatat 2460
cggaggggg gagcccaata gttccaggcct ttccagagcg gttgcctacca tggcggtaat 2520
tggaggtctc ttctctctct cattttcat aatcataatt caattcaact caattctcaa gttggaatgt 2580
gggatgtggc atgtgctgtag gtcgac 2606

<210> 27
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
   oligonucleotide

<220>
<221> misc_feature
<222> (4) ...(9)
<223> EcoRI restriction site

<400> 27
gggaattca tacttatatta tcattagaa g 31

<210> 28
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
   oligonucleotide

<220>
<221> misc_feature
<222> (4) ...(9)
<223> XbaI restriction site

<400> 28
gcgtctagaa agaactaaat actatattttc ac 32

<210> 29
<211> 30
<212> DNA
Artificial Sequence

Description of Artificial Sequence:
oligonucleotide

misc_feature
(4) ...(9)
NcoI restriction site

29
gogccatggt aatgaaga aagaactaaa

30

30
30
DNA
Artificial Sequence

Description of Artificial Sequence:
oligonucleotide

misc_feature
(4) ...(9)
XbaI restriction site

30
gogtctagat caaccgaat tcaattaagy
Description of Artificial Sequence:
oligonucleotide

misc_feature
(4) .. (9)
HindIII restriction site

cgcaagcttc aatggaagca atgataa 27

gggagtccct gatgattaaa taaccaaga ttttac 36

ttaa

Artificial Sequence

Description of Artificial Sequence:
oligonucleotide

Artificial Sequence
catggtaaaa tcttggttta ttaatcttc agggactccc

<210> 34
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
    oligonucleotide

<220>
<221> misc_feature
<222> (5)...(10)
<223> NcoI restriction site/ ATG start codon

<400> 34
    gggaccatgg attgtgtgat tgctggcyga gg

<210> 35
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
    oligonucleotide

<400> 35
    ctcogctctc cagcttagtg atac
sugar cane

CDS

(3)...(305)

sugar cane protox-1 partial coding sequence

<400> 36

tt tcc aag act gaa agt gag ctg gta gaa gca gtt gac cgt gac ctc 47
    Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu
    1      5      10      15

cgg aad atg ctt ata aat cct cca gca gtg gac cct tta gtc ctt ggt 95
    Arg Lys Met Leu Ile Asn Pro Thr Ala Val Asp Pro Leu Val Leu Gly
    20     25      30

gtc cga gtt tgt cca caa gcc ata cct cag ttc ctg gta gga cat ctt 143
    Val Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu
    35     40      45

gat ctt ctg gag gcc gca aaa tct gcc ctg gac cga ggt gcc tac gat 191
    Asp Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Arg Gly Gly Tyr Asp
    50     55      60

ggg ctg ttc cta gga ggg aac tat gtt gca gga gtt gcc cta ggc aga 239
    Gly Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg
    65     70      75

tgc gtt gag gcc gcg tat gag agt gcc tct cga ata tat gac ttc ttg 287
    Cys Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Tyr Asp Phe Leu
    80     85      90      95

acc aag tat gcc tac aag tgaatgaaga agtggagtgc tgcttgtttaa 335
    Thr Lys Tyr Ala Tyr Lys
    100

ttgttatgtt gcatacgatga ggtgagccca ggaatatgaa aaggctttacg agtatatatctc 395
atctcttttt gtaaatttgc acttcttgtt ttttcctgtc agtaattagt tagatcttttag 455
tttctgtagga gattgtttctg ttcactgcc tacaagaaga ttttatatat gcattgttttt 515
atgagagctg tgcagacatga tgtagcttttt ttctgtaagtt accaacaaaa tcaaatata 575
tttctgtaagagctacagaaagttgcactgctgctttt gatgaaaaaa aaaaaaaa 633

<210> 37
<211> 101
<212> PRT
<213> sugar cane

<400> 37
Ser Lys Thr Glu Ser Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg
1 5 10 15
Lys Met Leu Ile Asn Pro Thr Ala Val Asp Pro Leu Val Leu Gly Val
20 25 30
Arg Val Trp Pro Gln Ala Ile Pro Gln Phe Leu Val Gly His Leu Asp
35 40 45
Leu Leu Glu Ala Ala Lys Ser Ala Leu Asp Arg Gly Gly Tyr Asp Gly
50 55 60
Leu Phe Leu Gly Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys
65 70 75 80
Val Glu Gly Ala Tyr Glu Ser Ala Ser Gln Ile Tyr Asp Phe Leu Thr
85 90 95
Lys Tyr Ala Tyr Lys
100

- 89 -
<210> 38
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: sub-sequence 1

<220>
<221> misc_feature
<222> (3)
<223> Xaa = amino acid other than Arg

<400> 38
Ala Pro Xaa Phe
 1

<210> 39
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: sub-sequence 8

<220>
<221> misc_feature
<222> (5)
<223> Xaa = amino acid other than Ala or Ser

<400> 39
Tyr Ile Gly Gly Xaa
 1  5
<210> 40
<211> 4
<212> PRF
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: sub-sequence
    12

<220>
<221> misc_feature
<222> (4)
<223> Xaa = amino acid other than Thr

<400> 40
Ile Gly Gly Xaa
    1

<210> 41
<211> 5
<212> PRF
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: sub-sequence
    13

<220>
<221> misc_feature
<222> (3)
<223> Xaa = any amino acid

<220>
<221> misc_feature
<222> (5)
<223> Xaa = amino acid other than Ser or Thr
<400> 41
Ser Trp Xaa Leu Xaa
  1   5

<210> 42
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: sub-sequence
  15

<220>
<221> misc_feature
<222> (2)
<223> Xaa = amino acid other than Tyr, His, Phe or Val

<220>
<221> misc_feature
<222> (3)
<223> Xaa = any amino acid

<400> 42
Gly Xaa Xaa Gly Leu
  1   5

<210> 43
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: sub-sequence
17

<220>
<221> misc_feature
<222> (3)
<223> Xaa = amino acid other than Ala or Ser

<400> 43
Tyr Val Xaa Gly
1

<210> 44
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: sub-sequence
18

<220>
<221> misc_feature
<222> (3)
<223> Xaa = amino acid other than Ala

<400> 44
Lys Ala Xaa Phe
1
## INTERNATIONAL SEARCH REPORT

**A. CLASSIFICATION OF SUBJECT MATTER**

<table>
<thead>
<tr>
<th>IPC</th>
<th>C12N15/82</th>
</tr>
</thead>
<tbody>
<tr>
<td>C12N9/02</td>
<td>A01H5/00</td>
</tr>
</tbody>
</table>

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

<table>
<thead>
<tr>
<th>IPC</th>
<th>C12N</th>
</tr>
</thead>
</table>

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>WO 97 32011 A (CIBA GEIGY AG ;VOLRATH SANDRA L (US); JOHNSON MARIE A (US); POTTER) 4 September 1997 (1997-09-04)</td>
<td>1-3, 7-16, 19, 22-25, 27,31-54</td>
</tr>
<tr>
<td></td>
<td>the whole document</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>WO 95 34659 A (CIBA GEIGY AG ;WARD ERIC RUSSELL (CH); VOLRATH SANDRA (US)) 21 December 1995 (1995-12-21) cited in the application</td>
<td>1-3, 7, 16, 22-24, 31-40, 42-54</td>
</tr>
<tr>
<td></td>
<td>the whole document</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>US 6 084 155 A (WARD ERIC R ET AL) 4 July 2000 (2000-07-04)</td>
<td>1-3, 6-19, 22-54</td>
</tr>
<tr>
<td></td>
<td>the whole document</td>
<td></td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of box C. Patent family members are listed in annex.

* Special categories of cited documents:
  *A* document defining the general state of the art which is not considered to be of particular relevance
  *E* earlier document but published on or after the international filing date
  *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or special reason (as specified)
  *O* document referring to an oral disclosure, use, exhibition or other means
  *P* document published prior to the international filing date but later than the priority date claimed
  *T* later document published after the international filing date of priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
  *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
  *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
  *Z* document member of the same patent family

**Date of the actual completion of the international search**

3 November 2000

**Date of mailing of the international search report**

10/11/2000

**Name and mailing address of the ISA**

European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epos nl, Fax: (+31-70) 340-3016

**Authorized officer**

Oderwald, H
<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family member(s)</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td>WO 9732011 A</td>
<td>04-09-1997</td>
<td>AU 1984697 A</td>
<td>16-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 724838 B</td>
<td>28-09-2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 2065497 A</td>
<td>16-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR 9707679 A</td>
<td>27-07-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR 9707783 A</td>
<td>27-07-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2247074 A</td>
<td>04-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2247797 A</td>
<td>04-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 1212724 A</td>
<td>31-03-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 1212725 A</td>
<td>31-03-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CZ 9802726 A</td>
<td>16-12-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CZ 9802727 A</td>
<td>11-11-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0883682 A</td>
<td>16-12-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0885305 A</td>
<td>23-12-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HU 9900623 A</td>
<td>28-06-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HU 9901044 A</td>
<td>28-07-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2000506724 T</td>
<td>06-06-2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2000506011 T</td>
<td>23-05-2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PL 328617 A</td>
<td>01-02-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PL 328651 A</td>
<td>15-02-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 6023012 A</td>
<td>08-02-2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 9732028 A</td>
<td>04-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 6084155 A</td>
<td>04-07-2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 6018105 A</td>
<td>25-01-2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 5939602 A</td>
<td>17-08-1999</td>
</tr>
<tr>
<td>WO 9534659 A</td>
<td>21-12-1995</td>
<td>AU 706763 B</td>
<td>24-06-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BG 2453895 A</td>
<td>05-01-1996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR 101115 A</td>
<td>31-10-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR 9508030 A</td>
<td>16-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 1150820 A</td>
<td>28-05-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CZ 9063674 A</td>
<td>18-03-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0795599 A</td>
<td>30-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FI 964958 A</td>
<td>11-12-1996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HU 76353 A</td>
<td>28-08-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 1050052 T</td>
<td>10-03-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PL 317759 A</td>
<td>28-04-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SK 161096 A</td>
<td>28-07-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 5767373 A</td>
<td>16-06-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 5939602 A</td>
<td>17-08-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 5767373 A</td>
<td>16-06-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 1984697 A</td>
<td>16-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 724838 B</td>
<td>28-09-2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU 2065497 A</td>
<td>16-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR 9707769 A</td>
<td>27-07-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR 9707783 A</td>
<td>27-07-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2247074 A</td>
<td>04-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2247797 A</td>
<td>04-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 1212724 A</td>
<td>31-03-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 1212725 A</td>
<td>31-03-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CZ 9802726 A</td>
<td>16-12-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CZ 9802727 A</td>
<td>11-11-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0883682 A</td>
<td>16-12-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 0885305 A</td>
<td>23-12-1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HU 9900623 A</td>
<td>28-06-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HU 9901044 A</td>
<td>28-07-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2000506724 T</td>
<td>06-06-2000</td>
</tr>
<tr>
<td>Patent document cited in search report</td>
<td>Publication date</td>
<td>Patent family member(s)</td>
<td>Publication date</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-----------------</td>
<td>-------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PL 328617 A</td>
<td>01-02-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PL 328651 A</td>
<td>15-02-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 9732011 A</td>
<td>04-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 9732028 A</td>
<td>04-09-1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 6018105 A</td>
<td>25-01-2000</td>
</tr>
</tbody>
</table>