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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 March 2001 (08.03.2001)

PCT

(10) International Publication Number
WO 01/16144 A2

(51) International Patent Classification[7]: C07H

(21) International Application Number: PCT/IL00/00511

(22) International Filing Date: 27 August 2000 (27.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
131633 29 August 1999 (29.08.1999) IL

(71) Applicant *(for all designated States except US)*: **YEDA RESEARCH AND DEVELOPMENT CO. LTD.** [IL/IL]; The Weizmann Institute Of Science, P.O. Box 95, 76100 Rehovot (IL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **FROMM, Hillel** [IL/IL]; 10/11 Bustenai Street, 76289 Rehovot (IL). **ARAZI, Tzahi** [IL/IL]; 34 Hamagal Street, 55402 Kiryat Ono (IL).

(74) Agent: **G. E. EHRLICH (1995) LTD.**; Gibor-Sport Building, 17th Floor, 28 Bezalel Street, 52521 Ramat Gan (IL).

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

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

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

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

WO 01/16144 A2

(54) Title: PLANT DERIVED CALCIUM-DEPENDENT CALMODULIN-BINDING POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAME

(57) Abstract: A method of modulating a tolerance of a plant to a metal cation is disclosed. The method is effected by overexpressing in the plant a recombinant protein which modulates uptake and concentration of the metal cation within the plant cells.

PLANT DERIVED CALCIUM-DEPENDENT CALMODULIN-BINDING POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAME

FIELD AND BACKGROUND OF THE INVENTION

The present invention relates to polynucleotides encoding calcium-dependent calmodulin-binding proteins from plants, to polypeptides which are the translation products of these polynucleotides, to expression and antisense vectors containing the polynucleotides or portions thereof, to cultured cells and transgenic or viral infected plants expressing the polynucleotides or portions thereof and to methods of using such plants to assist agriculture and/or phytoremediation by providing plants characterized by resistivity to toxic levels of metal ions or plants capable of accumulating high concentration of a metal ions.

Specifically, the present invention relates to several polynucleotides encoding calcium-dependent calmodulin-binding proteins isolated from tobacco and other plants and to functional homologs thereof.

Further specifically, the present invention relates to the use of these polynucleotides, functional portions thereof, functionally similar homologs, and/or portions of the functionally similar homologs to confer useful properties on plants expressing them. In particular, these useful properties include tolerance to heavy metal and other metal cations, such as, but not limited to Ni^{2+}, and increased uptake of heavy and other metal cations, such as, but not limited to, Pb^{2+}. Tolerance to Ni^{2+} cations is useful because it allows plants to grow in an environment which would normally be considered toxic to sustain plant growth. Increased uptake of other heavy metal cations, such as Pb^{2+}, has utility as a means of phytoremediation.

Citation or identification of any reference in this application shall not be construed as an admission that such reference is available as prior art to the present invention.

Toxic metal pollution, mainly from combustion of fossil fuels, from manufactured waste and the indiscriminate use of chemicals, has accelerated dramatically in recent years. The direct emission of trace metals such as Ni into the air, water and soil is by now measured in millions of tons per annum (Nriagu *et al.*, 1988). The levels of most toxic metals emitted into the atmosphere because of humans are several times higher than these released by natural sources. For example, the anthropogenic emission of Pb into the atmosphere is more than 300-fold that of its natural emission

(Ayres, 1992). Increased acidification of soils and fresh waters, which leads to enhanced solubility of toxic metals, further increases toxic metal levels (Rengel, 1996).

Plants have developed unique protective strategies to deal with this environmental danger (Braam *et al.*, 1990; Bowler *et al.*, 1994; Knight *et al.*, 1998; Snedden *et al.*, 1998). By understanding these strategies, improved plant tolerance can be engineered and implementation of biotechnologies aimed at cleaning the environment can be undertaken.

Genes encoding proteins that are involved in the transport of metal ions are crucial in this endeavor. Specifically, such genes form potential targets for improving plant tolerance to toxic metals or increasing the accumulation of metal ions in plants as a means of phytoremediation (Raskin, 1996).

There is thus a widely recognized need for, and it would be highly advantageous to have, a means of regulating uptake and storage of heavy and other metal ions by plants which would, in some cases facilitate cultivation of plants in an environment containing high concentrations of metal ions, and in other cases facilitate phytoremediation of an area for future use as a cultivation site. The present invention represents a significant step forward towards these goals.

SUMMARY OF THE INVENTION

According to one aspect of the present invention there is provided an isolated nucleic acid comprising a polynucleotide at least 60 % identical with any of SEQ ID NOs:1, 3, 5, 7, 9, 11 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3), in particular, SEQ ID NOs:1, 3, 5, 7, 9, 11 or portions thereof.

According to another aspect of the present invention there is provided an isolated nucleic acid comprising a polynucleotide encoding a polypeptide being at least 60 % homologous (identical + similar) with any of SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

According to yet another aspect of the present invention there is provided an isolated nucleic acid comprising a plant derived polynucleotide

encoding a transmembrane polypeptide having a cation channel activity when assembled in a plasmatic membrane of a plant cell, the polypeptide having overlapping cyclic nucleotide-binding domain site and calmodulin-binding site.

According to still another aspect of the present invention there is provided an isolated nucleic acid comprising a polynucleotide hybridizable with any of SEQ ID NOs:1, 3, 5, 7, 9, 11 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 ° C.

According to an additional aspect of the present invention there is provided an isolated nucleic acid comprising a polynucleotide hybridizable with a polynucleotide encoding a polypeptide as set forth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 °C.

According to yet an additional aspect of the present invention there is provided a nucleic acid construct comprising a polynucleotide as set forth herein downstream of a plant promoter. The polynucleotide can be in either sense or antisense orientation, whereas the plant promoter can be a constitutive promoter, a tissue specific promoter, an inducible promoter or a chimeric promoter.

According to still an additional aspect of the present invention there is provided a recombinant protein comprising a polypeptide encoded by any of the polynucleotides described herein.

According to a further aspect of the present invention there is provided a recombinant protein comprising a polypeptide at least 60 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3), SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof, in particular.

According to yet a further aspect of the present invention there is provided a genetically transformed or virus infected cell or plant comprising any of the isolated nucleic acids or constructs described herein and preferably expressing any of the recombinant proteins described herein.

According to still a further aspect of the present invention there is provided an antibody specifically recognizing any of the recombinant proteins as described herein and/or natural equivalents thereof. Such an antibody can be, for example, a polyclonal antibody produced by a non-human mammal or in an egg, or alternatively, such an antibody can be a monoclonal antibody produced by a cell such as a hybridoma.

According to yet a further aspect of the present invention there is provided a method of increasing a tolerance of a plant to a metal cation, the method comprising the step of overexpressing in the plant a recombinant protein which reduces uptake and concentration of the metal cation within the plant cells.

According to still a further aspect of the present invention there is provided a method for phytoremediation of an area polluted with a metal cation, the method comprising the steps of (a) providing a plant characterized in resistivity to elevated concentrations of the metal cation, the plant overexpressing a recombinant protein which facilitates uptake and concentration of the metal cation within the plant cells; (b) planting the plant in the area; (c) following a time period, in which at least a fraction of the metal cation in the area has been accumulated in the plant, harvesting the plant, thereby removing at least the fraction of the metal cation from the area; and optionally (d) repeating steps (b) - (c) until a sufficient amount of the metal cation has been removed from the area.

According to still further features in the described preferred embodiments the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide at least 60 % identical with any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

According to further features in preferred embodiments of the invention described below, the polynucleotide is as set forth in any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof.

According to still further features in the described preferred embodiments the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide encoding a polypeptide being at least 60 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer

Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

According to still further features in the described preferred embodiments the plant is genetically transformed or viral infected with a nucleic acid including a plant derived polynucleotide encoding a transmembrane polypeptide having a cation channel activity when assembled in a plasmatic membrane of a plant cell, the polypeptide having overlapping cyclic nucleotide-binding domain site and calmodulin-binding site.

According to still further features in the described preferred embodiments the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide hybridizable with any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 °C.

According to still further features in the described preferred embodiments the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide hybridizable with a polynucleotide encoding a polypeptide as set forth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 °C.

According to still further features in the described preferred embodiments the plant is genetically transformed or viral infected with a nucleic acid construct comprising a polynucleotide as set forth herein downstream of a plant promoter in a sense orientation.

According to still further features in the described preferred embodiments the recombinant protein includes a polypeptide at least 60 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

According to still further features in the described preferred embodiments the polypeptide is as set fourth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof.

According to still further features in the described preferred embodiments the recombinant protein includes a calmodulin binding domain.

According to still further features in the described preferred embodiments the recombinant protein further includes a cyclic nucleotide binding domain.

According to still further features in the described preferred embodiments the calmodulin binding domain and the cyclic nucleotide binding domain overlap.

According to still further features in the described preferred embodiments the metal cation is Ni, Na, Ba, Cd, Co, Cu, La, Mn, Zn and/or Pb cation.

According to still further features in the described preferred embodiments the area is selected from the group comprising of a terrestrial area and an aquatic area.

The present invention successfully addresses the shortcomings of the presently known configurations by providing and characterizing a new gene family which can be used for generating plants resistant to heavy metals and plants which accumulate heavy and other metals and which can be used for phytoremediation of soils and water reservoirs.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention is herein described, by way of example only, with reference to the accompanying drawings, wherein:

FIG. 1A illustrates an amino acid sequence homology of NtCBP4 and NtCBP7 (SEQ ID NOs:2 and 4). Numbers correspond to the deduced amino acids from the first ATG codon. Identical amino acids are boxed. The six hydrophobic domains (S1-S6) are indicated by a solid underline, the predicted pore region (P) by a dashed underline, and the cyclic nucleotide-binding site is underlined with asterisks.

FIG. 1B illustrates hydropathic profiles of NtCBP4 and NtCBP7. Mean hydropathy (ordinate) was plotted against amino acid residue number (abscissa) by using a moving window of 11 amino acids. Hydrophobic and hydrophilic regions are shown above and below the zero line, respectively. The presumptive membrane-spanning and pore-lining regions are shown as S1-S6 and P, respectively. Putative cyclic nucleotide-binding domain (cNBD) and calmodulin-binding domain (CaM) site are indicated by a horizontal bar.

FIG. 2 illustrates a comparison of NtCBP4 and NtCBP7 sequences with sequences of family members from other species. Amino-acid sequence alignment of tobacco (*Nicotiana tabacum*, NtCBP4), thale cress (*Arabidopsis thaliana*, AtCNGC1) and barley (*Hordeum vulgare*, HvCBT1) transporters (SEQ IDS NOs:2, 4, 18 and 16, respectively). Amino acids identical to NtCBP4 are marked by dashes (—) and gaps by dots. The six hydrophobic domains (S1-S6), predicted pore region (P), and putative cyclic nucleotide monophosphate binding domain are colored in red, blue and green, respectively. The CaM binding domain is underlined.

FIG. 3A illustrates a comparison of NtCBP4 and NtCBP7 sequences with that of other ion channels. Amino acid sequence alignment of the S4 regions of NtCBP4 and NtCBP7 (SEQ ID NOs:2 and 4) with *Arabidopsis* AKT1 (SEQ ID NO:35) and KAT1 (SEQ ID NO:36), *Drosophila* Shaker (SEQ ID NO:37) and EAG (SEQ ID NO:38), and the rat CNG2 olfactory channels (SEQ ID NO:34) (GenBank Accession Numbers X62907, X93022 and Swiss-Prot # P08510, Q02280, Q00195, respectively). Positively charged amino acids are highlighted by gray background.

FIG. 3B illustrates a comparison of NtCBP4 and NtCBP7 sequences with that of other ion channels. Amino acid sequence comparison of the putative pore region of NtCBP4 and NtCBP7 with the pore region of the ion channels of Figure 3A is shown. Residues identical or similar to the corresponding positions in NtCBP4 or NtCBP7 are highlighted by a black or gray background, respectively. The GYGD (SEQ ID NO:43) motif of K^+- selective channels is underlined by asterisks.

FIG. 3C illustrates a comparison of NtCBP4 sequence with that of other ion channels. Amino acid sequence comparison of the putative pore region and S6 region of NtCBP4 (SEQ ID NO:2) with similar regions of KcsA, Dmshaker, AtAKT1, OCNG2 (SEQ ID NOs:39, 37, 35, 34). Residues of the ion-selectivity filter, lining of the cavity and inner pore, and conserved amino acid residues, are colored in yellow, orange and gray, respectively.

FIG. 4 illustrates amino acid sequence alignment of tobacco channel-related proteins. Amino acid sequences spanning the S5 through S6 domains of four tobacco channel-related proteins (NtCBP4, NtCBP7, NtCBP10 and NtCBP63; SEQ ID NOs:2, 4, 6 and 8, respectively) were aligned using the CLUSTALW multiple sequence alignment program (Devereux et al., 1984). Positions with identical amino acids in at least three proteins are boxed.

Gaps in the alignment are indicated by dots. The S5, P, and S6 domains are underlined.

FIGs. 5A-B illustrate the purification of full-length NtCBP4 by affinity chromatography. Lysophosphatidylcholine-solubilized membrane proteins from Sf9-NtCBP4 cells were bound to the CaM-agarose column. Equal volumes of total protein loaded to the column (T), the effluent fraction (FT), and the first to fifth EGTA-eluted fractions (1-5, respectively) were analyzed by electrophoresis and were either stained with Coomassie Blue (5A) or transferred to nitrocellulose membrane and probed with anti-NtCBP4 polyclonal antibodies (5B). The positions of molecular-mass standards (kDa) are indicated on the left.

FIG. 6 illustrates interaction between calmodulin and NtCBP4 C-terminal in yeast. β–galactosidase activity in liquid cultures (graph) and X-gal plates (streaks) of yeast expressing NtCBP4 C-terminal as bait and CaM as prey (NtCBP4C), or a negative control of yeast expressing non-specific bait and CaM as prey (*Control*) are shown. LacZ-specific activity units were calculated as described in materials and experimental methods in the Examples section that below.

FIG. 7A illustrates CaM binding by a GST-NtCPB4 fusion protein. GST-NtCBP4 fusion proteins (see Figure 7B for details) or GST alone transferred to membranes after SDS-PAGE, were probed with either ^{35}S-CaM in the presence of 1 mM Ca^{2+} or anti-GST antibodies.

FIG. 7B is a schematic map of the CaM-binding domain of NtCBP4. Schematic presentation of the GST-NtCBP4 fusion proteins tested for CaM binding in Figure 7A. The numbers of the different fusion proteins correspond to the numbers of the samples loaded on the SDS-PAGE in Figure 7A.

FIG. 8A illustrates the α-helical portion of NtCBP4. The predicted α–helical wheel formed by NtCBP4 (SEQ ID NO:2) amino acid residues 595-611 (using the GCG protein analysis program) is shown. Hydrophobic amino acids are boxed and basic amino acids are marked by +.

FIG. 8B illustrates that an NtCBP4 synthetic peptide forms a stable complex with CaM. The complex formation between CaM and a synthetic peptide corresponding to NtCBP4 amino acids 595-617 was assayed in the presence of 0.1 mM $CaCl_2$. Increasing amounts of the peptide (peptide/CaM molar ratios indicated) were incubated with 100 pmole of bovine CaM, and then samples were separated by non-denaturing PAGE. Arrows indicate the positions of free CaM and the peptide/CaM complex.

FIG 9A demonstrates the interaction of NtCBP4 CaM-binding peptide with dansyl CaM by fluorescence excitation. Fluorescence emission spectra of dansyl CaM and its complex with NtCBP4 CaM-binding peptide. Fluorescence emission spectra of 300 nM dansyl CaM without (*empty circles*) and with (*filled circles*) 300 nM peptide was measured at 23 ^{o}C in 50 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl and 0.5 mM Ca^{2+}, using an excitation wavelength of 340 nm with a band pass of 8 nm.

FIG. 9B illustrates the binding kinetics of the interaction of NtCBP4 CaM-binding peptide with dansyl CaM. Titration of the dansyl CaM with NtCBP4 CaM-binding peptide monitored by fluorescence enhancement. The concentration of the dansyl CaM was 200 nM. Emitted fluorescence was measured at 480 nm. Data was fitted as described in the Examples section that follows.

FIG. 10 illustrates the alignment of the NtCBP4, NtCBP7 and olfactory channel cyclic nucleotide binding domains. Amino acid sequences of NtCBP4 (SEQ ID NO:2; residues 487-616), NtCBP7 (SEQ ID NO: 4; residues 484-614), and rat olfactory channel (residues 464-586) are pictured in alignment. Regions corresponding to *E. coli* CAP secondary structure (Shabb et al., 1992) are underlined. Identical (boxed), conserved (gray shaded), and invariant (highlighted in black) amino acid residues are indicated. Residues that fit to the PROSITE cAMP/cGMP binding domain signature motif (PROSITE Accession Number PS50042) are overlined with asterisks. The NtCBP4 region that binds CaM is presented in bold letters.

FIG. 11A illustrates immunodetection of NtCBP4 expressed in insect cells. Non soluble fractions of SF9 cells (50 μg total protein) were extracted from Sf9 cells infected with recombinant NtCBP4 (NtCBP4) or WT (Con) baculoviruses 1-3 days post infection. Proteins were separated by SDS-PAGE, stained by Coomassie Blue or transferred to a nitrocellulose membrane. The membrane was probed with [^{35}S]calmodulin and then with anti-NtCBP4 polyclonal antibodies. Immunoreactive proteins were detected by chemiluminescence. The positions of molecular-mass standards (kDa) are indicated.

FIG. 11B illustrates immunodetection of NtCBP4 expressed in transgenic tobacco. Tobacco proteins (25 μg/lane) from soluble and membrane (microsome) fractions, and a sample of the full-length recombinant NtCBP4 protein (Rec) expressed in Sf9 insect cells were separated by SDS-PAGE and blotted to a nitrocellulose membrane. The

membrane was probed with anti-NtCBP4 polyclonal antibodies. Immunoreactive proteins were detected by chemiluminescence. The positions of molecular-mass standards (kDa) are indicated.

FIG. 12A illustrates detection of NtCBP4 in tobacco plasma membranes with antibodies. Total microsomes from WT and transgenic tobacco seedlings were fractionated on a non-continuous sucrose gradient consisting of 20%, 30%, 34%, 38%, and 45% (w/w) sucrose. Membrane fractions were collected from interfaces between different sucrose concentrations. Proteins from each interface were electrophoresed, blotted and immunodetected with anti-NtCBP4 antibodies, the 60-kDa subunit of vacuolar H^+-ATPase from oat roots (V-ATPas), endoplasmic reticulum BIP (ER-BIP), plasma membrane H^+-ATPase (P-ATPase), mitochondrial prohibitin (M-Prohibitin) and endoplasmic reticulum NADPH cytochrome-p450 reductase (ER-Cyt-p450).

FIG. 12B illustrated detection of NtCBP4 in tobacco plasma membranes with antibodies. Microsomal membranes (Mic) from roots of WT and transgenic plants (line 49-79) were fractionated by the aqueous two-phase partitioning method into a lower phase (L) enriched with intracellular membranes and an upper phase (U) enriched with plasma membranes. Proteins from each fraction (25μg/lane) were immunoblotted or stained with Coomassie Blue. Blots were probed with antibodies as described for Figure 12A.

FIG. 13A illustrates relative expression levels of NtCBP4 in transgenic tobacco lines. Immunodetection of NtCBP4 in microsomes of WT and transgenic plants with NtCBP4 cDNA in either antisense (AS), or sense (transgenic lines 15-89, 13-93, 49-79, 10-2) orientation. Immunodetection of vacuolar membrane H^+-ATPase served as an internal standard. The relative levels of NtCBP4 in transgenic compared to WT (set as 1.00) plants were determined by densitometric scanning of the blot and are given below the immunoblot.

FIG. 13B demonstrates tolerance of transgenic seedlings to Ni^{2+}. $NiCl_2$-tolerance test of representative transgenic tobacco lines. WT and transgenic seedlings (line numbers indicated) were germinated in half-strength Hoagland's liquid medium in the absence or presence of $NiCl_2$ (concentration indicated). Photographs were taken after 12 days of exposure to $NiCl_2$.

FIG. 13C demonstrates tolerance of transgenic seedlings to Ni^{2+}. Photo enlargements of WT and transgenic seedlings (line 49-79) as in Figure 13B in the presence of 0.2 mM $NiCl_2$.

FIG. 14A is a graphic representation of the effects of Ni^{2+} on transgenic chlorophyll accumulation in transgenic seedlings. Relative chlorophyll content (%) (y axis) for WT and transgenic seedlings (line numbers on x axis). Data are given as the mean ± SD of three independent experiments. The values determined in seedlings grown without $NiCl_2$ were set as 100%. Chlorophyll accumulation corresponds to NtCBP4 expression level.

FIG. 14B is a graphic representation of the effects of Ni^{2+} on root growth of transgenic seedlings. Root length of seedlings of WT and transgenic line 49-79 grown for 12 days in the absence or presence of 0.1 mM $NiCl_2$. Root length was calculated as the mean ± SD of 25 seedlings for each measurement. The values determined in seedlings grown without $NiCl_2$ were set as 100%.. Root growth corresponds to NtCBP4 expression level.

FIG. 14C is a graphic representation of the effects of Ni^{2+} on transgenic seedling growth. Relative fresh weight of WT or transgenic (lines 49-79, 10-2) seedlings grown in the absence or presence of different concentrations of $NiCl_2$. Data are given as the mean ± SD of three independent experiments (150 seedlings were used for each point). In (a), (b), and (c), The values determined in seedlings grown without $NiCl_2$ were set as 100%. Resistance to NI^{2+} increases as NtCBP4 expression level increases.

FIG. 15A demonstrates tolerance of developed transgenic plants to Ni^{2+}. WT and transgenic plants were grown for four weeks in half-strength Hoagland's liquid medium and then transferred to the same medium supplemented with 0.2 mM $NiCl_2$. Photographs were taken after two weeks of exposure to $NiCl_2$. Immunodetection of the vacuolar membrane H^+-ATPase served as an internal standard.

FIG. 15B illustrates tolerance of developed transgenic calli to Ni^{2+}. Calli were regenerated from leaves of WT and transgenic line 10-2 on B5+0.8% agar medium. Tolerance of calli to Ni^{2+} was studied as described in the experimental procedures. Photographs were taken two weeks after recovery of calli on Ni^{2+}-free medium (bar = 10 mm). Immunodetection of the vacuolar membrane H^+-ATPase served as an internal standard.

FIG. 15C illustrates tolerance of developed transgenic calli to Ni^{2+}. Immunodetection of NtCBP4 in microsomes from calli of WT and transgenic line 10-2 is demonstrated. Immunodetection of the vacuolar membrane H^+-ATPase served as an internal standard.

FIG. 16 illustrates Ni^{2+} accumulation in WT and transgenic plants. Plants were grown for four weeks before exposing them to different concentrations of $NiCl_2$ (0-0.2 mM) for 24 hours. Ni content was determined by ICP-AES as described in the Experimental procedures. Data are the mean ± SD of three independent experiments.

FIG. 17A illustrates Pb^{2+} hypersensitivity and accumulation in transgenic plants. WT and transgenic seedlings grown with or without 1 mM $Pb(NO_3)_2$ for 12 days. Immunodetection of the vacuolar membrane H^+-ATPase served as an internal standard.

FIG. 17B illustrates Pb^{2+} hypersensitivity and accumulation in transgenic plants. Relative chlorophyll content (%) in WT and transgenic seedlings (line numbers indicated). Chlorophyll content without $Pb(NO_3)_2$ was set as 100 % for each line. Data are the mean ± SD of three independent experiments. Pb^{2+} accumulation correlates to NtCBP4 expression levels.

FIG. 17C illustrates Pb^{2+} hypersensitivity and accumulation in transgenic plants. Pb accumulation in shoots of WT and transgenic line 49-79. The plants were grown for four weeks before exposing them to 0.1 mM $Pb(NO_3)_2$ for 24 hours. Pb content was determined as described in the Examples section that follows. Data are the mean ± SD of three independent experiments. Pb^{2+} accumulation correlates to NtCBP4 expression levels.

FIGs 18A-B show schematic presentations of DNA constructs for expression of the full-length and C-terminal truncated NtCBP4 in transgenic tobacco. Figure 18A shows the full-length NtCBP4 cDNA in a binary Ti plant transformation vector, as described by Arazi *et al.* (1999). Regions with the 35S CaMV promoter and the transcription termination sequence are shown in hatched boxes. The NtCBP4 six-transmembrane core, putative cyclic-nucleotide monophosphate-binding domain (cNBD, in grey; Arazi *et al.*, 1999) and the overlapping calmodulin-binding domain (CaMBD, cross hatched; Arazi *et al.*, 2000) are indicated. The amino acid sequence of part of the cNBD is shown on top. The CaMBD (bold letters) and the αB and αC predicted helices (underlined) of the cNBD are according to Arazi *et al.* (2000). The downward arrow shows the site of the deletion used to create

the C-terminal truncated protein designated NtCBP4ΔC (amino acids Met_1-Ser_{593}). Numbers denote the terminal amino acid residues shown in the sequence, based on Arazi *et al.* (1999; 2000). Figure 18B shows the NtCBP4ΔC cDNA in a binary Ti-plant transformation vector. The NtCBP4ΔC cDNA, as described in Figure 18A, was prepared by cloning a corresponding PCR-amplified DNA fragment into the *Xho*I and *Eco*RI sites of the same vector used for the full-length NtCBP4.

FIGs 19A-B demonstrates Transgenic seedlings expressing NtCBP4ΔC are tolerant to Pb^{2+}. Figure 19A - seedlings of WT and transgenic tobacco expressing either the full-length NtCBP4 (NtCBP4FL; Arazi *et al.*, 1999) or the NtCBP4ΔC mRNA (transgenic lines ΔC-22-11 and ΔC-42-11, ΔC-29) were germinated and grown in the presence of the indicated concentrations of $Pb(NO_3)_2$ in modified Blaydes solution (pH 4.5) for 12 days and then photographed. Figure 19B - photo enlargements of seedlings of WT and representative transgenic lines in the presence of 0 and 0.75 mM $Pb(NO_3)_2$.

FIGs. 20A-B show expression analysis in transgenic plants. Figure 20A - twenty μg of total RNA samples from WT and the transgenic plants indicated were separated by gel electrophoresis, blotted, and hybridised with an NtCBP4-specific probe (upper panel), or with an actin-specific probe (lower panel). The gel positions of the 25S and 18S ribosomal RNA bands are indicated. Arrows point to the full-length NtCBP4 (FL) and the C-terminal truncated NtCBP4 (ΔC) mRNAs in the upper panel. Figure 20B - expression analysis of the full length NtCBP4 mRNA in WT, NtCBP4FL, and NtCBP4ΔC (line ΔC-42-11) plants by RT-PCR. Poly-A^+ mRNA samples from the lines indicated were treated as described in Experimental procedures and a 404-bp DNA fragment corresponding to a region of the full-length NtCBP4 mRNA was amplified with the primers designated NtCBP4. Amplification of the corresponding region from the NtCBP4 cDNA clone with the same primers served as a positive control (cDNA control). The amounts of poly-A^+ RNA were normalised using the expression of the tobacco β-ATPase gene as a standard. Amplified DNA samples and DNA size markers (indicated in bp on the left) were fractionated by agarose gel electrophoresis, stained with ethidium bromide and photographed.

FIGs 21A-B show that Pb^{2+} tand accumulation in transgenic seedlings. Figure 21 A - relative fresh weight of WT and transgenic tobacco seedlings expressing the full-length NtCBP4 mRNA (NtCBP4FL;

Arazi *et al.*, 1999), and two transgenic lines expressing the truncated NtCBP4 mRNA (ΔC-22-11 and ΔC-42-11) in the presence of different concentrations of $Pb(NO_3)_2$. Data are given as the mean fresh weight of 150 seedlings for each concentration ±SD of three independent experiments. Fresh weight of seedlings of each line grown without $Pb(NO_3)_2$ was set as 100%. Figure 21B - lead accumulation in 12-day-old seedlings grown in the presence of 0.2 mM $Pb(NO_3)_2$. Seedlings were dried at 80°C for 3 days and lead content was determined by ICP-AES as described (Arazi *et al.*, 1999). Data are the mean ±SD of three independent experiments.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is of (i) polynucleotides encoding calcium-dependent calmodulin-binding proteins from plants; (ii) polypeptides which are the translation products of these polynucleotides; (iii) expression and antisense vectors containing the polynucleotides or portions thereof; (iv) cultured cells and transgenic or viral infected plants expressing the polynucleotides or portions thereof; and methods of (v) using such plants to assist agriculture and/or (vi) phytoremediation, by providing plants characterized by heavy and other metal resistivity or plants capable of accumulating high concentration of a heavy or other toxic metal and thereby removing the heavy or other toxic metal from the environment.

As used herein the term "heavy metal" refers to a metal having a density above 5 g per cubic cm (see, Prasad M.N.V. and Hagemeyer J. 1999, Heavy Metal Stress in Plants, Springer-Verlag, Berlin). As used herein the term "toxic metal" refers to toxic concentrations of that metal.

The principles and operation of a according to the present invention may be better understood with reference to the drawings and accompanying descriptions.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction and the arrangement of the components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments or of being practiced or carried out in various ways. Also, it is to be understood that the phraseology and terminology employed herein is for the purpose of description and should not be regarded as limiting.

Isolation and analysis of tobacco cDNAs designated NtCBP4 and NtCBP7 and other family members are herein described. These cDNAs

encode Ca^{2+}/Calmodulin (CaM)-binding proteins with sequences similar to K^+-selective channels. NtCBP4 family members contain transmembrane domains, a pore motif, and a highly conserved cyclic nucleotide-binding domain in their C-terminus. Sequence analysis indicates that NtCBP4 and its family members (NtCBP7, NtCBP10, NtCBP63, StCBP18 and MsCBP15) are most similar to the olfactory cyclic nucleotide-gated channel (Dhallan et al., 1990). Differences between NtCBP4 family members and other dated ion channel proteins suggest a selectivity for cations other than potassium. Differences in amino acid sequences among family members, imply differences in ion permeability between members of this novel protein family.

The full-length NtCBP4 binds CaM in a Ca^{2+}-dependent manner via a single CaM-binding domain of approximately 23 amino acid (Phe^{595}-Trp^{617}, SEQ ID NO:2). This domain is characterized by aromatic and long-chain aliphatic residues separated by positively charged amino acids. Only NtCBP4 and its family members, and the olfactory channel, contain both CaM- and cyclic nucleotide-binding domains. In contrast to NtCBP4, in which the CaM-binding domain coincides with the cyclic nucleotide-binding domain, the CaM-binding domain is separate from the cyclic nucleotide-binding domain in the olfactory channel.

NtCBP4 resembles recently cloned ion transporters from *Arabidopsis* (Kohler *et al.*, 1999) and *Hordeum vulgare* (Schuurink *et al.*, 1998). Kohler *et al.* (1999) demonstrated that AtCNGC1 partially complements a yeast mutant deficient in K^+ uptake and suggested that this protein is a non-selective cation channels. However, a major study of the function of this or other members of this plant protein family was not done (Leng, et al. 1999). Transgenic techniques were used to investigate the functions of NtCBP4 in tobacco. Independently derived transgenic lines differing in the level of expressed NtCBP4 were assayed. The NtCBP4 protein was found in tobacco root and shoot microsomes, indicating its presence in the membranes of both organs. As predicted from its sequence, membrane solubilization studies revealed that NtCBP4 is an integral membrane protein. To determine the subcellular localization of NtCBP4, separation of membranes by sucrose gradients and by aqueous two-phase partitioning was performed. These experiments demonstrated that both WT and transgenic NtCBP4 co-fractionate with the plasma membrane. In conclusion, data submitted as part of the present invention support the postulated function of

NtCBP4 and its family members as a component of a plasma membrane ion channel.

. Analysis of the effect of the expression of NtCBP4 on metal uptake and sensitivity to toxic metals in transgenic plants was performed. Transgenic plants were tested with different toxic metal ions. Transgenic plants with higher levels of NtCBP4 are relatively tolerant to Ni^{2+} but hypersensitive to Pb^{2+}. Because of the dual role of Ni^{2+} as an essential microelement on the one hand and as a toxic environmental factor on the other, complete exclusion is not possible. Therefore, genetically engineering plants to have a reduced metal uptake can contribute to improving their tolerance to Ni^{2+}. Exclusion of toxic metal ions by the plasma membrane is conceptually the best mechanism for preventing damage to cellular functions under metal stress conditions.

The tolerance of transgenic lines expressing relatively high levels of NtCBP4 to Ni^{2+} was apparent in very young seedlings and in plants grown for four weeks and then transferred to a Ni^{2+}-containing medium for two weeks, as well as in transgenic calli. Thus, resistance to Ni^{2+} is both cell autonomous and independent of developmental stage. Determination of the concentration-dependent Ni accumulation in shoots of wild type and transgenic plants revealed that overexpression of NtCBP4 restricts the accumulation of Ni.

Because Pb is a non-essential element for plants and is extremely toxic, plant cells are not likely to possess specific Pb transporters. However, certain metal uptake transporters in plants are relatively non-selective, such that both metal nutrients and non-essential toxic metals are taken up (Rubio *et al.,* 1995; Huang *et al.,* 1994). Consequently, the uptake of heavy metal ions, including Pb, by crop plants is a major cause for the accumulation of these toxic ions in the human body (Foy *et al.*, 1978). On the other hand, plants that hyperaccumulate heavy and other metals should be useful for phytoremediation. Use of plants to extract Pb from contaminated soils requires a better understanding of the mechanisms of Pb^{2+} uptake, translocation and accumulation by plants. As part of the present invention, enhancement of the accumulation of Pb in plants by NtCBP4 is demonstrated. In animals, voltage-gated Ca^{2+} channels are permeable to Pb^{2+} (Simons and Pocock, 1987; Tomsing and Suszokiw, 1991). However, to date, plant transporters mediating Pb^{2+} uptake into plants have not been characterized at the molecular level. Therefore, NtCBP4 is the first example of a plant protein that can modulate Pb

tolerance and accumulation in plants. As such, it has unique value for improving phytoremediation strategies.

In spite of the fact that the physiological roles of NtCBP4 and the related genes from barley, *Arabidopsis,* and other plants, remains to be resolved, the remarkable effects of NtCBP4 on heavy and other metal tolerance and accumulation in transgenic plants can not be ignored. The transgenic plants prepared and characterized as part of the present invention, as well as other transgenic plants expressing modified NtCBP4 proteins with changes in their ion selectivity and regulatory properties will undoubtedly find myriad uses in sustaining commercial agriculture in an environment increasingly prone to contamination with toxic metal cations.

In summary, a channel-like protein family in plants with a new structural motif in which a phylogenetically conserved helix in the cyclic nucleotide-binds CaM with high affinity. This makes members of the new gene family communication points for cross-talk between Ca^{2+} and cyclic nucleotide signal transduction pathways in plants. Transgenic plants overexpressing NtCBP4 exhibit increased Ni^{2+} tolerance and hypersensitivity to Pb^{2+}. Different family members are expected to influence tolerance or sensitivity to other metal cations. These qualities make the new gene family useful for cultivating plants in soil contaminated with toxic metal cations on the one hand, and for phytoremediation on the other hand.

By expressing a modified version of NtCBP4, lacking presumed regulatory domains, the phenotype, with respect to Pb^{2+} tolerance and accumulation, is the opposite of that exhibited by plants expressing the full-length protein. The contrasting phenotypes obtained by expressing two variants of the same gene strongly supports the likelihood that NtCBP4 is a component of an ion transport system that is responsible for Pb^{2+} entry into plant cells.

Because Pb^{2+} is a non-essential toxic metal, the presumed NtCBP4-associated ion transport mechanism is likely to have other yet unknown physiological roles. Of particular interest is the possible involvement of NtCBP4 and related proteins in Ca^{2+} signal transduction, either by the regulation of NtCBP4 by Ca^{2+} signals through calmodulin, by being permeable to Ca^{2+}, or by both, like the mammalian cyclic nucleotide-gated non-selective cation channels. This would be consistent with earlier reports identifying Ca^{2+}-permeable channels as a pathway for Pb^{2+} entry into animal cells (Simons and Pocock, 1987; Tomsig and Suszkiw, 1991) and plants

(Huang and Cunningham, 1996), and with the recent demonstration of cyclic nucleotide-dependent Ca^{2+} entry into human embryonic kidney cells transfected with the *Arabidopsis AtCNGC2* gene (Leng *et al.*, 1999).

The mammalian cyclic nucleotide-gated channels function as tetrameric complexes (Liu et al., 1996). Therefore, the inhibitory effect of NtCBP4ΔC on Pb^{2+} accumulation, and the concomitant improved tolerance to Pb^{2+} may have resulted from the formation of non functional NtCBP4/NtCBP4ΔC heteromeric complexes, since the endogenous native NtCBP4 gene was expressed in the background of the NtCBP4ΔC transgene. Regardless of the mechanism, an inhibitory effect of NtCBP4ΔC on the activity of NtCBP4-associated channels might influence certain physiological and developmental processes. One interesting difference between the plants expressing the full-length NtCBP4 and the plants expressing the truncated protein is the apparent male sterility in the latter. Although the reason for this phenotype is at present unknown, further analysis of the two types of transgenic plants described here should be useful for elucidating the physiological role of NtCBP4. Based on the phenotypes that was characterised regarding Pb^{2+} uptake into plant cells, it is expect that the physiological responses that are mediated by NtCBP4-associated channels will be enhanced in transgenic lines overexpressing the full-length channel but attenuated in plants expressing the truncated one.

Thus, according to one aspect of the present invention there is provided an isolated nucleic acid comprising a polynucleotide at least 60 %, preferably at least 65 %, more preferably at least 70 %, still preferably at least 75 %, still preferably at least 80 %, yet preferably at least 90-100 % identical with any of SEQ ID NOs:1, 3, 5, 7, 9, 11 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3), SEQ ID NOs:1, 3, 5, 7, 9, 11 or portions thereof, in particular.

According to another aspect of the present invention there is provided an isolated nucleic acid comprising a polynucleotide encoding a polypeptide being at least 60 %, preferably at least 65 %, more preferably at least 70 %, still preferably at least 75 %, still preferably at least 80 %, yet preferably at least 90-100 % homologous (similar + identical acids) with any of SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package

developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

According to yet another aspect of the present invention there is provided an isolated nucleic acid comprising a plant derived polynucleotide encoding a transmembrane polypeptide having a cation channel activity when assembled in a plasmatic membrane of a plant cell, the polypeptide having overlapping (coinciding) cyclic nucleotide-binding domain site and calmodulin-binding site. The polypeptide, when incorporated in a plasmatic membrane functions as a metal channel, i.e., in some embodiments it transports metal ions into the cell cytoplasm, wherein in other embodiments it transports metal ions out of the cell cytoplasm.

According to still another aspect of the present invention there is provided an isolated nucleic acid comprising a polynucleotide hybridizable with any of SEQ ID NOs:1, 3, 5, 7, 9, 11 or portions thereof.

According to an additional aspect of the present invention there is provided an isolated nucleic acid comprising a polynucleotide hybridizable with a polynucleotide encoding a polypeptide as set forth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof.

Hybridization for long nucleic acids (e.g., above 200 bp in length) is effected according to preferred embodiments of the present invention by stringent or moderate hybridization, wherein stringent hybridization is effected by a hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 0.2 x SSC and 0.1 % SDS and final wash at 65°C; whereas moderate hybridization is effected by a hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 °C.

It will be appreciated that by using the above hybridization conditions and/or other molecular techniques, such as those further exemplified in the Examples section that follows, one of ordinary skills, without undue experimentation, would readily isolate additional members of the new gene family. The scope of the present invention also encompasses such sequences which are the result of man induced variation, such as site directed or non-specific mutagenesis. Methods of man induced variation of nucleic acids are well known in the art and are further described in detail Sambrook *et al.*, molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989), which is

incorporated herein by reference. The scope of the present invention also encompasses such sequences which are the result of naturally occurring variations. Such variations can be characterized and such sequence alterations isolated by one ordinarily skilled in the art using the assays and procedures described hereinunder in the Examples section that follows. In addition, the scope of the present invention also encompasses both cDNA and genomic sequences. cDNA sequences have been isolated while reducing the present invention to practice. By employing the stringent hybridization conditions described herein one of skills in the art would be able, without undue experimentation, to readily isolate genomic sequences corresponding to the isolated or isolatable cDNA sequences.

According to yet an additional aspect of the present invention there is provided a nucleic acid construct comprising a polynucleotide as set forth herein downstream of a plant promoter.

As used herein in the specification and in the claims section that follows the phrase "plant promoter" includes a promoter which can direct gene expression in plant cells (including DNA containing organelles). Such a promoter can be derived from a plant, bacterial, viral, fungal or animal origin. Such a promoter can be constitutive, i.e., capable of directing high level of gene expression in a plurality of plant tissues, tissue specific, i.e., capable of directing gene expression in a particular plant tissue or tissues, inducible, i.e., capable of directing gene expression under a stimulus, or chimeric, i.e., formed of portions of at least two different promoters.

Thus, the plant promoter employed can be a constitutive promoter, a tissue specific promoter, an inducible promoter or a chimeric promoter.

Examples of constitutive plant promoters include, without being limited to, CaMV35S and CaMV19S promoters, FMV34S promoter, sugarcane bacilliform badnavirus promoter, CsVMV promoter, *Arabidopsis* ACT2/ACT8 actin promoter, *Arabidopsis* ubiquitin UBQ1 promoter, barley leaf thionin BTH6 promoter, and rice actin promoter.

Examples of tissue specific promoters include, without being limited to, bean phaseolin storage protein promoter, DLEC promoter, PHSβ promoter, zein storage protein promoter, conglutin gamma promoter from soybean, AT2S1 gene promoter, ACT11 actin promoter from *Arabidopsis*, napA promoter from *Brassica napus* and potato patatin gene promoter.

The inducible promoter is a promoter induced by a specific stimuli such as stress conditions comprising, for example, light, temperature, chemicals, drought, high salinity, osmotic shock, oxidant conditions or in

case of pathogenicity and include, without being limited to, the light-inducible promoter derived from the pea rbcS gene, the promoter from the alfalfa rbcS gene, the promoters DRE, MYC and MYB active in drought; the promoters INT, INPS, prxEa, Ha hsp17.7G4 and RD21 active in high salinity and osmotic stress, and the promoters hsr203J and str246C active in pathogenic stress.

According to still an additional aspect of the present invention there is provided a recombinant protein comprising a polypeptide encoded by any of the polynucleotides described herein.

According to a further aspect of the present invention there is provided a recombinant protein comprising a polypeptide at least 60 %, preferably at least 65 %, more preferably at least 70 %, still preferably at least 75 %, still preferably at least 80 %, yet preferably at least 90-100 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3), SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof, in particular.

As used herein in the specification and in the claims section that follows the term "polypeptide" refers also to a protein, in particular a transmembrane protein, which may include a transit peptide, and further to a post translationally modified protein, such as, but not limited to, a phosphorylated protein, glycosylated protein, ubiquitinylated protein, acetylated protein, methylated protein, etc.

According to a preferred embodiment of the present invention, the polypeptide includes an N terminal transit peptide fused thereto which serves for directing the polypeptide to a specific membrane. Such a membrane can be, for example, the cell membrane, wherein the polypeptide will serve to transport metal ions from the apoplast into the cytoplasm or vice versa, or, such a membrane can be the outer and preferably the inner chloroplast membrane, wherein the polypeptide will serve to transport metal ions from the cytoplasm to the intermembranal space and the stroma, respectively, or vice versa. Transit peptides which function as herein described are well known in the art. Further description of such transit peptides is found in, for example, Johnson *et al.* The Plant Cell (1990) 2:525-532; Sauer *et al.* EMBO J. (1990) 9:3045-3050; Mueckler *et al.* Science (1985) 229:941-945; Von Heijne, Eur. J. Biochem. (1983) 133:17-21; Yon Heijne, J. Mol. Biol. (1986) 189:239-242; Iturriaga *et al.* The

Plant Cell (1989) 1:381-390; McKnight *et al.*, Nucl. Acid Res. (1990) 18:4939-4943; Matsuoka and Nakamura, Proc. Natl. Acad. Sci. USA (1991) 88:834-838. A recent text book entitled "Recombinant proteins from plants", Eds. C. Cunningham and A.J.R. Porter, 1998 Humana Press Totowa, N.J. describe methods for the production of recombinant proteins in plants and methods for targeting the proteins to different compartments in the plant cell. The book by Cunningham and Porter is incorporated herein by reference. It will however be appreciated by one of skills in the art that a large number of membrane integrated proteins fail to poses a removable transit peptide. It is accepted that in such cases a certain amino acid sequence in said proteins serves not only as a structural portion of the protein, but also as a transit peptide.

According to yet a further aspect of the present invention there is provided a genetically transformed or virus infected cell or plant comprising any of the isolated nucleic acids or constructs described herein and preferably expressing any of the recombinant proteins described herein.

As used herein in the specification and in the claims section that follows the term "plant" includes organisms, both unicellular or multicellular, both prokaryotes or eukaryotes, both soil grown or aquatic, capable of producing complex organic materials, especially carbohydrates, from carbon dioxide using light as the source of energy and with the aid of chlorophyll and optionally associated pigments.

As used herein in the specification and in the claims section that follows the term "transformed" and its conjugations such as transformation, transforming and transform, all relate to the process of introducing heterologous nucleic acid sequences into a cell or an organism, which nucleic acid are propagatable to the offspring. The term thus reads on, for example, "genetically modified", "transgenic" and "transfected", which may be used herein to further described the present invention. The term relates both to introduction of a heterologous nucleic acid sequence into the genome of an organism and/or into the genome of a nucleic acid containing organelle thereof, such as into a genome of chloroplast or a mitochondrion.

As used herein in the specification and in the claims section that follows the phrase "viral infected" includes infection by a virus carrying a heterologous nucleic acid sequence. Such infection typically results in transient expression of the nucleic acid sequence, which nucleic acid sequence is typically not integrated into a genome and therefore not

propagatable to offspring, unless further infection of such offspring is experienced.

There are various methods of introducing foreign genes into both monocotyledonous and dicotyledenous plants (Potrykus, I., Annu. Rev. Plant. Physiol., Plant. Mol. Biol. (1991) 42:205-225; Shimamoto *et al.*, Nature (1989) 338:274-276). The principle methods of causing stable integration of exogenous DNA into plant genomic DNA include two main approaches:

(i) *Agrobacterium*-mediated gene transfer: Klee *et al.* (1987) Annu. Rev. Plant Physiol. 38:467-486; Klee and Rogers in Cell Culture and Somatic Cell Genetics of Plants, Vol. 6, Molecular Biology of Plant Nuclear Genes, eds. Schell, J., and Vasil, L. K., Academic Publishers, San Diego, Calif. (1989) p. 2-25; Gatenby, in Plant Biotechnology, eds. Kung, S. and Arntzen, C. J., Butterworth Publishers, Boston, Mass. (1989) p. 93-112.

(ii) direct DNA uptake: Paszkowski *et al.*, in Cell Culture and Somatic Cell Genetics of Plants, Vol. 6, Molecular Biology of Plant Nuclear Genes eds. Schell, J., and Vasil, L. K., Academic Publishers, San Diego, Calif. (1989) p. 52-68; including methods for direct uptake of DNA into protoplasts, Toriyama, K. *et al.* (1988) Bio/Technology 6:1072-1074. DNA uptake induced by brief electric shock of plant cells: Zhang *et al.* Plant Cell Rep. (1988) 7:379-384. Fromm *et al.* Nature (1986) 319:791-793. DNA injection into plant cells or tissues by particle bombardment, Klein *et al.* Bio/Technology (1988) 6:559-563; McCabe *et al.* Bio/Technology (1988) 6:923-926; Sanford, Physiol. Plant. (1990) 79:206-209; by the use of micropipette systems: Neuhaus *et al.*, Theor. Appl. Genet. (1987) 75:30-36; Neuhaus and Spangenberg, Physiol. Plant. (1990) 79:213-217; or by the direct incubation of DNA with germinating pollen, DeWet *et al.* in Experimental Manipulation of Ovule Tissue, eds. Chapman, G. P. and Mantell, S. H. and Daniels, W. Longman, London, (1985) p. 197-209; and Ohta, Proc. Natl. Acad. Sci. USA (1986) 83:715-719.

The *Agrobacterium* system includes the use of plasmid vectors that contain defined DNA segments that integrate into the plant genomic DNA. Methods of inoculation of the plant tissue vary depending upon the plant species and the *Agrobacterium* delivery system. A widely used approach is the leaf disc procedure which can be performed with any tissue explant that provides a good source for initiation of whole plant differentiation. Horsch

et al. in Plant Molecular Biology Manual A5, Kluwer Academic Publishers, Dordrecht (1988) p. 1-9. A supplementary approach employs the *Agrobacterium* delivery system in combination with vacuum infiltration. The *Agrobacterium* system is especially viable in the creation of transgenic dicotyledenous plants.

There are various methods of direct DNA transfer into plant cells. In electroporation, the protoplasts are briefly exposed to a strong electric field. In microinjection, the DNA is mechanically injected directly into the cells using very small micropipettes. In microparticle bombardment, the DNA is adsorbed on microprojectiles such as magnesium sulfate crystals or tungsten particles, and the microprojectiles are physically accelerated into cells or plant tissues.

Following transformation plant propagation is exercised. The most common method of plant propagation is by seed. Regeneration by seed propagation, however, has the deficiency that due to heterozygosity there is a lack of uniformity in the crop, since seeds are produced by plants according to the genetic variances governed by Mendelian rules. Basically, each seed is genetically different and each will grow with its own specific traits. Therefore, it is preferred that the transformed plant be produced such that the regenerated plant has the identical traits and characteristics of the parent transgenic plant. Therefore, it is preferred that the transformed plant be regenerated by micropropagation which provides a rapid, consistent reproduction of the transformed plants.

Micropropagation is a process of growing new generation plants from a single piece of tissue that has been excised from a selected parent plant or cultivar. This process permits the mass reproduction of plants having the preferred tissue expressing the protein. The new generation plants which are produced are genetically identical to, and have all of the characteristics of, the original plant. Micropropagation allows mass production of quality plant material in a short period of time and offers a rapid multiplication of selected cultivars in the preservation of the characteristics of the original transgenic or transformed plant. The advantages of cloning plants are the speed of plant multiplication and the quality and uniformity of plants produced.

Micropropagation is a multi-stage procedure that requires alteration of culture medium or growth conditions between stages. Thus, the micropropagation process involves four basic stages: Stage one, initial tissue culturing; stage two, tissue culture multiplication; stage three,

differentiation and plant formation; and stage four, greenhouse culturing and hardening. During stage one, initial tissue culturing, the tissue culture is established and certified contaminant-free. During stage two, the initial tissue culture is multiplied until a sufficient number of tissue samples are produced to meet production goals. During stage three, the tissue samples grown in stage two are divided and grown into individual plantlets. At stage four, the transformed plantlets are transferred to a greenhouse for hardening where the plants' tolerance to light is gradually increased so that it can be grown in the natural environment.

The basic bacterial/plant vector construct according to the present invention will preferably provide a broad host range prokaryote replication origin; a prokaryote selectable marker; and, for *Agrobacterium* transformations, T DNA sequences for *Agrobacterium*-mediated transfer to plant chromosomes. Where the heterologous sequence is not readily amenable to detection, the construct will preferably also have a selectable marker gene suitable for determining if a plant cell has been transformed. A general review of suitable markers for the members of the grass family is found in Wilmink and Dons, Plant Mol. Biol. Reptr. (1993) 11:165-185.

Sequences suitable for permitting integration of the heterologous sequence into the plant genome are also recommended. These might include transposon sequences and the like for homologous recombination as well as Ti sequences which permit random insertion of a heterologous expression cassette into a plant genome.

Suitable prokaryote selectable markers include resistance toward antibiotics such as ampicillin or tetracycline. Other DNA sequences encoding additional functions may also be present in the vector, as is known in the art.

The constructs of the subject invention will include an expression cassette for expression of the protein of interest. Usually, there will be only one expression cassette, although two or more are feasible. The recombinant expression cassette will contain in addition to the heterologous sequence one or more of the following sequence elements, a promoter region, plant 5' untranslated sequences, initiation codon depending upon whether or not the structural gene comes equipped with one, and a transcription and translation termination sequence. Unique restriction enzyme sites at the 5' and 3' ends of the cassette allow for easy insertion into a pre-existing vector.

Viruses are a unique class of infectious agents whose distinctive features are their simple organization and their mechanism of replication. In fact, a complete viral particle, or virion, may be regarded mainly as a block of genetic material (either DNA or RNA) capable of autonomous replication, surrounded by a protein coat and sometimes by an additional membranous envelope such as in the case of alpha viruses. The coat protects the virus from the environment and serves as a vehicle for transmission from one host cell to another.

Viruses that have been shown to be useful for the transformation of plant hosts include CaV, TMV and BV. Transformation of plants using plant viruses is described in U.S. Pat. No. 4,855,237 (BGV), EP-A 67,553 (TMV), Japanese Published Application No. 63-14693 (TMV), EPA 194,809 (BV), EPA 278,667 (BV); and Gluzman, Y. *et al.*, Communications in Molecular Biology: Viral Vectors, Cold Spring Harbor Laboratory, New York, pp. 172-189 (1988). Pseudovirus particles for use in expressing foreign DNA in many hosts, including plants, is described in WO 87/06261.

Construction of plant RNA viruses for the introduction and expression of non-viral foreign genes in plants is demonstrated by the above references as well as by Dawson, W. O. *et al.*, Virology (1989) 172:285-292; Takamatsu *et al.* EMBO J. (1987) 6:307-311; French *et al.* Science (1986) 231:1294-1297; and Takamatsu *et al.* FEBS Letters (1990) 269:73-76.

When the virus is a DNA virus, the constructions can be made to the virus itself. Alternatively, the virus can first be cloned into a bacterial plasmid for ease of constructing the desired viral vector with the foreign DNA. The virus can then be excised from the plasmid. If the virus is a DNA virus, a bacterial origin of replication can be attached to the viral DNA, which is then replicated by the bacteria. Transcription and translation of this DNA will produce the coat protein which will encapsidate the viral DNA. If the virus is an RNA virus, the virus is generally cloned as a cDNA and inserted into a plasmid. The plasmid is then used to make all of the constructions. The RNA virus is then produced by transcribing the viral sequence of the plasmid and translation of the viral genes to produce the coat protein(s) which encapsidate the viral RNA.

Construction of plant RNA viruses for the introduction and expression of non-viral foreign genes in plants is demonstrated by the above references as well as in U.S. Pat. No. 5,316,931

In one embodiment, a plant viral nucleic acid is provided in which the native coat protein coding sequence has been deleted from a viral nucleic acid, a non-native plant viral coat protein coding sequence and a non-native promoter, preferably the subgenomic promoter of the non-native coat protein coding sequence, capable of expression in the plant host, packaging of the recombinant plant viral nucleic acid, and ensuring a systemic infection of the host by the recombinant plant viral nucleic acid, has been inserted. Alternatively, the coat protein gene may be inactivated by insertion of the non-native nucleic acid sequence within it, such that a protein is produced. The recombinant plant viral nucleic acid may contain one or more additional non-native subgenomic promoters. Each non-native subgenomic promoter is capable of transcribing or expressing adjacent genes or nucleic acid sequences in the plant host and incapable of recombination with each other and with native subgenomic promoters. Non-native (foreign) nucleic acid sequences may be inserted adjacent the native plant viral subgenomic promoter or the native and a non-native plant viral subgenomic promoters if more than one nucleic acid sequence is included. The non-native nucleic acid sequences are transcribed or expressed in the host plant under control of the subgenomic promoter to produce the desired products.

In a second embodiment, a recombinant plant viral nucleic acid is provided as in the first embodiment except that the native coat protein coding sequence is placed adjacent one of the non-native coat protein subgenomic promoters instead of a non-native coat protein coding sequence.

In a third embodiment, a recombinant plant viral nucleic acid is provided in which the native coat protein gene is adjacent its subgenomic promoter and one or more non-native subgenomic promoters have been inserted into the viral nucleic acid. The inserted non-native subgenomic promoters are capable of transcribing or expressing adjacent genes in a plant host and are incapable of recombination with each other and with native subgenomic promoters. Non-native nucleic acid sequences may be inserted adjacent the non-native subgenomic plant viral promoters such that said sequences are transcribed or expressed in the host plant under control of the subgenomic promoters to produce the desired product.

In a fourth embodiment, a recombinant plant viral nucleic acid is provided as in the third embodiment except that the native coat protein coding sequence is replaced by a non-native coat protein coding sequence.

The viral vectors are encapsidated by the coat proteins encoded by the recombinant plant viral nucleic acid to produce a recombinant plant virus. The recombinant plant viral nucleic acid or recombinant plant virus is used to infect appropriate host plants. The recombinant plant viral nucleic acid is capable of replication in the host, systemic spread in the host, and transcription or expression of foreign gene(s) in the host to produce the desired protein.

Thus, according to a preferred embodiment of the present invention the polynucleotide or nucleic acid molecule of the present invention further includes one or more sequence elements, such as, but not limited to, a nucleic acid sequence encoding a transit peptide, an origin of replication for propagation in bacterial cells, at least one sequence element for integration into a plant's genome, a polyadenylation recognition sequence, a transcription termination signal, a sequence encoding a translation start site, a sequence encoding a translation stop site, plant RNA virus derived sequences, plant DNA virus derived sequences, tumor inducing (Ti) plasmid derived sequences and a transposable element derived sequence.

Thus, according to preferred embodiments of the present invention, the step of transforming the cell or plant with a polynucleotide encoding polypeptide according to the present invention is effected by a method such as genetic transformation and transient transformation (viral infection). Genetic transformation can be effected by, for example, *Agrobaterium* mediated transformation, whereas transient transformation can be effected by, for example, viral infection. Both transient and genetic transformation can be effected by electroporation, particle bombardment or any of the other methods listed and further described hereinabove.

A technique for introducing heterologous nucleic acid sequences to the genome of the chloroplasts is known. This technique involves the following procedures. First, plant cells are chemically treated so as to reduce the number of chloroplasts per cell to about one. Then, the heterologous nucleic acid is introduced via particle bombardment into the cells with the aim of introducing at least one heterologous nucleic acid molecule into the chloroplasts. The heterologous nucleic acid is selected such that it is integratable into the chloroplast's genome via homologous recombination which is readily effected by enzymes inherent to the chloroplast. To this end, the heterologous nucleic acid includes, in addition to a gene of interest, at least one nucleic acid stretch which is derived from the chloroplast's genome. In addition, the heterologous nucleic acid

includes a selectable marker, which serves by sequential selection procedures to ascertain that all or substantially all of the copies of the chloroplast genomes following such selection will include the heterologous nucleic acid. Further details relating to this technique are found in U.S. Pat. Nos. 4,945,050; and 5,693,507 which are incorporated herein by reference. A polypeptide can thus be produced by the protein expression system of the chloroplast and become integrated into the chloroplast's inner membrane.

According to still a further aspect of the present invention there is provided an antibody specifically recognizing any of the recombinant proteins as described herein and/or natural equivalents thereof. Such an antibody can be, for example, a polyclonal antibody produced by a non-human mammal or in an egg, or alternatively, such an antibody can be a monoclonal antibody produced by a cell such as a hybridoma.

Thus, the present invention can utilize serum immunoglobulins, polyclonal antibodies or fragments thereof, (i.e., immunoreactive derivative of an antibody), or monoclonal antibodies or fragments thereof. Monoclonal antibodies or purified fragments of the monoclonal antibodies having at least a portion of an antigen binding region, including such as Fv, F(abl)2, Fab fragments (Harlow and Lane, 1988 Antibody, Cold Spring Harbor), single chain antibodies (U.S. Patent 4,946,778), chimeric or humanized antibodies and complementarily determining regions (CDR) may be prepared by conventional procedures. Purification of these serum immunoglobulins antibodies or fragments can be accomplished by a variety of methods known to those of skill including, precipitation by ammonium sulfate or sodium sulfate followed by dialysis against saline, ion exchange chromatography, affinity or immunoaffinity chromatography as well as gel filtration, zone electrophoresis, etc. (see Goding in, Monoclonal Antibodies: Principles and Practice, 2nd ed., pp. 104-126, 1986, Orlando, Fla., Academic Press). Under normal physiological conditions antibodies are found in plasma and other body fluids and in the membrane of certain cells and are produced by lymphocytes of the type denoted B cells or their functional equivalent. Antibodies of the IgG class are made up of four polypeptide chains linked together by disulfide bonds. The four chains of intact IgG molecules are two identical heavy chains referred to as H-chains and two identical light chains referred to as L-chains. Additional classes includes IgD, IgE, IgA, IgM and related proteins.

Methods for the generation and selection of monoclonal antibodies are well known in the art, as summarized for example in reviews such as

Tramontano and Schloeder, Methods in Enzymology 178, 551-568, 1989. A recombinant protein of the present invention may be used to generate antibodies *in vitro*. More preferably, the recombinant protein of the present invention is used to elicit antibodies *in vivo*. In general, a suitable host animal or egg is immunized with the recombinant protein of the present invention. Advantageously, the animal host used is a mouse of an inbred strain. Animals or eggs are typically immunized with a mixture comprising a solution of the recombinant protein of the present invention in a physiologically acceptable vehicle, and any suitable adjuvant, which achieves an enhanced immune response to the immunogen. By way of example, the primary immunization conveniently may be accomplished with a mixture of a solution of the recombinant protein of the present invention and Freund's complete adjuvant, said mixture being prepared in the form of a water in oil emulsion. Typically the immunization may be administered to the animals intramuscularly, intradermally, subcutaneously, intraperitoneally, into the footpads, or by any appropriate route of administration. The immunization schedule of the immunogen may be adapted as required, but customarily involves several subsequent or secondary immunizations using a milder adjuvant such as Freund's incomplete adjuvant. Antibody titers and specificity of binding to the protein can be determined during the immunization schedule by any convenient method including by way of example radioimmunoassay, or enzyme linked immunosorbant assay, which is known as the ELISA assay. When suitable antibody titers are achieved, antibody producing lymphocytes from the immunized animals are obtained, and these are cultured, selected and cloned, as is known in the art. Typically, lymphocytes may be obtained in large numbers from the spleens of immunized animals, but they may also be retrieved from the circulation, the lymph nodes or other lymphoid organs. Lymphocytes are then fused with any suitable myeloma cell line, to yield hybridomas, as is well known in the art. Alternatively, lymphocytes may also be stimulated to grow in culture, and may be immortalized by methods known in the art including the exposure of these lymphocytes to a virus, a chemical or a nucleic acid such as an oncogene, according to established protocols. After fusion, the hybridomas are cultured under suitable culture conditions, for example in multiwell plates, and the culture supernatants are screened to identify cultures containing antibodies that recognize the hapten of choice. Hybridomas that secrete antibodies that recognize the recombinant protein of the present invention are cloned by limiting dilution

and expanded, under appropriate culture conditions. Monoclonal antibodies are purified and characterized in terms of immunoglobulin type and binding affinity.

According to yet a further aspect of the present invention there is provided a method of increasing a tolerance of a plant to a metal cation. The method according to this aspect of the present invention comprising the step of overexpressing in the plant a recombinant protein which reduces uptake and concentration of the metal cation within the plant cells.

According to still a further aspect of the present invention there is provided a method for phytoremediation of an area polluted with a metal cation. The method according to this aspect of the present invention is effected by implementing the following method steps, in which, in a first step, a plant is provided characterized in resistivity to elevated concentrations of the metal cation, the plant overexpressing a recombinant protein which facilitates uptake and concentration of the metal cation within the plant cells. In a second step of the method according to this aspect of the present invention the plant in planted in the polluted area. Then, following a time period, in which at least a fraction of the metal cation in the polluted area has been accumulated in the plant, the plant is harvested, thereby at least the fraction of the metal cation is removed from the area. These steps are optionally repeated until a sufficient amount of the metal cation has been removed from the polluted area. Methods well known in the art can be employed to monitor the phytoremediation process of a polluted area. Such methods include both chemical/physical methods which are used to directly determine the concentration of the metal cation in the polluted area, and biological methods in which a series of plants with variable sensitivity to the metal cation are planted and their growth or growth inhibition monitored to thereby provide an insight to the level of phytoremediation so far achieved. It will be appreciated that according to the present invention terrestrial or aquatic phytoremediation can be employed, provided that suitable terrestrial or aquatic plants are selected for such purposes. It will further be appreciated that removing plant harvest can be effected both from terrestrial and aquatic environments. In the latter case nets or filters can be employed to effect plant harvesting, depending on the size of the plants employed.

The following provides several preferred embodiments employed when implementing any one of the two methods according to the present invention described above.

Thus, according to a preferred embodiment of the present invention, and in order to implement the methods according to the present invention, the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide at least 60 %, preferably at least 70 %, more preferably at least 80 %, still preferably at least 90 %, yet preferably at least 100 % identical with any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3), SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof, in particular.

According to anther preferred embodiment of the present invention, and in order to implement the methods according to the present invention, the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide encoding a polypeptide being at least 60 %, preferably at least 70 %, more preferably at least 80 %, still preferably at least 90 %, yet preferably at least 100 % homologous (identical + similar) with any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

According to still anther preferred embodiment of the present invention, and in order to implement the methods according to the present invention, the plant is genetically transformed or viral infected with a nucleic acid including a plant derived polynucleotide encoding a transmembrane polypeptide having a cation channel activity when assembled in a plasmatic membrane of a plant cell, the polypeptide having overlapping cyclic nucleotide-binding domain site and calmodulin-binding site.

According to still anther preferred embodiment of the present invention, and in order to implement the methods according to the present invention, the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide hybridizable with any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 °C.

According to still anther preferred embodiment of the present invention, and in order to implement the methods according to the present invention, the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide hybridizable with a polynucleotide encoding a polypeptide as set forth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 °C.

According to a further preferred embodiment of the present invention, and in order to implement the methods according to the present invention, the plant is genetically transformed or viral infected with a nucleic acid construct comprising a polynucleotide as set forth herein downstream of a plant promoter in a sense orientation. Plant promoters and plant transformation/viral infection methods are further described hereinabove.

According to a further preferred embodiment of the present invention, and in order to implement the methods according to the present invention, the recombinant protein includes a polypeptide at least 60 %, preferably at least 70 %, more preferably at least 80 %, still preferably at least 90 %, yet preferably at least 100 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

According to a further preferred embodiment of the present invention, and in order to implement the methods according to the present invention, the recombinant protein includes a calmodulin binding domain.

According to a further preferred embodiment of the present invention, and in order to implement the methods according to the present invention, the recombinant protein further includes a cyclic nucleotide binding domain.

According to a further preferred embodiment of the present invention, and in order to implement the methods according to the present invention, the calmodulin binding domain and the cyclic nucleotide binding domain overlap (coincide).

According to still further features in the described preferred embodiments the metal cation is Ni, Na, Ba, Cd, Co, Cu, La, Mn, Zn and/or

Pb. Using the molecular, physical and physiological methods described in the Examples section that follows, one of ordinary skills would be able without undue experimentation to characterize additional sequences of the new gene family according to the present invention in order to find sequences associated with accumulation of or resistance to any of the above listed metal cations.

According to another aspect of the present invention there is provided a male sterile plant expressing a C termainl truncated form of a calcium-dependent calmodulin-binding protein as herein described.

Additional objects, advantages, and novel features of the present invention will become apparent to one ordinarily skilled in the art upon examination of the following examples, which are not intended to be limiting. Additionally, each of the various embodiments and aspects of the present invention as delineated hereinabove and as claimed in the claims section below finds experimental support in the following examples.

EXAMPLES

Reference is now made to the following examples, which together with the above descriptions, illustrate the invention in a non limiting fashion.

Generally, the nomenclature used herein and the laboratory procedures in recombinant DNA technology described below are these well known and commonly employed in the art. Standard techniques are used for cloning, DNA and RNA isolation, amplification and purification. Generally enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like are performed according to the manufacturers' specifications. These techniques and various other techniques are generally performed according to Sambrook *et al.*, molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989). The procedures described therein and in other citations are believed to be well known in the art and are provided for the convenience of the reader. All the information contained therein is incorporated herein by reference.

MATERIALS AND EXPERIMENTAL METHODS

The following experimental methods were employed in conjunction with the examples described hereinbelow.

Library Screenings and Sequence Analysis: *Nicotiana tabacum* (tobacco var. Samsun NN) leaf cDNA library was obtained from R. Fluhr (Weizmann Institute of Science). Library screening was performed using high specific activity (approximately 2 x 10^6 cpm μg^{-1}) ^{35}S-labeled recombinant CaM as a probe (Fromm, *et al.*, 1992). Full-length clones were obtained by screening 1 x 10^6 plaque forming units of the above cDNA library with ^{32}P-labeled partial NtCBP4 DNA as a probe. Both strands of the NtCBP4 and NtCBP7 full-length cDNAs were sequenced at the Weizmann Institute Sequencing Unit (SEQ ID NOs: 1 and 3). Sequence analysis was performed using the GCG (Genetics Computer Group, Inc., Wisconsin, USA), version 9.0 (Devereux *et al.*, 1984) and the PROSITE server (European Bioinformatics Institute) software. Hydropathy was calculated by published methods (Lockow *et al.*, 1988).

Additional screenings of Potato guard cell cDNA library (given by Berned Muller-Rober, Max Planck Institute of Molecular Plant Physiology, Karl-Liebknecht-Strabetae 25, Haus 20, D-14476 Golm/Potsdam, Germany) and Alfalfa root cDNA library (given by Aviha Zilberstein, Department of Plant Sciences, Tel-Aviv university) produced clones StCBP18 (SEQ ID NO: 9) and MsCBP15 (SEQ ID NO: 11) respectively.

Isolation of cDNAs Encoding Channel-Like Proteins Using RT-PCR: Degenerate oligonucleotides corresponding to highly conserved sequences flanking the S5 and S6 transmembrane domains were used as sense and antisense primers as follows:
Sense - 5'-TGGAA(T/C)AA(A/G)AT(/C/A)TT(T/C)GT-3' (SEQ ID NO:13) and antisense - 5'-GA(A/G)ACIGCITGGG(C/G)IGG-3' (SEQ ID NO:14, I = inosine). As a template served cDNA that was synthesized from tobacco leaf poly $(A)^+$ RNA using Superscript II reverse transcriptase (GIBCO-BRL). The amplified DNA fragments were subcloned into pGEM vector (Promega) and their nucleotide sequences were determined. PCR conditions were: 1 minute at 94 °C, followed by 40 cycles of: 1 minute at 94 °C, 1 minute at 48 °C and 1 minute at 72 °C, in the presence an appropriate PCR buffer containing 2.5 mM Mg.

Bacterial expression of GST fusion proteins and antibody preparation: Extracts of *Escherichia coli* cells expressing a glutathione S-transferase (GST)-NtCBP4 (amino acids 394-708 of SEQ ID NO:2) fusion protein were separated by SDS-PAGE. An acrylamide band containing the recombinant fusion protein was excised from the gel, crushed, and mixed (1:1) with complete Freund's adjuvant. Three ml of the mixture containing

100 μg of recombinant protein were injected into two rabbits. Each rabbit was given two booster injections about two weeks apart and bled 10 days after each injection. The serum was depleted from antibodies recognizing GST by passage through a GST column (Pierce, USA).

In order to express NtCBP4 in insect cells, a transfer plasmid pFBNtCBP4 was constructed by inserting the NtCBP4 cDNA into a pFastBac vector, downstream of the baculovirus polyhedrin gene promoter. The resultant plasmids were then transformed into DH10BAC *E. coli* cells (GibcoBRL, UK) for transposition into the bacamid. The screening and isolation of recombinant bacamid DNA were according to the manufacturer's instructions.

Sf9 insect cells were transfected with recombinant bacamid DNA using CellFECTIN (GibcoBRL, United Kingdom). Recombinant baculoviruses were harvested 72 hours after the start of transfection. Subsequently, the Sf9 cells were layered at a density of 5 x 10^6 cells per 90-mm plate and infected with high recombinant baculoviruses. After the indicated days of incubation at 27 °C, cells were harvested by centrifugation at 500 x *g* for 10 minutes, washed once with PBS and the total protein was extracted. Proteins were separated by SDS-PAGE, blotted to nitrocellulose membranes and subjected to either immunodetection or [^{35}S]calmodulin-overlay by published methods (Arazi *et al.*, 1995).

Expression of NtCBP4 in Spodoptera frugiperda Sf9 Insect Cells and its Affinity Purification: Cells were cultured as previously described (Lockow *et al.*, 1988) in Graces medium supplemented with 10 % fetal calf serum and 10 μg/ml gentamycin. Recombinant proteins were expressed using the Bac-to-Bac system (Gibco-BRL). Transfer plasmid pFBNtCBP4 was constructed by inserting the NtCBP4 cDNA into a pFastBac vector. Plasmid pFBNtCBP4 was recombined *in vivo* with Bacmid DNA, resulting in a recombinant virus DNA termed vir-NtCBP4. Protein expression was carried out by infecting Sf9 cells with vir-NtCBP4. The cells were collected at 72 hours post-infection and microsomal fractions containing recombinant protein were prepared by breaking cells with a polytron in hypotonic buffer (5 mM Tris, 2 mM EDTA) at 4 °C. The homogenate was subjected to differential centrifugation (100 x *g* for 10 minutes and 100,000 x *g* for 1 hour), and the resultant microsomal fraction was collected. Recombinant NtCBP4 was solubilized by resuspending NtCBP4-containing microsomes with CaM-binding buffer (0.1 % lysophosphatidylcholine, 50 mM Hepes-NaOH (pH 7.4), 150 mM KCl and 1 mM $CaCl_2$), incubated overnight with

constant slow shaking, and centrifuged at 100,000 x *g* for 1 hour. The supernatant containing solubilized NtCBP4 was collected and loaded on a CaM-agarose (Sigma) column, which was pre-equilibrated with CaM-binding buffer. The column was washed with 10 volumes of CaM-binding buffer, and subsequently CaM-binding proteins were eluted with CaM-Binding buffer containing 2 mM EGTA instead of 1 mM $CaCl_2$.

Construction of DNA Templates Coding for Truncated NtCBP4 Proteins: Templates for amino- and carboxy-termini-deleted forms of NtCBP4 were produced by PCR amplification of the corresponding NtCBP4 coding regions, with specific oligonucleotides as primers:

AS	5'-GAAGGAATTCTAATTATCTTCAGCAGTAAAATC-3'	SEQ ID NO:19
S	5'-CTAGCGGATCCAGATAAAGCGTCATCTTTGTTTAG-3'	SEQ ID NO:20
S	5'-GTCTAGCGGATCCAAGCACTCTCAGAAGTTGAAG-3'	SEQ ID NO:21
AS	5'- GAAGGAATTCTTATGCTTGGGCAGTTCTAGTTGAG -3'	SEQ ID NO:22
AS	5'-GAAGGAATTCTTAACGAAGAGACTCTTCTACATTC-3'	SEQ ID NO:23
S	5'-GTCTAGCGGATCCGTGATGAAGAAAACAGGTTGC-3'	SEQ ID NO:24
S	5'-CCGCTCGAGCCCGGGATCAATCACCGCCAAGACGAG-3'	SEQ ID NO:25
AS	5'-CGGAATTCCCCGGGTTACCTGAGGTCTTTCGGAAGGTTG-3'	SEQ ID NO:26

S = Sense; AS = Antisense

The sense primers used for PCR amplification contained a *Bam*HI site and the antisense primers contained an *Eco*RI site. The PCR fragments were digested with *Eco*RI and *Bam*HI and then cloned into the *Bam*HI-*Eco*RI sites of a pGEX-3X vector (Pharmacia), creating an in-frame fusion of the coding sequence for GST and the deleted forms of NtCBP4. The nucleotide sequences of all cloned fragments derived by PCR amplification were confirmed by sequencing.

^{35}S-CaM Overlay Assay: Total *E. coli* extracts containing recombinant proteins were separated by SDS-PAGE, electrotransferred to nitrocellulose and overlayed with ^{35}S-recombinant CaM as previously described (Arazi *et al.*, 1995). After autoradiography, specific proteins were identified by immunodetection assays. Nitrocellulose membranes were washed for 30 minutes in PBS-T (PBS containing 0.05 % Tween 20) and 1 % non-fat milk. All additional immunostaining steps were performed as previously described (Arazi *et al.*, 1995).

CaM/Peptide Complex Detection on Non-Denaturing PAGE: Samples containing 100 pmole (1 µg) of bovine CaM (Sigma) and different

quantities of a HPLC-purified synthetic peptides (kindly provided by Prof. M. Fridkin, Department of Organic Chemistry, Weizmann Institute of Science, amino acids 595 - 617 and 1-13 of SEQ ID NO:2) in 100 mM Tris-HCl (pH 7.2), and 0.1 mM $CaCl_2$, making a total volume of 30 µl, were incubated for 1 hour at room temperature. Samples were analyzed by non-denaturing gel electrophoresis as previously described (Arazi *et al.*, 1995).

Fluorescence Measurements of the Interaction Between NtCBP4 CaM-Binding Peptide and Dansyl CaM: Dansylated bovine CaM (300 nM; Sigma) was incubated alone or with different concentrations of the C terminal derived synthetic peptide (amino acids 595 - 617 of SEQ ID NO:2) or a single concentration of the N terminal derived synthetic peptide (amino acids 1-13 of SEQ ID NO:2) in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.5 mM $CaCl_2$. After each addition of peptide, the CaM/peptide solution was mixed and incubated for 5 minutes at 23 °C. Emission fluorescence at 480 nm was then measured using a SLM AMINCO 8000 fluorimeter (SLM instruments); excitation wavelength was at 340 nm. Each measurement was the average of 3 readings. The dissociation constant (K_d) was determined by directly fitting fluorescence measurements of dansyl CaM at different peptide concentrations as previously described (Faiman *et al.*, 1998; equation-1) using Kaleidagraph (version 2.1; Synergy Software).

Detecting Protein-Protein Interactions with a Yeast Two-Hybrid System: The cDNA corresponding to the NtCBP4 C-terminal (amino acids 401-708, SEQ ID NO:2) was amplified by PCR using specific primers: 5'-CGGAATTCCAATCCTCAACATTAAGATTAGAG-3' (SEQ ID NO: 27) and 5'-TACCGCTCGAGCTAATTATCTTCAGCAGTAAAATCT-3' (SEQ ID NO:28) and cloned into pEG202 vector (bait) *Eco*RI-*Xho*I sites, to fuse with the DNA-binding domain of the bacterial repressor LexA. The cDNA of a full length plant CaM (Swiss-Prot No. P27162) was amplified by PCR using specific primers: 5'-CCGCTCGAGTCACTTGGCCA TCATGACCT-3'(SEQ ID NO: 29) and 5'-CGGAATTCATGGCGGA TCAGCTTACAGA-3'(SEQ ID NO:30), and cloned into pJG4-5 vector (prey) *Eco*RI-*Xho*I sites to create in-frame fusions with the bacterial B42 transcriptional-activator domain. Interaction trap assay was performed as previously described (Golemis *et al.*, 1997). Briefly, the yeast strain EGY48 was transformed with the LacZ reporter vector pSH18-34. The pRFHM1 vector, which encodes LexA fused to the N-terminus of the *Drosophila* protein bicoid, served as a non specific bait. Protein-protein interactions were identified by the ability of the yeast transformed with bait

and prey plasmids to grow on galactose plates lacking leucine, and by the appearance of blue and white color on the galactose-X-gal and glucose-X-gal plates, respectively. β-galactosidase activity in liquid cultures was determined as previously described (Lundblad, 1997).

Preparation of transgenic plants: DNA constructs with NtCBP4 in the sense or anti-sense orientation were prepared by cloning an *Eco*RI fragment of the entire NtCBP4 cDNA into the unique *Eco*RI site of a binary Ti plant transformation vector (Cuozzo *et al.*, 1988) downstream of the Cauliflower Mosaic Virus 35S promoter (Guilley *et al.*, 1982). The orientation of the cDNA was determined by PCR amplification using gene-specific primers corresponding to NtCBP4 and vector sequences (SEQ ID NOs:31, 32 and 33). Transgenic tobacco plants (Nicotiana tabacum var. Samsun, NN) were prepared by the leaf-disk transformation procedure (Horsch *et al.*, 1985) and selected on kanamycin (50 μg per ml). Two-month-old primary transformants were initially screened for NtCBP4 mRNA by Northern analysis and later by immunodetection of NtCBP4. NtCBP4 expression in each transgenic line was confirmed at the T2 generation stage.

Preparation of tobacco membranes, fractionation and immunodetection of proteins: Tobacco leaves or roots were homogenized using a pestle and mortar in a buffer (1.5 ml g^{-1} fresh weight) containing 0.5 M sucrose, 50 mM HEPES-KOH, pH 7.5, 5 mM ascorbic acid, 1 mM DTT, 0.6 % (w/v) polyvinylpyrrolidone, and the protease inhibitors phenylmethylsulfonyl fluoride (2 mM), aprotinin (1 μg ml^{-1}), leupeptin (1 μg ml^{-1}), and pepstatin (1 μg ml^{-1}). The homogenate was filtered through one layer of Miracloth (Calbiochem, LaJolla, CA), and the resulting filtrate centrifuged at 10,000 x *g* for 15 minutes. Microsomal membranes were pelleted from the supernatant by centrifugation at 50,000 x *g* for 30 minutes. Membrane proteins were separated by SDS-PAGE, blotted and immunodetected with polyclonal antibodies against NtCBP4 or plant vacuolar H^+-ATPase (Ward *et al.*, 1992). For fractionation by non-continuous sucrose gradients, microsomes (7.4 mg protein) were gently resuspended in 0.5 ml resuspension buffer (5 mM KPO_4^-, pH 7.8, 0.25 M sucrose) and layered onto a non-continuous gradient containing 20 %, 30 %, 34 %, 38 %, 45 % (W/W) sucrose in centrifugation buffer (10 mM Tris-MES pH 7.2, 2.5 mM DTT, 1 mM $MgSO_4$). The gradient was then centrifuged at 100,000 x *g* for 2 hours and the membranes were collected from the interface between different sucrose concentrations, frozen in liquid

nitrogen, and stored at -80 °C. For aqueous two-phase partitioning, microsomes from tobacco roots (57 grams) were gently resuspended in 0.33 M sucrose and 5 mM potassium phosphate, pH 7.8, and then fractionated by the aqueous two-phase partitioning method as described (Larsson, 1985). Phase separations were carried out in a series of 5-g phase systems with a final composition of 6.2 % (w/w) dextran T500, 6.2 % (w/w) polyethylene glycol 3350, 0.33 M sucrose and 5 mM potassium phosphate, pH 7.8. Three successive rounds of partitioning yielded a colorless upper phase enriched in plasma membranes and a brown lower phase containing intracellular membranes. After washing, pelleting, and resuspending in homogenization buffer, the partitioned membranes were stored in liquid nitrogen for further analysis. Antibodies used for immunodetection on blots were against NtCBP4 (described above), the 60-kDa subunit of vacuolar H^+-ATPase from oat roots (V-ATPase, from H. Sze, University of Maryland; (Ward et al., 1992)), endoplasmic reticulum BIP (ER-BIP, from G. Galili, The Weizmann Institute), plasma membrane H^+-ATPase (P-ATPase, from Dr. Orit Shaul, of the Weizmann Institue of Science, Israel), mitochondrial prohibitin (M-Prohibitin; Snedden and Fromm, 1997) and endoplasmic reticulum NADPH cytochrome-p450 reductase (ER-Cyt-p450, from I. Benveniste, CNRS, Institute de Biologie moleculaire des plants, Strasbourg, France).

Plant culture and metal toxicity assays: Seeds were sterilized with 70 % ethanol for 10 seconds, NaOCl (10 % active chlorine) for 15 minutes, then rinsed with distilled water and placed in plates containing either half-strength Hoagland's nutrient solution (Hoagland *et al.*, 1950), pH 5.7, for Ni treatment, or in modified Blaydes solution (Parrot *et al.*, 1990), pH 4.5, for Pb treatment. The seedlings were grown in a controlled growth room for 12 to 13 days, with a day/night cycle of 16h/8h at 25 °C. Chlorophyll content was measured as described (Arnon, 1949). For the analysis of older plants, seeds were mixed with vermiculite that was placed on nylon nets and half-strength Hoagland's nutrient solution, in the greenhouse. After four weeks, plants of uniform size were subjected to $NiCl_2$ treatments for two weeks. Calli were generated from leaves on B5 medium (Gamborg *et al.*, 1968) supplemented with 2-4-D (0.5 $\mu g\ ml^{-1}$) and kinetin (0.2 $\mu g\ ml^{-1}$), mixed with 0.8 % agar. Tolerance to Ni^{2+} was studied by transferring calli to B5 medium supplemented with $NiCl_2$ for two weeks and then to Ni-free B5 medium for another two weeks. Statistical analysis of results was by single factor ANOVA .

Metal accumulation analysis: To measure Ni accumulation, plants were grown for four weeks on 0.5 strength Hoagland's solution (pH 5.7) and then transferred to the same medium supplemented with a range of concentrations of $NiCl_2$ (0.025, 0.050. 0.1, 0.15 and 0.2 mM). For the analysis of Pb accumulation, plants were grown for 4 weeks on 0.5 strength Hoagland's solution (pH 5.7), transferred to modified Blaydes medium, pH 4.5 (Ca^{2+} concentration was 100 μM and Pi concentration was 10 μM), supplemented with 100 μM Pb. Shoots were harvested at the times indicated, rinsed briefly in deionized water, blotted and dried at 80 °C for 3 days. The dried plant material was digested as described (Kramer *et al.*, 1997) and metal content was measured by Inductively Coupled Plasma Atomic Emission Spectrometer (ICP-AES) with a spectroflame-ICP (Spectro Analytical Instruments, Germany).

Preparation of transgenic plants expressing NtCBP4ΔC: The DNA construct for the expression of the C-terminal truncated NtCBP4 (NtCBP4ΔC) was prepared by cloning a PCR-amplified DNA fragment using the gene-specific oligonucleotides 5'-CCGCTGAGCTATGAATCACCGCCAAGACGAG-3' (sense) (SEQ ID NO:44) and 5'-GAAGGAATTCTTAAGAGGCTACAAGCTTTAAATCAT-3' (antisense) (SEQ ID NO:45), spanning the coding region Met_1-Ser_{593} of NtCBP4 (Arazi *et al.*, 1999), and containing *Xho*I and *Eco*RI restriction sites on the 5′ and 3′ termini, respectively. The *Xho*I-*Eco*RI-digested DNA fragment was ligated into the corresponding sites of the binary vector described by Arazi *et al.* (1999). Transgenic lines expressing the truncated NtCBP4ΔC produced very few seeds by self-pollination and were therefore routinely hand pollinated with WT pollen. Transfer of the transgene to progeny was verified by germion kanamycin, and then by Northern hybridizations to confirm the expression of the NtCNBP4ΔC mRNA.

Analysis of transgene expression by Northern hybridizations and RT-PCR: Total RNA was isolated from 3-week-old seedlings as described (Logemann *et al.*, 1987). RNA fractionation, blotting and hybridisation conditions were as described (Chen *et al.*, 1994). A *Hind*III fragment from the cDNA of NtCBP4 (nucleotides 1113-2024, GenBank accession AF079872) was labelled with [^{32}P]dCTP by random priming and used as a probe. The same membrane was hybridised with a [^{32}P]dCTP-labelled actin probe prepared by amplifying an actin gene fragment using the tomato actin gene (GenBank accession number U60480) specific primers 5'-

TTGTGTTGGACTCTGGTGATGG-3' (sense) (SEQ ID NO:46) and 5'-AGCCAAGATAGAGCCTCCAATC-3' (antisense) (SEQ ID NO:47). Total RNA was passed through an oligo-dT column to enrich for poly-A mRNA, and reverse transcriptase was used to prepare the corresponding cDNA templates. PCR amplification of a 404-bp NtCBP4 full-length-specific sequence was performed with the primers 5'-GACGACTTCACAGTAAGCAGC-3' (SEQ ID NO:48) and 5'-CACGACTAAAAATGCACTCAATC-3' (SEQ ID NO:49) (sense and antisense, respectively), which amplify a region spanning nucleotides 2044-2447 of the NtCBP4 cDNA (GenBank accession number AF079872). The β-ATPase cDNA amplification served as a control for RT-PCR with the primers 5'-CTTACAGGTTTGACCGTGGCTGAGC-3' (sense) (SEQ ID NO:50) and 5'-TAGTGATCCTCTCCCAAAATGTGAGG-3' (antisense) (SEQ ID NO:51) designed for the gene from *Nicotiana plumbaginifolia* (GenBank accession number X02868).

EXPERIMENTAL RESULTS

Example 1

Cloning of CaM-binding Channel-Like Proteins from Nicotiana tabacum and characterization based on sequence

In order to determine the molecular components involved in Ca^{2+}/CaM-mediated signaling pathways in plants, a tobacco cDNA expression library was screened with radiolabeled recombinant CaM as a probe. Preliminary analysis revealed several partial clones containing a region with a marked similarity to the cyclic nucleotide-binding domains of mammalian cyclic AMP/GMP dependent kinases and cyclic nucleotide-gated cation channels (Shabb *et al.*, 1992). The full-length clones of two such cDNAs were obtained and their sequences revealed open reading frames of 2124 and 2106 nucleotides encoding proteins of 708 and 702 amino acid residues, respectively (see SEQ ID NOs:1, 2, 3 and 4). These clones were named NtCBP4 and NtCBP7 (for *Nicotiana tabacum* CaM-Binding Protein). The occurrence of an in-frame termination codon 5' to the translational initiation codon in each clone indicates that the NtCBP4 and NtCBP7 cDNAs contain the complete coding regions of their respective proteins. A sequence comparison of NtCBP4 and NtCBP7 revealed an overall amino acid sequence homology (identical + similar acids of 62 %. The highest level of homology (73 %) was found in their C-termini

(corresponding to NtCBP4 amino acids 348 - 708, SEQ ID NO:2) and a lower homology (50 %) in their N-termini (corresponding to NtCBP4 amino acids 1-347, SEQ ID NO:2) (Figure 1a).

Evidence for homology to ion channels at the amino acid sequence level of NtCBP4 and NtCBP7 came from several types of analyses. BLASTP searches of all databases, using the amino acid sequence of NtCBP4 and NtCBP7 as query sequences, revealed sequences similar to cyclic nucleotide-gated non selective cation channels and voltage-gated K^+ channels, all belonging to a group of ion channels that contain a cyclic nucleotide-binding domain in their C-terminus (Warmke *et al.*, 1994). The rat cyclic nucleotide-gated olfactory channel (Dhallan., 1990) received the most significant score (23 % identity and 48.5 % total homology (similarity + identity) over the whole NtCBP4 sequence. Furthermore, a conserved putative cyclic nucleotide binding site (cNBD) was located at the carboxy termini of NtCBP4 and NtCBP7 (Figures 1a-b). This domain shows similarities with cyclic nucleotide binding domains of different eukaryotic and prokaryotic proteins (Shabb *et al.*, 1992).

Some common features of a group of eukaryotic ion channels are the presence of six transmembrane domains (S1-S6) and a pore region (P) (Jan *et al.*, 1992) (Figures 2; 3 a-c). Analysis of the hydrophobic profile of NtCBP4 and NtCBP7 showed the presence of several hydrophobic stretches along the sequence (Figure 1b). Only six hydrophobic stretches, which lie in the N-terminal half of the protein, are presumably long enough (19-26 amino acids) to span the membrane and indeed, they are predicted to be transmembrane domains (according to Eisenberg, 1984). In addition, these domains are nicely aligned with the core regions (S1-S6) of cyclic nucleotide-binding domain-containing channels. The equivalent membrane-spanning regions of the *Drosophila* EAG voltage-activated K^+ channel (Warmke *et al.*, 1991) showed the highest similarity (sequence alignments not shown).

A group of ion channels that show intrinsic voltage-induced gating contain five to seven positively charged residues located within a stretch of hydrophobic amino acids that constitute the fourth transmembrane domain (S4) (Liman *et al.*, 1991). In contrast, cyclic nucleotide-gated channels have less positive charges in the S4 domain and are insensitive to voltage (Fin *et al.*, 1996). The fourth hydrophobic domain (S4) of NtCBP4 and NtCBP7 contains only 4 positively charged residues. Hence the S4 domains

of NtCBP4 and NtCBP7 are more related to the S4 domains of cyclic nucleotide-gated channels.

Thus, the predicted channel-like structure of NtCBP4 and NtCBP7 is corroborated by the presence of a putative pore region preceding the S6 hydrophobic domain (Figures 2; 3b-c). This region shares similarities with pore-lining regions of K^+-selective (Nakamura *et al.*, 1997) and cyclic nucleotide-gated non selective cation channels (Fin *et al.*, 1996) (Figures 3 a-c). However, in contrast to K^+-selective channels, the NtCBP4 and NtCBP7 pore regions lack the consensus amino acid pair YG (SEQ ID NO:40) that is essential for K^+ selectivity (Nakamura *et al.*, 1997 and Doyle *et al.*, 1998). Moreover, in both NtCBP4 and NtCBP7 the aspartate residue typically following the GYG (acids 1-3 of SEQ ID NO:43) signature is replaced by leucine, and the cluster of threonine residues that co-determines K^+ selectivity (Doyle *et al.*, 1998) is missing. This suggests that their permeability may differ from that of known K^+-selective channels. A few amino acid differences between the predicted pore regions of NtCBP4 and NtCBP7 (e.g., GQN (SEQ ID NO:41) versus GQG (SEQ ID NO:42), respectively) may suggest differences in ion permeability between these two putative channels.

Similar to these channels, NtCBP4 and NtCBP7 contain six putative transmembrane domains, a pore motif between the S5 and S6 domains, and a highly conserved cyclic nucleotide-binding domain in their C-terminus. Sequence analysis indicated that NtCBP4 and NtCBP7 are most similar to the olfactory cyclic nucleotide-gated channel (Dhallan *et al.*, 1990). Like the olfactory channel, NtCBP4 and NtCBP7 contain a 90 amino acid spacer between the end of S6 and the predicted cyclic nucleotide-binding domains. Recently, this region was shown to contain amino acid residues important for channel gating (Varnum *et al.*, 1997). In addition, the S4 domains of NtCBP4 and NtCBP7 are more closely related to the S4 domains of non selective cation channels that are voltage-insensitive (Fin *et al.*, 1996). NtCBP4 and NtCBP7-putative P domains share similarities with pore regions of non selective cation channels and K^+-selective channels (Nakamura *et al.*, 1997). However, both lack some of the residues found to be important for K^+ selectivity such as the GYG (SEQ ID NO: 43) motif (Doyle *et al.*, 1998) and thus may be selective for cations other than potassium.

Therefore, the two full-length tobacco cDNAs designated NtCBP4 and NtCBP7, encode Ca^{2+}/CaM-binding proteins with sequences similar to

K^+-selective channels like *Drosophila* EAG, *At*KAT1, and AtAKT1 (Figures 3a-c) inward rectifiers, as well as to animal cyclic nucleotide-gated non selective cation channels such as the rat olfactory channel (OCNG2).

NtCBP4 has an amino acid sequence which resembles recently cloned ion transporters from *Arabidopsis* (Kohler *et al.*, 1999; SEQ ID NOs:17 and 18, AtCNGC1) and *Hordeum vulgare* (Schuurink *et al.*, 1998; SEQ ID NOs: 15 and 16, HvCBT1). Kohler *et al.* recently demonstrated that AtCNGC1 (Figure 2, SEQ ID NO:17) can partially complement a yeast mutant that is deficient in K^+ uptake and therefore suggested that this protein belongs to a family of non-selective cation channels. However, neither of these earlier works studied the function of this or other members of this plant protein family in plants. Characterization of the function of NtCBP4, and other family members, is described herein for the first time.

The following Table provides nucleic acid identity and amino acid homology (in brackets, identical + similar acids) between sequences described herein.

GENE	NtCBP4	NtCBP7	NtCBP10	NtCBP63	StCBP18	MsCBP15	HvCBT1	AtCNGC1
NtCBP4	100 (100)	****	****	****	****	****	****	****
NtCBP7	69 (72)	100 (100)	****	****	****	****	****	****
NtCBP10	71 (76)	81 (84)	100(100)	****	****	****	****	****
NtCBP63	63 (66)	63 (62)	66 (66)	100 (100)	****	****	****	****
StCBP18	68 (71)	76 (72)	84 (83)	61 (57)	100 (100)	****	****	****
MsCBP15	72 (81)	69 (72)	70 (77)	64 (69)	67 (71)	100 (100)	****	****
HvCBT1	68 (73)	64 (68)	67 (73)	65 (61)	63 (67)	68 (71)	100 (100)	****
AtCNGC1	72 (82)	68 (72)	71 (78)	64 (64)	66 (73)	72 (81)	68 (72)	100 (100)

Analysis was performed using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

Example 2

Discovery of a Family of Proteins Related to NtCBP4 and NtCBP7 in Tobacco and other plants

In addition to NtCBP4 and NtCBP7, another related partial clone designated NtCBP10 (SEQ ID NOs:5 and 6) was isolated from the tobacco library by protein-interaction screening. To explore the possible occurrence

of other tobacco proteins related to NtCBP4 and NtCBP7, RT-PCR with degenerated PCR primers was used to isolate related cDNAs based on similarities among NtCBP4, NtCBP7, and NtCBP10 in the regions flanking the S5 through S6 transmembrane domains. These screenings produced another related cDNA, designated NtCBP63 (SEQ ID NOs: 7 and 8). Figure 4 shows the sequences of the four related tobacco proteins in the region spanning the S5 through S6 transmembrane domains. The major differences between the clones, including insertions or deletions, occur in the region between S5 and the pore (Figure 4), whereas the S5, pore, and S6 domains are relatively conserved in sequence. Thus tobacco possesses a family of proteins related to NtCBP4 and NtCBP7.

Within this family, two of the four tobacco clones (NtCBP4 and NtCBP10; Figure 4; SEQ ID NOs:2 and 6) have a GQN (SEQ ID NO:41) sequence instead of GYG (acids 1-3 of SEQ ID NO:43), whereas the two other tobacco proteins (NtCBP7 and NtCBP63; Figure 3; SEQ ID NOs: 4 and 8) have a GQG (SEQ ID NO:42) sequence at the same position. This, as well as other differences in amino acid sequences (Figure 4), demonstrate potential differences in ion permeability between members of this new tobacco protein family which suggest that different family members may be suited to different future applications.

Additional screenings of Potato guard cell cDNA library and Alfalfa root cDNA library produced additional NtCBP4 family members (SEQ ID NOs:9, 10, 11, and 12), showing that the family is widespread in the plant kingdom.

Example 3
Demonstration of the Ability of CaM to Bind to the Full-Length NtCBP4

Elucidation of the potential role of CaM in regulating NtCBP4 and NtCBP7 focused first on NtCBP4. Initially, examination of whether the full-length NtCBP4 binds CaM was undertaken. The baculovirus-infected Sf9 insect cell system (Lockow *et al.*, 1988) was employed to express NtCBP4. As predicted, the expressed NtCBP4 was found solely in the insect cell membrane fraction (data not shown). Membrane proteins were solubilized with 0.1 % LPC as described hereinabove and loaded on a CaM-agarose column in the presence of 1 mM Ca^{2+}. After washing the column, CaM binding proteins were eluted by exchanging Ca^{2+} with 2 mM EGTA. An equal volume from each fraction was analyzed by SDS-PAGE. Staining

with Coomassie-Blue revealed two major bands at a relative mobility of approximately 80 kDa (Figure 5A), consistent with the predicted M_r of NtCBP4 (81 kDa). Western blot analysis with anti-NtCBP4 antibodies confirmed that these bands represent the recombinant NtCBP4 (Figure 5B). Proteins from control insect cells infected with wild type baculovirus showed no cross reactivity with anti-NtCBP4 antibodies (data not shown). These data indicate that the full-length NtCBP4 is capable of binding CaM in its phospholipid-solubilized form.

To test the capacity of NtCBP4 to bind CaM *in vivo,* the heterologous yeast two-hybrid system described hereinabove was employed. A full-length CaM coding region was used as the prey with NtCBP4 C-terminal (amino acids 401-708, SEQ ID NO:2) as bait. Protein-protein interactions were identified by yeast growing on galactose plates lacking leucine and by the appearance of blue on the galactose-X-gal plates as opposed to white on glucose-X-gal plates (Figure 6). The results indicate that a CaM-binding domain is present in the C-terminal part of NtCBP4. In addition, these results show the ability of NtCBP4 to interact with CaM *in vivo*, in the yeast system.

Example 4

Demonstration that NtCBP4 Contains a Single CaM-Binding Domain

CaM-binding domains are not conserved in amino acid sequences (O'Neil *et al.*, 1990). Therefore, to determine the number and positions of CaM binding sites in NtCBP4, different regions of NtCBP4 were fused to GST (Figure 7B constructs Nos. 2, 3, 4, 5, 6 and 7) and their binding to ^{35}S-CaM was tested using a CaM-overlay assay (Arazi *et al.*, 1995). A minimal region of 66 amino acids (SEQ ID NO: 2 residues Q^{573}-R^{639}; Figure 6*B*: construct No. 6) was sufficient for CaM binding on a blot (Figure 7A; lane 5). This binding was Ca^{2+} dependent as evidenced by the fact that it could be blocked with 2 mM EGTA (data not shown). Thus, the NtCBP4 Q^{573}-R^{639} domain contains the CaM binding site.

CaM binds to peptides that tend to form amphipathic α helices with one face of the helix positively charged (O'Neil *et al.*, 1990). Sequence analysis of the 66-amino acid CaM binding sequence, revealed a region with typical CaM binding characteristics between amino acid residues Phe^{595}-Gly^{617}. When drawn in the form of an α-helical wheel (Devereux *et al.*, 1984), it exhibits an amphipathic structure with a positively charged binding face and an opposite hydrophobic face (Figure 8A). Subsequently,

a synthetic peptide corresponding to NtCBP4 amino acids Phe^{595}-Trp^{617} (SEQ ID NO:2; acids 595-617) was synthesized, incubated with bovine CaM and complex formation was determined by non-denaturing PAGE (Arazi *et al.*, 1995) with or without 0.1 mM Ca^{2+}. In the presence of a peptide derived from the N-terminal part of NtCBP4 (SEQ ID NO: 2; amino acids 1-13), which was used as a negative control (Figure 8B, control lane), or in the absence of the peptide (Figure 8B, lane 0), only a single band corresponding to free CaM was apparent. When the NtCBP4 peptide corresponding to amino acids Phe^{595}-Gly^{617} was added in the presence of Ca^{2+}, another band of slower mobility appeared, representing a peptide/CaM complex. When the ratio of peptide to CaM was unity, most of the free CaM was no longer apparent, and the intensity of the peptide/CaM complex increased. At a peptide-to-CaM molar ratio of 2, very low free CaM was detected. At a still higher ratio of 4, no new band appeared on the gel, nor did the peptide/CaM complex change its intensity. The conclusion is that multivalent complexes did not form. No mobility shift of CaM was apparent in the absence of free Ca^{2+} (data not shown). These results demonstrate that this peptide is capable of forming a stable complex with CaM in the presence of Ca^{2+} and suggest that NtCBP4 binds Ca^{2+}/CaM with a one-to-one stochiometry.

As mentioned above, a protein (designated HvCBT1, Figure 2, SEQ ID NO:16) which is similar to NtCBP4 was isolated recently from barley (Schuurink *et al.*, 1998). In that study the CaM-binding site was not determined experimentally and, based on sequence analysis, it was assumed to be closer to the C-terminal end of the protein (Schuurink *et al.*, 1998), where no CaM-binding activity occurs in NtCBP4. The full-length NtCBP4 can bind CaM in a Ca^{2+}-dependent manner and this binding is via a single CaM-binding domain of approximately 23 amino acid (Phe^{595}-Trp^{617}). This domain is characterized by aromatic and long-chain aliphatic residues separated by positively charged amino acids. These features are present in known CaM-binding sites of several proteins such as myosin light chain kinase, and peptides such as melittin and mastoparan (O'Neil *et al.*, 1990). However, except for NtCBP4 and its family members, the olfactory channel (SEQ ID NO:34) is the only identified protein containing both CaM- and cyclic nucleotide-binding domains. In contrast to NtCBP4, in which the CaM-binding domain coincides with α–helix C of its cyclic nucleotide binding domain, in the olfactory channel the CaM-binding domain is found in the N-terminal part of the protein, separated from the cyclic nucleotide

binding domain by 377 amino acid residues. Binding of Ca^{2+}/CaM to this domain in the olfactory channel reduces the effective affinity for cyclic nucleotides by up to 20 fold (Liu *et al.*, 1994).

Example 5

NtCBP4 Binds Ca^{2+}/CaM with an Apparent K_d in the Low nM Range

Demonstrating high affinity interactions between CaM and a protein *in vitro* can indicate its physiological importance. Therefore, investigations of the affinity between the NtCBP4 CaM-binding peptide and Ca^{2+}/CaM using fluorescence measurements of dansyl-CaM (Liu *et al.*, 1994) with or without the NtCBP4 CaM binding peptide were undertaken. Without the peptide, the photoexcited emission spectrum of dansyl-CaM (300 nM) peaked at 500 nm (Figure 9A). Adding 300 nM peptide, which should convert all CaM to the bound form, increased the fluorescence intensity of dansyl-CaM 1.96 times, and shifted the emission peak to 482 nm (Figure 9B). These experiments demonstrate that the dansyl moiety enters a more hydrophobic environment upon binding of the peptide to dansyl-CaM (Kinkaid *et al.*, 1982). With the dansyl-CaM concentration at 200 nM, the fraction of bound CaM increased linearly with total peptide concentration until the signal saturated at 300 nM peptide, 1.5 times the concentration of dansyl-CaM (Figure 9B). These data are consistent with the non-denaturing gel results estimating a one-to-one binding ratio between CaM and NtCBP4 (Figure 8B). An apparent K_d of 7.9 ± 2.75 nM was calculated from fitting these data using a method described hereinabove (Faiman *et al.*, 1998). These data indicate that NtCBP4 binds CaM with a binding constant that is in the physiological range known for other plant- (Snedden *et al.*, 1996) and animal- (James *et al.*, 1995) CaM-regulated proteins.

Example 6

Demonstration that the CaM-Binding Domain of NtCBP4 Coincides with a Conserved Helix of its Cyclic Nucleotide Binding Domain

Proteins that bind cyclic nucleotides (cAMP or cGMP) share a structural domain of about 100-130 residues. The best studied is the prokaryotic catabolite gene activator (CAP). Other proteins that are known to contain these domains are animal cAMP- and cGMP-regulated protein kinases, vertebrate cyclic nucleotide-gated ion channels (Shabb *et al.*, 1992) and several EAG-related K^+ channels (Warmke *et al.*, 1994). X-ray crystallography of CAP showed that this domain is composed of three α-

helices and has a distinctive eight-stranded, antiparallel beta-barrel structure (Weber *et al.*, 1987). In addition, in the cyclic nucleotide-binding domains of all eukaryotic and prokaryotic proteins, there are six invariant amino acid residues, three of which are Gly residues, thought to be essential for maintaining the structural integrity of the β-barrel. The cyclic nucleotide binds within a pocket formed by the α–helix C and the β–barrel, where the remaining invariant residues are located (Weber *et al.*, 1987). The NtCBP4 and NtCBP7 cyclic nucleotide-binding domains contain all the invariant amino acid residues and they fit perfectly the PROSITE data bank cAMP/cGMP binding-domain signature motif (Figure 10). The only variation from the signature is the number of amino acids separating NtCBP4 Leu^{553} and Arg^{570} (i.e., 16 amino acids instead of 5 or 11 in the signature motif). However, this region is included in a loop that is not important for the binding of the cyclic nucleotide (Shabb *et al.*, 1992).

When the NtCBP4 and NtCBP7 putative cyclic nucleotide binding sites are aligned with the olfactory channel cyclic nucleotide binding site (Figure 10) and both are superimposed with elements of a secondary structure taken from the *E. coli* CAP crystal structure (Weber *et al.*, 1987), a unique feature of the NtCBP4 cyclic nucleotide binding domain is revealed. The stretch of amino acids that form the α–helix C (ARG^{597}-Thr^{616}; see Figure 1A), a key element in the structure of cyclic nucleotide-binding domains, is exactly the region that forms the NtCBP4 CaM-binding domain (Figure 10, bold letters). Thus, the same structure functions both as a CaM-binding helix and as a key structure in binding cyclic nucleotides in NtCBP4 and NtCBP7.

In summary, Examples 1 through 6 demonstrate identification of a channel-like protein family in plants with a new structural motif not previously described in any other protein. In this motif, a phylogenetically conserved helix in the cyclic nucleotide-binding domain has the ability to bind CaM with high affinity, and thus makes NtCBP4 and related proteins a potential communication point for cross-talk between Ca^{2+} and cyclic nucleotide signal transduction pathways in plants.

Example 7

Demonstration that NtCBP4 is Associated with the Tobacco Plasma Membrane

Ion channels have been discovered in various intracellular membrane compartments of plant cells (Tester, 1990). In order to study the function of

NtCBP4 *in planta*, its subcellular localization was determined. To facilitate this determination, polyclonal antibodies against the recombinant NtCBP4 were prepared as described hereinabove. Confirmation that these antibodies detect the full-length NtCBP4 came from both a heterologous insect-cell expression system (Figure 11A) and from plant membranes (Figure 11B). In insect-cell extracts, anti-NtCBP4 antibodies (Figure 11A) detected two close protein bands (80-81 kDa). The intensity of the immunoreactive proteins increased with time post-infection (Figure 11A). No protein was detected with the anti-NtCBP4 antibodies in the control extract infected with the WT baculovirus vector (Figure 11A). Incubation of the same blot with [^{35}S]calmodulin confirmed that the full-length recombinant NtCBP4 protein binds calmodulin (Figure 11A). Moreover, a protein of the same electrophoretic mobility as the full-length recombinant NtCBP4 was detected in membranes of transgenic plants expressing NtCBP4 under the transcriptional control of the Cauliflower Mosaic Virus (CaMV) 35S gene-promoter (Figure 11B). No protein was detected in the soluble fraction of these plants. Thus, this analysis confirmed the specificity of the antibodies against NtCBP4.

Subsequently, the subcellular localization of NtCBP4 in wild type (WT) and in transgenic plants (transgenic line designated 49-79 was chosen for this analysis) was determined. First, microsomal membranes were fractionated on sucrose gradients and analyzed by Western blots using antibodies against NtCBP4 and against known marker proteins from plasma membrane (P), vacuole (V), and endoplasmic reticulum (ER) (Figure 12A). NtCBP4 co-fractionated with the plasma membrane markers as indicated by sedimentation profiles overlapping with P-ATPase (Figure 12A). In contrast, the vacuolar marker peaked at the 20/30 % sucrose fraction. BiP, a commonly used ER marker (Haas, 1994) also peaked at the 34/38 % sucrose fraction; however, unlike NtCBP4, it was also present in the 20/30 % fraction. These results demonstrate that NtCBP4 is associated with the plasma membrane and not with the vacuole or ER. Immunodetection of NtCBP4 in membrane fractions from transgenic tobacco overexpressing NtCBP4 showed a similar pattern of distribution on the sucrose gradient, except that NtCBP4 levels were higher (Figure 12A). This indicates that the overexpressed NtCBP4 is localized in same membrane compartment as the endogenous NtCBP4 in WT plants.

To confirm the localization of NtCBP4 in the plasma membrane, separation of tobacco root plasma membranes from intracellular membranes

using the aqueous two-phase partitioning method (Larson, 1985) was undertaken. Membrane purity was confirmed by immunodetection of the marker proteins. As shown in Figure 12B, NtCBP4 in both WT and transgenic plants was highly enriched in the upper-phase fraction (U) and largely depleted from lower-phase fraction (L). A similar pattern was detected for P-ATPase. The purity of the plasma membrane preparation was also confirmed by the relative absence of intracellular membrane markers like V-ATPase (Ward *et al.*, 1992) and mitochondrial prohibitin (Snedden *et al.*, 1997), which were highly enriched in the lower-phase fraction (Figure 12B). The ER marker was also detected in the upper-phase fraction, but unlike the P-ATPase, was also present in the lower phase. These results are consistent with analysis by sucrose gradients and confirm that NtCBP4 is a plasma membrane protein.

The nature of the interaction of NtCBP4 with the plasma membrane was also investigated using various detergents and non-detergent chemicals (data not shown). Specifically, NtCBP4 was removed from WT and transgenic plant membranes only by detergents that completely solubilize the membranes (e.g., 1 % Triton X-100) and not by treatments known to remove membrane-associated or peripheral proteins (e.g., 1 M Urea, 0.1 M Na_2CO_3 pH 11.5, and 0.75 M NaCl). This confirms the results of immunologic assays and sequence analysis that predicted six transmembrane domains. Together, these data constitute a demonstration that NtCBP4 is an integral plasma membrane protein.

These results concur with earlier findings of Schuurink *et al.* (Schuurink *et al.*, 1998), which demonstrated that an overexpressed GFP/HvCBT1 fusion protein was associated with the plasma membrane of barley aleurone cells. In conclusion, these data support the postulated function of NtCBP4 as a component of a plasma membrane ion channel.

Example 8

Transgenic Seedlings Exhibit Improved Tolerance to Ni^{2+}

To elucidate the function of NtCBP4 *in planta*, several transgenic tobacco lines were established that express the NtCBP4 RNA in the sense orientation under the transcriptional control of the promoter of the CaMV 35S gene. Twenty-five independent transgenic lines were screened for the presence of NtCBP4 RNA, and these showing relatively high levels were selected for further investigation at the protein level. As a control, transgenic tobacco lines were prepared with the NtCBP4 cDNA in the

antisense orientation. Several independently derived transgenic lines were obtained, differing in their level of NtCBP4. Figure 13A shows an example of some of these lines. Two lines (49-79 and 10-2) had about double the amount of NtCBP4 protein compared to the WT plants. One line (13-93) had intermediate levels of NtCBP4 and one line (15-89) exhibited levels of NtCBP4 similar to WT (Figure 13A). The transgenic line with the antisense construct (AS) had about the same levels of NtCBP4 as WT and the 15-89 transgenic line. Transgenic lines overexpressing NtCBP4 and control plants were then subjected to different physiological studies.

Under normal growth conditions in the greenhouse, none of the transgenic plants showed any aberrant phenotype compared to the WT plants, and NtCBP4 overexpression was genetically transmitted to the progeny. Because NtCBP4 was assumed to function as a plasma membrane cation channel, the WT and transgenic plants were challenged with exposure to different toxic metals including Na^{+}, Ba^{2+}, Cd^{2+}, Co^{2+}, Cu^{2+}, La^{3+} and Mn^{2+} as chloride salts; Pb^{2+} as $Pb(NO_3)_2$ and Zn^{2+} as $ZnSO_4$. Interestingly, transgenic plants overexpressing NtCBP4 were specifically more tolerant to $NiCl_2$ but more sensitive to $Pb(NO_3)_2$. No differences were found between WT and the transgenic lines in their responses to the other metal salts. This demonstrates that transgenic plants that overexpress NtCBP4 have an improved tolerance to Ni^{2+} and a hypersensitivity to Pb^{2+}. These observations were followed by more detailed analysis of transgenic plants, as described hereinbelow.

The relationship between the levels of NtCBP4 in transgenic lines and the effects of Ni^{2+} on early seedling development was investigated. Without Ni^{2+}, seedling development was similar in WT and in the transgenic lines (Figure 13B). Increasing $NiCl_2$ concentrations adversely affected seed development of WT, AS, and line 15-89, in which the NtCBP4 protein levels were indistinguishable. Specifically, with increasing $NiCl_2$ concentrations, seedlings of these lines became more chlorotic and their development was severely retarded. These symptoms are known to occur in plants with phytotoxic levels of Ni^{2+} (Woolhouse, 1983; Brune *et al.*, 1995; and Brune *et al.*, 1995b). In contrast, the seedlings of transgenic lines 49-79 and 10-2 that expressed higher levels of NtCBP4 (Figure 13A) showed fewer toxic symptoms, and grew and developed even at a concentration of 200 μM $NiCl_2$ (Figure 13B and 13C). Four additional transgenic lines with relatively high levels of NtCBP4 had phenotypes similar to that of lines 49-79 and 10-2 (not shown). Seedlings of line 13-93,

which had intermediate level of NtCBP4 (Figure 13A), showed intermediate symptoms of toxicity (Figure 13B).

For a quantitative evaluation of Ni^{2+} tolerance, three parameters typically used to describe metal toxicity in plants were monitored: chlorophyll content, root length and seedling biomass. In the presence of $NiCl_2$, chlorophyll content was the highest in seedlings of transgenic lines that expressed higher levels of NtCBP4 (i.e., lines 49-79 and 10-2; Figure 14A). Chlorophyll levels in these two lines were double that of WT or in transgenic seedlings with the NtCBP4 antisense construct (AS; Figure 14A). These differences were statistically significant at $p<0.001$. Root length of 12-day-old seedlings grown in the presence or absence of 100 μM $NiCl_2$ was also measured. WT root growth was inhibited by 50 % (compared to the control without $NiCl_2$), whereas root growth of lines 49-79 and 10-2 was inhibited by only 30 % and 15 %, respectively (Figure 14B). Analysis of whole seedling fresh weight (Figure 14C) revealed a 50 % inhibition (LC_{50}) at 0.09 mM Ni^{2+} for WT, versus 0.20 mM for transgenic lines 10-2 and 49-79 (Figure 14C). These differences were statistically significant ($p<0.001$) at Ni^{2+} concentrations above 0.05 mM. Taken together, these results show that transgenic lines with the highest levels of NtCBP4 were the most tolerant to Ni^{2+} stress.

Example 9

Demonstration that Tolerance to Ni^{2+} is not Dependent on the Developmental Stage of the Plant

In order to demonstrate that Ni^{2+} tolerance is not stage specific, the Ni^{2+} tolerance of 4-week-old plants was tested. After the plants had been grown in a Ni^{2+} free medium for 4 weeks, they were transferred to the same medium supplemented with 200 μM $NiCl_2$. WT plants showed severe toxicity symptoms after two weeks of exposure; they became necrotic and ceased to grow (Figure 15A). In contrast, transgenic plants showed attenuated symptoms and continued to grow (Figure 15A). Without $NiCl_2$, both WT and transgenic plants grew at the same rate (data not shown).

The response of calli derived from WT and transgenic plants to Ni^{2+} stress was also examined. Calli were exposed to different concentrations of $NiCl_2$ for two weeks, and then transferred to a Ni-free medium for two weeks to recover. Without $NiCl_2$, both WT and transgenic calli grew to the same size (Figure 15B). After an exposure to 0.5 mM $NiCl_2$, the recovery of WT calli was poor compared with the transgenic calli, which showed

only slight growth retardation. WT calli exposed to 1.0 mM $NiCl_2$, became brown and could not recover on the Ni-free medium (Figure 15B), whereas transgenic calli exposed to the same concentration of $NiCl_2$ showed only about 50 % growth retardation compared with unstressed calli (Figure 15B). Western blot analysis confirmed that transgenic calli had high levels of NtCBP4 compared with that in WT calli (Figure 15C). Therefore, Ni^{2+} tolerance is not dependent upon a specific developmental stage of the plant and is exhibited even in undifferentiated transgenic cells with high levels of NtCBP4.

Example 10

Demonstration that Ni^{2+} Tolerance is Associated with Reduced Ni^{2+} Accumulation in Transgenic Plants

One of the possible explanations for the improved tolerance of transgenic plants to Ni^{2+} may be the reduced uptake rates and slower accumulation of Ni. To test this hypothesis, the concentration-dependent accumulation of Ni in shoots of WT and transgenic plants was determined. Plants were exposed to different concentrations of $NiCl_2$, and Ni content in shoots was measured by ICP-AES (Experimental procedures). At concentrations up to 100 μM $NiCl_2$, shoots of WT and transgenic plants accumulated Ni at about the same rate. However, at higher concentrations of $NiCl_2$ (150 and 200 μM), Ni^{2+} accumulated more rapidly in the WT compared with transgenic plants (Figure 16). At even higher concentrations (300 μM and above) WT but not transgenic plants wilted after one day. Therefore, no measurements were taken at these concentrations. This analysis revealed slower accumulation of Ni in transgenic plants compared with WT. These results are consistent with previous reports on the critical concentrations of Ni necessary to confer toxic symptoms in plant tissues (Chaney *et al.*, 1997) and validate the hypothesis that increased tolerance to high concentrations of environmental Ni is due to reduced rates of Ni uptake and accumulation.

Example 11

Demonstration that Pb^{2+} Hypersensitivity is Associated with Enhanced Accumulation in Transgenic Plants

As indicated previously, preliminary analysis showed that transgenic seedlings overexpressing NtCBP4 have a greater sensitivity to Pb^{2+} than do WT plants. In the presence of 1 mM $Pb(NO_2)_2$, transgenic seedlings (line

49-79) were chlorotic and their development was retarded compared with WT seedlings (Figure 17A), whereas the growth and chlorophyll content in WT and transgenic plants were indistinguishable in the absence of $Pb(NO_2)_2$ (Figure 17A). Subsequently, chlorophyll was measured in WT and different transgenic lines after 12-days of growth in the presence of 1 mM $Pb(NO_2)_2$. The chlorophyll content was the lowest in the transgenic lines with relatively high levels of NtCBP4 (e.g., lines 49-79 and 10-2; Figure 17B). These values were significantly different ($p<0.001$) from these of WT or AS. Dose-response growth studies based on fresh weight (data not shown) revealed that transgenic line 49-79 had a 50 % greater sensitivity to $Pb(NO_2)_2$ than did WT (LC50 of 0.6 and 0.9 mM, respectively).

One explanation for the Pb^{2+} hypersensitivity of transgenic plants is enhanced Pb^{2+} accumulation. To test this possibility, 4-week-old WT and transgenic lines were exposed to 0.1 mM $Pb(NO_2)_2$ for 1-2 days and determined Pb content in shoots by ICP-AES. This analysis revealed that Pb accumulated faster in shoots of the transgenic plants (Figure 17C). This observation confirms the hypothesis that increased Pb sensitivity is due to increased uptake and accumulation. This means that transgenic plants overexpressing NtCBP4 are uniquely suited to phytoremediation of lead from soil.

Example 12

Transgenic plants expressing a truncated NtCBP4 exhibit improved tolerance to Pb^{2+}

The C-terminal half of NtCBP4 contains a structurally conserved putative cyclic nucleotide binding domain and a high-affinity calmodulin-binding site (Arazi *et al.*, 2000). In contrast to the mammalian cyclic nucleotide-gated channel, where the calmodulin-binding domain is near the N-terminus of the protein (Varnum and Zagotta, 1997), in NtCBP4 the calmodulin-binding site coincides with the predicted αC-helix structure of the cyclic nucleotide-binding domain (Arazi *et al.*, 2000). This helix is essential for cyclic nucleotide binding in related proteins from other organisms (Shabb and Corbin, 1992). It is reasoned that removal of this important helix should disrupt the binding of both cyclic nucleotides and calmodulin and could possibly render the channel protein inactive. Moreover, if plant NtCBP4-containing channels function as tetramers, like the mammalian cyclic nucleotide-gated channels (Liu *et al.*, 1996),

expression of an inactive subunit might have inhibitory effects on endogenous native NtCBP4-associated channel activities. As a first step to test this possibility, a truncated version of NtCBP4 was constructed from which a major part of the C-terminal half including the αC helix, was removed (designated NtCBP4ΔC; Figures 18A-B).

Transgenic plants expressing NtCBP4ΔC were prepared as described in Experimental procedures, selected on kanamycin, and transferred to the greenhouse for seed production. Most of the putative transgenic lines were found to be male sterile and were therefore pollinated manually with pollen from WT plants. Seeds were collected from the maternal plants and plated on kanamycin-containing medium to select for progeny carrying the transgene. Apparently, all lines that were resistant to kanamycin were derived from male sterile plants. These plants expressed the truncated NtCBP4ΔC transgene, as will be shown. This suggests that NtCBP4ΔC had detrimental effects on functions crucial for plant reproduction.

To investigate the possible effects of NtCBP4ΔC expression on plant response to Pb^{2+}, NtCBP4ΔC transgenic lines were germinated in the presence of $Pb(NO_3)_2$ and their growth was compared with that of WT plants and of the transgenic line 49-79, which overexpresses the full-length NtCBP4 (designated herein NtCBP4FL or NtCBP4) and is associated with Pb^{2+} hypersensitivity (Arazi *et al.*, 1999). Several independent NtCBP4ΔC lines designated 42-11, 22-11, 29, 33-11, 16-6, 45-16, 52-36, 55-16, 9-2, 36-16, 55-61 were found to be more tolerant of Pb^{2+} than WT plants. Figures 19A-B show three of these NtCBP4ΔC lines, with the WT and NtCBP4FL control lines, grown with or without Pb^{2+}. The NtCBP4ΔC plants were clearly more tolerant of the toxic metal than the WT seedlings, whereas the NtCBP4FL plants were more sensitive, as previously reported (Arazi *et al.*, 1999). The extent of tolerance was not identical in all NtCBP4ΔC transgenic lines, which could be attributed to differences in expression levels of the transgene or other reasons. Nevertheless, several NtCBP4ΔC independent transgenic lines exhibited a phenotype contrasting that of NtCBP4FL transgenic plants with respect to their response to Pb^{2+}. No differences were found between WT and the transgenic lines in their response to other metals including Na^{+}, Zn^{2+}, Mn^{2+}, Cd^{2+} and La^{3+}.

Although the expression of NtCBP4FL and NtCBP4ΔC in transgenic plants is driven by the same 35S CaMV promoter, it was desired to confirm that their expression levels were comparable. In addition, it was desired to exclude the possibility that the expression of the NtCBP4ΔC mRNA

abolished the expression of the endogenous native gene (e.g., by a gene-silencing mechanism). Northern hybridisations with total-RNA samples were used for analysing the steady-state levels of the NtCBP4FL and NtCBP4ΔC mRNAs in representative transgenic lines (Figure 20A). The mRNAs of both transgenes were easily detected and distinguishable by their different lengths (727 bases difference) in transgenic lines expressing the full-length mRNA versus that in NtCBP4ΔC lines (ΔC-42-11, ΔC-22-11 and ΔC-29). These results indicate that the steady-state levels of the mRNAs for both transgenes are comparable, although some variation in mRNA levels between lines is apparent. Thus, the contrasting phenotypes of the NtCBP4FL and NtCBP4ΔC transgenic plants are certainly not the result of differences in expression levels. The endogenous native NtCBP4 mRNA in WT plants was not detected in the total RNA sample due to its very low abundance (T. Arazi and H. Fromm, unpublished results) (Figure 20A).

To ascertain that the expression of the endogenous native gene in NtCBP4ΔC transgenic plants was not silenced, RT-PCR amplification was performed with a set of primers that amplify a 404-bp region of the full-length NtCBP4 mRNA but not of the truncated mRNA (see, Experimental procedures). The RT-PCR results (Figure 20B) show that a 404-bp DNA band was amplified from mRNA of the NtCBP4ΔC plants, indicating that the endogenous native NtCBP4 mRNA was indeed present in these plants. This amplified band had the same mobility as the control band amplified from the NtCBP4 cDNA clone. Although this RT-PCR analysis did not allow quantitative assessment, the results show that with the same amount of mRNA used for each sample and with the same primers, the DNA band amplified from the mRNA of NtCBP4ΔC plants was similar in intensity to that amplified from the WT plants (Figure 19B), and both were substantially less intense than the amplified band from the NtCBP4FL plants. Where reverse transcriptase was omitted (-RT), no amplification was detected, indicating that the observed amplifications described above originated from mRNA templates rather than from residual genomic DNA or from contaminating plasmid DNA in the samples. Amplification with β-ATPase-specific primers indicated that similar amounts of poly-A^+ mRNA were used in all reactions. Thus, it is concluded that expression of NtCBP4ΔC does not result in silencing of the endogenous NtCBP4 gene. Rather, the apparent phenotype conferred by NtCBP4ΔC occurs in the presence of the full-length native endogenous NtCBP4.

Example 13
Quantitative assessment of Pb tolerance in NtCBP4ΔC transgenic seedlings, and Pb accumulation

The extent of growth inhibition of seedlings by Pb^{2+} was evaluated by determining the fresh weight of seedlings grown in different concentrations of $Pb(NO_3)_2$ (Figure 21A). Whole-seedling fresh weight analysis revealed 50% inhibition at 0.34 mM, 0.47 mM, 0.58 mM, 0.62 mM Pb^{2+} for NtCBP4FL, WT, NtCBP4ΔC-22-11 and NtCBP4ΔC-42-11, respectively (Figure 21A). Moreover, statistical analysis revealed that the two NtCBP4ΔC lines were significantly more tolerant ($p<0.05$) than the WT seedlings at Pb^{2+} concentrations of 0.1 mM and more.

One of the possible explanations for the Pb^{2+}-tolerant phenotype of the NtCBP4ΔC plants may be the attenuated uptake of the metal, since is has already been demonstrated that Pb^{2+} hypersensitivity in NtCBP4FL plants is associated with enhanced accumulation of Pb^{2+} (Arazi *et al.*, 1999). To test this possibility, Pb^{2+} accumulation was determined by ICP-AES in 12-day-old seedlings grown in the presence of 0.2 mM $Pb(NO_3)_2$. NtCBP4ΔC plants (lines ΔC-42-11 and ΔC-22-11) showed a marked reduction in Pb^{2+} accumulation compared with WT seedlings, whereas NtCBP4FL plants accumulated substantially more Pb ($p<0.05$) than WT (Figure 21B).

Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims. All publications, patents, patent applications and sequences identified by an accession number mentioned in this specification are herein incorporated in their entirety by reference into the specification, to the same extent as if each individual publication, patent, patent application or sequence was specifically and individually indicated to be incorporated herein by reference. In addition, citation or identification of any reference in this application shall not be construed as an admission that such reference is available as prior art to the present invention.

REFERENCES:

1. Arazi, T., Baum, G., Snedden, W. A., Shelp, B. J., and Fromm H. (1995) *Plant Physiol.* 108, 551-561.
2. Arazi, T., Sunkar R., Kaplan, B. and Fromm, H. (1999) A tobacco plasma membrane calmodulin-binding transporter confers Ni^{2+} tolerance and Pb^{2+} hypersensitivity in transgenic plants. *Plant J.* 20, 171-182.
3. Arazi, T., Kaplan, B. and Fromm, H. (2000) A high-affinity calmodulin-binding site in a tobacco plasma membrane channel protein coincides with a characteristic element of cyclic nucleotide-binding domains. *Plant Mol. Biol.* (in press)
4. Arnon, D.I. (1949) *Plant Physiol.* 24, 1-15.
5. Ayres, R.U. (1992) *Proc. Natl. Acad. Sci. USA* 89, 815-820.
6. Baker, A.J.M. (1987) *New Phytol.* 106, 93-111.
7. Baum, G., Chen, Y., Arazi, T., Takatsuji, H. and Fromm, H. (1993) *J. Biol. Chem.* 268, 19610- 19617.
8. Baum, G., Lev-Yadun, S., Fridmann, Y., Arazi, T., Katsnelson, H., Zik, M., and Fromm, H. (1996) *EMBO J.* 15, 2988-2996
9. Bei, Q. and Luan, S. (1998) . *Plant J.* 13, 857-865.
10. Bowler, C., and Chua, N. H. (1994) *Plant Cell* 6, 1529-1541
11. Bowler, C., Neuhaus, G., Yamagata, H., and Chua, N.-H. (1994) *Cell* 77, 73-81
12. Braam, J. and Davis, R.W. (1990) . *Cell*, 60, 357-364.
13. Brune, A. and Dietz, K-J. (1995) *J. Plant Nutr.* 18, 85-868.
14. Brune, A., Urbac, W. and Dietz, K-J. (1995) *New Phytol.* 129, 403-409.
15. Bush, D. S. (1995) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 46, 95-122
16. Chaney, R.L., Malik, M., Li, Y.M., Brown, S.L., Brewer, E.P., Angle, J.S. and Baker, 52. A.J.M. (1997) *Curr. Opin. Biotechnol.* 8, 279-283.
17. Chen Y, Baum G. and Fromm, H. (1994) The 58-kilodalton calmodulin-binding glutamate decarboxylase is a ubiquitous protein in petunia organs and its expression is developmentally regulated. *Plant Physiol.* 106, 1381-1387.
18. Clapham, D. E. (1995) *Cell* 80, 259-268.
19. Cunningham, S.D. and Ow, D.W. (1996) *Plant Physiol.* 110, 715-719.
20. Cuozzo, M., O'Connell, K.M., Kaniewski, W., Fang, R.X., Chua, N.H. and Tumer, N.E. (1988) *Biotech.* 6, 549-557.
21. Dhallan, R. S., Yau, K. W., Schrader, K. A., and Reed, R. R. (1990) *Nature* 347, 184-187

22. Devereux, J., Haeberli, P., and Smithies O. (1984) *Nucl. Acids Res.* 12, 387-395
23. Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., and Mackinnon, R. (1998) *Science* 280, 69-77
24. Ehsan, H., Reichheld, J.-P., Roef, L., Witters,E., Lardon, F., Van Bockstaele, D., Van Montagu, M., Inze, D., and Van Onckelen, H. (1998) *FEBS Lett.* 422, 165-169
25. Eisenberg, D. (1984) *Annu. Rev. Biochem.* 53, 595-623
26. Faiman G.A., Levy R., Anglister J., and Horovitz A. (1998) *J. Biol. Chem.* 271, 13829-13833
27. Foy, C.D., Chaney, R.L. and White, M.C. (1978) *Ann. Rev. Plant Physiol.* 29, 511-566.
28. Fin, J. T., Grunwald, M. E., and Yau, K. W. (1996) *Annu. Rev. Physiol.* 58, 395-426
29. Fromm, H., and Chua, N. H. (1992) *Plant Mol. Biol. Rep.* 10, 199-206
30. Gamborg, O.L., Miller, R.A. and Ojima, K. (1968) *Exp. Cell Res.* 50, 151-158.
31. Gaymard, F., Cerutti, M., Horeau, C., Lemaillet, G., Urbach, S., Revallec, M., Devauchelle, G., Sentenac, H., and Thibaud, J.B. (1996) *J. Biol. Chem.* 271, 2283-2287
32. Golemis, E. A., Gyuris, J., and Brent, R. (1997) In: *Current Protocolsin Molecular Biology* . (Chanda, V.B., ed.) Chapter 20, unit 20.1. Wiley, New York
33. Guilley, H., Dudley, R., Jonard, G., Balas, E. and Richards, K. (1982) *Cell* 30, 763-773.
34. Haas, I.G. (1994) *Experientia* 50, 1012-1020.
35. Heginbotham, L., Lu, Z., Abramson, T., and Mackinnon, R. (1994) *Biophys. J.* 66, 1061-1067
36. Hoagland, D.R. and Arnon, D.I. (1950) California Agricultural Experiment Station Circular 371. *The College of Agriculture*, University of California, Berkeley, CA.
37. Horsch, R.B., Fry, J.E., Hoffman, N.L., Eichnoltz, D., Rogers, S.G. and Fraley, R.T. (1985) *Science* 227, 1229-1231.
38. Huang, J.W., Grunes, D.L. and Kochian, L.V. (1994) *Proc. Natl. Acad. Sci. USA* 91, 3473-3477.
39. Huang, J.W. and Cunningham, S.D. (1996) *New Phytol.* 134, 75-84.
40. Ichida, A.M., Pei, Z.M., Baizabal-Aguirre, V.M., Turner, K.J. and Scroeder, J.I. (1997) *Plant J.* 9, 1843-1857.

41. Jan, L. Y., and Jan, Y. N. (1992) *Annu. Rev. Physiol.* 54, 537-555
42. James, P., Vorherr, T., and Carafoli, E. (1995) *Trends Biochem. Sci.* 20, 38-42
43. Kincaid, R. L., Vaughan M., Osborne J. C. Jr., and Tkacheck, V. A. (1982) *J. Biol. Chem.* 257, 10638-10643
44. Kohler, C., Merkle, T. and Neuhaus, G. (1999) *Plant J.* 18, 97-104.
45. Knight, H., Brandt, S. and Knight, M.R. (1998) *Plant J.* 16, 681-687.
46. Kramer, U., Smith, R.D., Wenzel, W.W., Raskin, I. and Salt, D.E. (1997) *Plant Physiol.* 115, 1641-1650.
47. Kurosaki, F., Kaburaki, H., and Nishi, A. (1994) *FEBS Lett.* 340, 193-196
48. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* 157, 105-132
49. Larson, C.H. (1985) *Modern Methods in Plant Analysis*. (Linskens, H.F. and Jackson, J.F., eds). Berlin: Springer-Verlag.
50. Leng, Q., Mercier, R.W., Yao, W. and Berkowitz, G.A. (1999) Cloning and first functional characterization of a plant cyclic nucleotide-gated cation channel. *Plant Physiol.* 121, 753-761.
51. Liman, E. R., Hess, P., Weaver, F., and Koren, G. (1991) *Nature* 353, 752-756.
52. Liu, M., Chen, T. Y., Ahamed, B., Li, J., and Yau, K. W. (1994) *Science* 266, 1348-1354 Lockow, V. A, and Summers, M. D. (1988) *Bio/Technology* 6, 47-55.
53. Liu, D.T., Tibbs, G.R. and Siegelbaum, S.A. (1996) Subunit stochiometry of cyclic nucleotide-gated channels and effects of subunit order on channel function. *Neuron* 16, 983-90.
54. Logemann, J., Schell, J. and Willmitzer, L. (1987) Improved method for the isolation of RNA from plant tissues. *Anal Biochem.* 163, 16-20.
55. Lundblad, V. (1997) In: *Current Protocols in Molecular Biology* . Chanda VB, ed. Chapter 13 unit 13.6. Wiley, New York.
56. McAinsh, M. R., and Hetherington, A. M. (1998) *Trends Plant Sci.* 3, 32-36.
57. Moffat, A.S. (1995) *Science,* 269, 302-303.
58. Nakamura, R. L, Anderson, J. A., and Gaber, R. F. (1997) *J. Biol. Chem.* 272, 1011-1018.
59. Neuhaus, G., Bowler, C., Kern, R., and Chua, N.-H. (1993) *Cell* 73, 937-952.
60. Nriagu, J.O. and Pacyna, J.M. (1988) *Nature* 333, 134-139.

61. O'Neil, K. T., and DeGrado, W. F. (1990) *Trends Biochem. Sci.* 15, 59-64.

62. Norman, T. and Banuelos, G. (1999) *Phytoremediation of contaminated soil and water.* CRC Press, Boca Raton, FL.

63. O'Reilly, D.R., Miller, L.K. and Luckow, V.A. (1992) Baculovirus Expression Vectors, a Laboratory Manual. New York: WH Freeman and Co.

64. Ouyang, H. and Vogel, H.J. (1998) Metal ion binding to calmodulin: NMR and fluorescence studies. *BioMetals* 11, 213-222.

65. Parrot, W.A. and Bouton, J.H. (1990) *Crop Sci.* 30, 387-389.

66. Raskin, I.(1996) *Proc. Natl. Acad. Sci. USA* 93, 3164-3166.

67. Rengel, Z. (1996) *New Phytol.* 134, 389-406.

68. Rubio, F., Grassmann, W. and Scroeder, J.I. (1995) *Science* 270, 1660-1663.

69. Schachtman, D.P., Kumar, R., Schroeder, J.I. and Marsh, E.L. (1997) *Proc. Natl. Acad. Sci. USA* 94, 11079-11084.

70. Schuurink, R.C., Shartzer, S.F., Fath, A. and Jones, R.L. (1998) *Proc. Natl. Acad. Sci.* 95, 1944-1949.

71. Shabb, J. B., and Corbin, J. D. (1992) *J. Biol. Chem.* 267, 5723-5726

72. Simons, T.J.B. and Pocock G. (1987) *J. Neurochem.* 48, 383-389.

73. Snedden, W.A. and Fromm, H. (1997) *Plant Mol. Biol.* 33, 753-756.

74. Snedden, W. A., and Fromm, H. (1998) *Trends Plant Sci.* 3, 299-304

75. Snedden, W. A., Koutsia, N., Baum, G., and Fromm, H. (1996) *J. Biol. Chem.* 271, 4148-4153 Tester, M. (1990) *New Phytol.* 114, 305-340.

76. Tomsig, J.L. and Suszkiw, J.B. (1991) *Biochim. Biophys. Acta* 1069, 197-200.

77. Toshi, T. (1995) *J. Gen. Physiol.* 105, 309-328

78. Van der Zaal, B.J., Neuteboom, L.W., Pinas, J.E., Chardonnens, A.N., Schat, H., Verkleij, 82. J.A.C. and Hooykaas, P.J.J. (1999) *Plant Physiol.* 119, 1047-1055.

79. Van Eldik, L., and Watterson, M. D. (1998) *Calmodulin and Signal Transduction.* Academic Press, New York..

80. Varnum, M. D, and Zagotta, W. N. (1997) *Science* 278, 110-113.

81. Volotovski, I. D., Sokolovsky, S. G., Molchan, O. V., and Knight, M. R. (1998) *Plant Physiol.* 117, 1023-1030.

82. Ward, J., Reinders, A., Hsu, H. and Sze, H. (1992) *Plant Physiol.* 99, 161-169.

83. Warmke, J., and Ganetzky, B. (1994) *Proc. Natl. Acad. Sci.* 91, 3438-3442.

84. Warmke, J., Drysdale, R., and Ganetzky, B. (1991) *Science* 252, 1560-1562.

85. Weber, I. T., Gilliland, G. L., Harman, J. G., and Peterkofsky, A. (1987) *J. Biol. Chem.* 262, 5630-5636.

86. Welch, R.M. (1995) *Crit. Rev. Plant Sci.* 14, 49-82.

87. Woolhouse, H.W. (1983) *Physiological Plant Ecology, III. Responses to the chemical and biological environment.* Volume 12C (Lange, O.L., Nobel, P.S., Osmond, C.B. and Zeigler, H., eds). Berlin: Springer-Verlag, pp. 245-300.

88. Zielinski, R.N. (1998) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49, 697-725.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid comprising a polynucleotide at least 60 % identical with any of SEQ ID NOs:1, 3, 5, 7, 9, 11 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

2. The isolated nucleic acid of claim 1, wherein said polynucleotide is as set forth in any of SEQ ID NOs:1, 3, 5, 7, 9, 11 or portions thereof.

3. An isolated nucleic acid comprising a polynucleotide encoding a polypeptide being at least 60 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

4. An isolated nucleic acid comprising a plant derived polynucleotide encoding a transmembrane polypeptide having a cation channel activity when assembled in a plasmatic membrane of a plant cell, said polypeptide having overlapping cyclic nucleotide-binding domain site and calmodulin-binding site.

5. An isolated nucleic acid comprising a polynucleotide hybridizable with any of SEQ ID NOs:1, 3, 5, 7, 9, 11 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 °C.

6. An isolated nucleic acid comprising a polynucleotide hybridizable with a polynucleotide encoding a polypeptide as set forth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M

NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 °C.

7. A nucleic acid construct comprising a polynucleotide as set forth in any of claims 1-6 downstream of a plant promoter.

8. The nucleic acid construct of claim 7, wherein said polynucleotide is in a sense orientation.

9. The nucleic acid construct of claim 7, wherein said polynucleotide is in an antisense orientation.

10. The nucleic acid construct of claim 7, wherein said plant promoter is selected from the group consisting of a constitutive promoter, a tissue specific promoter, an inducible promoter and a chimeric promoter.

11. A genetically transformed or virus infected cell comprising an isolated nucleic acid as set forth in any of claims 1-6.

12. A genetically transformed or virus infected cell comprising a nucleic acid construct as set forth in any of claims 7-10.

13. A genetically transformed or virus infected plant comprising an isolated nucleic acid as set forth in any of claims 1-6.

14. A genetically transformed or virus infected plant comprising a nucleic acid construct as set forth in any of claims 7-10.

15. A recombinant protein comprising a polypeptide encoded by a polynucleotide as set forth in any of claims 1-6.

16. A recombinant protein comprising a polypeptide at least 60 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

17. The recombinant protein of claim 16, wherein said polypeptide is as set fourth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12 or portions thereof.

18. A genetically transformed or virus infected cell expressing a recombinant protein as set forth in any of claims 15-17.

19. A genetically transformed or virus infected cell expressing a recombinant protein as set forth in any of claims 15-17.

20. An antibody specifically recognizing a recombinant protein as set forth in any of claims 15-17 or a natural equivalent thereof.

21. The antibody of claim 20, selected from the group consisting of a polyclonal antibody and a monoclonal antibody.

22. An immunized non-human mammal producing the polyclonal antibody of claim 21.

23. An immunized egg producing the polyclonal antibody of claim 21.

24. A cell producing the monoclonal antibody of claim 21.

25. A method of increasing a tolerance of a plant to a metal cation, the method comprising the step of overexpressing in the plant a recombinant protein which reduces uptake and concentration of the metal cation within the plant cells.

26. The method of claim 25, wherein the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide at least 60 % identical with any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

27. The method of claim 26, wherein said polynucleotide is as set forth in any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof.

28. The method of claim 25, wherein the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide encoding a polypeptide being at least 60 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

29. The method of claim 25, wherein the plant is genetically transformed or viral infected with a nucleic acid including a plant derived polynucleotide encoding a transmembrane polypeptide having a cation channel activity when assembled in a plasmatic membrane of a plant cell, said polypeptide having overlapping cyclic nucleotide-binding domain site and calmodulin-binding site.

30. The method of claim 25, wherein the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide hybridizable with any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 °C.

31. The method of claim 25, wherein the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide hybridizable with a polynucleotide encoding a polypeptide as set forth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 ° C.

32. The method of any of claims 26-31, wherein the plant is genetically transformed or viral infected with a nucleic acid construct

comprising any of the polynucleotides downstream of a plant promoter in a sense orientation.

33. The method claim 32, wherein said plant promoter is selected from the group consisting of a constitutive promoter, a tissue specific promoter, an inducible promoter and a chimeric promoter.

34. The method of claim 25, wherein said recombinant protein includes a polypeptide at least 60 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

35. The method of claim 34, wherein said polypeptide is as set fourth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof.

36. The method of claim 25, wherein said recombinant protein includes a calmodulin binding domain.

37. The method of claim 36, wherein said recombinant protein further includes a cyclic nucleotide binding domain.

38. The method of claim 37, wherein said calmodulin binding domain and said cyclic nucleotide binding domain overlap.

39. The method of claim 25, wherein the metal cation is Ni, Na, Ba, Cd, Co, Cu, La, Mn, Zn and Pb.

40. A method for phytoremediation of an area polluted with a metal cation, the method comprising the steps of:

(a) providing a plant characterized in resistivity to elevated concentrations of the metal cation, said plant overexpressing a recombinant protein which facilitates uptake and concentration of the metal cation within the plant cells;

(b) planting said plant in the area;

(c) following a time period, in which at least a fraction of the metal cation in the area has been accumulated in said plant,

harvesting said plant, thereby removing at least said fraction of the metal cation from the area; and optionally

(d) repeating steps (b) - (c) until a sufficient amount of the metal cation has been removed from the area.

41. The method of claim 40, wherein said area is selected from the group comprising of a terrestrial area and an aquatic area.

42. The method of claim 40, wherein the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide at least 60 % identical with any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

43. The method of claim 24, wherein said polynucleotide is as set forth in any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof.

44. The method of claim 40, wherein the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide encoding a polypeptide being at least 60 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

45. The method of claim 40, wherein the plant is genetically transformed or viral infected with a nucleic acid including a plant derived polynucleotide encoding a transmembrane polypeptide having a cation channel activity when assembled in a plasmatic membrane of a plant cell, said polypeptide having overlapping cyclic nucleotide-binding domain site and calmodulin-binding site.

46. The method of claim 40, wherein the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide hybridizable with any of SEQ ID NOs:1, 3, 5, 7, 9, 11, 15, 17 or portions thereof under hybridization conditions of hybridization solution containing

10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 °C.

47. The method of claim 40, wherein the plant is genetically transformed or viral infected with a nucleic acid including a polynucleotide hybridizable with a polynucleotide encoding a polypeptide as set forth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof under hybridization conditions of hybridization solution containing 10 % dextrane sulfate, 1 M NaCl, 1 % SDS and 5 x 10^6 cpm ^{32}p labeled probe, at 65 °C, with a final wash solution of 1 x SSC and 0.1 % SDS and final wash at 50 ° C.

48. The method of claim 41, wherein the plant is genetically transformed or viral infected with a nucleic acid construct comprising any of the polynucleotides downstream of a plant promoter in a sense orientation.

49. The method claim 48, wherein said plant promoter is selected from the group consisting of a constitutive promoter, a tissue specific promoter, an inducible promoter and a chimeric promoter.

50. The method of claim 40, wherein said recombinant protein includes a polypeptide at least 60 % homologous with any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof as determined using the Bestfit procedure of the DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the university of Wisconsin (gap creation penalty - 50, gap extension penalty - 3).

51. The method of claim 50, wherein said polypeptide is as set fourth in any of SEQ ID NOs:2, 4, 6, 8, 10, 12, 16, 18 or portions thereof.

52. The method of claim 40, wherein said recombinant protein includes a calmodulin binding domain.

53. The method of claim 52, wherein said recombinant protein further includes a cyclic nucleotide binding domain.

54. The method of claim 53, wherein said calmodulin binding domain and said cyclic nucleotide binding domain overlap.

55. The method of claim 40, wherein the metal cation is selected from the group consisting of Ni, Na, Ba, Cd, Co, Cu, La, Mn, Zn and Pb.

56. A plant expressing a C termainl truncated form of a calcium-dependent calmodulin-binding protein said plant being male-sterile.

57. A method of producing a male sterile plant, the method comprising the step of expressing in the plant a C termainl truncated form of a calcium-dependent calmodulin-binding protein.

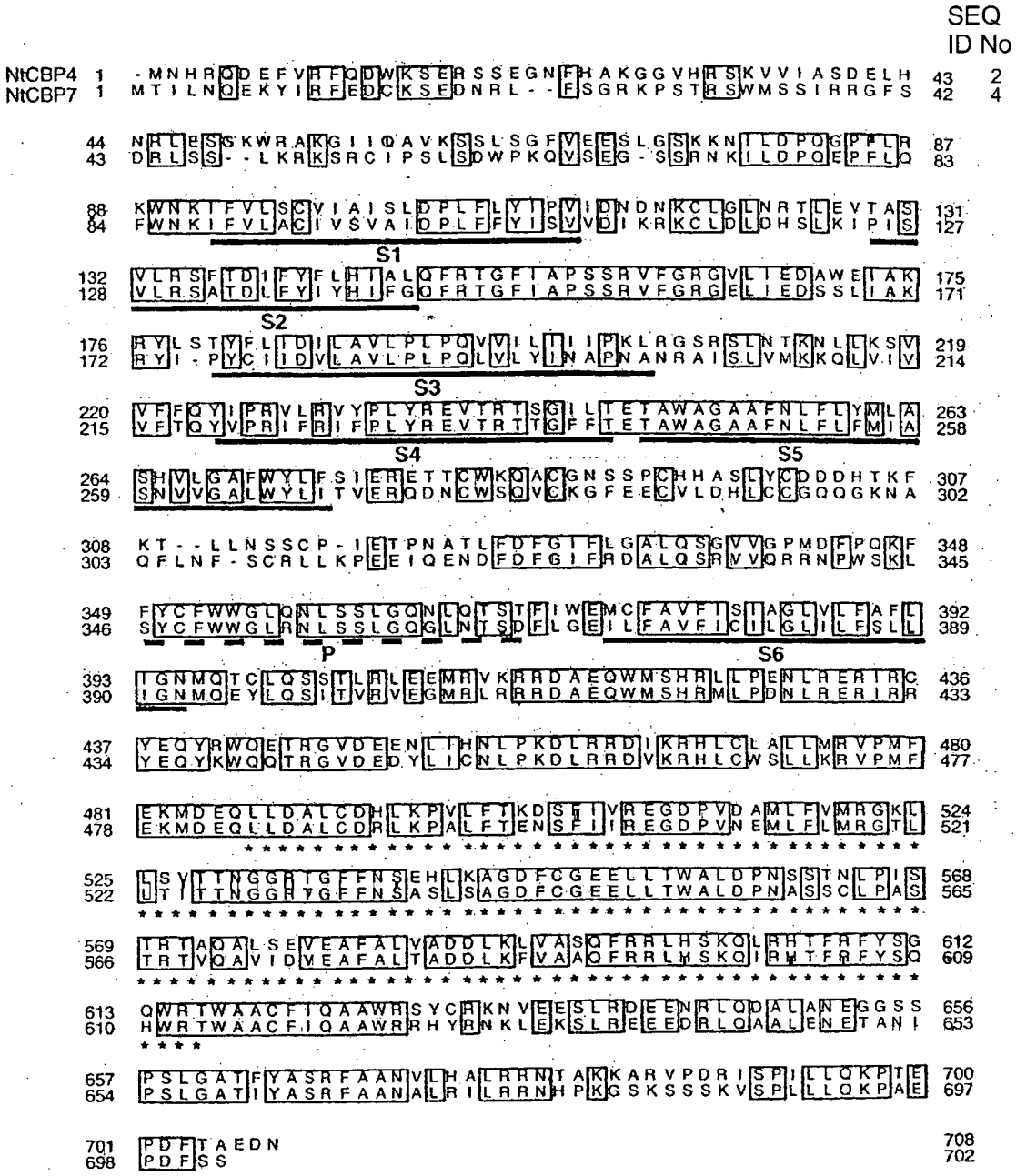


Fig. 1a

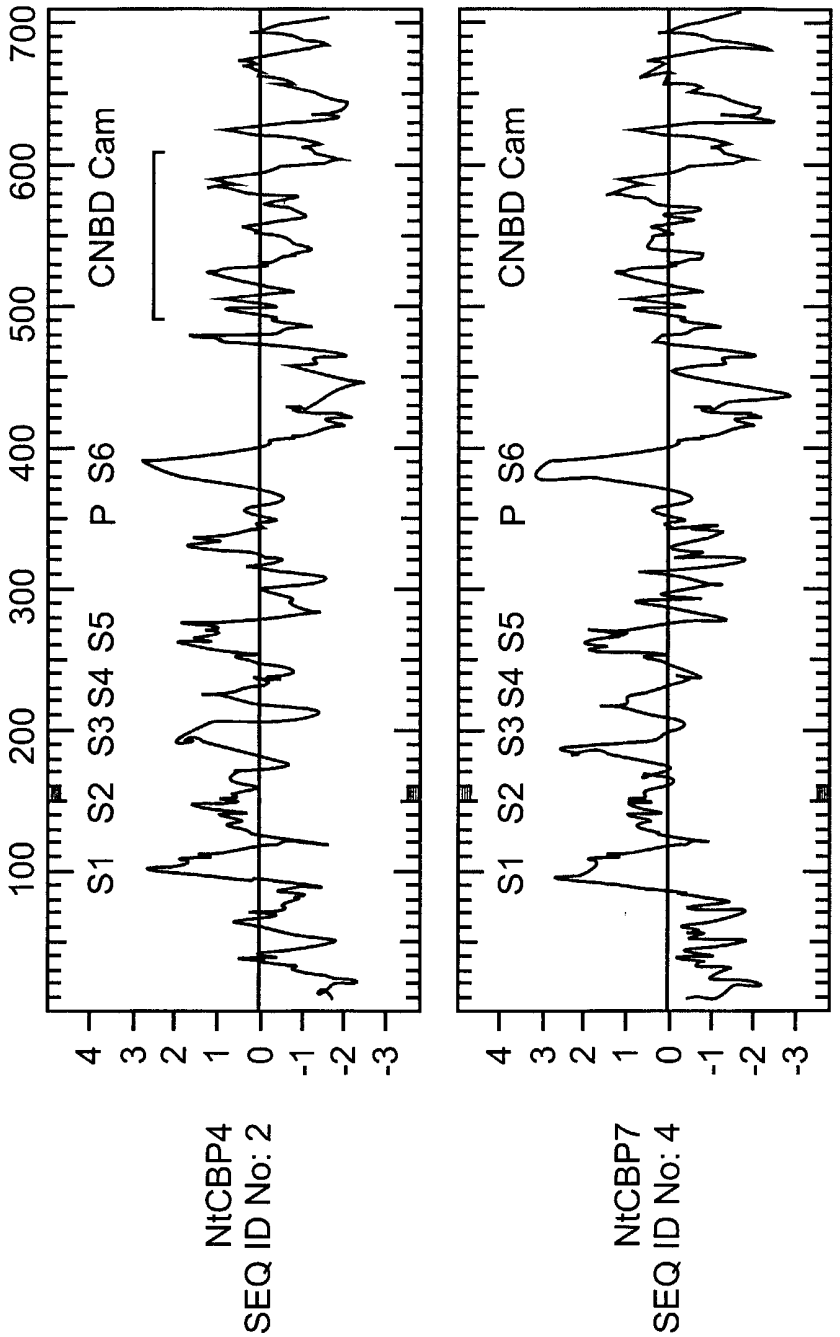

NtCBP4
SEQ ID No: 2
NtCBP7
SEQ ID No: 4
S1 S2 S3 S4 S5 P S6 CNBD Cam
100 200 300 400 500 600 700
4 3 2 1 0 -1 -2 -3

Fig. 1b

```
NtCBP4   .MNHRQDEPVRFQDWKSER.SSEGNFHAK....GGVHRSKVVIASDELHNRL.ESGKWRAKGIIQAVKS.SLSGPVEESLGSKKNILDPQGPFLRKWNKI  92
NtCBP7   MTILN-EKYI--E-C---.............DNRLFSGRKPSTRSWNSSIRRGFSDRLSSLKRKSRCIPSLSDWPKQ-S-G.S-RNK-----E---QF----  88
AtCNGC1  .--P--EK----------DKT--DVEYSG-NEIQT-IFQRTISSI--KFYRSF.--SSA-I-LFKRSY--Y-FKEA-SKGI--THK----------QR----  98
HvCBT1   ..ME-E-RY-----C---.....................QSVISENA-SPRRHCVFRS-K-RTGGLFAFLGNFLR-ET-KRSML-DRK-MQ-VFH------QR----  84

NtCBP4   FVLSCVIAISLDPLFLYIPVIDNDNKCLGLNRTLEVTASVLRSFTDIFYFLHIALQFRTGPIAPSSRVPGRGVLIEDAMEIAKRYLSTYFLIDILAVLPL 192
NtCBP7   ---A-IVSVAI----F--S-V-IKR---D-DHS-KIPI-----------A--L--IY--FG-------------------E----SSL-------.IP-CI--V------  187
AtCNGC1  ---A-I--V---------F-V-I--DAK----IDKKM-I-----------V--V---IP-------------------V--KR----------SH-I----------  198
HvCBT1   -----IP-V-V-----------NDK-L-WY-D-KMKI-------------I---IP-------TS--TN-------V--RYA-----------------VC-----  184
                      S1                    S2
NtCBP4   PQVVILIIIPKLRGSRSLNTKNLLKSVVFFQYIPRVLRVYPLYREVTRTSGILTETAWAGAAFNLFLYMLASHVLGAFWYLPSIERETTCWKQACG.NSS 291
NtCBP7   --L-LY-NA-NANRAI--VM-KQ-VI---T--V--IP-IP----------T-FF-------------------F-I--N-V--L---ITV--QDN--S-V-KGF.E 286
AtCNGC1  --M-------HN---S-------N--FI----------FI-I----K-------------------------P-----------------V-------ER-NP 298
HvCBT1   -----W-VL-T-QV-QFMKA--I-MLI-IC--V---I-IR---LQI--SA--I-----------II-----------V---L--Q-KDA----N-.SLTR 283
                      S3                    S4                    S5
NtCBP4   PCHHASLYCDDDHT..KFKTLLNSSCP.IETPNATL...FDFGIFLGALQSGVVGPNDFPQKFFYCFWWGLQNLSSLGQNLQTSTFIMEMCFAVFISIAG 385
NtCBP7   E-VLDH-..CCQQG..-NAQF--P--RLLKPEEIQE.ND----------RD-----R--QREN-WS-LS-------R----------G-N--D-LG-IL-----C-L-  382
AtCNGC1  --ISKL----PETA..GGNAF--E---.-Q---T-..-.---------D---------ESQ-------------------------K---Y---I----------  392
HvCBT1   D-NP-Y---GNGG-RAGNAF-Q-VWV-N-TKD-LPD...PL----VP-IN.N-SQST--FA-L---V-------------------K---YA--NL---V--S-  379
                                                                                                    P
NtCBP4   LVLFAFLIGNMQTYCLQSSTLRLEEMRVKRRDAEQWMSHRLLPENLRERIRCYEQYTWQETRGVDEENLIHNLPKDLRRDIKRHLCLALLARVPMFEKMDE 485
NtCBP7   -I--SL------EY---I-V-V-G--LR-----------M--D-------R----K--Q-------DY--C-----------V-----MS--K----------  482
AtCNGC1  ----S---------Y---T-T-----------------K---R----K-------------LS-------------------------------------  492
HvCBT1   ----SL------Y---A---I--T---S--TD---Y----D--K----R----------S-----H-LM---------A-------S--K---------D  479
                      S6
NtCBP4   QLLDALCDHLKPVLFTKDSFIVREGDPVDAMLFVMRGKLLSVTTNGGRTGFFNSEHLKAGDFCGEELLTWALDPN.SSTNLPISTRTAQALSEVEAFALV 584
NtCBP7   ---------R---A---EN---I------NE---L---T--TI----------AS-S-----------------.A-SC--A----V--VID------T  581
AtCNGC1  ---------R-Q---Y-EE-Y----------E---I------TI----------L---Y-G-----------------H.--S---------VR--M--------K  591
HvCBT1   ---N----C-----Y-EGGC---------NE-P-IT--N-N-M-----K------DV--S-----------------.-A-S--S----VKSN----------M  578

NtCBP4   ADDLKLVASQFRRLHSKQLRHTFRFYSGQWRTNAACFIQAAWRSYCRRNVEESLRDEENRLQDALANEG.GSSPSLGATFYASRFAANVLHALRRNTA.K 682
NtCBP7   -----P--A----------I-------QH-----------------RHY-NKL-K---E--D---A--E--T.ANI------I----------A-RI----HP.- 679
AtCNGC1  -----P-----------------------Y--Q--K-----------R-IK-KL----KE----------K-AC----------I---------I-RTI--SGSVR 691
HvCBT1   -E---P--M-----------------------Q-----P----------HRR---KM-DA--EK-E---L-IV-D-.ST-L-F--AI------R-MMRT---AT.R 676
                        CaM
NtCBP4   KARVFDRISPILLQKPTEPDFTAEDN 708
NtCBP7   GSKSSSKRV--L-----A----SS... 702
AtCNGC1  -P-M-E-MP-M-----A----NSD-. 716
HvCBT1   ---LQE-VPAR-----A--N-S--EQ 702
```

Fig. 2

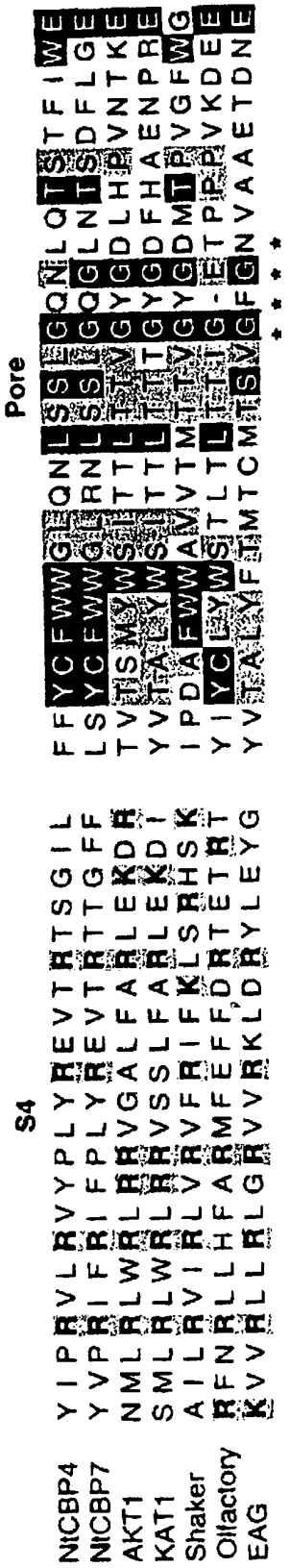


Fig. 3a

Fig. 3b

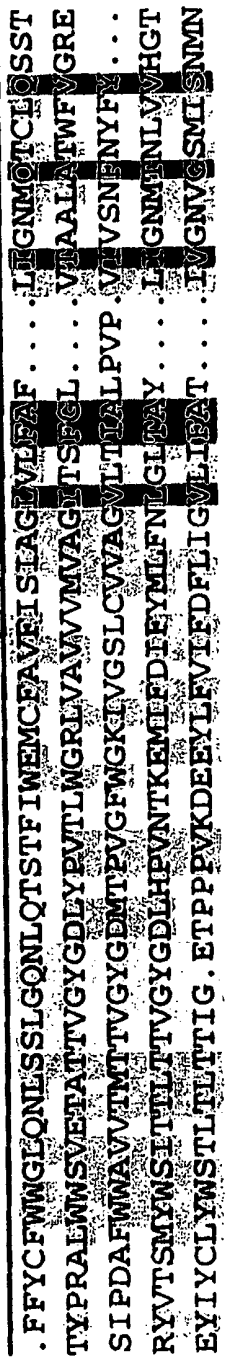


Fig. 3c

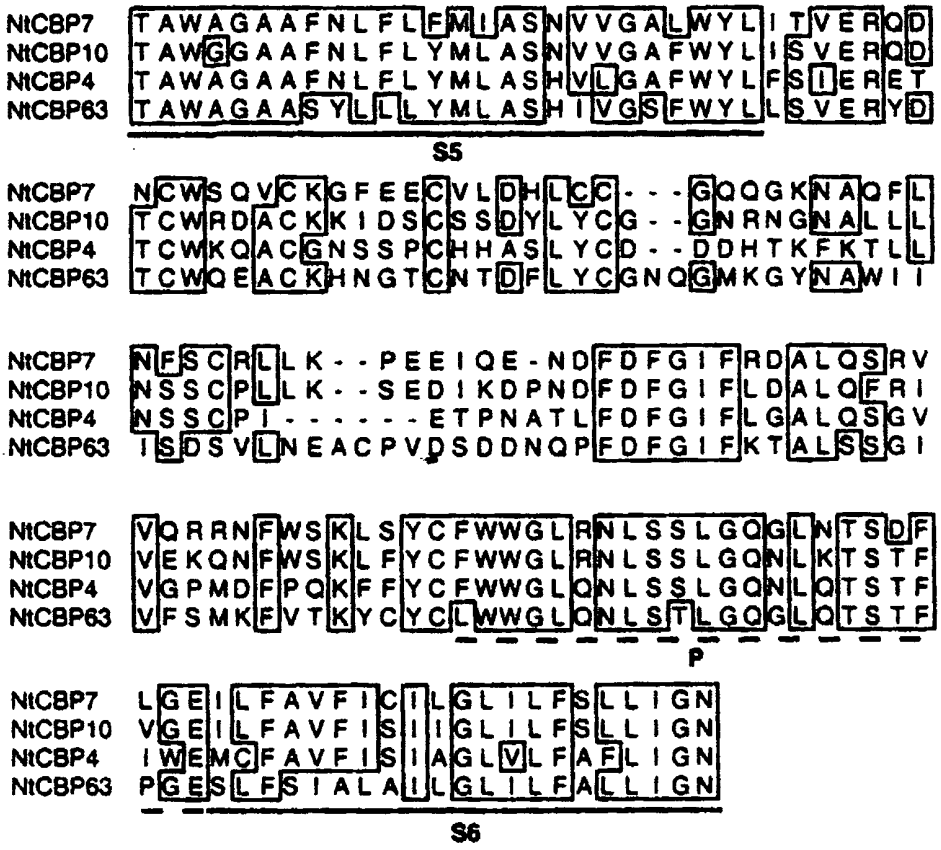


Fig. 4

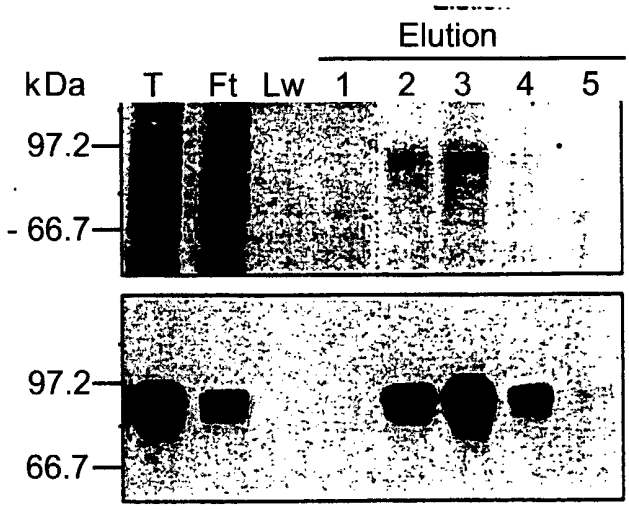


Fig. 5a

Fig. 5b

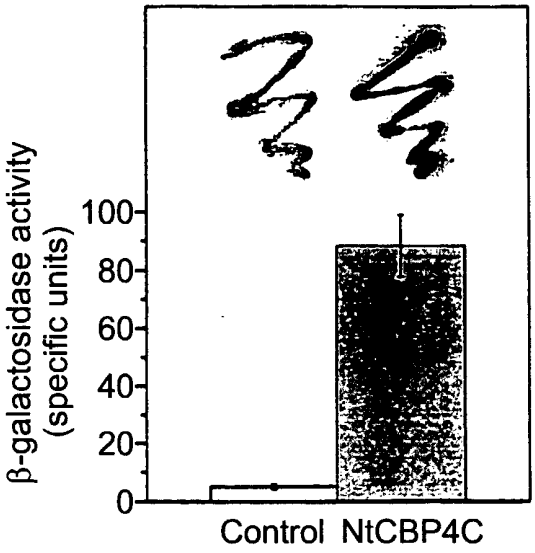

β-galactosidase activity
(specific units)
100
80
60
40
20
0
Control
NtCBP4C

Fig. 6

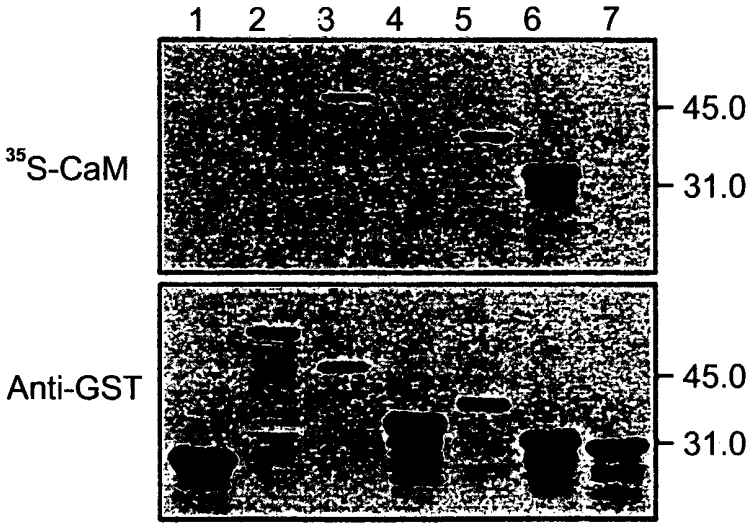

1 2 3 4 5 6 7
35S-CaM
45.0
31.0
Anti-GST
45.0
31.0

Fig. 7a

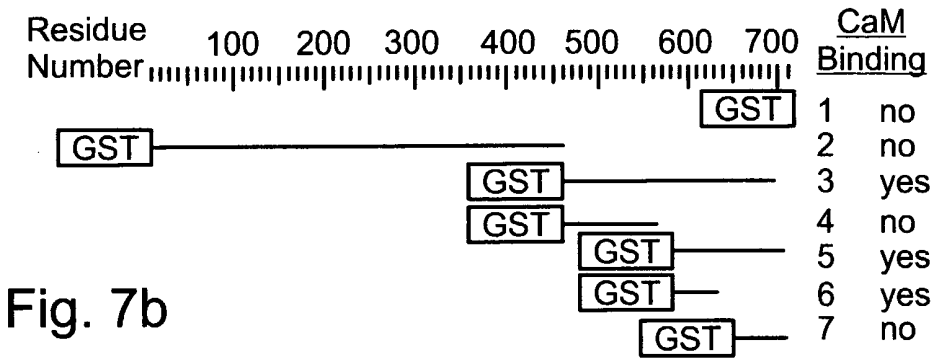


Fig. 7b

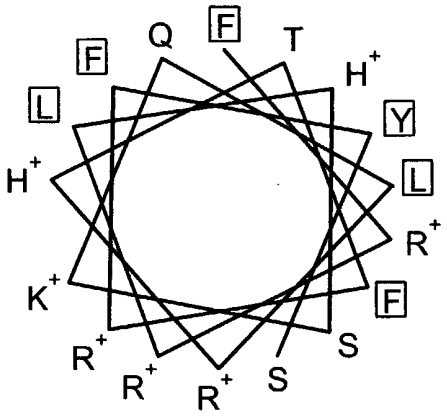


Fig. 8a

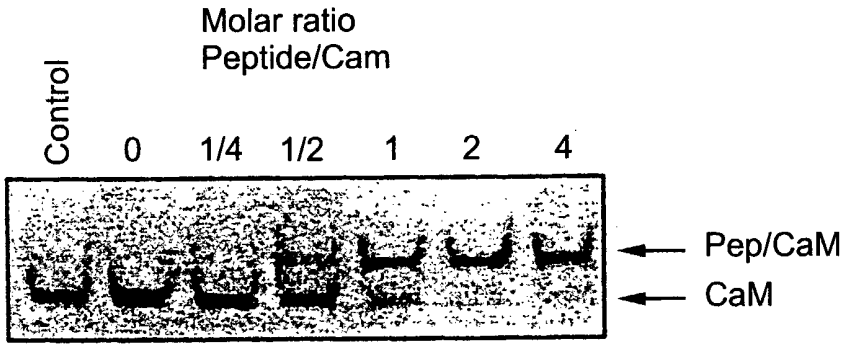


Fig. 8b

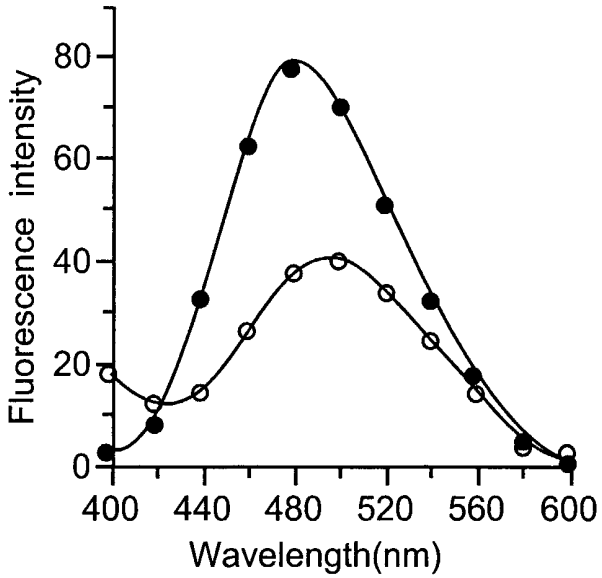

Fluorescence intensity
80
60
40
20
0
400
440
480
520
560
600
Wavelength(nm)

Fig. 9a

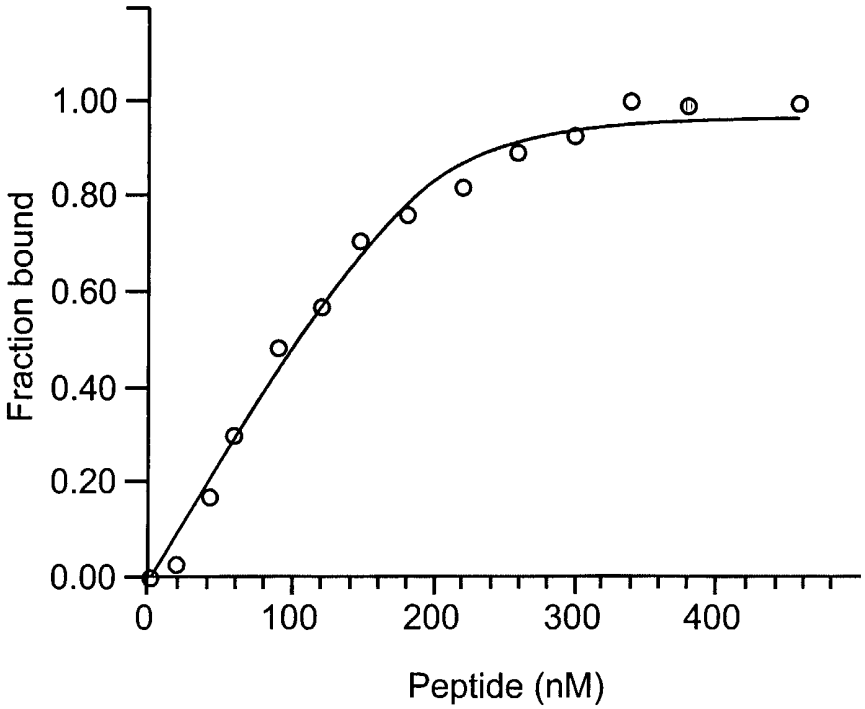

Fraction bound
1.00
0.80
0.60
0.40
0.20
0.00
0
100
200
300
400
Peptide (nM)

Fig. 9b

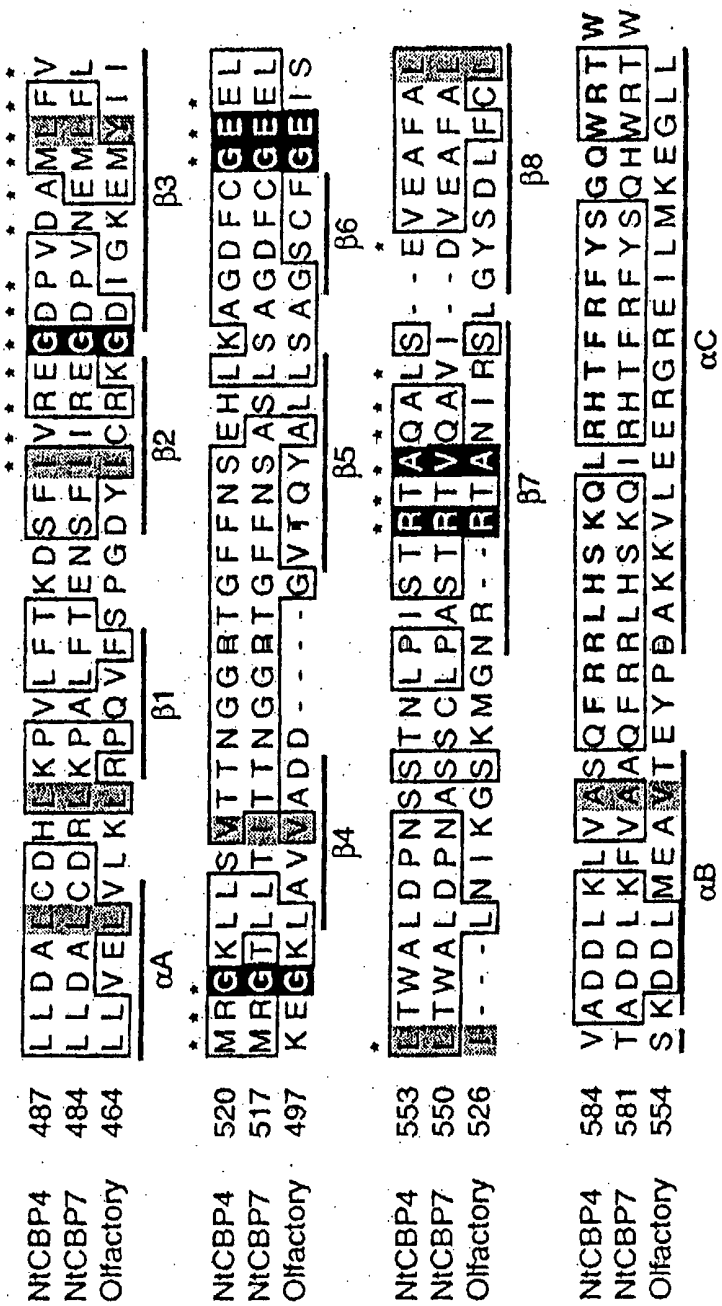


Fig. 10

Fig. 11a

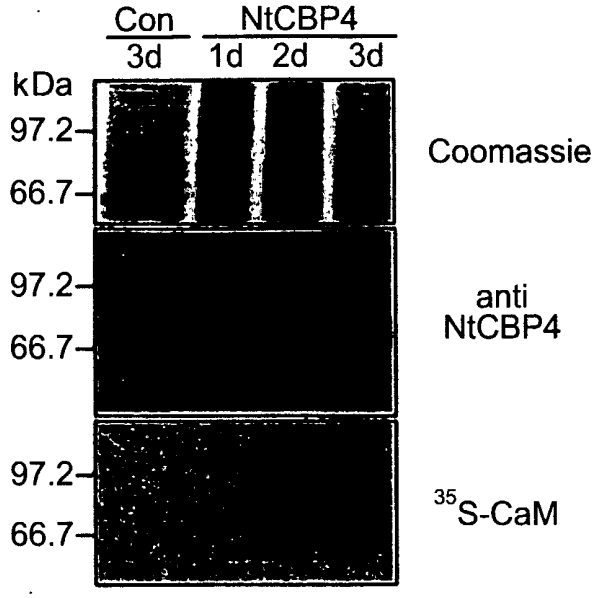
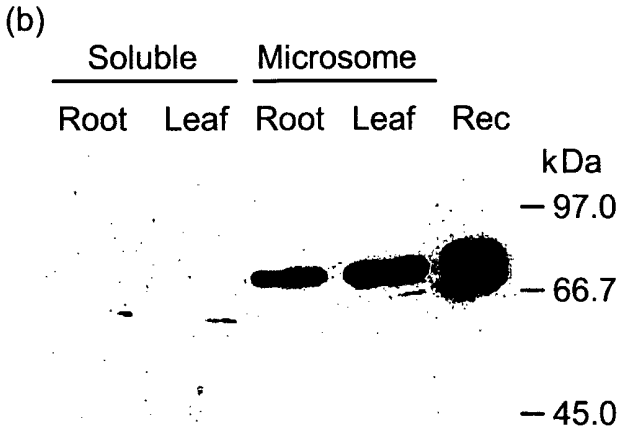


Fig. 11b



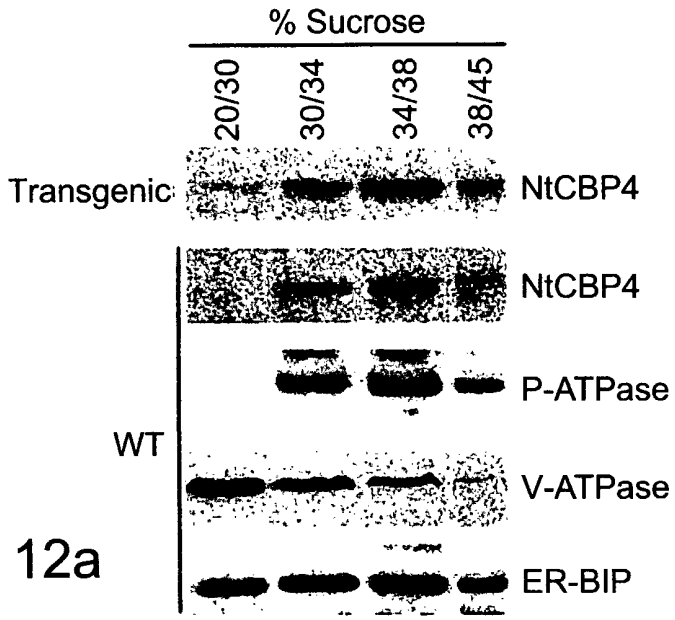


Fig. 12a

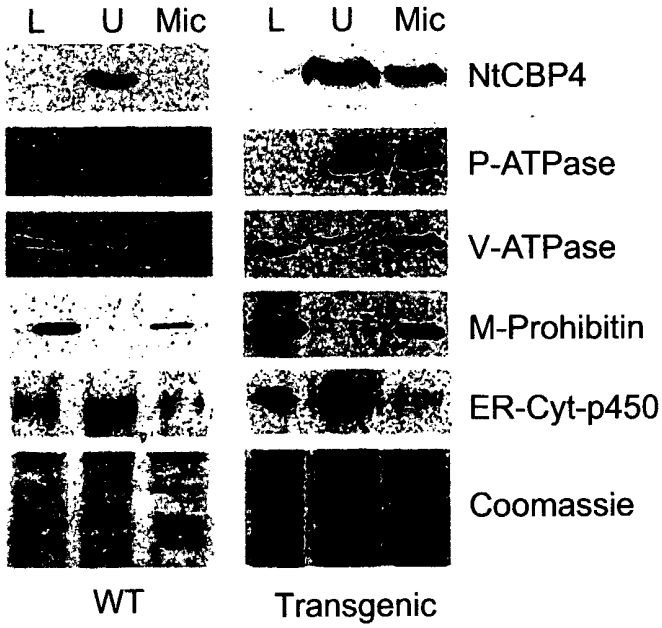


Fig. 12b

Fig. 13a

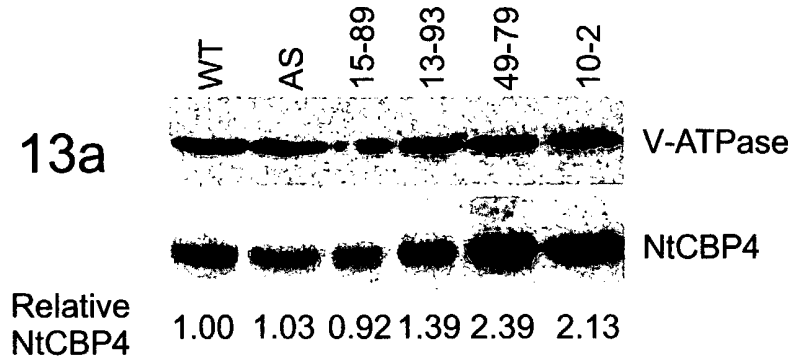


Fig. 13b

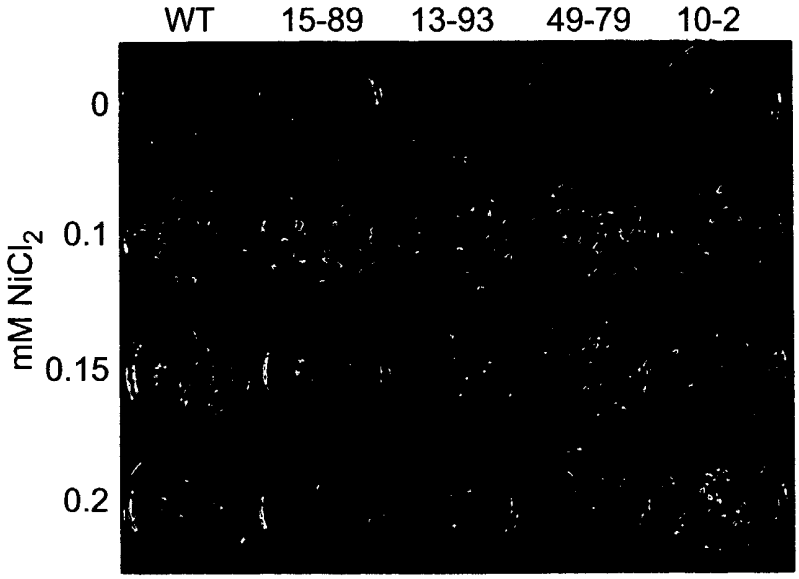
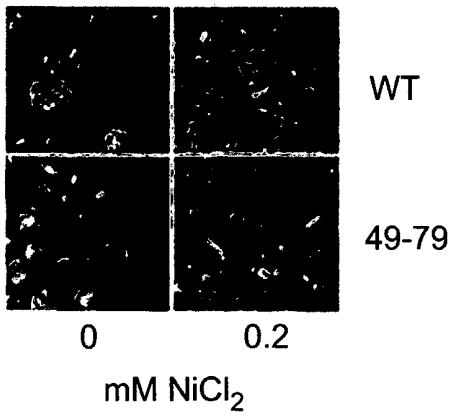


Fig. 13c



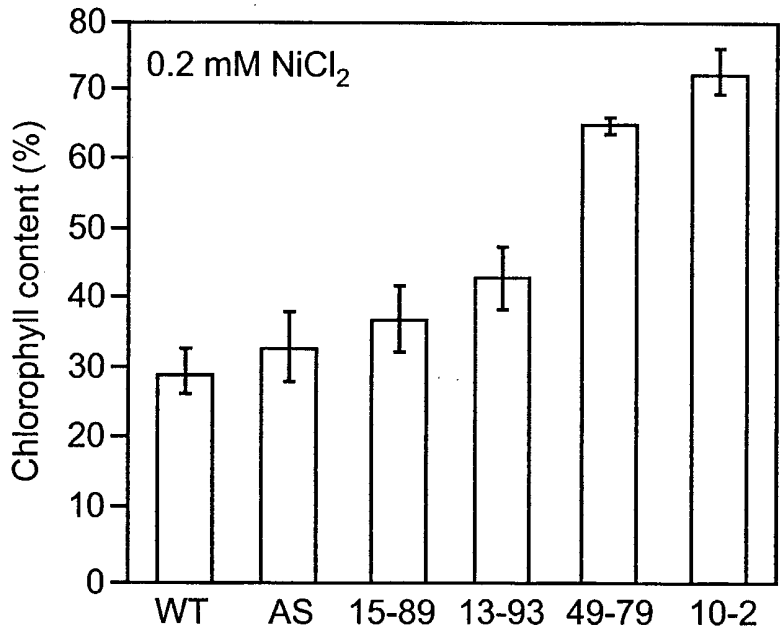


Fig. 14a

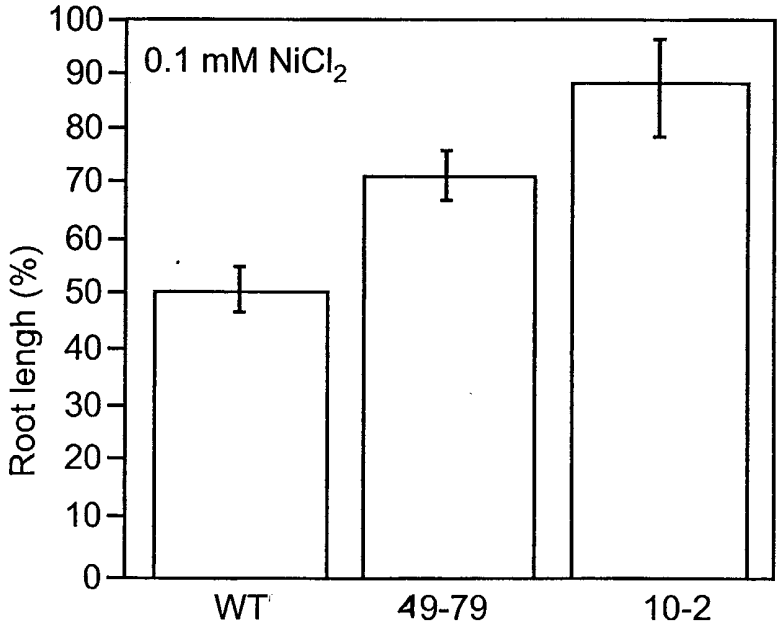


Fig. 14b

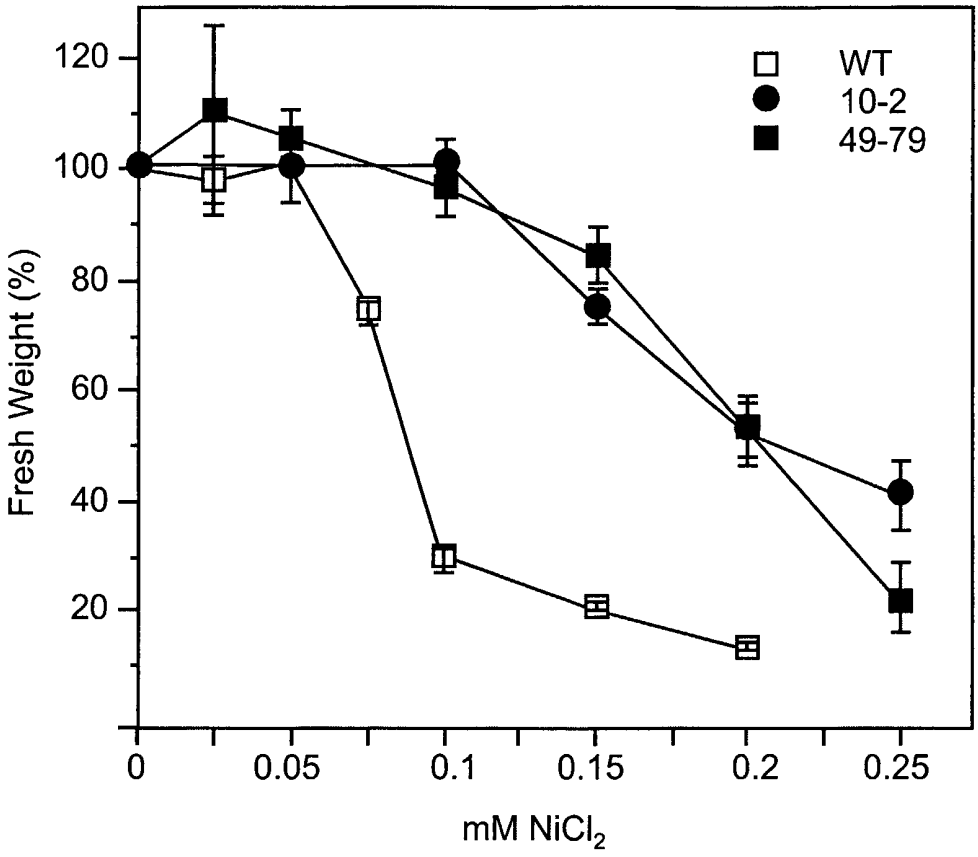

120
100
80
60
40
20
Fresh Weight (%)
0
0.05
0.1
0.15
0.2
0.25
mM NiCl2
WT
10-2
49-79

Fig. 14c

Fig. 15a

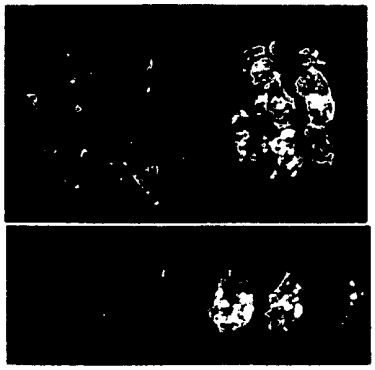


Fig. 15b

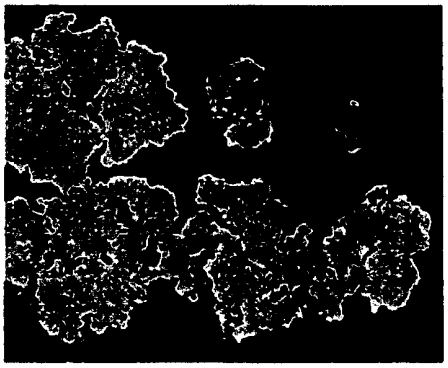
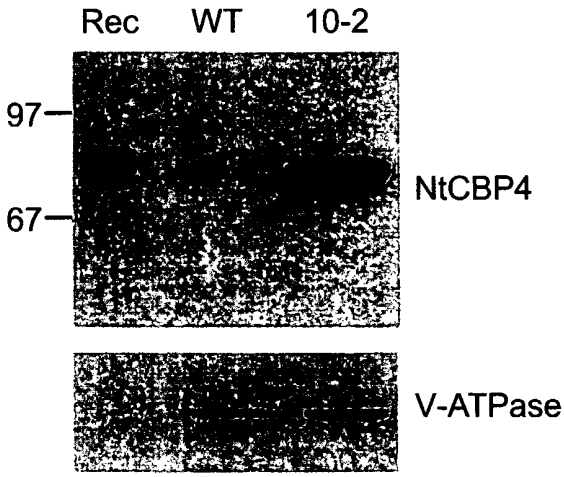


Fig. 15c



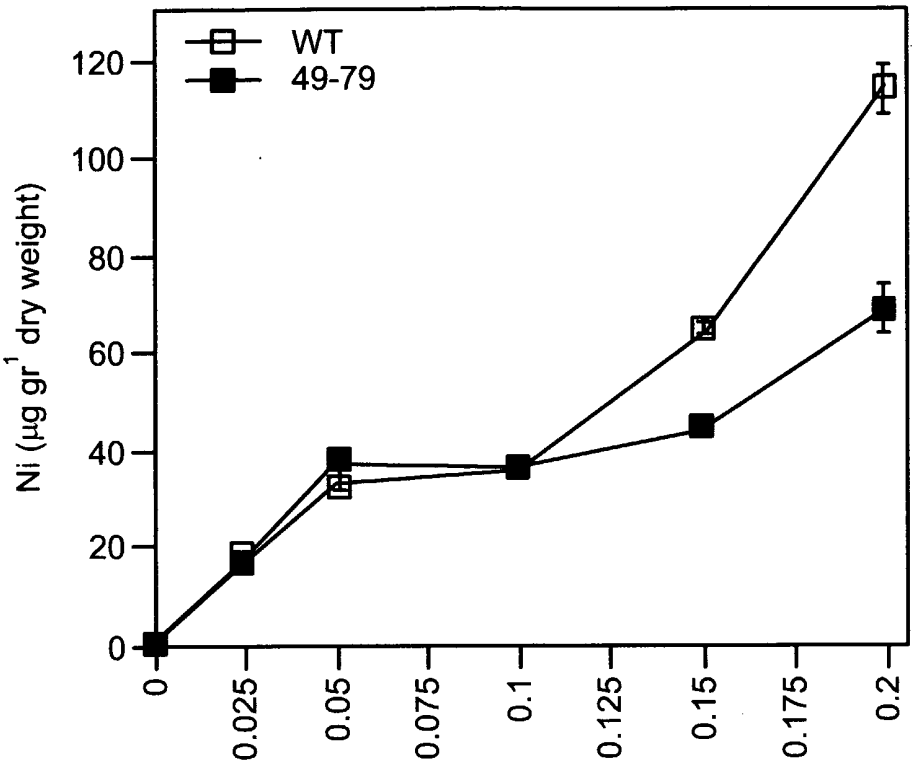


Fig. 16

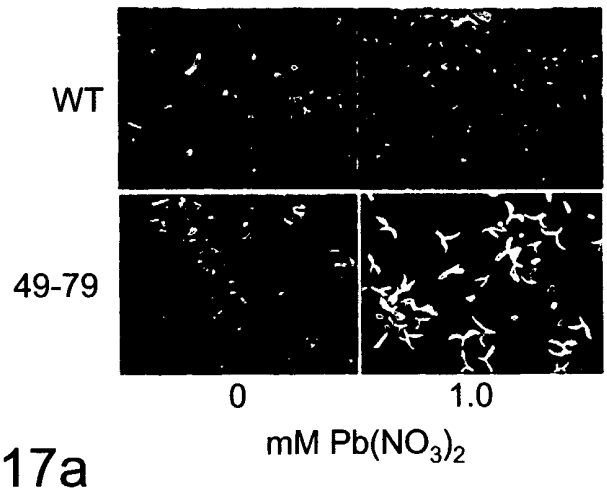


Fig. 17a

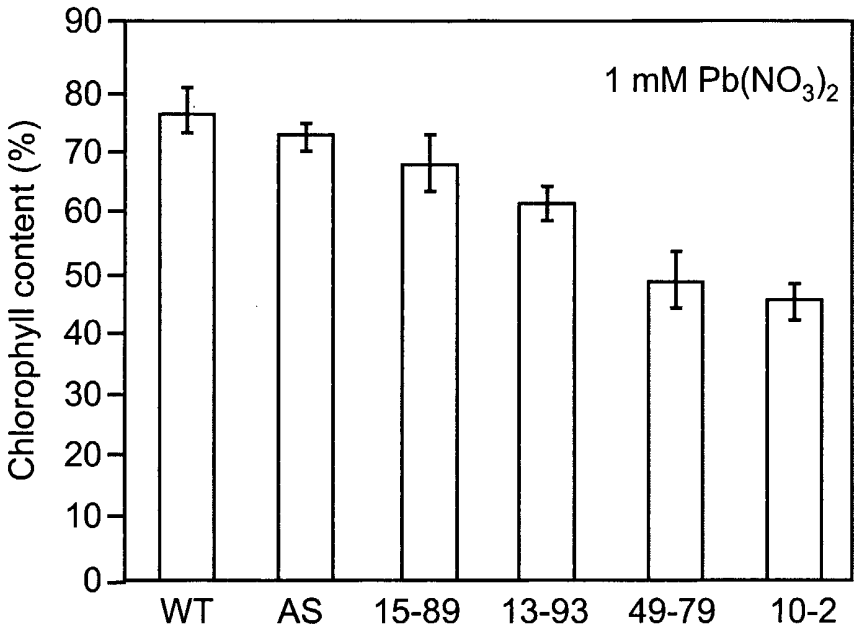


Fig. 17b

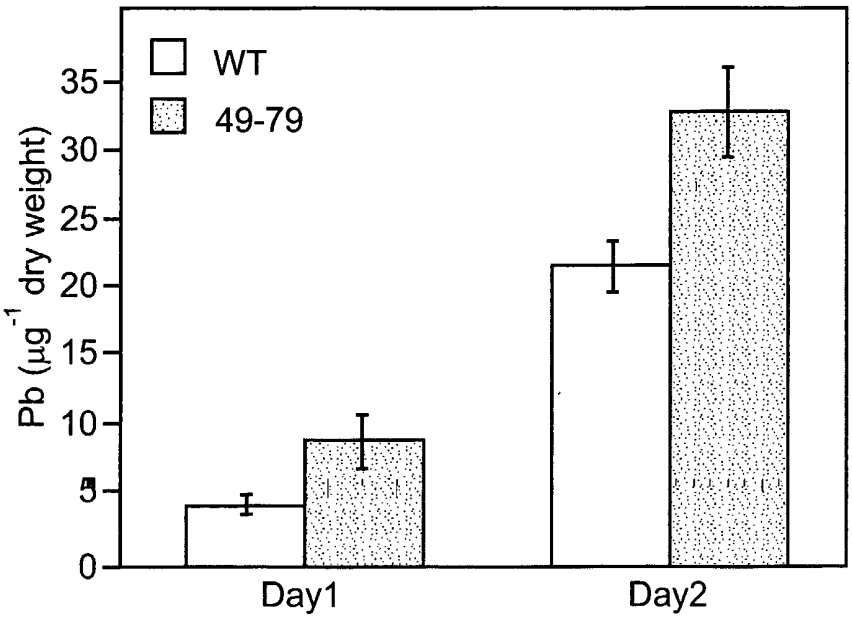


Fig. 17c

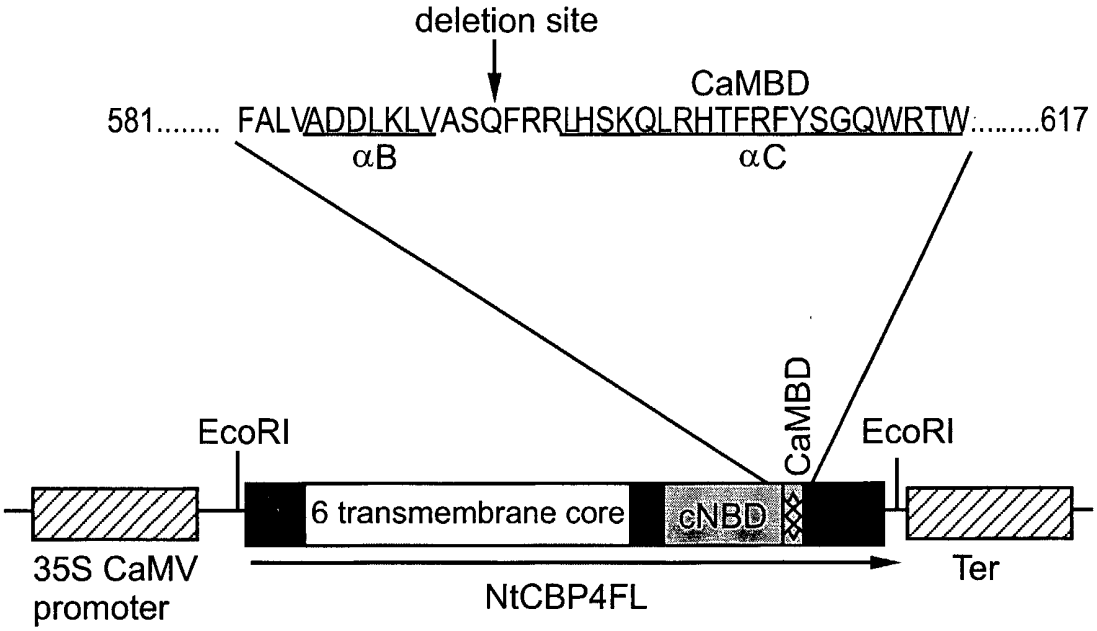


Fig. 18a

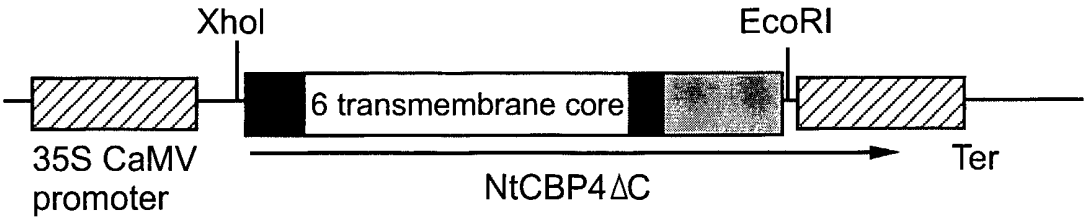


Fig. 18b

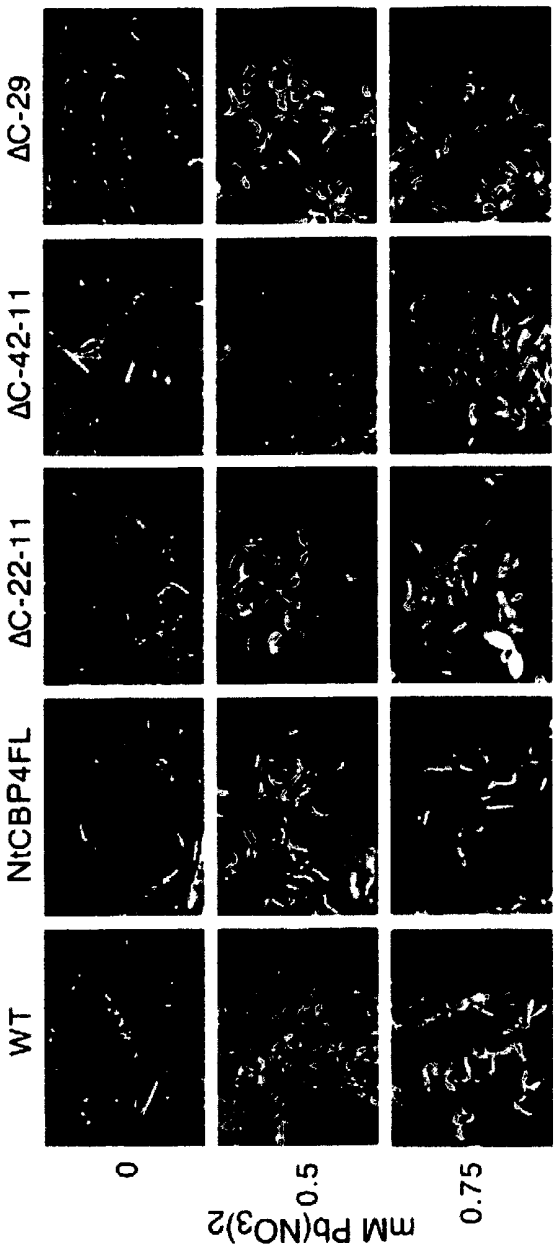


Fig. 19a

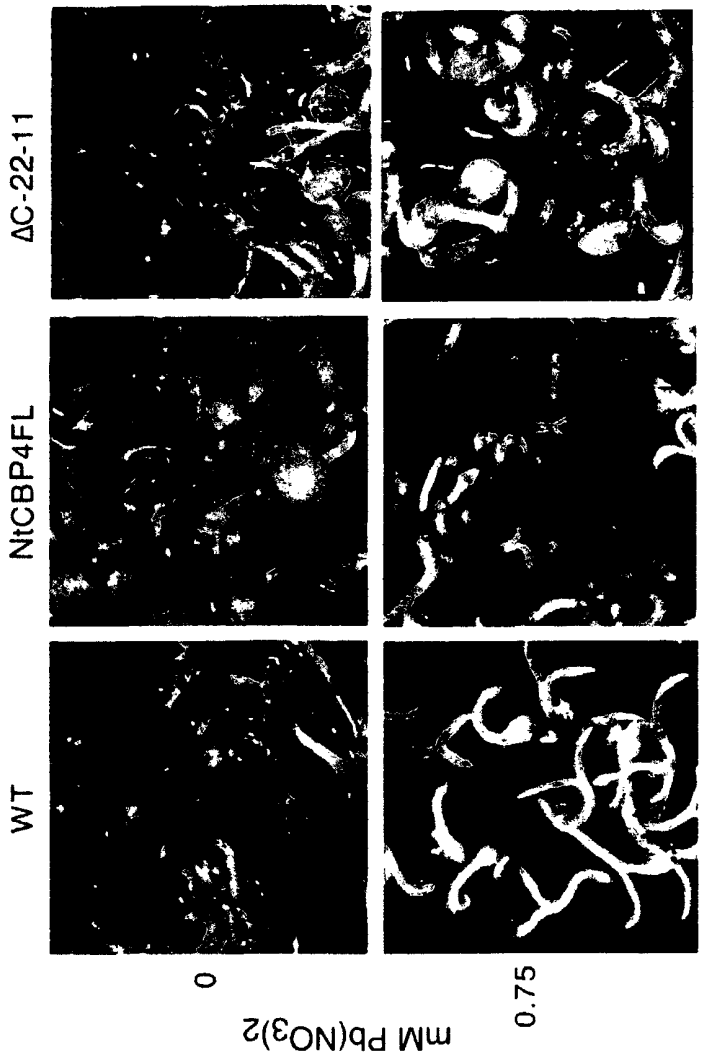


Fig. 19b

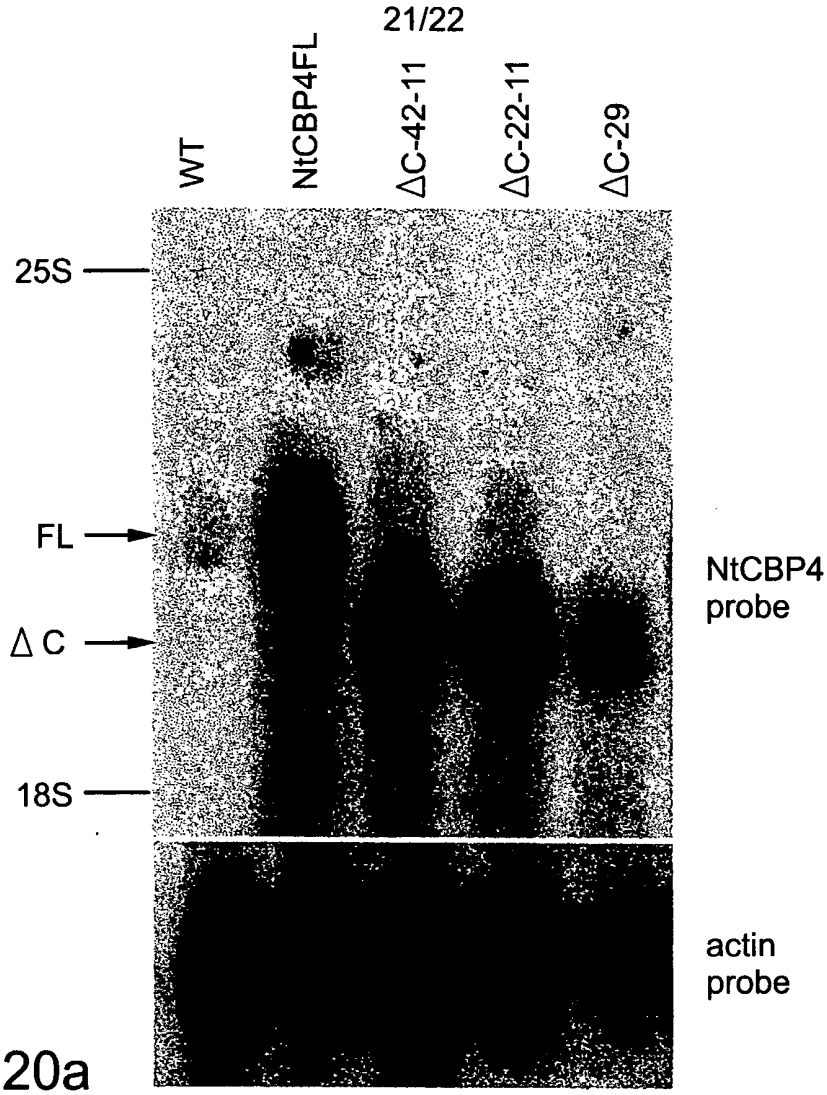


Fig. 20a

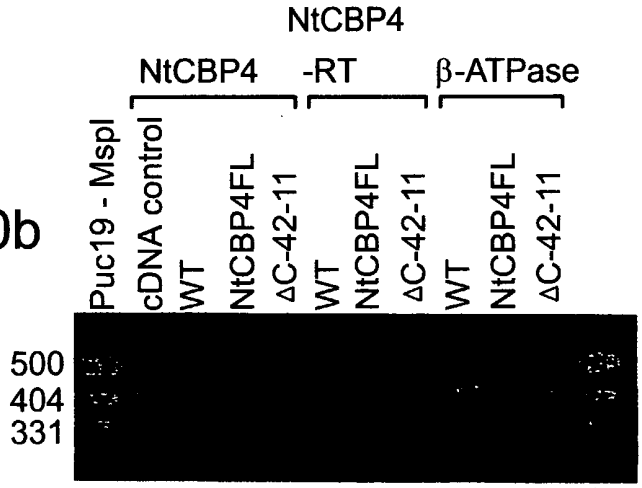


Fig. 20b

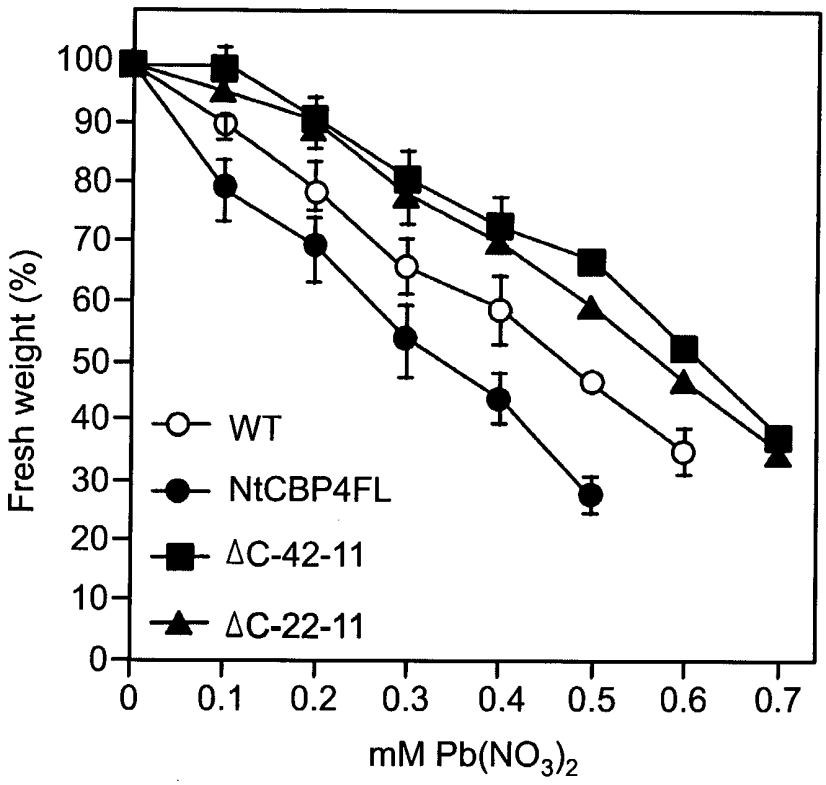


Fig. 21a

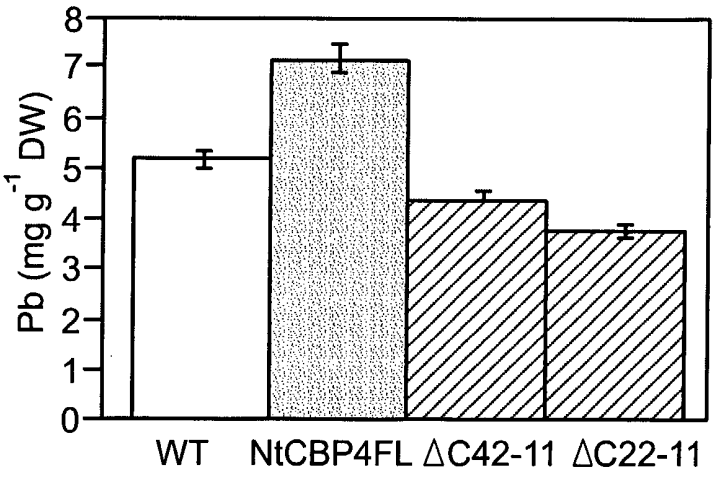


Fig. 21b

SEQUENCE LISTING

<110> YEDA RESEARCH AND DEVELOPMENT CO. LTD. et al.

<120> PLANT DERIVED CALCIUM-DEPENDENT CALMODULIN-BINDING POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAME

<130> 00/20880

<160> 51

<170> PatentIn version 3.0

<210> 1
<211> 2124
<212> DNA
<213> Nicotiana tabacum

<400> 1
atgaatcacc gccaagacga gtttgtgagg tttcaggatt ggaaatcaga gagaagctct 60

gagggaaatt ttcatgctaa aggtggagtc caccgaagta aagttgtaat agcttccgac 120

gaattacata atcgcttgga atctggtaaa tggagggcca aaggcataat acaagctgta 180

aaatctagtt taagtggttt tgtggaagaa agcttgggat ccaaaaagaa cattcttgat 240

cctcaaggac cttttcttcg aaagtggaac aaaatatttg tattatcctg tgtaattgct 300

atctccttgg atcctttgtt cctttatatt ccagtgattg ataacgacaa caaatgcctg 360

gggttaaaca gaacgttgga ggtgacagct agtgttttgc gttcttttac tgatatattt 420

tactttcttc atattgccct tcaatttcga actggtttta ttgctccttc gtcacgtgta 480

tttggaagag gtgttttaat tgaagatgct tgggagatag caaagagata cttatccact 540

tacttcttga tcgacattct tgcagttcta ccactaccac aggttgtaat tttaattata 600

attcctaagt tgcggggttc aagatctttg aataccaaga atttgctgaa gtctgttgtc 660

tttttccaat atattccaag ggtccttcga gtgtatcctt tatataggga agtcacaaga 720

acatctggca tactcactga aacagcatgg gctggagctg ctttcaatct cttcctttat 780

atgcttgcca gccatgtact tggagccttt tggtatctgt tttctataga acgcgaaact 840

acttgctgga aacaagcttg tggaaattct tctccctgtc atcatgcctc attgtattgt 900

gatgatgatc atacaaaatt taaaacattg ctgaacagtt cctgccctat agagacacca 960

aatgcaacgc tatttgattt tgggatattc cttggcgctc tccaatcagg tgttgtggga 1020

ccaatggatt ttccgcagaa gttcttttat tgtttctggt ggggtttaca aaatttaagt 1080

tcccttggtc aaaacctcca aacaagtacc ttcatctggg aaatgtgctt tgcagttttc 1140

atttccatcg ctggcctggt gctatttgcc tttctcattg gcaatatgca gacatgtctg 1200

caatcctcaa cattaagatt agaggagatg agggtgaaaa gacgagatgc agaacagtgg 1260

atgtcacatc ggttgctccc tgagaacctc agagagcgga ttaggtgcta tgagcaatac 1320

aggtggcagg aaactagggg tgttgatgaa gagaatctga ttcacaacct tccgaaagac 1380

ctcaggagag atataaagcg tcatctttgt ttagctttgc tcatgagagt gcccatgttt 1440

gaaaaaatgg atgaacaact cttagatgca ttatgcgacc atctgaagcc agtgctattc 1500

accaaggaca gcttcattgt tcgtgaaggt gatccagttg atgcgatgct gtttgtaatg 1560

aggggcaaac tactgtctgt aaccactaat ggagggagga caggtttttt taactctgag 1620

catctaaaag ctggtgattt ttgtggagag gagctactta cttgggctct tgaccctaat 1680

tcgtcaacta acctaccaat ctcaactaga actgcccaag cactctcaga agttgaagca 1740

tttgctttag tggctgatga tttaaagctt gtagcctctc agttcagacg acttcacagt 1800

aagcagctcc gccacacttt caggttttac tcaggccaat ggagaacttg ggcagcttgc 1860

ttcatacaag cagcatggcg gagttattgc aggaagaatg tagaagagtc tcttcgtgat 1920

gaagaaaaca ggttgcaaga tgcattggca aatgaaggtg gcagttcacc aagtttgggt 1980

gctacctttt atgcatcaag atttgctgct aatgtactcc atgctttgcg gcgcaatact 2040

gcaaaaaaag caagggtgcc agacaggata tcacccattc tgcttcagaa gccaacagag 2100

ccagatttta ctgctgaaga taat 2124

<210> 2
<211> 708
<212> PRT
<213> Nicotiana tabacum

<400> 2

Met Asn His Arg Gln Asp Glu Phe Val Arg Phe Gln Asp Trp Lys Ser
1 5 10 15

Glu Arg Ser Ser Glu Gly Asn Phe His Ala Lys Gly Gly Val His Arg
20 25 30

Ser Lys Val Val Ile Ala Ser Asp Glu Leu His Asn Arg Leu Glu Ser
35 40 45

Gly Lys Trp Arg Ala Lys Gly Ile Ile Gln Ala Val Lys Ser Ser Leu
50 55 60

Ser Gly Phe Val Glu Glu Ser Leu Gly Ser Lys Lys Asn Ile Leu Asp
65 70 75 80

Pro Gln Gly Pro Phe Leu Arg Lys Trp Asn Lys Ile Phe Val Leu Ser
85 90 95

Cys Val Ile Ala Ile Ser Leu Asp Pro Leu Phe Leu Tyr Ile Pro Val
100 105 110

Ile Asp Asn Asp Asn Lys Cys Leu Gly Leu Asn Arg Thr Leu Glu Val
115 120 125

Thr Ala Ser Val Leu Arg Ser Phe Thr Asp Ile Phe Tyr Phe Leu His
130 135 140

Ile Ala Leu Gln Phe Arg Thr Gly Phe Ile Ala Pro Ser Ser Arg Val
145 150 155 160

Phe Gly Arg Gly Val Leu Ile Glu Asp Ala Trp Glu Ile Ala Lys Arg
165 170 175

Tyr Leu Ser Thr Tyr Phe Leu Ile Asp Ile Leu Ala Val Leu Pro Leu
180 185 190

Pro Gln Val Val Ile Leu Ile Ile Ile Pro Lys Leu Arg Gly Ser Arg
195 200 205

Ser Leu Asn Thr Lys Asn Leu Leu Lys Ser Val Val Phe Phe Gln Tyr
210 215 220

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

Thr Ser Gly Ile Leu Thr Glu Thr Ala Trp Ala Gly Ala Ala Phe Asn
245 250 255

Leu Phe Leu Tyr Met Leu Ala Ser His Val Leu Gly Ala Phe Trp Tyr
260 265 270

Leu Phe Ser Ile Glu Arg Glu Thr Thr Cys Trp Lys Gln Ala Cys Gly
275 280 285

Asn Ser Ser Pro Cys His His Ala Ser Leu Tyr Cys Asp Asp Asp His
290 295 300

Thr Lys Phe Lys Thr Leu Leu Asn Ser Ser Cys Pro Ile Glu Thr Pro
305 310 315 320

Asn Ala Thr Leu Phe Asp Phe Gly Ile Phe Leu Gly Ala Leu Gln Ser

325 330 335

Gly Val Val Gly Pro Met Asp Phe Pro Gln Lys Phe Phe Tyr Cys Phe
340 345 350

Trp Trp Gly Leu Gln Asn Leu Ser Ser Leu Gly Gln Asn Leu Gln Thr
355 360 365

Ser Thr Phe Ile Trp Glu Met Cys Phe Ala Val Phe Ile Ser Ile Ala
370 375 380

Gly Leu Val Leu Phe Ala Phe Leu Ile Gly Asn Met Gln Thr Cys Leu
385 390 395 400

Gln Ser Ser Thr Leu Arg Leu Glu Glu Met Arg Val Lys Arg Arg Asp
405 410 415

Ala Glu Gln Trp Met Ser His Arg Leu Leu Pro Glu Asn Leu Arg Glu
420 425 430

Arg Ile Arg Cys Tyr Glu Gln Tyr Arg Trp Gln Glu Thr Arg Gly Val
435 440 445

Asp Glu Glu Asn Leu Ile His Asn Leu Pro Lys Asp Leu Arg Arg Asp
450 455 460

Ile Lys Arg His Leu Cys Leu Ala Leu Leu Met Arg Val Pro Met Phe
465 470 475 480

Glu Lys Met Asp Glu Gln Leu Leu Asp Ala Leu Cys Asp His Leu Lys
485 490 495

Pro Val Leu Phe Thr Lys Asp Ser Phe Ile Val Arg Glu Gly Asp Pro
500 505 510

Val Asp Ala Met Leu Phe Val Met Arg Gly Lys Leu Leu Ser Val Thr
515 520 525

Thr Asn Gly Gly Arg Thr Gly Phe Phe Asn Ser Glu His Leu Lys Ala
530 535 540

Gly Asp Phe Cys Gly Glu Glu Leu Leu Thr Trp Ala Leu Asp Pro Asn
545 550 555 560

Ser Ser Thr Asn Leu Pro Ile Ser Thr Arg Thr Ala Gln Ala Leu Ser
565 570 575

Glu Val Glu Ala Phe Ala Leu Val Ala Asp Asp Leu Lys Leu Val Ala
580 585 590

Ser Gln Phe Arg Arg Leu His Ser Lys Gln Leu Arg His Thr Phe Arg
595 600 605

Phe Tyr Ser Gly Gln Trp Arg Thr Trp Ala Ala Cys Phe Ile Gln Ala
610 615 620

Ala Trp Arg Ser Tyr Cys Arg Lys Asn Val Glu Glu Ser Leu Arg Asp
625 630 635 640

Glu Glu Asn Arg Leu Gln Asp Ala Leu Ala Asn Glu Gly Gly Ser Ser
645 650 655

Pro Ser Leu Gly Ala Thr Phe Tyr Ala Ser Arg Phe Ala Ala Asn Val
660 665 670

Leu His Ala Leu Arg Arg Asn Thr Ala Lys Lys Ala Arg Val Pro Asp
675 680 685

Arg Ile Ser Pro Ile Leu Leu Gln Lys Pro Thr Glu Pro Asp Phe Thr
690 695 700

Ala Glu Asp Asn
705

<210> 3
<211> 2106
<212> DNA
<213> Nicotiana tabacum

<400> 3

atgaccattc tcaatcaaga aaaatacata aggtttgagg actgcaagtc agaagataat 60

aggctgtttt caggcagaaa accatcaacc agatcatgga tgagtagtat tagaagagga 120

ttttccgata gactaagtag cttaaaaaga aaatcgcgct gcatcccttc tctaagtgac 180

tggccaaaac aagtgtctga aggatcatca agaaataaaa ttctggatcc tcaggagcct 240

tttcttcaat tctggaacaa aatatttgtc ttggcttgta tagtttcagt ggccattgat 300

ccattgttct tctacatttc tgtcgttgat atcaagagga aatgccttga tttggaccac 360

tcattaaaga tccctatttc tgttctgcga tcagccaccg atcttttcta tatatatcat 420

atttttgggc aatttcgaac tggctttatc gctccttctt ctcgagtatt tggtaggggt 480

gagttgattg aggactcttc ccttatagcc aagcgataca tcccttattg catcattgac 540

gttctagctg ttcttccact cccacagctc gtgctctata tcaacgctcc taacgcgaat 600

cgtgccatat ctctggtgat gaaaaaacag ttggtgattg ttgttttcac acaatatgtt 660

ccaagaatct ttaggatatt tccgttatac agagaagtaa caaggacaac aggtttttt 720

actgaaacgg cgtgggctgg agctgctttc aacttattcc tattcatgat cgccagtaat 780

gtagttggag ccctgtggta tttgatcaca gtggaaagac aagataactg ctggagccaa 840

gtatgtaagg gttttgaaga atgtgtgtta gaccacttat gttgcgggca acaagggaaa 900

aacgctcaat ttcttaattt ttcttgtcgt ctcctaaaac cggaagaaat acaagagaat 960

gactttgatt ttggtatatt ccgtgatgct cttcagtctc gagttgtaca aagaagaaac 1020

ttctggtcca aactcagcta ttgcttctgg tgggggctgc gaaatttaag ttctctaggc 1080

caaggcctta acacaagcga ttttctaggg gagattctct ttgccgtctt catctgcatt 1140

cttgggctga ttctgttttc gttgcttatt ggcaacatgc aggaatattt gcagtctatc 1200

acagttagag tagaagggat gaggttgaga aggcgggatg cagaacagtg gatgtctcac 1260

cgcatgcttc ctgacaacct tagagagcgc attcgaagat atgagcagta taaatggcaa 1320

caaactagag gtgttgatga agattatcta atttgtaacc ttcccaaaga cttgaggaga 1380

gacgtaaagc gccatctctg ctggtctttg ctcaaaagag tccccatgtt tgagaaaatg 1440

gatgaacagt tactggatgc attgtgtgat cggctgaaac cagctctttt cacggagaat 1500

agcttcataa tccgagaagg agatccagtg aatgagatgc tttttctgat gagaggtact 1560

ttgttgacta tcaccaccaa tggtggaaga actggattct tcaactctgc atctctctcg 1620

gctggtgatt tctgtggaga agagcttctt acatgggctt tggaccccaa tgcttcctcg 1680

tgtcttcccg cctcaacaag aacagtacaa gctgttatag atgttgaagc ttttgcactt 1740

acagctgatg atctcaagtt tgtagctgcc caatttagac gccttcatag taagcagatt 1800

cggcatactt ttaggttcta ttcacagcac tggaggacat gggcagcatg ctttatacaa 1860

gcagcatggc gcagacacta caggaacaag cttgagaaat ctttgagaga ggaagaagat 1920

aggctgcagg ctgcattgga aaatgagact gcaaatatac caagtcttgg ggctaccatt 1980

tatgcatcaa gatttgctgc caacgcattg cgtatcttgc ggcgcaacca cccaaagggt 2040

tccaaatcat ctagcaaagt aagtccctta ctgcttcaga agccagctga acctgatttt 2100

agctcc 2106

<210> 4
<211> 702
<212> PRT
<213> Nicotiana tabacum

<400> 4

```
Met Thr Ile Leu Asn Gln Glu Lys Tyr Ile Arg Phe Glu Asp Cys Lys
1               5                  10                  15

Ser Glu Asp Asn Arg Leu Phe Ser Gly Arg Lys Pro Ser Thr Arg Ser
            20                  25                  30

Trp Met Ser Ser Ile Arg Arg Gly Phe Ser Asp Arg Leu Ser Ser Leu
        35                  40                  45

Lys Arg Lys Ser Arg Cys Ile Pro Ser Leu Ser Asp Trp Pro Lys Gln
    50                  55                  60

Val Ser Glu Gly Ser Ser Arg Asn Lys Ile Leu Asp Pro Gln Glu Pro
65                  70                  75                  80

Phe Leu Gln Phe Trp Asn Lys Ile Phe Val Leu Ala Cys Ile Val Ser
                85                  90                  95

Val Ala Ile Asp Pro Leu Phe Phe Tyr Ile Ser Val Val Asp Ile Lys
            100                 105                 110

Arg Lys Cys Leu Asp Leu Asp His Ser Leu Lys Ile Pro Ile Ser Val
        115                 120                 125

Leu Arg Ser Ala Thr Asp Leu Phe Tyr Ile Tyr His Ile Phe Gly Gln
    130                 135                 140

Phe Arg Thr Gly Phe Ile Ala Pro Ser Ser Arg Val Phe Gly Arg Gly
145                 150                 155                 160

Glu Leu Ile Glu Asp Ser Ser Leu Ile Ala Lys Arg Tyr Ile Pro Tyr
                165                 170                 175

Cys Ile Ile Asp Val Leu Ala Val Leu Pro Leu Pro Gln Leu Val Leu
            180                 185                 190

Tyr Ile Asn Ala Pro Asn Ala Asn Arg Ala Ile Ser Leu Val Met Lys
        195                 200                 205

Lys Gln Leu Val Ile Val Val Phe Thr Gln Tyr Val Pro Arg Ile Phe
    210                 215                 220

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

Thr Glu Thr Ala Trp Ala Gly Ala Ala Phe Asn Leu Phe Leu Phe Met
                245                 250                 255

Ile Ala Ser Asn Val Val Gly Ala Leu Trp Tyr Leu Ile Thr Val Glu
            260                 265                 270

Arg Gln Asp Asn Cys Trp Ser Gln Val Cys Lys Gly Phe Glu Glu Cys
        275                 280                 285

Val Leu Asp His Leu Cys Cys Gly Gln Gln Gly Lys Asn Ala Gln Phe
    290                 295                 300

Leu Asn Phe Ser Cys Arg Leu Leu Lys Pro Glu Glu Ile Gln Glu Asn
305                 310                 315                 320

Asp Phe Asp Phe Gly Ile Phe Arg Asp Ala Leu Gln Ser Arg Val Val
                325                 330                 335

Gln Arg Arg Asn Phe Trp Ser Lys Leu Ser Tyr Cys Phe Trp Trp Gly
            340                 345                 350

Leu Arg Asn Leu Ser Ser Leu Gly Gln Gly Leu Asn Thr Ser Asp Phe
        355                 360                 365

Leu Gly Glu Ile Leu Phe Ala Val Phe Ile Cys Ile Leu Gly Leu Ile
    370                 375                 380

Leu Phe Ser Leu Leu Ile Gly Asn Met Gln Glu Tyr Leu Gln Ser Ile
385                 390                 395                 400

Thr Val Arg Val Glu Gly Met Arg Leu Arg Arg Arg Asp Ala Glu Gln
                405                 410                 415

Trp Met Ser His Arg Met Leu Pro Asp Asn Leu Arg Glu Arg Ile Arg
            420                 425                 430
```

```
Arg Tyr Glu Gln Tyr Lys Trp Gln Gln Thr Arg Gly Val Asp Glu Asp
        435                 440                 445

Tyr Leu Ile Cys Asn Leu Pro Lys Asp Leu Arg Arg Asp Val Lys Arg
    450                 455                 460

His Leu Cys Trp Ser Leu Leu Lys Arg Val Pro Met Phe Glu Lys Met
465                 470                 475                 480

Asp Glu Gln Leu Leu Asp Ala Leu Cys Asp Arg Leu Lys Pro Ala Leu
                485                 490                 495

Phe Thr Glu Asn Ser Phe Ile Ile Arg Glu Gly Asp Pro Val Asn Glu
            500                 505                 510

Met Leu Phe Leu Met Arg Gly Thr Leu Leu Thr Ile Thr Thr Asn Gly
        515                 520                 525

Gly Arg Thr Gly Phe Phe Asn Ser Ala Ser Leu Ser Ala Gly Asp Phe
    530                 535                 540

Cys Gly Glu Glu Leu Leu Thr Trp Ala Leu Asp Pro Asn Ala Ser Ser
545                 550                 555                 560

Cys Leu Pro Ala Ser Thr Arg Thr Val Gln Ala Val Ile Asp Val Glu
                565                 570                 575

Ala Phe Ala Leu Thr Ala Asp Asp Leu Lys Phe Val Ala Ala Gln Phe
            580                 585                 590

Arg Arg Leu His Ser Lys Gln Ile Arg His Thr Phe Arg Phe Tyr Ser
        595                 600                 605

Gln His Trp Arg Thr Trp Ala Ala Cys Phe Ile Gln Ala Ala Trp Arg
    610                 615                 620

Arg His Tyr Arg Asn Lys Leu Glu Lys Ser Leu Arg Glu Glu Glu Asp
625                 630                 635                 640

Arg Leu Gln Ala Ala Leu Glu Asn Glu Thr Ala Asn Ile Pro Ser Leu
                645                 650                 655

Gly Ala Thr Ile Tyr Ala Ser Arg Phe Ala Ala Asn Ala Leu Arg Ile
            660                 665                 670

Leu Arg Arg Asn His Pro Lys Gly Ser Lys Ser Ser Ser Lys Val Ser
        675                 680                 685

Pro Leu Leu Leu Gln Lys Pro Ala Glu Pro Asp Phe Ser Ser
    690                 695                 700
```

<210> 5
<211> 1479
<212> DNA
<213> Nicotiana tabacum

<400> 5

```
atttttgtcc aatatgttcc aagaattttt agaatttatc cattgtataa agaagtaaca     60

agaactgcag gcttatttac agaaacagca tggggtggag ctgctttcaa ccttttcctc    120

tacatgttag ccagtaatgt agtcggagcc ttctggtatt tgatctcagt agaacgccaa    180

gatacatgtt ggcgcgacgc atgtaagaag attgactcat gttcttcaga ctacttatac    240

tgtggaggaa acagaaacgg aaacgctttg cttctaaatt cttcttgccc tctcctgaaa    300

tcagaagata taaaagatcc aaatgacttt gattttggta tatttcttga tgctcttcag    360

tttcggatag tagaaaagca aaatttctgg tccaaactct tctattgctt ttggtggggg    420

ctgagaaact taagttctct tggccaaaac cttaagacaa gcacatttgt tggggagatt    480

ctctttgctg tattcatttc aattattggg ctaatcttgt tttccttgct tattggcaat    540

atgcagaaat atttgcaatc tatcacagtt agagtagaag agatgagagt gagaaggcgg    600

gacgcagagc aatggatgtc tcatcggatg cttcctgata atctgagagg gcgtattaga    660
```

cgacatgaac aatacaaatg gcaagaaacc agaggtgttg aagaagattt acttattcat 720

aatcttccta gagatttgag aagagattta aagcgccatc tctgttggtc tttggtcaaa 780

agagttccaa tattcgataa aatggatgaa cagttactag atgcaatgtg tgatcgactt 840

aaaccagcac tctacacaga gaagagtttc ataatccgag aaggcgatcc agtagatgaa 900

atgctctttc taatgagagg tactctatta actatgacaa caaatggtgg aagaactggt 960

tttttcaact ctgtatcact aaaagctggt gatttctgtg gagaggagct tcttacttgg 1020

gctttagacc ctcatacttc ctctagtctt cctatttcaa caagaacagt tcaagctgaa 1080

actgatatag aagcttttgc cctcactgct gatgatctca agtttgttgc ctcacagttt 1140

cgacgccttc atagcaaaca gcttcaacat agtttcaggt tttactcaca gcaatggagg 1200

acatggggtg catgctttat acaagtagca tggcgtcgac attgtaggaa caagcttgag 1260

aaatctttaa gagaggaaga agatagattg caggctgcgt tagcaaaaga gagcacaaat 1320

gcaccaagtc ttggagctac catttatgca tcaagatttg ctgctaatgc actacgagcc 1380

ttacgacgca accatacaac tggggctaaa ttatctccaa cattacctct actgcttcag 1440

aaaccagctg aaccaaattt cagtgaggaa aatcattca 1479

<210> 6
<211> 493
<212> PRT
<213> Nicotiana tabacum

<400> 6

Ile Phe Val Gln Tyr Val Pro Arg Ile Phe Arg Ile Tyr Pro Leu Tyr
1 5 10 15

Lys Glu Val Thr Arg Thr Ala Gly Leu Phe Thr Glu Thr Ala Trp Gly
20 25 30

Gly Ala Ala Phe Asn Leu Phe Leu Tyr Met Leu Ala Ser Asn Val Val
35 40 45

Gly Ala Phe Trp Tyr Leu Ile Ser Val Glu Arg Gln Asp Thr Cys Trp
50 55 60

Arg Asp Ala Cys Lys Lys Ile Asp Ser Cys Ser Ser Asp Tyr Leu Tyr
65 70 75 80

Cys Gly Gly Asn Arg Asn Gly Asn Ala Leu Leu Leu Asn Ser Ser Cys
85 90 95

Pro Leu Leu Lys Ser Glu Asp Ile Lys Asp Pro Asn Asp Phe Asp Phe
100 105 110

Gly Ile Phe Leu Asp Ala Leu Gln Phe Arg Ile Val Glu Lys Gln Asn
115 120 125

Phe Trp Ser Lys Leu Phe Tyr Cys Phe Trp Trp Gly Leu Arg Asn Leu
130 135 140

Ser Ser Leu Gly Gln Asn Leu Lys Thr Ser Thr Phe Val Gly Glu Ile
145 150 155 160

Leu Phe Ala Val Phe Ile Ser Ile Ile Gly Leu Ile Leu Phe Ser Leu
165 170 175

Leu Ile Gly Asn Met Gln Lys Tyr Leu Gln Ser Ile Thr Val Arg Val
180 185 190

Glu Glu Met Arg Val Arg Arg Arg Asp Ala Glu Gln Trp Met Ser His
195 200 205

Arg Met Leu Pro Asp Asn Leu Arg Gly Arg Ile Arg Arg His Glu Gln
210 215 220

Tyr Lys Trp Gln Glu Thr Arg Gly Val Glu Glu Asp Leu Leu Ile His
225 230 235 240

```
Asn Leu Pro Arg Asp Leu Arg Arg Asp Leu Lys Arg His Leu Cys Trp
                245                 250                 255

Ser Leu Val Lys Arg Val Pro Ile Phe Asp Lys Met Asp Glu Gln Leu
            260                 265                 270

Leu Asp Ala Met Cys Asp Arg Leu Lys Pro Ala Leu Tyr Thr Glu Lys
        275                 280                 285

Ser Phe Ile Ile Arg Glu Gly Asp Pro Val Asp Glu Met Leu Phe Leu
    290                 295                 300

Met Arg Gly Thr Leu Leu Thr Met Thr Thr Asn Gly Gly Arg Thr Gly
305                 310                 315                 320

Phe Phe Asn Ser Val Ser Leu Lys Ala Gly Asp Phe Cys Gly Glu Glu
                325                 330                 335

Leu Leu Thr Trp Ala Leu Asp Pro His Thr Ser Ser Ser Leu Pro Ile
            340                 345                 350

Ser Thr Arg Thr Val Gln Ala Glu Thr Asp Ile Glu Ala Phe Ala Leu
        355                 360                 365

Thr Ala Asp Asp Leu Lys Phe Val Ala Ser Gln Phe Arg Arg Leu His
    370                 375                 380

Ser Lys Gln Leu Gln His Ser Phe Arg Phe Tyr Ser Gln Gln Trp Arg
385                 390                 395                 400

Thr Trp Gly Ala Cys Phe Ile Gln Val Ala Trp Arg Arg His Cys Arg
                405                 410                 415

Asn Lys Leu Glu Lys Ser Leu Arg Glu Glu Glu Asp Arg Leu Gln Ala
            420                 425                 430

Ala Leu Ala Lys Glu Ser Thr Asn Ala Pro Ser Leu Gly Ala Thr Ile
        435                 440                 445

Tyr Ala Ser Arg Phe Ala Ala Asn Ala Leu Arg Ala Leu Arg Arg Asn
    450                 455                 460

His Thr Thr Gly Ala Lys Leu Ser Pro Thr Leu Pro Leu Leu Leu Gln
465                 470                 475                 480

Lys Pro Ala Glu Pro Asn Phe Ser Glu Glu Asn His Ser
                485                 490
```

<210> 7
<211> 543
<212> DNA
<213> Nicotiana tabacum

<400> 7

```
gcgtgggcgg gcgctgcatc ctatttgctg ttgtacatgc ttgctagtca catagttggt     60

tccttttggt atttgctttc tgtggaacgc tatgatacat gctggcaaga agcttgtaaa    120

cataatggaa catgcaacac tgatttcttg tactgtggca atcaaggaat gaaaggatat    180

aatgcttgga tcattatcag tgattcggtt ttaaatgaag catgccctgt agatagtgac    240

gacaatcaac catttgactt tggaattttt aagtacgctt tgtcatctgg catagttttc    300

tcgatgaagt tcgtgactaa atattgctac tgtttgtggt ggggactcca gaatttaagt    360

acccttggac agggacttca aactagcaca tttcccgggg agtctctttt ctctattgca    420

cttgcaatac tcgggctcat tctctttgca ctgttgattg gtaacatgca gacatacctt    480

cagtcactta ctattcgact tgaggagatg agagttaaaa ggcgcgacgc cgaccagtgg    540

atg                                                                  543
```

<210> 8
<211> 181
<212> PRT
<213> Nicotiana tabacum

<400> 8

Ala Trp Ala Gly Ala Ala Ser Tyr Leu Leu Leu Tyr Met Leu Ala Ser
1 5 10 15

His Ile Val Gly Ser Phe Trp Tyr Leu Leu Ser Val Glu Arg Tyr Asp
20 25 30

Thr Cys Trp Gln Glu Ala Cys Lys His Asn Gly Thr Cys Asn Thr Asp
35 40 45

Phe Leu Tyr Cys Gly Asn Gln Gly Met Lys Gly Tyr Asn Ala Trp Ile
50 55 60

Ile Ile Ser Asp Ser Val Leu Asn Glu Ala Cys Pro Val Asp Ser Asp
65 70 75 80

Asp Asn Gln Pro Phe Asp Phe Gly Ile Phe Lys Tyr Ala Leu Ser Ser
85 90 95

Gly Ile Val Phe Ser Met Lys Phe Val Thr Lys Tyr Cys Tyr Cys Leu
100 105 110

Trp Trp Gly Leu Gln Asn Leu Ser Thr Leu Gly Gln Gly Leu Gln Thr
115 120 125

Ser Thr Phe Pro Gly Glu Ser Leu Phe Ser Ile Ala Leu Ala Ile Leu
130 135 140

Gly Leu Ile Leu Phe Ala Leu Leu Ile Gly Asn Met Gln Thr Tyr Leu
145 150 155 160

Gln Ser Leu Thr Ile Arg Leu Glu Glu Met Arg Val Lys Arg Arg Asp
165 170 175

Ala Asp Gln Trp Met
180

<210> 9
<211> 2028
<212> DNA
<213> Solanum tuberosum

<400> 9
atgagaagcg ttaaaagaag gattgagagg ggttctgaaa tgataagtag caggaggaaa 60

ccagttcgcg tccataatgt aactaataag ctaacaaaaa atcaatctga tgcatcaaac 120

aagaaagttc ttgatcctca gagtcaattt cttcagaaat ggaacaaagt atctatattg 180

gtttgtgtat ttgcagtgtc acttgatccg ttgttcttct acattcctgt tattgataac 240

aaaaataagt gcctagattt ggacaagacg ttaaagatca ctgcttgcgt tctgcgttct 300

atcactgatc tattctatat ctttcacatt atcttgaaat ttcgtactgg cttcattact 360

ccttcttctc gagtttttgg aaggggtgag ttgattgaag attcctctgc tatagccaag 420

cgatatttgt tatcttattt cattgttgac gtcgtagcag tccttccact cccacagatt 480

gtgatattga ttatcgctcc caacatgaat ggccccattt ctctggcgac gacagaaatg 540

ttgaagattg tagtttttgc tcaatatgct ccaagacttt ttagaattat tccattgtac 600

aaagaagtag aaaggattac aggcttcttt agtggaagta catggggtgg agctgttttt 660

taccttttcc tcttcatgtg gtgcagtaat gtaactggag ccttttggta tctgttctca 720

atagaacgcc aagatgcatg ctggcgtagt gcgtgtgaca agatccccaa ctgtttgtca 780

gactacttac attgtggagg aaaaaggaat ggaaatactt ttttactgaa ttcttcttgc 840

cctctcctgc aacaagagga tataaaagat ccaaatgact ttgattttgg aattgcactt 900

gatgctcttc agtttcaggt tgtggaaaaa agaaaattcc ggaccaaact cttgtattgc 960

ttttggtggg ggctgagaaa cttaagttct cttggccaaa accttaagac aagcaccttc 1020

gatggagata ttatcttcgc catctgcatc tcaattatgg ggcttatctt gttttccttg 1080

attattggca atatgcagaa attattgcaa tttgacctgg ttagagtaga agagatgaga 1140

gcaaggaggt gggatgtaga acagtggatg tctaatcgca tgcttcctga caatctgaga 1200

gagcaaatca gaagacatga acagtacaaa tggcaacaaa ccaagggagt tgaggaagat 1260

tcattcattc aaaatcttcc tagagacttg aggagaaatt taaagcgaca tctctgttgg 1320

tctttgctct atagagttcc catatttgag aaaatggacg aacaatcact gcatgtgttg 1380

tgtgatagac tgaaaccagc acgcttcaca gagaagagct acataatccg agaaggcgaa 1440

ccagttgaag aaatgcactt tctaatgaga ggtgcagtgt tatctatgac cacaaatggt 1500

ggaaggactg gttttttcaa ctctgttcat ctcaaggctg gtgacttctg tggagatgag 1560

cttttgacat gggttataag ccctcattct tcctccagta gtcttccagt ttcgactaga 1620

acagtgcagg ctgtaactga tatagaagct tttgccctca ctgctgatga cctcaagttt 1680

gtcgtctcac aattgagacg tctgcatagc aagcaacttc aacatacttt caagttctac 1740

tcacaagaat ggaggacatg ggctgcatgc tttatacaag tagcatggcg tcgacactgc 1800

aggaataagc ttgagaaatc cttaagagag gaagaagata aactgcaggc tacattagct 1860

aaagagagca caaatgcacc aagtcttgga gataccatat atgcatcaag gtttgcagct 1920

aatatgctat gcgccttgcg acgcaacaat acaactggta ccaaatcatc tcctacatta 1980

cgtcgactgc ttcataaacc agctgaacca gattttagtg agaaaaaa 2028

<210> 10
<211> 676
<212> PRT
<213> Solanum tuberosum

<400> 10

Met Arg Ser Val Lys Arg Arg Ile Glu Arg Gly Ser Glu Met Ile Ser
1 5 10 15

Ser Arg Arg Lys Pro Val Arg Val His Asn Val Thr Asn Lys Leu Thr
20 25 30

Lys Asn Gln Ser Asp Ala Ser Asn Lys Lys Val Leu Asp Pro Gln Ser
35 40 45

Gln Phe Leu Gln Lys Trp Asn Lys Val Ser Ile Leu Val Cys Val Phe
50 55 60

Ala Val Ser Leu Asp Pro Leu Phe Phe Tyr Ile Pro Val Ile Asp Asn
65 70 75 80

Lys Asn Lys Cys Leu Asp Leu Asp Lys Thr Leu Lys Ile Thr Ala Cys
85 90 95

Val Leu Arg Ser Ile Thr Asp Leu Phe Tyr Ile Phe His Ile Ile Leu
100 105 110

Lys Phe Arg Thr Gly Phe Ile Thr Pro Ser Ser Arg Val Phe Gly Arg
115 120 125

Gly Glu Leu Ile Glu Asp Ser Ser Ala Ile Ala Lys Arg Tyr Leu Leu
130 135 140

Ser Tyr Phe Ile Val Asp Val Val Ala Val Leu Pro Leu Pro Gln Ile
145 150 155 160

Val Ile Leu Ile Ile Ala Pro Asn Met Asn Gly Pro Ile Ser Leu Ala
165 170 175

Thr Thr Glu Met Leu Lys Ile Val Val Phe Ala Gln Tyr Ala Pro Arg
180 185 190

Leu Phe Arg Ile Ile Pro Leu Tyr Lys Glu Val Glu Arg Ile Thr Gly
195 200 205

```
Phe Phe Ser Gly Ser Thr Trp Gly Gly Ala Val Phe Tyr Leu Phe Leu
    210                 215                 220

Phe Met Trp Cys Ser Asn Val Thr Gly Ala Phe Trp Tyr Leu Phe Ser
225                 230                 235                 240

Ile Glu Arg Gln Asp Ala Cys Trp Arg Ser Ala Cys Asp Lys Ile Pro
                245                 250                 255

Asn Cys Leu Ser Asp Tyr Leu His Cys Gly Gly Lys Arg Asn Gly Asn
            260                 265                 270

Thr Phe Leu Leu Asn Ser Ser Cys Pro Leu Leu Gln Gln Glu Asp Ile
        275                 280                 285

Lys Asp Pro Asn Asp Phe Asp Phe Gly Ile Ala Leu Asp Ala Leu Gln
    290                 295                 300

Phe Gln Val Val Glu Lys Arg Lys Phe Arg Thr Lys Leu Leu Tyr Cys
305                 310                 315                 320

Phe Trp Trp Gly Leu Arg Asn Leu Ser Ser Leu Gly Gln Asn Leu Lys
                325                 330                 335

Thr Ser Thr Phe Asp Gly Asp Ile Ile Phe Ala Ile Cys Ile Ser Ile
            340                 345                 350

Met Gly Leu Ile Leu Phe Ser Leu Ile Ile Gly Asn Met Gln Lys Leu
        355                 360                 365

Leu Gln Phe Asp Leu Val Arg Val Glu Glu Met Arg Ala Arg Arg Trp
    370                 375                 380

Asp Val Glu Gln Trp Met Ser Asn Arg Met Leu Pro Asp Asn Leu Arg
385                 390                 395                 400

Glu Gln Ile Arg Arg His Glu Gln Tyr Lys Trp Gln Gln Thr Lys Gly
                405                 410                 415

Val Glu Glu Asp Ser Phe Ile Gln Asn Leu Pro Arg Asp Leu Arg Arg
            420                 425                 430

Asn Leu Lys Arg His Leu Cys Trp Ser Leu Leu Tyr Arg Val Pro Ile
        435                 440                 445

Phe Glu Lys Met Asp Glu Gln Ser Leu His Val Leu Cys Asp Arg Leu
    450                 455                 460

Lys Pro Ala Arg Phe Thr Glu Lys Ser Tyr Ile Ile Arg Glu Gly Glu
465                 470                 475                 480

Pro Val Glu Glu Met His Phe Leu Met Arg Gly Ala Val Leu Ser Met
                485                 490                 495

Thr Thr Asn Gly Gly Arg Thr Gly Phe Phe Asn Ser Val His Leu Lys
            500                 505                 510

Ala Gly Asp Phe Cys Gly Asp Glu Leu Leu Thr Trp Val Ile Ser Pro
        515                 520                 525

His Ser Ser Ser Ser Ser Leu Pro Val Ser Thr Arg Thr Val Gln Ala
    530                 535                 540

Val Thr Asp Ile Glu Ala Phe Ala Leu Thr Ala Asp Asp Leu Lys Phe
545                 550                 555                 560

Val Val Ser Gln Leu Arg Arg Leu His Ser Lys Gln Leu Gln His Thr
                565                 570                 575

Phe Lys Phe Tyr Ser Gln Glu Trp Arg Thr Trp Ala Ala Cys Phe Ile
            580                 585                 590

Gln Val Ala Trp Arg Arg His Cys Arg Asn Lys Leu Glu Lys Ser Leu
        595                 600                 605

Arg Glu Glu Glu Asp Lys Leu Gln Ala Thr Leu Ala Lys Glu Ser Thr
    610                 615                 620

Asn Ala Pro Ser Leu Gly Asp Thr Ile Tyr Ala Ser Arg Phe Ala Ala
625                 630                 635                 640
```

Asn Met Leu Cys Ala Leu Arg Arg Asn Asn Thr Thr Gly Thr Lys Ser
645 650 655

Ser Pro Thr Leu Arg Arg Leu Leu His Lys Pro Ala Glu Pro Asp Phe
660 665 670

Ser Glu Lys Lys
675

<210> 11
<211> 2148
<212> DNA
<213> Medicago sativa

<400> 11
atgagtacct ttcagcaaga gaagtttgtt aggtttcagg attctaattc agaggtttct 60

aagagaaatt atcctgcatt taagacaaca cagtcaggga gatttagaag tacaataagt 120

tcatttctg agaagtttca aaggggactt gaatctggtt ccgagaggat aaagaggttc 180

aaaacatcaa tcgaccatca tcactatgac aacgctcttt ctcgaaattt tggttcgaaa 240

aggaaaatcc ttgatccgca gggtaccttc cttcaaaagt ggaacaagat atttgtcttg 300

ttatgcgtta ttgcagtatc attggatcct ttgttctttt atgtccctgt gattgatggt 360

gaaaataaat gcctttcact ggacagaggg atggagatta cagccacttt tctgagaact 420

ttctctgatg ttttttatat aatccacatg attttccagt ttcgtactgg attcattgct 480

ccttcttctc gagtatttgg aagaggtgtt ttgatcgatg attcttgggt aattgctaag 540

aggtatctgt catcatattt cttagttgac attcttgcga ttcttccact cccacaggtg 600

gtgattctag ttatcatccc aaagatgagc ggctttaaat cactcaatac aaagaacttg 660

ctgaaagtta ttgttgtctt ccaatatgtg cctcgtttat tacggataat tccattatat 720

aaagaagtta caagaacatc tggcattctc actgaaacag cttgggcggg agctgcattc 780

aatcttttac tttacatgct tgcaagtcat gtccttggtg ccttttggta cttgttttcc 840

atagaacgag aaaccacatg ctggcaagaa gcgtgtcgaa gtaataacat gtgcaacacg 900

gcagatatgt attgcaatcc tcggggaggt tttaagtata ttgcgaaatt cttgaatgat 960

tcttgcccaa tacaggagga agataaaaaa ctatttgatt ttggaatttt cctcgatgcc 1020

cttcaatccg gtgttgtgga gtcgagagat tttccctcaa aactcttcta ctgtttttgg 1080

tggggcctaa aaaatttaag ctctcttggt cagaacctcg caacaagtac cgatttttgg 1140

gaaatctgtt ttgcaatttt cattgccata gctggtttgg tgttattttc attcctcatt 1200

ggaaatatgc agacatattt gcagtcaaca caaacaagat tggaggagat gagagtaaaa 1260

aggagggatg cagaacagtg gatgtctcat cgattgcttc ctgacgacct tagagagcga 1320

atcaggcggt atgatcagta caaatggcaa gaaaccagag gtgtaaatga agacaattta 1380

gttcgtgatc ttccaaaaga tttacgaagg gacattaagc gtcatctttg cttggatttg 1440

ttgatgaggg tgccaatgtt tgagaaaatg gacgaacaac ttttggacgc aatgtgtgac 1500

cgtcttaaac cagtgcttta cacagaacaa agctatgttg tccgagaagg agatcctgtt 1560

gatgagatgc tcttcataat gcgaggcaag ctattgacca taacaactaa cggtggacga 1620

accggcttct tcaattccga gtacttaaaa gccggcgact tctgtggaga ggagcttctc 1680

acatgggcct tagatcctcg tacctcatac aaccttccca tatccactag gacagtcaga 1740

tcaatcgtag aagttgaagc ctttgcgcta aaagccgatg atctgaagtt tgtggcgtcc 1800

caattccggc gccttcatag caagcagctt cgccacactt tcagattcta ctcgcaacac 1860

tggcgcacat gggctgcatg ttttatacaa gctgcatgga ggcgctacag caagaagaaa 1920

cttgaagagt ctcttgttga agaggagaac aggctgaaag atgcattggc aaagactggt 1980

ggtaacacaa caagtttagg tgccaccata tatgcttcta ggtttgctgc caatgcactc 2040

agattattgc gacgaagtgg caccgcaagg aagacaaggg cgctagagag aatgcctgct 2100

atattgcttc agaagcctga agaacctaac tttactgatg atgaacaa 2148

<210> 12
<211> 716
<212> PRT
<213> Medicago sativa

<400> 12

Met Ser Thr Phe Gln Gln Glu Lys Phe Val Arg Phe Gln Asp Ser Asn
1 5 10 15

Ser Glu Val Ser Lys Arg Asn Tyr Pro Ala Phe Lys Thr Thr Gln Ser
20 25 30

Gly Arg Phe Arg Ser Thr Ile Ser Ser Phe Ser Glu Lys Phe Gln Arg
35 40 45

Gly Leu Glu Ser Gly Ser Glu Arg Ile Lys Arg Phe Lys Thr Ser Ile
50 55 60

Asp His His His Tyr Asp Asn Ala Leu Ser Arg Asn Phe Gly Ser Lys
65 70 75 80

Arg Lys Ile Leu Asp Pro Gln Gly Thr Phe Leu Gln Lys Trp Asn Lys
85 90 95

Ile Phe Val Leu Leu Cys Val Ile Ala Val Ser Leu Asp Pro Leu Phe
100 105 110

Phe Tyr Val Pro Val Ile Asp Gly Glu Asn Lys Cys Leu Ser Leu Asp
115 120 125

Arg Gly Met Glu Ile Thr Ala Thr Phe Leu Arg Thr Phe Ser Asp Val
130 135 140

Phe Tyr Ile Ile His Met Ile Phe Gln Phe Arg Thr Gly Phe Ile Ala
145 150 155 160

Pro Ser Ser Arg Val Phe Gly Arg Gly Val Leu Ile Asp Asp Ser Trp
165 170 175

Val Ile Ala Lys Arg Tyr Leu Ser Ser Tyr Phe Leu Val Asp Ile Leu
180 185 190

Ala Ile Leu Pro Leu Pro Gln Val Val Ile Leu Val Ile Ile Pro Lys
195 200 205

Met Ser Gly Phe Lys Ser Leu Asn Thr Lys Asn Leu Leu Lys Val Ile
210 215 220

Val Val Phe Gln Tyr Val Pro Arg Leu Leu Arg Ile Ile Pro Leu Tyr
225 230 235 240

Lys Glu Val Thr Arg Thr Ser Gly Ile Leu Thr Glu Thr Ala Trp Ala
245 250 255

Gly Ala Ala Phe Asn Leu Leu Leu Tyr Met Leu Ala Ser His Val Leu
260 265 270

Gly Ala Phe Trp Tyr Leu Phe Ser Ile Glu Arg Glu Thr Thr Cys Trp
275 280 285

Gln Glu Ala Cys Arg Ser Asn Asn Met Cys Asn Thr Ala Asp Met Tyr
290 295 300

Cys Asn Pro Arg Gly Gly Phe Lys Tyr Ile Ala Lys Phe Leu Asn Asp
305 310 315 320

Ser Cys Pro Ile Gln Glu Glu Asp Lys Lys Leu Phe Asp Phe Gly Ile
325 330 335

Phe Leu Asp Ala Leu Gln Ser Gly Val Val Glu Ser Arg Asp Phe Pro

```
            340                 345                 350

Ser Lys Leu Phe Tyr Cys Phe Trp Trp Gly Leu Lys Asn Leu Ser Ser
        355                 360                 365

Leu Gly Gln Asn Leu Ala Thr Ser Thr Asp Phe Trp Glu Ile Cys Phe
    370                 375                 380

Ala Ile Phe Ile Ala Ile Ala Gly Leu Val Leu Phe Ser Phe Leu Ile
385                 390                 395                 400

Gly Asn Met Gln Thr Tyr Leu Gln Ser Thr Gln Thr Arg Leu Glu Glu
                405                 410                 415

Met Arg Val Lys Arg Arg Asp Ala Glu Gln Trp Met Ser His Arg Leu
            420                 425                 430

Leu Pro Asp Asp Leu Arg Glu Arg Ile Arg Arg Tyr Asp Gln Tyr Lys
        435                 440                 445

Trp Gln Glu Thr Arg Gly Val Asn Glu Asp Asn Leu Val Arg Asp Leu
    450                 455                 460

Pro Lys Asp Leu Arg Arg Asp Ile Lys Arg His Leu Cys Leu Asp Leu
465                 470                 475                 480

Leu Met Arg Val Pro Met Phe Glu Lys Met Asp Glu Gln Leu Leu Asp
                485                 490                 495

Ala Met Cys Asp Arg Leu Lys Pro Val Leu Tyr Thr Glu Gln Ser Tyr
            500                 505                 510

Val Val Arg Glu Gly Asp Pro Val Asp Glu Met Leu Phe Ile Met Arg
        515                 520                 525

Gly Lys Leu Leu Thr Ile Thr Thr Asn Gly Gly Arg Thr Gly Phe Phe
    530                 535                 540

Asn Ser Glu Tyr Leu Lys Ala Gly Asp Phe Cys Gly Glu Glu Leu Leu
545                 550                 555                 560

Thr Trp Ala Leu Asp Pro Arg Thr Ser Tyr Asn Leu Pro Ile Ser Thr
                565                 570                 575

Arg Thr Val Arg Ser Ile Val Glu Val Glu Ala Phe Ala Leu Lys Ala
            580                 585                 590

Asp Asp Leu Lys Phe Val Ala Ser Gln Phe Arg Arg Leu His Ser Lys
        595                 600                 605

Gln Leu Arg His Thr Phe Arg Phe Tyr Ser Gln His Trp Arg Thr Trp
    610                 615                 620

Ala Ala Cys Phe Ile Gln Ala Ala Trp Arg Arg Tyr Ser Lys Lys Lys
625                 630                 635                 640

Leu Glu Glu Ser Leu Val Glu Glu Glu Asn Arg Leu Lys Asp Ala Leu
                645                 650                 655

Ala Lys Thr Gly Gly Asn Thr Thr Ser Leu Gly Ala Thr Ile Tyr Ala
            660                 665                 670

Ser Arg Phe Ala Ala Asn Ala Leu Arg Leu Leu Arg Arg Ser Gly Thr
        675                 680                 685

Ala Arg Lys Thr Arg Ala Leu Glu Arg Met Pro Ala Ile Leu Leu Gln
    690                 695                 700

Lys Pro Glu Glu Pro Asn Phe Thr Asp Asp Glu Gln
705                 710                 715
```

<210> 13
<211> 17
<212> DNA
<213> synthetic oligonucleotide

<400> 13
tggaayaara tmttygt

<210> 14
<211> 17
<212> DNA
<213> synthetic oligonucleotide

<220>
<221> modified_base
<222> (6)..(6)
<223> i

<220>
<221> modified_base
<222> (9)..(9)
<223> i

<220>
<221> modified_base
<222> (15)..(15)
<223> i

<400> 14
garacngcnt gggsngg 17

<210> 15
<211> 2106
<212> DNA
<213> Hordeum vulgare

<400> 15
atggaaagag aggacagata tgttagattt caggactgca aatcagagca atctgttatc 60

tcagagaacg cagtttcacc acgcagacat tgtgtattta gatcacttaa agaaagaact 120

ggtggattat ttgcatttct tggaaacttt ttgcgttctg aaaccttgaa gagatcaatg 180

ctggaagacc ggaagtccat gcaaaatgtt tttcaccctc aaggaccatt tctgcaaaga 240

tggaacaaga tatttgtatt atcatgtatt tttgcagtct ctgtggaccc attgttcctc 300

tatatcccgg ttatcaatga taaaaatctt tgctggtatt tggataggaa gatgaaaatc 360

actgcaagtg tcctgcgctc tttcacagat atattttata tactccatat catattccaa 420

tttcggacag gctttatcac gtcttcctct acaaattttg gtaggggtgt attggtcgag 480

gatagatatg ctatagcaaa gcggtatctg tcaacatatt ttctgattga tgtctgtgct 540

gttttacccc tccctcaggt ggtcatttgg attgtcctgc caactcttca ggtctcacag 600

ttcatgaaag ccaaaaacat cttgatgctt atagttatat gccaatatgt acctcgggta 660

atccgaataa ggccattata tcttcaaatc actagatctg ctgggataat tacagagaca 720

gcatgggctg gtgctgcctt caatctaata atttacatgc ttgccagcca tgtccttgga 780

gcagtttggt acttgctttc tattcaacgc aaagatgcct gctggaaaca gaattgcagt 840

cttactagag actgtaatcc tgcatatcta tactgtggga atggtggtac aaatgctgga 900

aatgctttct tgcagaatgt ttgcgttcca aatataacca aagacaatct tccggaccca 960

ctctttggaa tctttgtgcc agctataaat aatgtttcgc agtcaacaga tttcttcgcg 1020

aaattatttt actgtgtctg gtggggtctg caaaatctca gttcgcttgg ccaaaaccta 1080

aaaacaagca cttatgcttg ggagaatttg tttgcagttt ttgtctcaat atccggctta 1140

gttttgttct cacttctgat cggtaatatg cagacatatt tgcagtcagc tactctgaga 1200

atagaagaaa cgagagtgaa aagccgcgac acagaccaat ggatgtcata tcggcttctt 1260

cctgataacc tcaaggagcg aatacggcgt tatgaacaat atagatggca agagacaagc 1320

ggggttgatg aagagcacct cctgatgaac ctccccaagg atcttagaag ggctataaaa 1380

cgacatcttt gtttatcact tctcaagagg gttccaatgt ttgaaaaaat ggatgatcag 1440

ctcttgaatg ccttgtgcga ctgcctaaag cctgttttgt acacagaagg tggctgcatt 1500

gttcgtgagg gtgatccagt aaatgaaatg ttctttatca cgagaggaaa cctcatgagc 1560

atgacaacaa acggtggaaa aactggcttt ttcaactctg acgttctgaa aagcggagat 1620

ttctgcggcg aggagcttct tacctgggct cttgatccca actcggcaac cagtctcccc 1680

agctcaacta ggacggtgaa gtcaatgtct gaagtcgaag cctttgcttt gatggctgaa 1740

gacctgaagt ttgtggccat gcaatttcga cggctacaca gcaaacagct ccggcacacc 1800

ttcaggttct attcacagca gtggagaacc tgggccgcct gtttcataca agcagcctgg 1860

cacaggcgct gcagaaagaa gatggaagat gctctgcgcg agaaggagga gagactgcaa 1920

ctcgcgattg tgaacgatgg ctccacttcg ctcagcttcg gtgcggccat atacgcttca 1980

cgttttgctc gcaacatgat gcggaccttg aggagaaatg ccacccggaa ggcccggctg 2040

caggaaaggg tgcctgcaag gctgttacag aagccagcag aacccaactt ctccgcagag 2100

gagcag 2106

<210> 16
<211> 702
<212> PRT
<213> Hordeum vulgare

<400> 16

Met Glu Arg Glu Asp Arg Tyr Val Arg Phe Gln Asp Cys Lys Ser Glu
1 5 10 15

Gln Ser Val Ile Ser Glu Asn Ala Val Ser Pro Arg Arg His Cys Val
20 25 30

Phe Arg Ser Leu Lys Glu Arg Thr Gly Gly Leu Phe Ala Phe Leu Gly
35 40 45

Asn Phe Leu Arg Ser Glu Thr Leu Lys Arg Ser Met Leu Glu Asp Arg
50 55 60

Lys Ser Met Gln Asn Val Phe His Pro Gln Gly Pro Phe Leu Gln Arg
65 70 75 80

Trp Asn Lys Ile Phe Val Leu Ser Cys Ile Phe Ala Val Ser Val Asp
85 90 95

Pro Leu Phe Leu Tyr Ile Pro Val Ile Asn Asp Lys Asn Leu Cys Trp
100 105 110

Tyr Leu Asp Arg Lys Met Lys Ile Thr Ala Ser Val Leu Arg Ser Phe
115 120 125

Thr Asp Ile Phe Tyr Ile Leu His Ile Ile Phe Gln Phe Arg Thr Gly
130 135 140

Phe Ile Thr Ser Ser Ser Thr Asn Phe Gly Arg Gly Val Leu Val Glu
145 150 155 160

Asp Arg Tyr Ala Ile Ala Lys Arg Tyr Leu Ser Thr Tyr Phe Leu Ile
165 170 175

Asp Val Cys Ala Val Leu Pro Leu Pro Gln Val Val Ile Trp Ile Val
180 185 190

Leu Pro Thr Leu Gln Val Ser Gln Phe Met Lys Ala Lys Asn Ile Leu
195 200 205

Met Leu Ile Val Ile Cys Gln Tyr Val Pro Arg Val Ile Arg Ile Arg
210 215 220

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

Ala Trp Ala Gly Ala Ala Phe Asn Leu Ile Ile Tyr Met Leu Ala Ser
245 250 255

```
His Val Leu Gly Ala Val Trp Tyr Leu Leu Ser Ile Gln Arg Lys Asp
            260                 265                 270

Ala Cys Trp Lys Gln Asn Cys Ser Leu Thr Arg Asp Cys Asn Pro Ala
        275                 280                 285

Tyr Leu Tyr Cys Gly Asn Gly Gly Thr Asn Ala Gly Asn Ala Phe Leu
    290                 295                 300

Gln Asn Val Cys Val Pro Asn Ile Thr Lys Asp Asn Leu Pro Asp Pro
305                 310                 315                 320

Leu Phe Gly Ile Phe Val Pro Ala Ile Asn Asn Val Ser Gln Ser Thr
                325                 330                 335

Asp Phe Phe Ala Lys Leu Phe Tyr Cys Val Trp Trp Gly Leu Gln Asn
            340                 345                 350

Leu Ser Ser Leu Gly Gln Asn Leu Lys Thr Ser Thr Tyr Ala Trp Glu
        355                 360                 365

Asn Leu Phe Ala Val Phe Val Ser Ile Ser Gly Leu Val Leu Phe Ser
    370                 375                 380

Leu Leu Ile Gly Asn Met Gln Thr Tyr Leu Gln Ser Ala Thr Leu Arg
385                 390                 395                 400

Ile Glu Glu Thr Arg Val Lys Ser Arg Asp Thr Asp Gln Trp Met Ser
                405                 410                 415

Tyr Arg Leu Leu Pro Asp Asn Leu Lys Glu Arg Ile Arg Arg Tyr Glu
            420                 425                 430

Gln Tyr Arg Trp Gln Glu Thr Ser Gly Val Asp Glu Glu His Leu Leu
        435                 440                 445

Met Asn Leu Pro Lys Asp Leu Arg Arg Ala Ile Lys Arg His Leu Cys
    450                 455                 460

Leu Ser Leu Leu Lys Arg Val Pro Met Phe Glu Lys Met Asp Asp Gln
465                 470                 475                 480

Leu Leu Asn Ala Leu Cys Asp Cys Leu Lys Pro Val Leu Tyr Thr Glu
                485                 490                 495

Gly Gly Cys Ile Val Arg Glu Gly Asp Pro Val Asn Glu Met Phe Phe
            500                 505                 510

Ile Thr Arg Gly Asn Leu Met Ser Met Thr Thr Asn Gly Gly Lys Thr
        515                 520                 525

Gly Phe Phe Asn Ser Asp Val Leu Lys Ser Gly Asp Phe Cys Gly Glu
    530                 535                 540

Glu Leu Leu Thr Trp Ala Leu Asp Pro Asn Ser Ala Thr Ser Leu Pro
545                 550                 555                 560

Ser Ser Thr Arg Thr Val Lys Ser Met Ser Glu Val Glu Ala Phe Ala
                565                 570                 575

Leu Met Ala Glu Asp Leu Lys Phe Val Ala Met Gln Phe Arg Arg Leu
            580                 585                 590

His Ser Lys Gln Leu Arg His Thr Phe Arg Phe Tyr Ser Gln Gln Trp
        595                 600                 605

Arg Thr Trp Ala Ala Cys Phe Ile Gln Ala Ala Trp His Arg Arg Cys
    610                 615                 620

Arg Lys Lys Met Glu Asp Ala Leu Arg Glu Lys Glu Glu Arg Leu Gln
625                 630                 635                 640

Leu Ala Ile Val Asn Asp Gly Ser Thr Ser Leu Ser Phe Gly Ala Ala
                645                 650                 655

Ile Tyr Ala Ser Arg Phe Ala Arg Asn Met Met Arg Thr Leu Arg Arg
            660                 665                 670

Asn Ala Thr Arg Lys Ala Arg Leu Gln Glu Arg Val Pro Ala Arg Leu
```

675 680 685

Leu Gln Lys Pro Ala Glu Pro Asn Phe Ser Ala Glu Glu Gln
690 695 700

<210> 17
<211> 2148
<212> DNA
<213> Arabidopsis thaliana

<400> 17
atgaatttcc gacaagagaa gtttgtaagg tttcaagatt ggaagtcgga taagacttcc 60

tccgacgtgg aatattccgg taaaaacgag attcaaactg gaatattcca gagaacaata 120

agctcaatct ccgacaagtt ttacagaagc tttgaatcaa gctctgcaag gatcaaacta 180

ttcaaaagat cttacaagtc ttactctttt aaagaagctg tttcaaaagg gattggttct 240

actcacaaaa ttcttgaccc acaaggacct tttcttcaga gatggaacaa gatctttgtt 300

ttagcttgta tcatcgctgt ctcgcttgac cctttgttct tctacgtgcc tatcatcgat 360

gatgctaaga aatgtcttgg tattgacaag aaaatggaaa taacagcaag cgttttgcgc 420

tctttcactg atgtttttta tgtccttcac atcattttcc agttccgtac tggctttatc 480

gctccttcgt ctcgtgtttt tgggagaggt gttcttgttg aggacaagcg agagatcgct 540

aaacgttact tgtcctcaca tttcataatt gacattcttg ctgttcttcc gcttccgcag 600

atggtgattt tgataatcat tccacatatg agaggttcat cgtctttgaa cacgaagaat 660

atgttgaagt ttattgtttt cttccaatat ataccgaggt ttataagaat atatccgctc 720

tacaaggaag ttacaagaac ttcaggcata ctcactgaga cagcttgggc tggagctgct 780

ttcaatctct tcctctacat gcttgctagt catgtgtttg gtgctttctg gtatttgttc 840

tcaatcgaac gcgaaacggt gtgctggaaa caagcttgtg agaggaataa ccctccgtgc 900

atttcgaagt tgttgtactg tgaccctgaa actgcaggag gcaatgcttt cctcaatgag 960

tcttgtccga ttcagacacc aaacacaaca ctcttcgact ttgggatatt ccttgacgca 1020

cttcaatccg gtgtagtgga atctcaagat ttccctcaaa agtttttta ctgtttctgg 1080

tggggtctgc agaacctcag ttcgctcggt cagaacctta aaacaagtac atatatttgg 1140

gaaatctgtt tcgcggtgtt catttctatt gcggggctgg ttttgttctc cttcttgatt 1200

ggaaatatgc agacgtatct gcaatccact accacgagat tggaggagat gagggtaaag 1260

agaagagacg cagaacaatg gatgtcacac cgtttgctac ctgagaactt gagaaaaaga 1320

atccggcgat acgagcagta caaatggcaa gagacaagag gtgttgacga agagaatctt 1380

cttagtaatc ttcccaaaga tcttagacgc gacatcaaac gtcatctctg tctcgccctt 1440

ctcatgcggg tccccatgtt tgagaaaatg gacgaacagc ttcttgatgc gctctgtgac 1500

cgtttgcaac ctgtgttata cacagaggaa agctacatag taagagaagg agatccggta 1560

gacgaaatgc tcttcataat gcgcgggaag cttctaacaa tcacaacaaa cggtggaaga 1620

accggttttt taaattccga gtatcttgga gccggtgatt tctgtggtga ggagctttta 1680

acctgggctt tagacccaca ctcatcctca aacctcccaa tctcaacaag aactgttcga 1740

gctctcatgg aagttgaagc tttcgcactt aaagctgatg acctcaaatt cgtggcttcc 1800

cagttcagac gtcttcacag caaacagcta agacatactt tcaggtacta ctcacaacaa 1860

tggaagactt gggccgcttg cttcatacaa gccgcttgga gaagatacat taagaagaaa 1920

ctcgaagagt ctcttaaaga agaagagaat cggttgcagg atgctttggc taaagaagct 1980

tgtggaagtt ccccaagcct cggtgctaca atatacgcat cacggtttgc tgcaaatatc 2040

ttgcgcacaa tacgtaggag cggatcagta aggaaaccaa ggatgccgga acgaatgcca 2100

cctatgctac ttcagaaacc agcagagcca gatttcaaca gtgatgat 2148

<210> 18
<211> 716
<212> PRT
<213> Arabidopsis thaliana

<400> 18

Met Asn Phe Arg Gln Glu Lys Phe Val Arg Phe Gln Asp Trp Lys Ser
1 5 10 15

Asp Lys Thr Ser Ser Asp Val Glu Tyr Ser Gly Lys Asn Glu Ile Gln
20 25 30

Thr Gly Ile Phe Gln Arg Thr Ile Ser Ser Ile Ser Asp Lys Phe Tyr
35 40 45

Arg Ser Phe Glu Ser Ser Ser Ala Arg Ile Lys Leu Phe Lys Arg Ser
50 55 60

Tyr Lys Ser Tyr Ser Phe Lys Glu Ala Val Ser Lys Gly Ile Gly Ser
65 70 75 80

Thr His Lys Ile Leu Asp Pro Gln Gly Pro Phe Leu Gln Arg Trp Asn
85 90 95

Lys Ile Phe Val Leu Ala Cys Ile Ile Ala Val Ser Leu Asp Pro Leu
100 105 110

Phe Phe Tyr Val Pro Ile Ile Asp Asp Ala Lys Lys Cys Leu Gly Ile
115 120 125

Asp Lys Lys Met Glu Ile Thr Ala Ser Val Leu Arg Ser Phe Thr Asp
130 135 140

Val Phe Tyr Val Leu His Ile Ile Phe Gln Phe Arg Thr Gly Phe Ile
145 150 155 160

Ala Pro Ser Ser Arg Val Phe Gly Arg Gly Val Leu Val Glu Asp Lys
165 170 175

Arg Glu Ile Ala Lys Arg Tyr Leu Ser Ser His Phe Ile Ile Asp Ile
180 185 190

Leu Ala Val Leu Pro Leu Pro Gln Met Val Ile Leu Ile Ile Ile Pro
195 200 205

His Met Arg Gly Ser Ser Ser Leu Asn Thr Lys Asn Met Leu Lys Phe
210 215 220

Ile Val Phe Phe Gln Tyr Ile Pro Arg Phe Ile Arg Ile Tyr Pro Leu
225 230 235 240

Tyr Lys Glu Val Thr Arg Thr Ser Gly Ile Leu Thr Glu Thr Ala Trp
245 250 255

Ala Gly Ala Ala Phe Asn Leu Phe Leu Tyr Met Leu Ala Ser His Val
260 265 270

Phe Gly Ala Phe Trp Tyr Leu Phe Ser Ile Glu Arg Glu Thr Val Cys
275 280 285

Trp Lys Gln Ala Cys Glu Arg Asn Asn Pro Pro Cys Ile Ser Lys Leu
290 295 300

Leu Tyr Cys Asp Pro Glu Thr Ala Gly Gly Asn Ala Phe Leu Asn Glu
305 310 315 320

Ser Cys Pro Ile Gln Thr Pro Asn Thr Thr Leu Phe Asp Phe Gly Ile
325 330 335

Phe Leu Asp Ala Leu Gln Ser Gly Val Val Glu Ser Gln Asp Phe Pro
340 345 350

Gln Lys Phe Phe Tyr Cys Phe Trp Trp Gly Leu Gln Asn Leu Ser Ser
355 360 365

Leu Gly Gln Asn Leu Lys Thr Ser Thr Tyr Ile Trp Glu Ile Cys Phe
370 375 380

Ala Val Phe Ile Ser Ile Ala Gly Leu Val Leu Phe Ser Phe Leu Ile
385 390 395 400

Gly Asn Met Gln Thr Tyr Leu Gln Ser Thr Thr Thr Arg Leu Glu Glu
405 410 415

Met Arg Val Lys Arg Arg Asp Ala Glu Gln Trp Met Ser His Arg Leu
420 425 430

Leu Pro Glu Asn Leu Arg Lys Arg Ile Arg Arg Tyr Glu Gln Tyr Lys
435 440 445

Trp Gln Glu Thr Arg Gly Val Asp Glu Glu Asn Leu Leu Ser Asn Leu
450 455 460

Pro Lys Asp Leu Arg Arg Asp Ile Lys Arg His Leu Cys Leu Ala Leu
465 470 475 480

Leu Met Arg Val Pro Met Phe Glu Lys Met Asp Glu Gln Leu Leu Asp
485 490 495

Ala Leu Cys Asp Arg Leu Gln Pro Val Leu Tyr Thr Glu Glu Ser Tyr
500 505 510

Ile Val Arg Glu Gly Asp Pro Val Asp Glu Met Leu Phe Ile Met Arg
515 520 525

Gly Lys Leu Leu Thr Ile Thr Thr Asn Gly Gly Arg Thr Gly Phe Leu
530 535 540

Asn Ser Glu Tyr Leu Gly Ala Gly Asp Phe Cys Gly Glu Glu Leu Leu
545 550 555 560

Thr Trp Ala Leu Asp Pro His Ser Ser Ser Asn Leu Pro Ile Ser Thr
565 570 575

Arg Thr Val Arg Ala Leu Met Glu Val Glu Ala Phe Ala Leu Lys Ala
580 585 590

Asp Asp Leu Lys Phe Val Ala Ser Gln Phe Arg Arg Leu His Ser Lys
595 600 605

Gln Leu Arg His Thr Phe Arg Tyr Tyr Ser Gln Gln Trp Lys Thr Trp
610 615 620

Ala Ala Cys Phe Ile Gln Ala Ala Trp Arg Arg Tyr Ile Lys Lys Lys
625 630 635 640

Leu Glu Glu Ser Leu Lys Glu Glu Glu Asn Arg Leu Gln Asp Ala Leu
645 650 655

Ala Lys Glu Ala Cys Gly Ser Ser Pro Ser Leu Gly Ala Thr Ile Tyr
660 665 670

Ala Ser Arg Phe Ala Ala Asn Ile Leu Arg Thr Ile Arg Arg Ser Gly
675 680 685

Ser Val Arg Lys Pro Arg Met Pro Glu Arg Met Pro Pro Met Leu Leu
690 695 700

Gln Lys Pro Ala Glu Pro Asp Phe Asn Ser Asp Asp
705 710 715

<210> 19
<211> 33
<212> DNA
<213> synthetic oligonucleotide;

<400> 19
gaaggaattc taattatctt cagcagtaaa atc 33

<210> 20
<211> 35
<212> DNA
<213> synthetic oligonucleotide;

<400> 20
ctagcggatc cagataaagc gtcatctttg tttag 35

<210> 21
<211> 34
<212> DNA
<213> synthetic oligonucleotide;

<400> 21
gtctagcgga tccaagcact ctcagaagtt gaag 34

<210> 22
<211> 35
<212> DNA
<213> synthetic oligonucleotide;

<400> 22
gaaggaattc ttatgcttgg gcagttctag ttgag 35

<210> 23
<211> 35
<212> DNA
<213> synthetic oligonucleotide;

<400> 23
gaaggaattc ttaacgaaga gactcttcta cattc 35

<210> 24
<211> 34
<212> DNA
<213> synthetic oligonucleotide;

<400> 24
gtctagcgga tccgtgatga agaaaacagg ttgc 34

<210> 25
<211> 36
<212> DNA
<213> synthetic oligonucleotide;

<400> 25
ccgctcgagc ccgggatcaa tcaccgccaa gacgag 36

<210> 26
<211> 39
<212> DNA
<213> synthetic oligonucleotide;

<400> 26
cggaattccc cgggttacct gaggtctttc ggaaggttg 39

<210> 27
<211> 32
<212> DNA
<213> synthetic oligonucleotide;

<400> 27
cggaattcca atcctcaaca ttaagattag ag 32

<210> 28
<211> 36
<212> DNA
<213> synthetic oligonucleotide;

<400> 28
taccgctcga gctaattatc ttcagcagta aaatct 36

<210> 29
<211> 29
<212> DNA
<213> synthetic oligonucleotide;

<400> 29
ccgctcgagt cacttggcca tcatgacct 29

<210> 30
<211> 28
<212> DNA
<213> synthetic oligonucleotide;

<400> 30
cggaattcat ggcggatcag cttacaga 28

<210> 31
<211> 20
<212> DNA
<213> synthetic oligonucleotide;

<400> 31
gacgagatgc agaacagtgg 20

<210> 32
<211> 22
<212> DNA
<213> synthetic oligonucleotide;

<400> 32
catttcattt ggagaggaca cg 22

<210> 33
<211> 21
<212> DNA
<213> synthetic oligonucleotide;

<400> 33
ccagcaagta gtttcgcgtt c 21

<210> 34
<211> 664
<212> PRT
<213> Rattus rattus

<400> 34

Met Met Thr Glu Lys Ser Asn Gly Val Lys Ser Ser Pro Ala Asn Asn
1 5 10 15

His Asn His His Pro Pro Pro Ser Ile Lys Ala Asn Gly Lys Asp Asp
20 25 30

His Arg Ala Gly Ser Arg Pro Gln Ser Val Ala Ala Asp Asp Asp Thr
35 40 45

Ser Pro Glu Leu Gln Arg Leu Ala Glu Met Asp Thr Pro Arg Arg Gly
50 55 60

Arg Gly Gly Phe Gln Arg Ile Val Arg Leu Val Gly Val Ile Arg Asp
65 70 75 80

Trp Ala Asn Lys Asn Phe Arg Glu Glu Glu Pro Arg Pro Asp Ser Phe
85 90 95

Leu Glu Arg Phe Arg Gly Pro Glu Leu Gln Thr Val Thr Thr His Gln
100 105 110

Gly Asp Asp Lys Gly Gly Lys Asp Gly Glu Gly Lys Gly Thr Lys Lys
115 120 125

Lys Phe Glu Leu Phe Val Leu Asp Pro Ala Gly Asp Trp Tyr Tyr Arg
130 135 140

Trp Leu Phe Val Ile Ala Met Pro Val Leu Tyr Asn Trp Cys Leu Leu
145 150 155 160

Val Ala Arg Ala Cys Phe Ser Asp Leu Gln Arg Asn Tyr Phe Val Val
165 170 175

```
Trp Leu Val Leu Asp Tyr Phe Ser Asp Thr Val Tyr Ile Ala Asp Leu
            180                 185                 190

Ile Ile Arg Leu Arg Thr Gly Phe Leu Glu Gln Gly Leu Leu Val Lys
        195                 200                 205

Asp Pro Lys Lys Leu Arg Asp Asn Tyr Ile His Thr Leu Gln Phe Lys
    210                 215                 220

Leu Asp Val Ala Ser Ile Ile Pro Thr Asp Leu Ile Tyr Phe Ala Val
225                 230                 235                 240

Gly Ile His Ser Pro Glu Val Arg Phe Asn Arg Leu Leu His Phe Ala
                245                 250                 255

Arg Met Phe Glu Phe Phe Asp Arg Thr Glu Thr Arg Thr Ser Tyr Pro
            260                 265                 270

Asn Ile Phe Arg Ile Ser Asn Leu Val Leu Tyr Ile Leu Val Ile Ile
        275                 280                 285

His Trp Asn Ala Cys Ile Tyr Tyr Val Ile Ser Lys Ser Ile Gly Phe
    290                 295                 300

Gly Val Asp Thr Trp Val Tyr Pro Asn Ile Thr Asp Pro Glu Tyr Gly
305                 310                 315                 320

Tyr Leu Ala Arg Glu Tyr Ile Tyr Cys Leu Tyr Trp Ser Thr Leu Thr
                325                 330                 335

Leu Thr Thr Ile Gly Glu Thr Pro Pro Pro Val Lys Asp Glu Glu Tyr
            340                 345                 350

Leu Phe Val Ile Phe Asp Phe Leu Ile Gly Val Leu Ile Phe Ala Thr
        355                 360                 365

Ile Val Gly Asn Val Gly Ser Met Ile Ser Asn Met Asn Ala Thr Arg
    370                 375                 380

Ala Glu Phe Gln Ala Lys Ile Asp Ala Val Lys His Tyr Met Gln Phe
385                 390                 395                 400

Arg Lys Val Ser Lys Asp Met Glu Ala Lys Val Ile Lys Trp Phe Asp
                405                 410                 415

Tyr Leu Trp Thr Asn Lys Lys Thr Val Asp Glu Arg Glu Val Leu Lys
            420                 425                 430

Asn Leu Pro Ala Lys Leu Arg Ala Glu Ile Ala Ile Asn Val His Leu
        435                 440                 445

Ser Thr Leu Lys Lys Val Arg Ile Phe Gln Asp Cys Glu Ala Gly Leu
    450                 455                 460

Leu Val Glu Leu Val Leu Lys Leu Arg Pro Gln Val Phe Ser Pro Gly
465                 470                 475                 480

Asp Tyr Ile Cys Arg Lys Gly Asp Ile Gly Lys Glu Met Tyr Ile Ile
                485                 490                 495

Lys Glu Gly Lys Leu Ala Val Val Ala Asp Asp Gly Val Thr Gln Tyr
            500                 505                 510

Ala Leu Leu Ser Ala Gly Ser Cys Phe Gly Glu Ile Ser Ile Leu Asn
        515                 520                 525

Ile Lys Gly Ser Lys Met Gly Asn Arg Arg Thr Ala Asn Ile Arg Ser
    530                 535                 540

Leu Gly Tyr Ser Asp Leu Phe Cys Leu Ser Lys Asp Asp Leu Met Glu
545                 550                 555                 560

Ala Val Thr Glu Tyr Pro Asp Ala Lys Lys Val Leu Glu Glu Arg Gly
                565                 570                 575

Arg Glu Ile Leu Met Lys Glu Gly Leu Leu Asp Glu Asn Glu Val Ala
            580                 585                 590

Ala Ser Met Glu Val Asp Val Gln Glu Lys Leu Glu Gln Leu Glu Thr
```

595 600 605
Met Asp Thr Leu Tyr Thr Arg Phe Ala Arg Leu Leu Ala Glu Tyr
610 615 620

Thr Gly Ala Gln Gln Lys Leu Lys Gln Arg Ile Thr Val Leu Glu Thr
625 630 635 640

Lys Met Lys Gln Asn His Glu Asp Asp Tyr Leu Ser Asp Gly Ile Asn
645 650 655

Thr Pro Glu Pro Thr Ala Ala Glu
660

<210> 35
<211> 857
<212> PRT
<213> Arabidopsis thaliana

<400> 35

Met Arg Gly Gly Ala Leu Leu Cys Gly Gln Val Gln Asp Glu Ile Glu
1 5 10 15

Gln Leu Ser Arg Glu Ser Ser His Phe Ser Leu Ser Thr Gly Ile Leu
20 25 30

Pro Ser Leu Gly Ala Arg Ser Asn Arg Arg Val Lys Leu Arg Arg Phe
35 40 45

Val Val Ser Pro Tyr Asp His Lys Tyr Arg Ile Trp Glu Ala Phe Leu
50 55 60

Val Val Leu Val Val Tyr Thr Ala Trp Val Ser Pro Phe Glu Phe Gly
65 70 75 80

Phe Leu Arg Lys Pro Arg Pro Pro Leu Ser Ile Thr Asp Asn Ile Val
85 90 95

Asn Ala Phe Phe Ala Ile Asp Ile Ile Met Thr Phe Phe Val Gly Tyr
100 105 110

Leu Asp Lys Ser Thr Tyr Leu Ile Val Asp Asp Arg Lys Gln Ile Ala
115 120 125

Phe Lys Tyr Leu Arg Ser Trp Phe Leu Leu Asp Leu Val Ser Thr Ile
130 135 140

Pro Ser Glu Ala Ala Met Arg Ile Ser Ser Gln Ser Tyr Gly Leu Phe
145 150 155 160

Asn Met Leu Arg Leu Trp Arg Leu Arg Arg Val Gly Ala Leu Phe Ala
165 170 175

Arg Leu Glu Lys Asp Arg Asn Phe Asn Tyr Phe Trp Val Arg Cys Ala
180 185 190

Lys Leu Val Cys Val Thr Leu Phe Ala Val His Cys Ala Ala Cys Phe
195 200 205

Tyr Tyr Leu Ile Ala Ala Arg Asn Ser Asn Pro Ala Lys Thr Trp Ile
210 215 220

Gly Ala Asn Val Ala Asn Phe Leu Glu Glu Ser Leu Trp Met Arg Tyr
225 230 235 240

Val Thr Ser Met Tyr Trp Ser Ile Thr Thr Leu Thr Thr Val Gly Tyr
245 250 255

Gly Asp Leu His Pro Val Asn Thr Lys Glu Met Ile Phe Asp Ile Phe
260 265 270

Tyr Met Leu Phe Asn Leu Gly Leu Thr Ala Tyr Leu Ile Gly Asn Met
275 280 285

Thr Asn Leu Val Val His Gly Thr Ser Arg Thr Arg Asn Phe Arg Asp
290 295 300

Thr Ile Gln Ala Ala Ser Asn Phe Ala His Arg Asn His Leu Pro Pro
305 310 315 320

```
Arg Leu Gln Asp Gln Met Leu Ala His Leu Cys Leu Lys Tyr Arg Thr
                325                 330                 335

Asp Ser Glu Gly Leu Gln Gln Gln Glu Thr Leu Asp Ala Leu Pro Lys
            340                 345                 350

Ala Ile Arg Ser Ser Ile Ser His Phe Leu Phe Tyr Ser Leu Met Asp
        355                 360                 365

Lys Val Tyr Leu Phe Arg Gly Val Ser Asn Asp Leu Leu Phe Gln Leu
    370                 375                 380

Val Ser Glu Met Lys Ala Glu Tyr Phe Pro Pro Lys Glu Asp Val Ile
385                 390                 395                 400

Leu Gln Asn Glu Ala Pro Thr Asp Phe Tyr Ile Leu Val Asn Gly Thr
                405                 410                 415

Ala Asp Leu Val Asp Val Asp Thr Gly Thr Glu Ser Ile Val Arg Glu
            420                 425                 430

Val Lys Ala Gly Asp Ile Ile Gly Glu Ile Gly Val Leu Cys Tyr Arg
        435                 440                 445

Pro Gln Leu Phe Thr Val Arg Thr Lys Arg Leu Cys Gln Leu Leu Arg
    450                 455                 460

Met Asn Arg Thr Thr Phe Leu Asn Ile Ile Gln Ala Asn Val Gly Asp
465                 470                 475                 480

Gly Thr Ile Ile Met Asn Asn Leu Leu Gln His Leu Lys Glu Met Asn
                485                 490                 495

Asp Pro Val Met Thr Asn Val Leu Leu Glu Ile Glu Asn Met Leu Ala
            500                 505                 510

Arg Gly Lys Met Asp Leu Pro Leu Asn Leu Cys Phe Ala Ala Ile Arg
        515                 520                 525

Glu Asp Asp Leu Leu Leu His Gln Leu Leu Lys Arg Gly Leu Asp Pro
    530                 535                 540

Asn Glu Ser Asp Asn Asn Gly Arg Thr Pro Leu His Ile Ala Ala Ser
545                 550                 555                 560

Lys Gly Thr Leu Asn Cys Val Leu Leu Leu Leu Glu Tyr His Ala Asp
                565                 570                 575

Pro Asn Cys Arg Asp Ala Glu Gly Ser Val Pro Leu Trp Glu Ala Met
            580                 585                 590

Val Glu Gly His Glu Lys Val Val Lys Val Leu Leu Glu His Gly Ser
        595                 600                 605

Thr Ile Asp Ala Gly Asp Val Gly His Phe Ala Cys Thr Ala Ala Glu
    610                 615                 620

Gln Gly Asn Leu Lys Leu Leu Lys Glu Ile Val Leu His Gly Gly Asp
625                 630                 635                 640

Val Thr Arg Pro Arg Ala Thr Gly Thr Ser Ala Leu His Thr Ala Val
                645                 650                 655

Cys Glu Glu Asn Ile Glu Met Val Lys Tyr Leu Leu Glu Gln Gly Ala
            660                 665                 670

Asp Val Asn Lys Gln Asp Met His Gly Trp Thr Pro Arg Asp Leu Ala
        675                 680                 685

Glu Gln Gln Gly His Glu Asp Ile Lys Ala Leu Phe Arg Glu Lys Leu
    690                 695                 700

His Glu Arg Arg Val His Ile Glu Thr Ser Ser Ser Val Pro Ile Leu
705                 710                 715                 720

Lys Thr Gly Ile Arg Phe Leu Gly Arg Phe Thr Ser Glu Pro Asn Ile
                725                 730                 735

Arg Pro Ala Ser Arg Glu Val Ser Phe Arg Ile Arg Glu Thr Arg Ala
            740                 745                 750
```

```
Arg Arg Lys Thr Asn Asn Phe Asp Asn Ser Leu Phe Gly Ile Leu Ala
        755                 760                 765

Asn Gln Ser Val Pro Lys Asn Gly Leu Ala Thr Val Asp Glu Gly Arg
    770                 775                 780

Thr Gly Asn Pro Val Arg Val Thr Ile Ser Cys Ala Glu Lys Asp Asp
785                 790                 795                 800

Ile Ala Gly Lys Leu Val Leu Leu Pro Gly Ser Phe Lys Glu Leu Leu
                805                 810                 815

Glu Leu Gly Ser Asn Lys Phe Gly Ile Val Ala Thr Lys Val Met Asn
            820                 825                 830

Lys Asp Asn Asn Ala Glu Ile Asp Asp Val Asp Val Ile Arg Asp Gly
        835                 840                 845

Asp His Leu Ile Phe Ala Thr Asp Ser
    850                 855
```

<210> 36
<211> 677
<212> PRT
<213> Arabidopsis thaliana

<400> 36

```
Met Ser Ile Ser Cys Thr Arg Asn Phe Phe Glu Arg Phe Cys Val Glu
1               5                   10                  15

Glu Tyr Asn Ile Asp Thr Ile Lys Gln Ser Ser Phe Leu Ser Ala Asp
            20                  25                  30

Leu Leu Pro Ser Leu Gly Ala Arg Ile Asn Gln Ser Thr Lys Leu Arg
        35                  40                  45

Lys His Ile Ile Ser Pro Phe Asn Pro Arg Tyr Arg Ala Trp Glu Met
    50                  55                  60

Trp Leu Val Leu Leu Val Ile Tyr Ser Ala Trp Ile Cys Pro Phe Gln
65                  70                  75                  80

Phe Ala Phe Ile Thr Tyr Lys Lys Asp Ala Ile Phe Ile Ile Asp Asn
                85                  90                  95

Ile Val Asn Gly Phe Phe Ala Ile Asp Ile Ile Leu Thr Phe Phe Val
            100                 105                 110

Ala Tyr Leu Asp Ser His Ser Tyr Leu Leu Val Asp Ser Pro Lys Lys
        115                 120                 125

Ile Ala Ile Arg Tyr Leu Ser Thr Trp Phe Ala Phe Asp Val Cys Ser
    130                 135                 140

Thr Ala Pro Phe Gln Pro Leu Ser Leu Leu Phe Asn Tyr Asn Gly Ser
145                 150                 155                 160

Glu Leu Gly Phe Arg Ile Leu Ser Met Leu Arg Leu Trp Arg Leu Arg
                165                 170                 175

Arg Val Ser Ser Leu Phe Ala Arg Leu Glu Lys Asp Ile Arg Phe Asn
            180                 185                 190

Tyr Phe Trp Ile Arg Cys Thr Lys Leu Ile Ser Val Thr Leu Phe Ala
        195                 200                 205

Ile His Cys Ala Gly Cys Phe Asn Tyr Leu Ile Ala Asp Arg Tyr Pro
    210                 215                 220

Asn Pro Arg Lys Thr Trp Ile Gly Ala Val Tyr Pro Asn Phe Lys Glu
225                 230                 235                 240

Ala Ser Leu Trp Asn Arg Tyr Val Thr Ala Leu Tyr Trp Ser Ile Thr
                245                 250                 255

Thr Leu Thr Thr Thr Gly Tyr Gly Asp Phe His Ala Glu Asn Pro Arg
            260                 265                 270
```

```
Glu Met Leu Phe Asp Ile Phe Phe Met Met Phe Asn Leu Gly Leu Thr
        275                 280                 285

Ala Tyr Leu Ile Gly Asn Met Thr Asn Leu Val Val His Trp Thr Ser
    290                 295                 300

Arg Thr Arg Thr Phe Arg Asp Ser Val Arg Ala Ala Ser Glu Phe Ala
305                 310                 315                 320

Ser Arg Asn Gln Leu Pro His Asp Ile Glu Asp Gln Met Leu Ser His
                325                 330                 335

Ile Cys Leu Lys Phe Lys Thr Glu Gly Leu Lys Gln Gln Glu Thr Leu
            340                 345                 350

Asn Asn Leu Pro Lys Ala Ile Arg Ser Ser Ile Ala Asn Tyr Leu Phe
        355                 360                 365

Phe Pro Ile Val His Asn Ile Tyr Leu Phe Gln Gly Val Ser Arg Asn
    370                 375                 380

Phe Leu Phe Gln Leu Val Ser Asp Ile Asp Ala Glu Tyr Phe Pro Pro
385                 390                 395                 400

Lys Glu Asp Ile Ile Leu Gln Asn Glu Ala Pro Thr Asp Leu Tyr Ile
                405                 410                 415

Leu Val Ser Gly Ala Val Asp Phe Thr Val Tyr Val Asp Gly His Asp
            420                 425                 430

Gln Phe Gln Gly Lys Ala Val Ile Gly Glu Thr Phe Gly Glu Val Gly
        435                 440                 445

Val Leu Tyr Tyr Arg Pro Gln Pro Phe Thr Val Arg Thr Thr Glu Leu
    450                 455                 460

Ser Gln Ile Leu Arg Ile Ser Arg Thr Ser Leu Met Ser Ala Met His
465                 470                 475                 480

Ala His Ala Asp Asp Gly Arg Val Ile Met Asn Asn Leu Phe Met Lys
                485                 490                 495

Leu Arg Gly Gln Gln Ser Ile Ala Ile Asp Asp Ser Asn Thr Ser Gly
            500                 505                 510

His Glu Asn Arg Asp Phe Lys Ser Met Gly Trp Glu Glu Trp Arg Asp
        515                 520                 525

Ser Arg Lys Asp Gly Tyr Gly Leu Asp Val Thr Asn Pro Thr Ser Asp
    530                 535                 540

Thr Ala Leu Met Asp Ala Ile His Lys Glu Asp Thr Glu Met Val Lys
545                 550                 555                 560

Lys Ile Leu Lys Glu Gln Lys Ile Glu Arg Ala Lys Glu Glu Arg Ser
                565                 570                 575

Ser Ser Glu Ser Ala Gly Arg Ser Tyr Ala Asn Asp Ser Ser Lys Lys
            580                 585                 590

Asp Pro Tyr Cys Ser Ser Ser Asn Gln Ile Ile Lys Pro Cys Lys Arg
        595                 600                 605

Glu Glu Lys Arg Val Thr Ile His Met Met Ser Glu Ser Lys Asn Gly
    610                 615                 620

Lys Leu Ile Leu Val Pro Ser Ser Ile Glu Glu Leu Leu Arg Leu Ala
625                 630                 635                 640

Ser Glu Lys Phe Gly Gly Cys Asn Phe Thr Lys Ile Thr Asn Ala Asp
                645                 650                 655

Asn Ala Glu Ile Asp Asp Leu Asn Val Ile Trp Asp Gly Asp His Leu
            660                 665                 670

Tyr Phe Ser Ser Asn
        675
```

<210> 37
<211> 616

<212> PRT
<213> Drosophila melanogaster

<400> 37

Met Ala Ala Val Ala Gly Leu Tyr Gly Leu Gly Glu Asp Arg Gln His
1 5 10 15

Arg Lys Lys Gln Gln Gln Gln Gln His Gln Lys Glu Gln Leu Glu
20 25 30

Gln Lys Glu Glu Gln Lys Lys Ile Ala Glu Arg Lys Leu Gln Leu Arg
35 40 45

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

Lys Leu Ser Ser Gln Asp Glu Glu Gly Gly Ala Gly His Gly Phe Gly
65 70 75 80

Gly Gly Pro Gln His Phe Glu Pro Ile Pro His Asp His Asp Phe Cys
85 90 95

Glu Arg Val Val Ile Asn Val Ser Gly Leu Arg Phe Glu Thr Gln Leu
100 105 110

Arg Thr Leu Asn Gln Phe Pro Asp Thr Leu Leu Gly Asp Pro Ala Arg
115 120 125

Arg Leu Arg Tyr Phe Asp Pro Leu Arg Asn Glu Tyr Phe Phe Asp Arg
130 135 140

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

Arg Leu Arg Arg Pro Val Asn Val Pro Leu Asp Val Phe Ser Glu Glu
165 170 175

Ile Lys Phe Tyr Glu Leu Gly Asp Gln Ala Ile Asn Lys Phe Arg Glu
180 185 190

Asp Glu Gly Phe Ile Lys Glu Glu Glu Arg Pro Leu Pro Asp Asn Glu
195 200 205

Lys Gln Arg Lys Val Trp Leu Leu Phe Glu Tyr Pro Glu Ser Ser Gln
210 215 220

Ala Ala Arg Val Val Ala Ile Ile Ser Val Phe Val Ile Leu Leu Ser
225 230 235 240

Ile Val Ile Phe Cys Leu Glu Thr Leu Pro Glu Phe Lys His Tyr Lys
245 250 255

Val Phe Asn Thr Thr Thr Asn Gly Thr Lys Ile Glu Glu Asp Glu Val
260 265 270

Pro Asp Ile Thr Asp Pro Phe Phe Leu Ile Glu Thr Leu Cys Ile Ile
275 280 285

Trp Phe Thr Phe Glu Leu Thr Val Arg Phe Leu Ala Cys Pro Asn Lys
290 295 300

Leu Asn Phe Cys Arg Asp Val Met Asn Val Ile Asp Ile Ile Ala Ile
305 310 315 320

Ile Pro Tyr Phe Ile Thr Leu Ala Thr Val Val Ala Glu Glu Glu Asp
325 330 335

Thr Leu Asn Leu Pro Lys Ala Pro Val Ser Pro Gln Asp Lys Ser Ser
340 345 350

Asn Gln Ala Met Ser Leu Ala Ile Leu Arg Val Ile Arg Leu Val Arg
355 360 365

Val Phe Arg Ile Phe Lys Leu Ser Arg His Ser Lys Gly Leu Gln Ile
370 375 380

Leu Gly Arg Thr Leu Lys Ala Ser Met Arg Glu Leu Gly Leu Leu Ile
385 390 395 400

Phe Phe Leu Phe Ile Gly Val Val Leu Phe Ser Ser Ala Val Tyr Phe
405 410 415

Ala Glu Ala Gly Ser Glu Asn Ser Phe Phe Lys Ser Ile Pro Asp Ala
420 425 430

Phe Trp Trp Ala Val Val Thr Met Thr Thr Val Gly Tyr Gly Asp Met
435 440 445

Thr Pro Val Gly Phe Trp Gly Lys Ile Val Gly Ser Leu Cys Val Val
450 455 460

Ala Gly Val Leu Thr Ile Ala Leu Pro Val Pro Val Ile Val Ser Asn
465 470 475 480

Phe Asn Tyr Phe Tyr His Arg Glu Ala Asp Arg Glu Glu Met Gln Ser
485 490 495

Gln Asn Phe Asn His Val Thr Ser Cys Ser Tyr Leu Pro Gly Ala Leu
500 505 510

Gly Gln His Leu Lys Lys Ser Ser Leu Ser Glu Ser Ser Ser Asp Ile
515 520 525

Met Asp Leu Asp Asp Gly Ile Asp Ala Thr Thr Pro Gly Leu Thr Asp
530 535 540

His Thr Gly Arg His Met Val Pro Phe Leu Arg Thr Gln Gln Ser Phe
545 550 555 560

Glu Lys Gln Gln Leu Gln Leu Gln Leu Gln Leu Gln Gln Gln Ser Gln
565 570 575

Ser Pro His Gly Gln Gln Met Thr Gln Gln Gln Gln Leu Gly Gln Asn
580 585 590

Gly Leu Arg Ser Thr Asn Ser Leu Gln Leu Arg His Asn Asn Ala Met
595 600 605

Ala Val Ser Ile Glu Thr Asp Val
610 615

<210> 38
<211> 1174
<212> PRT
<213> Drosophila melanogaster

<400> 38

Met Pro Gly Gly Arg Arg Gly Leu Val Ala Pro Gln Asn Thr Phe Leu
1 5 10 15

Glu Asn Ile Ile Arg Arg Ser Asn Ser Gln Pro Asp Ser Ser Phe Leu
20 25 30

Leu Ala Asn Ala Gln Ile Val Asp Phe Pro Ile Val Tyr Cys Asn Glu
35 40 45

Ser Phe Cys Lys Ile Ser Gly Tyr Asn Arg Ala Glu Val Met Gln Lys
50 55 60

Ser Cys Arg Tyr Val Cys Gly Phe Met Tyr Gly Glu Leu Thr Asp Lys
65 70 75 80

Glu Thr Val Gly Arg Leu Glu Tyr Thr Leu Glu Asn Gln Gln Gln Asp
85 90 95

Gln Phe Glu Ile Leu Leu Tyr Lys Lys Asn Asn Leu Gln Cys Gly Cys
100 105 110

Ala Leu Ser Gln Phe Gly Lys Ala Gln Thr Gln Glu Thr Pro Leu Trp
115 120 125

Leu Leu Leu Gln Val Ala Pro Ile Arg Asn Glu Arg Asp Leu Val Val
130 135 140

Leu Phe Leu Leu Thr Phe Arg Asp Ile Thr Ala Leu Lys Gln Pro Ile
145 150 155 160

Asp Ser Glu Asp Thr Lys Gly Val Leu Gly Leu Ser Lys Phe Ala Lys

```
                165                 170                 175

Leu Ala Arg Ser Val Thr Arg Ser Arg Gln Phe Ser Ala His Leu Pro
            180                 185                 190

Thr Leu Lys Asp Pro Thr Lys Gln Ser Asn Leu Ala His Met Met Ser
        195                 200                 205

Leu Ser Ala Asp Ile Met Pro Gln Tyr Arg Gln Glu Ala Pro Lys Thr
    210                 215                 220

Pro Pro His Ile Leu Leu His Tyr Cys Ala Phe Lys Ala Ile Trp Asp
225                 230                 235                 240

Trp Val Ile Leu Cys Leu Thr Phe Tyr Thr Ala Ile Met Val Pro Tyr
                245                 250                 255

Asn Val Ala Phe Lys Asn Lys Thr Ser Glu Asp Val Ser Leu Leu Val
            260                 265                 270

Val Asp Ser Ile Val Asp Val Ile Phe Phe Ile Asp Ile Val Leu Asn
        275                 280                 285

Phe His Thr Thr Phe Val Gly Pro Gly Gly Glu Val Val Ser Asp Pro
    290                 295                 300

Lys Val Ile Arg Met Asn Tyr Leu Lys Ser Trp Phe Ile Ile Asp Leu
305                 310                 315                 320

Leu Ser Cys Leu Pro Tyr Asp Val Phe Asn Ala Phe Asp Arg Asp Glu
                325                 330                 335

Asp Gly Ile Gly Ser Leu Phe Ser Ala Leu Lys Val Val Arg Leu Leu
            340                 345                 350

Arg Leu Gly Arg Val Val Arg Lys Leu Asp Arg Tyr Leu Glu Tyr Gly
        355                 360                 365

Ala Ala Met Leu Ile Leu Leu Leu Cys Phe Tyr Met Leu Val Ala His
    370                 375                 380

Trp Leu Ala Cys Ile Trp Tyr Ser Ile Gly Arg Ser Asp Ala Asp Asp
385                 390                 395                 400

Gly Ile Gln Tyr Ser Trp Leu Trp Lys Leu Ala Asn Val Thr Gln Ser
                405                 410                 415

Pro Tyr Ser Tyr Ile Trp Ser Asn Asp Thr Gly Pro Glu Leu Val Asn
            420                 425                 430

Gly Pro Ser Arg Lys Ser Met Tyr Val Thr Ala Leu Tyr Phe Thr Met
        435                 440                 445

Thr Cys Met Thr Ser Val Gly Phe Gly Asn Val Ala Ala Glu Thr Asp
    450                 455                 460

Asn Glu Lys Val Phe Thr Ile Cys Met Met Ile Ile Ala Ala Leu Leu
465                 470                 475                 480

Tyr Ala Thr Ile Phe Gly His Val Thr Thr Ile Ile Gln Gln Met Thr
                485                 490                 495

Ser Ala Thr Ala Lys Tyr His Asp Met Leu Asn Asn Val Arg Glu Phe
            500                 505                 510

Met Lys Leu His Glu Val Pro Lys Ala Leu Ser Glu Arg Val Met Asp
        515                 520                 525

Tyr Val Val Ser Thr Trp Ala Met Thr Lys Gly Leu Asp Thr Glu Lys
    530                 535                 540

Val Leu Asn Cys Cys Pro Lys Asp Met Lys Ala Asp Ile Cys Val His
545                 550                 555                 560

Leu Asn Arg Lys Val Phe Asp Glu His Pro Thr Phe Arg Leu Ala Ser
                565                 570                 575

Asp Gly Cys Leu Arg Ala Leu Ala Met His Phe Met Met Ser His Ser
            580                 585                 590
```

```
Ala Pro Gly Asp Leu Leu Tyr His Thr Gly Glu Ser Ile Asp Ser Leu
        595                 600                 605

Cys Phe Ile Val Thr Gly Ser Leu Glu Val Ile Gln Asp Asp Glu Val
    610                 615                 620

Val Ala Ile Leu Gly Lys Gly Asp Val Phe Gly Asp Gln Phe Trp Lys
625                 630                 635                 640

Asp Ser Ala Val Gly Gln Ser Ala Ala Asn Val Arg Ala Leu Thr Tyr
                645                 650                 655

Cys Asp Leu His Ala Ile Lys Arg Asp Lys Leu Leu Glu Val Leu Asp
            660                 665                 670

Phe Tyr Ser Ala Phe Ala Asn Ser Phe Ala Arg Asn Leu Val Leu Thr
        675                 680                 685

Tyr Asn Leu Arg His Arg Leu Ile Phe Arg Arg Val Ala Asp Val Lys
    690                 695                 700

Arg Glu Lys Glu Leu Ala Glu Arg Arg Lys Asn Glu Pro Gln Leu Pro
705                 710                 715                 720

Gln Asn Gln Asp His Leu Val Arg Lys Ile Phe Ser Lys Phe Arg Arg
                725                 730                 735

Thr Pro Gln Val Gln Ala Gly Ser Lys Glu Leu Val Gly Gly Ser Gly
            740                 745                 750

Gln Ser Asp Val Glu Lys Gly Asp Gly Glu Val Glu Arg Thr Lys Val
        755                 760                 765

Leu Pro Lys Ala Pro Lys Leu Gln Ala Ser Gln Ala Thr Leu Ala Arg
    770                 775                 780

Gln Asp Thr Ile Asp Glu Gly Gly Glu Val Asp Ser Ser Pro Pro Ser
785                 790                 795                 800

Arg Asp Ser Arg Val Val Ile Glu Gly Ala Ala Val Ser Ser Ala Thr
                805                 810                 815

Val Gly Pro Ser Pro Pro Val Ala Thr Thr Ser Ser Ala Ala Ala Gly
            820                 825                 830

Ala Gly Val Ser Gly Gly Pro Gly Ser Gly Gly Thr Val Val Ala Ile
        835                 840                 845

Val Thr Lys Ala Asp Arg Asn Leu Ala Leu Glu Arg Glu Arg Gln Ile
    850                 855                 860

Glu Met Ala Ser Ser Arg Ala Thr Thr Ser Asp Thr Tyr Asp Thr Gly
865                 870                 875                 880

Leu Arg Glu Thr Pro Pro Thr Leu Ala Gln Arg Asp Leu Val Ala Thr
                885                 890                 895

Val Leu Asp Met Lys Val Asp Val Arg Leu Glu Leu Gln Arg Met Gln
            900                 905                 910

Gln Arg Ile Gly Arg Ile Glu Asp Leu Leu Gly Glu Leu Val Lys Arg
        915                 920                 925

Leu Ala Pro Gly Ala Ser Ser Gly Gly Asn Ala Pro Asp Asn Ser Ser
    930                 935                 940

Gly Gln Thr Thr Pro Gly Asp Glu Ile Cys Ala Gly Cys Gly Ala Gly
945                 950                 955                 960

Gly Gly Gly Thr Pro Thr Thr Gln Ala Pro Pro Thr Ser Ala Val Thr
                965                 970                 975

Ser Pro Val Asp Thr Val Ile Thr Ile Ser Ser Pro Gly Ala Ser Gly
            980                 985                 990

Ser Gly Ser Gly Thr Gly Ala Gly  Ala Gly Ser Ala Val  Ala Gly Ala
        995                 1000                1005

Gly Gly  Ala Gly Leu Leu Asp  Pro Gly Ala Thr Val  Val Ser Ser
    1010                1015                1020
```

```
Ala Gly  Gly Asn Gly Leu Gly  Pro Leu Met Leu Lys  Lys Arg Arg
    1025                 1030                 1035

Ser Lys  Ser Gly Lys Ala Pro  Ala Pro Pro Glu Gln  Thr Leu Ala
    1040                 1045                 1050

Ser Thr  Ala Gly Thr Ala Thr  Ala Ala Pro Ala Gly  Val Ala Gly
    1055                 1060                 1065

Ser Gly  Met Thr Ser Ser Ala  Pro Ala Ser Ala Asp  Gln Gln Gln
    1070                 1075                 1080

Gln His  Gln Ser Ala Ala Asp  Gln Ser Pro Thr Thr  Pro Gly Ala
    1085                 1090                 1095

Glu Leu  Leu His Leu Arg Leu  Leu Glu Glu Asp Phe  Thr Ala Ala
    1100                 1105                 1110

Gln Leu  Pro Ser Thr Ser Ser  Gly Gly Ala Gly Gly  Gly Gly Gly
    1115                 1120                 1125

Ser Gly  Ser Gly Ala Thr Pro  Thr Thr Pro Pro Pro  Thr Ile Ala
    1130                 1135                 1140

Gly Gly  Ser Gly Ser Gly Thr  Pro Thr Ser Thr Thr  Ala Thr Thr
    1145                 1150                 1155

Thr Pro  Thr Gly Ser Gly Thr  Ala Thr Arg Gly Lys  Leu Asp Phe
    1160                 1165                 1170

Leu
```

<210> 39
<211> 160
<212> PRT
<213> Nicotiana tabacum

<400> 39

```
Met Pro Pro Met Leu Ser Gly Leu Leu Ala Arg Leu Val Lys Leu Leu
1               5                  10                  15

Leu Gly Arg His Gly Ser Ala Leu His Trp Arg Ala Ala Gly Ala Ala
            20                  25                  30

Thr Val Leu Leu Val Ile Val Leu Leu Ala Gly Ser Tyr Leu Ala Val
        35                  40                  45

Leu Ala Glu Arg Gly Ala Pro Gly Ala Gln Leu Ile Thr Tyr Pro Arg
    50                  55                  60

Ala Leu Trp Trp Ser Val Glu Thr Ala Thr Thr Val Gly Tyr Gly Asp
65                  70                  75                  80

Leu Tyr Pro Val Thr Leu Trp Gly Arg Leu Val Ala Val Val Val Met
                85                  90                  95

Val Ala Gly Ile Thr Ser Phe Gly Leu Val Thr Ala Ala Leu Ala Thr
            100                 105                 110

Trp Phe Val Gly Arg Glu Gln Glu Arg Arg Gly His Phe Val Arg His
        115                 120                 125

Ser Glu Lys Ala Ala Glu Glu Ala Tyr Thr Arg Thr Thr Arg Ala Leu
    130                 135                 140

His Glu Arg Phe Asp Arg Leu Glu Arg Met Leu Asp Asp Asn Arg Arg
145                 150                 155                 160
```

<210> 40
<211> 2
<212> PRT
<213> Arabidopsis thaliana

<400> 40

```
Tyr Gly
1
```

<210> 41
<211> 3
<212> PRT
<213> Nicotiana tabacum

<400> 41

Gly Gln Asn
1

<210> 42
<211> 3
<212> PRT
<213> Nicotiana tabacum

<400> 42

Gly Gln Gly
1

<210> 43
<211> 4
<212> PRT
<213> Arabidopsis thaliana

<400> 43

Gly Tyr Gly Asp
1

<210> 44
<211> 31
<212> DNA
<213> synthetic oligonucleotide;

<400> 44
ccgctgagct atgaatcacc gccaagacga g 31

<210> 45
<211> 36
<212> DNA
<213> synthetic oligonucleotide;

<400> 45
gaaggaattc ttaagaggct acaagcttta aatcat 36

<210> 46
<211> 22
<212> DNA
<213> synthetic oligonucleotide;

<400> 46
ttgtgttgga ctctggtgat gg 22

<210> 47
<211> 22
<212> DNA
<213> synthetic oligonucleotide;

<400> 47
agccaagata gagcctccaa tc 22

<210> 48
<211> 21
<212> DNA
<213> synthetic oligonucleotide;

<400> 48
gacgacttca cagtaagcag c 21

<210> 49
<211> 23
<212> DNA
<213> synthetic oligonucleotide;

<400> 49
cacgactaaa aatgcactca atc 23

<210> 50
<211> 25
<212> DNA
<213> synthetic oligonucleotide;

<400> 50
cttacaggtt tgaccgtggc tgagc 25

<210> 51
<211> 26
<212> DNA
<213> synthetic oligonucleotide;

<400> 51
tagtgatcct ctcccaaaat gtgagg 26