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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 August 2006 (03.08.2006)

PCT

(10) International Publication Number
WO 2006/081301 A2

(51) International Patent Classification:
A01H 5/00 (2006.01)

(21) International Application Number:
PCT/US2006/002661

(22) International Filing Date: 24 January 2006 (24.01.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/647,708 26 January 2005 (26.01.2005) US

(71) Applicant *(for all designated States except US)*: **WASHINGTON STATE UNIVERSITY RESEARCH FOUNDATION** [US/US]; 1610 NE Eastgate Boulevard, Pullman, Washington 99163 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **RYAN, Clarence, A.** [US/US]; NE 1280 Cove Way, Pullman, Washington 99163 (US). **PEARCE, Gregory, L.** [US/US]; 620 W. Illinois Street, Palouse, Washington 99163 (US). **HUFFAKER, Alisa** [US/US]; SE 865 Edge Knoll Drive, Pullman, Washington 99163 (US). **YAMAGUCHI, Yube** [JP/US]; 1610 NE Wheatland Drive, Pullman, Washington 99163 (US).

(74) Agent: **DOW, Alan**; BioTechnology Law Group, C/O Portfolio IP, P.O.Box 52050, Minneapolis, MN 55402 (US).

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

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

Published:
— *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 2006/081301 A2

(54) Title: PLANT DEFENSE SIGNAL PEPTIDES

(57) Abstract: A 23 amino acid peptide, *At*Ptpl, plays an important role as a signaling component of the innate immune system of Arabidopsis. The peptide precursor gene is transcribed in response to elicitors generated by pathogens, and *At*Pepl is produced to amplify the signaling pathways. Seven paralogs of the *AtproPepl* gene have been identified in the Arabidopsis genome, and orthologs have been identified in species of several agriculturally important families. *At*Pcpl and its paralogs and orthologs play important roles as endogenous signals to amplify innate immunity. The sequence of two AtPepl receptors from Arabidopsis are also provided.

PLANT DEFENSE SIGNAL PEPTIDES

Statement of Government Rights

The invention was supported, at least in part, by a grant from the Government of the United States of America (grant no. IBN 0090766 from the National Science Foundation). The Government may have certain rights to the invention.

Cross-reference to Related Cases

This application claims priority from U.S. provisional patent application Ser. No. 06/647,708, filed January 26, 2005, which is incorporated herein by reference.

Background

Technical Field

The present invention relates to materials and methods for enhancing plant disease resistance.

Background Information

Plants are exposed to numerous denizens of their environment, including bacteria, viruses, fungi, and nematodes. Although many of the interactions between these organisms and plants, particularly via the roots of the plants, are beneficial, many of the interactions are harmful to the plants. The decimation of agricultural crops, ornamental plants, and other plants by diseases caused by plant pathogens is a worldwide problem that has enormous economic impact.

Damage to plants is caused by pathogens of multiple genera. These genera include Alternaria, Ascochyta, Aspergillus, Botrytis, Cercospora, Colletotrichum, Diplodia, Erwinia, Erysiphe, Fusarium, Gaeumanomyces, Helminthosporium, Macrophomina, Magnaporthe, Mycosphaerella, Nectria, Peronospora, Phoma, Phymatotrichum, Phytophthora, Plasmopara, Podosphaera, Pseudomonas, Puccinia, Puthium, Pyrenophora, Pyricularia, Pythium, Rhizoctonia, Scerotium, Sclerotinia, Septoria, Thielaviopsis, Uncinula, Venturia, Verticillium, and Xanthomonas.

Many chemical compounds have been developed to combat these various pathogens. The activity of these compounds is typically limited to several species. As a consequence of the large number and diversity of plant pathogens, these compounds have not provided an effective solution to limiting infections in plants.

An alternative approach to controlling pathogenic infections in plants involves exploiting the natural defense mechanisms of plants to confer resistance. Many plants have developed natural resistance to some pathogens. However, resistance may be limited to certain genera of pathogens, or crops of agronomic interest may not exhibit sufficient resistance. Thus, natural plant defenses often do not provide sufficient protection against pathogens. By broadening the spectrum of pathogen defense or strengthening the defense response, it may be possible to enhance existing resistance mechanisms and promote pathogen defense in otherwise susceptible plants.

When present and active, the natural defense mechanisms of plants can be highly effective in preventing pathogen colonization and disease. Resistance is multi-tiered, with passive and active, constitutive and inducible elements.

Following the invasion of a plant by a potential pathogen, the pathogen either successfully proliferates in the host, causing associated disease symptoms, or its growth is halted by the defenses of the host plant. One such defense is the hypersensitive response (HR), rapid apoptotic cell death near the site of the infection that correlates with the generation of activated oxygen species, production of antimicrobial compounds, and reinforcement of host cell walls (Dixon and Lamb, Annu. Rev. Plant Physiol. Plant Mol. Biol. 41:339-367, 1990). Other defenses include systemic acquired resistance, which effectively protects the plant against subsequent attack by a broad range of pathogens (Ryals et al., Proc. Natl. Acad. Sci. USA 92:4202-4205, 1995).

Pathogens that elicit an HR on a given host are "avirulent" on that host, the host is "resistant," and the plant-pathogen interaction is "incompatible." If a pathogen proliferates and causes disease on the host, the pathogen is "virulent," the host is "susceptible," and the plant-pathogen interaction is "compatible."

In many cases in which a strains ("races") of a particular fungal or bacterial pathogen differ regarding virulence on a various cultivars (or wild accessions) of a particular host species, avirulent strains of the pathogen, but not virulent strains, possess one or more avirulence (avr) genes corresponding to "resistance" genes in the host. This observation is the basis for the "gene-for-gene" model of plant disease resistance (Crute et al., pp. 197-309 in Mechanisms of Resistance to Plant Disease, Fraser, ed., 1985; Ellingboe, Annu. Rev. Phytopathol. 19:125-143, 1981; Flor, Annu. Rev. Phytopathol. 9:275-296, 1971; and Keen et al., in Application of

Biotechnology to Plant Pathogen Control, Chet, ed., John Wiley & Sons, 1993, pp. 65-88).

Normally avirulence and resistance genes are organized in functional pairs. A given resistance gene is generally effective only against pathogen strains that express a specific cognate avirulence gene (Flor, Annu. Rev. Phytopathol. 9:275-296, 1971; Keen, Annu. Rev. Genet. 24:447-463, 1990). However, exceptions to this rule exist. For example the Arabidopsis RPM1 gene product (Grant et al., Science 269:843-846, 1995) is involved in the recognition of elicitors produced by *P. syringae* expressing the avirulence genes avrRpm1 or avrB (Bisgrove et al., Plant Cell 6:927-933, 1994), suggesting that resistance gene products may function as common points in transduction of distinct pathogen signals.

Resistance gene products are activated in response to pathogen signal molecules termed elicitors, production of which is controlled by pathogen avirulence genes.

A number of avirulence genes have been cloned. Many cloned avirulence genes have been shown to correspond to individual resistance genes in the cognate host plants and confer an avirulent phenotype when transferred to an otherwise virulent strain. A number of plant disease resistance genes have also been cloned. Similar features have been discovered among many of these resistance genes, in spite of the diversity of pathogens against which they act. These features include a leucine-rich-repeat (LRR), a motif found in a multitude of eukaryotic proteins with roles in signal transduction (Kobe and Deisenhofer, Trends Biochem. Sci. 19:415-421, 1994). The LRR motif is thought to be involved in protein-protein interactions and may allow interaction with other proteins that are involved in plant disease resistance. In addition, sequences predicted to encode nucleotide binding sites and leucine zippers are shared among many resistance genes (Dangl, Cell 80:383-386, 1995; Staskawicz et al., Science 268:661-667, 1995). These motifs are present and similarly organized among resistance gene products from plants as diverse as tobacco, tomato, rice, flax, and Arabidopsis, suggesting a common mechanism underlying disease resistance signal transduction throughout the plant kingdom.

The local perception of pathogen attack is conveyed to distant tissues via a transmissible signal that involves salicylic acid (SA), further activating gene expression and conditioning a state known as systemic acquired resistance (SAR). It has subsequently been found that resistance can be expressed near the region of

pathogen attack, as local acquired resistance, or can be induced systemically, depending on triggering signal and plant species. Thus the systemic and local responses collectively are referred to as acquired resistance (AR). Establishment of AR is a powerful line of plant defense because it can provide broad-spectrum resistance against viral, bacterial, and fungal challenges that would otherwise cause disease. The AR response triggers the transcriptional activation of a suite of genes encoding pathogenesis-related (PR) proteins. Included among these are hydrolases, cell-wall strengthening proteins, proteins involved in oxidative burst, the combination of which are believed to promote heightened resistance. Biochemical and genetic analyses have identified genes and molecular signals associated with acquired resistance. The Npr1/Nim1 gene plays a key regulatory role in the AR defense in Arabidopsis against a broad spectrum of fungal and bacterial pathogens (WO 98/06748; WO 94/16077; WO 98/26082). The central importance of Npr1 in dicots was further substantiated by transgenic overexpression of the cloned gene, which led to heightened disease resistance in Arabidopsis against both fungal and bacterial pathogens (WO 98/06748).

Although the bulk of AR research has defined the pathway in dicotyledonous plants, monocotyledonous plants, such as wheat, rice, and barley, have an inducible pathway that protects against pathogen attack. Acquired resistance can be conditioned by different external stimuli, including avirulent pathogen challenge, pathogen elicitor exposure, and chemical treatments, including application of SA or SA analogs, such as 2,6-dichloroisonicotinic acid (INA) or benzo(1,2,3) thiodiazole-7-carbothioic acid S-methyl ester (BTH). Given the inducibility of the AR pathway by the same classes of activating compounds in monocot and dicot plants, there is likely to be partial conservation of signaling pathways, as subsets of PR genes appear to be induced in both groups. In monocots, induced acquired resistance is broad-spectrum, extending to fungal and bacterial pests, irrespective of pathogen race, with activated resistance persisting for weeks to months. Thus, manipulation of the AR pathway in plants may promote resistance to pathogens for which there exists no genetic source of resistance.

Thus, there is a need to identify genes that may play key roles in disease defense. Expression of these genes in transgenic plants may enhance the level of disease resistance against certain pathogens.

Within the past decade, the mechanisms by which plants activate innate immunity have been found to share a number of similarities with the innate immune responses of animals (Nimchuk et al., Annu. Rev. Gen. 37:579-609, 2003; Jones and Takemoto, Curr. Opin. Immun. 16:48-62, 2004; Nürnberger and Scheel, Trends Plant Sci. 8:372-379, 2001; Nümberger et al., Immun. Rev. 198:249-266, 2004; Guttman, Biotech. Adv. 22:363-382, 2004; Staskawicz et al., Science 292:2285-2289, 2001; Nürnberger and Brunner, Curr. Opin. Plant Biol. 5:1-7, 2002). Innate immunity is initiated in animals and plants through the recognition of a variety of pathogen associated molecules that in animals are called "pathogen-associated molecular patterns," or PAMPS, and in plants are called elicitors. Peptides derived from pathogens can be powerful elicitors of plant defense responses (Hahlbrock et al., Proc. Natl. Acad. Sci. USA 92:4150-4157, 1995; van den Askerveken et al., Plant Physiol. 103:91-96, 1993; Kammpren, Curr. Opin. Plant Biol. 4:295-300, 2001; Kunze et al., Plant Cell, 16:3496-3507, 2004; Navarro et al., Plant Physiol. 135:1113-1128, 2004; Fellbrich et al., Plant J. 32:375-390, 2002); He et al., Cell 73:1255-1266, 1993).

However, there remains a need to enhance plant resistance against various biotic or abiotic stresses, including, but not limited to, disease resistance. The present invention meets these and other needs.

Summary of the Invention

We have identified a novel defense signal peptide from Arabidopsis and a variety of other plants, including many crop plants of commercial importance. In addition, we have identified genes encoding polypeptides that are processed in plant cells to produce the shorter peptides, as well as receptors for the peptides. Transgenic plants in which these defense signal peptides are expressed under the control of a heterologous promoter, such as a constitutive promoter, exhibit improved yield of plant product, reduced disease symptoms, and/or enhanced resistance to disease infestation. Peptides that comprise as few as ten amino acid residues from the carboxy-terminus of AtPep1, a 23 amino acid defense signal peptide from *Arabidopsis thaliana*, retain significant activity. In addition, individual amino acid residues that are important for activity have been defined by amino acid substitutions. Peptides and other substances that have defense signal peptide activity can be readily screened using an alkalinization assay that is

described herein, and their identity can be confirmed by exogenous application to plants or transgenic expression in plants.

According to one aspect of the present invention compositions are provided that comprise one or more isolated defense signal peptides that are 10 or more amino acid residues in length and that have substantial defense signal peptide activity. Such defense signal peptides may be longer, e.g., 15, 20, 23 or more amino acid residues in length. For example, they are more easily synthesized. Accordingly, according to various embodiments of the invention, the defense signal peptide is between about 10 and about 50 amino acid residues in length, or between about 15 and about 50 amino acid residues in length. However, longer defense signal peptides may be made and used in the practice of the invention.

According to another embodiment of the invention, compositions are provided that comprise one or more polypeptides that are processed in a plant cell to produce defense signal peptides, for example, a defense signal peptide that comprises a sequence having at least 75 percent homology, or 80 percent homology, or 85 percent homology, or 90 percent homology, or complete homology with a polypeptide selected from the group consisting of AtPep1, AtPep2, AtPep3, AtPep4, AtPep5, AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2. Alternatively, the defense signal peptide comprises a sequence having at least 90 percent homology, or complete homology, with a dicot or monocot defense signal peptide consensus sequence, as discussed herein.

Such compositions comprising peptide or polypeptide compositions may further comprise biologically acceptable carriers and/or other substances used in formulating peptides and polypeptides. For example, such compositions may be agricultural formulations that are suitable for application to plants. Accordingly, in another embodiment of the invention, plants or seeds of plants are provided that comprise such a composition applied to a plant or seed surface, respectively. When applied to plants under suitable conditions, such compositions induce the plants' innate immunity and enhancing their defense against attack by pathogens.

According to another aspect of the invention, polynucleotides that express defense signal peptides (or polypeptides, including, for example, pro-forms of defense signal peptides that are processed in plant cells to produce defense signal peptides) in plants are provided. Transgenic expression of such defense signal

peptides induces the plants' innate immunity. Accordingly, one embodiment of the present invention is an isolated polynucleotide comprising a sequence that encodes a defense signal peptide (as described above) operably linked to a plant promoter. Expression of the polynucleotide in a cell of a plant causes the plant to exhibit an improvement compared to a control plant lacking the polynucleotide that is selected from the group consisting of improved yield of plant product, reduced disease symptoms, and enhanced resistance to disease infestation. The encoded defense signal peptide is 10 or more, or 15 or more, or 20 or more, 23 or more amino acid residues in length. Alternatively, such a polynucleotide comprises a sequence that encodes a polypeptide that is processed in a plant cell to produce the defense signal peptide.

According to one embodiment of the invention, such polynucleotides encoding defense signal peptides have at least 80 percent, or at least 90 percent, or at least 95 percent, or complete sequence similarity to a polynucleotide sequence selected from the group consisting of *AtproPep1*, *AtproPep2*, *AtproPep3*, *AtproPep4*, *AtproPep5*, *AtproPep6*, *AtproPep7*, *BnproPep1*, *StproPep1*, *PbproPep1*, *GmproPep1*, *MsproPep1*, *VvproPep1*, *OsproPep1*, *OsproPep2*, *TaproPep1*, *TaproPep2*, *ZmproPep1*, and *HvproPep2*.

According to another embodiment of the invention, an isolated polynucleotide is provided that comprises a sequence that encodes a defense signal peptide operably linked to a heterologous promoter. Polynucleotides for expression in plant, bacterial, fungal (including yeast), insect, and other types of cells are contemplated. In one embodiment, expression of the polynucleotide in a cell of a plant causes the plant to exhibit an improvement compared to a control plant lacking the polynucleotide that is selected from the group consisting of improved yield of plant product, reduced disease symptoms, and enhanced resistance to disease infestation. The heterologous promoter may, for example, be a constitutive promoter or a non-constitutive promoter, including, but not limited to, an organ- or tissue-specific promoter or an inducible promoter.

According to another embodiment of the invention, cells are provided that comprise one or more of the above-mentioned polynucleotides, including, but not limited to, plant, bacterial, fungal (including yeast), and insect cells. According to another embodiment, plants that comprise such cells are provided, including, but not limited to, plants such as: Acacia, alfalfa, aneth, apple, apricot, artichoke,

arugula, asparagus, avocado, banana, barley, beans, beet, blackberry, blueberry, broccoli, brussels sprouts, cabbage, cantaloupe, carrot, cassava, castorbean, cauliflower, celery, cherry, chicory, cilantro, citrus, clementines, clover, coconut, coffee, corn, cotton, cucumber, Douglas fir, eggplant, endive, escarole, eucalyptus, fennel, figs, garlic, gourd, grape, grapefruit, honey dew, jicama, kiwifruit, lettuce, leeks, lemon, lime, Loblolly pine, linseed, mango, melon, mushroom, nectarine, nut, oat, oil palm, oil seed rape, okra, olive, onion, orange, an ornamental plant, palm, papaya, parsley, parsnip, pea, peach, peanut, pear, pepper, persimmon, pine, pineapple, plantain, plum, pomegranate, poplar, potato, pumpkin, quince, radiata pine, radicchio, radish, rapeseed, raspberry, rice, rye, sorghum, Southern pine, soybean, spinach, squash, strawberry, sugarbeet, sugarcane, sunflower, sweet potato, sweetgum, tangerine, tea, tobacco, tomato, triticale, turf grass, turnip, a vine, watermelon, wheat, yams, and zucchini. According to another embodiment, such a plant exhibits reduced symptoms from, or enhanced resistance to, a disease caused by an organism of a genus selected from the group consisting of Alternaria, Ascochyta, Aspergillus, Botrytis, Cercospora, Colletotrichum, Diplodia, Erwinia, Erysiphe, Fusarium, Gaeumanomyces, Helminthosporium, Macrophomina, Magnaporthe, Mycosphaerella, Nectria, Peronospora, Phoma, Phymatotrichum, Phytophthora, Plasmopara, Podosphaera, Pseudomonas, Puccinia, Puthium, Pyrenophora, Pyricularia, Pythium, Rhizoctonia, Scerotium, Sclerotinia, Septoria, Thielaviopsis, Uncinula, Venturia, Verticillium, and Xanthomonas. The invention further encompasses parts of such plants, including, but not limited to, seeds, seed pods, flowers, fruit, tubers, stems, cuttings, and pollen. The invention also encompasses products resulting from processing of such plants or parts thereof.

Formulations of such polynucleotides are also provided. Therefore, according to another aspect of the invention, a composition is provided that comprises one or more of the above-described polynucleotides and a biologically acceptable carrier.

According to another embodiment of the invention, methods are provided for making a defense signal peptide comprising expressing in a cell a polynucleotide as described above. Included are, for example, plant cells, bacterial cells, fungal cells, and insect cells. Such methods may further comprise purifying the defense signal peptide.

According to another embodiment of the invention, methods are provided for making a transgenic plant, comprising introducing into a cell of a plant one or more of the above-described polynucleotides of the invention, thereby producing a transformed cell, and regenerating a transgenic plant from the transformed cell, wherein, compared to a control plant lacking the polynucleotide, the transgenic plant exhibits a characteristic selected from the group consisting of substantially improved yield of plant product, substantially reduced disease symptoms, and substantially enhanced resistance to disease infestation.

According to another embodiment of the invention, methods are provided for making a plant that comprises a transgene comprising a sequence that encodes a defense signal peptide operably linked to a plant promoter, such methods comprising sexually crossing a plant that comprises the transgene with a plant that lacks the transgene, thereby producing a plurality of progeny plants, and selecting a progeny plant comprising the transgene.

According to another embodiment of the invention, methods are provided for making a plant that comprises a transgene comprising a sequence that encodes a defense signal peptide operably linked to a plant promoter, the method comprising asexually reproducing a plant that comprises the transgene, thereby producing a plurality of progeny plants, and selecting a progeny plant comprising the transgene.

According to another embodiment of the invention, methods are provided for growing a plant comprising planting a seed that comprises one or more of the above-mentioned polynucleotides of the invention, and growing the seed to produce a plant, wherein, compared to a control plant lacking said polynucleotide sequence, the plant grown from the seed exhibits a characteristic selected from the group consisting of substantially improved yield of plant product, substantially reduced disease symptoms, and substantially enhanced resistance to disease infestation.

According to another embodiment of the invention, methods are provided for detecting a plant cell comprising one or more of the above-mentioned polynucleotides of the invention in a biological sample, the method comprising contacting the biological sample with a probe that binds specifically to the polynucleotide, and detecting said binding. One such probe is a PCR primer, in which case the method comprises performing PCR on the sample and detecting

said binding by detecting an amplification product diagnostic of the presence of the polynucleotide in the sample.

According to another embodiment of the invention, kits are provided for detecting a plant cell comprising a polynucleotide according to the present invention in a biological sample, the kit comprising one or more probes that bind specifically to the polynucleotide, or to the defense signal peptide encoded by the polynucleotide, and instructions for use.

According to another embodiment of the invention, methods are provided for detecting a plant cell comprising a polynucleotide according to the invention in a biological sample, the method comprising contacting the biological sample with a probe that binds specifically to the polynucleotide, or with a probe that binds to the defense signal peptide encoded by the polypeptide (such as, for example, an antibody probe), and detecting said binding.

According to another embodiment of the invention, plant cells are provided that comprise an insertion of a foreign promoter upstream of a coding sequence for a defense signal protein, wherein the foreign promoter is operably linked to the coding sequence for the defense signal protein and the plant is characterized by a substantially enhanced resistance to a disease compared to a control plant lacking the insertion of the foreign promoter.

According to another embodiment of the invention, methods of making a transgenic plant are provided that comprise (a) introducing into cells of a plant a polynucleotide that comprises a heterologous promoter, thereby producing a cell comprising an insertion of the heterologous promoter upstream of a coding sequence for a defense signal protein, wherein expression of the defense signal protein is controlled by the foreign promoter, and (b) regenerating a transgenic plant from said cell comprising the insertion.

According to another embodiment of the invention, methods of identifying a defense signal peptide are provided, such methods comprising: (a) providing a plurality of candidate peptides having a length of at least 10 amino acids; (b) assaying said plurality of candidate peptides for defense signal peptide activity in an alkalinization assay; and (c) selecting a candidate peptide that has substantial defense signal peptide activity. The candidate peptides may be provided for such methods by, for example, chemically synthesizing the candidate peptides. Such methods may further comprise administering the candidate peptide to a plant by

applying a composition comprising the candidate peptide to the plant. Alternatively, such methods may comprise administering the candidate peptide to a plant by expressing within a cell of the plant a polynucleotide that comprises a sequence that encodes the candidate peptide, thereby producing the candidate peptide within the cell of the plant.

According to another embodiment of the invention, methods are provided for identifying a substance that enhances defense of a plant against a disease comprising (a) contacting an isolated AtPep1 receptor with a plurality of candidate substances (e.g., peptides or non-peptide compounds); (b) selecting a candidate substance that has a detectable interaction with the isolated AtPep1 receptor; and (c) applying the selected candidate substance to a plant to determine whether the selected candidate substance enhances defense of the plant against a disease.

The foregoing and other aspects of the invention will become more apparent from the following detailed description, accompanying drawings, and the claims.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Brief Description of the Figures

Fig. 1 provides the nucleotide and deduced amino acid sequences for the precursor protein for AtPep1, AtproPep1 and a number of paralogs and orthologs. Sequences that correspond to AtPep1 are underlined.

(a) AtproPep1 (Locus tag At5g64900; Gene: GenBank GeneID: 836613, Arabidopsis thaliana Chromosome V [GenBank ID#AB019236], region 25954396-25955302; mRNA: NCBI RefSeq ID#NM_125888 – 499 bp; Protein: NCBI RefSeq ID#NP_569001 – 92 aa)

Paralogs (*Arabidopsis thaliana*):

(b) AtproPep3 (Locus tag At5g64890; Gene: GenBank GeneID: 836612, Arabidopsis thaliana Chromosome V [GenBank ID#AB019236], region 25951798-25952735; mRNA: NCBI RefSeq ID#NM_125887 – 568 bp; Protein: NCBI RefSeq ID#NP_569000 – 109 aa)

(c) AtproPep4 (Locus tag At5g64905; Gene: GenBank GeneID: 836614, Arabidopsis thaliana Chromosome V [GenBank ID#AB019236], region 25956800-25957285; mRNA: NCBI RefSeq ID#NM_125889 – 486 bp; Protein: NCBI RefSeq ID#NP_569002 – 96 aa)

(d) AtproPep5 (Locus tag At5g09980; Gene: GenBank GeneID: 830859, Arabidopsis thaliana Chromosome V [GenBank ID#AB019236], region 3122757-3123909; mRNA: NCBI RefSeq ID#NM_121035 – 460 bp; Protein: NCBI RefSeq ID#NP_568223 – 81 aa)

(e) AtproPep2 (Locus tag At5g09990; Gene: GenBank GeneID: 830860, Arabidopsis thaliana Chromosome V [GenBank ID#AB019236], region 3124569-3125073; mRNA: NCBI RefSeq ID#NM_121036 – 412 bp; Protein: NCBI RefSeq ID#NP_568224 – 86 aa)

(f) AtproPep6 (Locus tag At2g22000; Gene: GenBank GeneID: 816736, Arabidopsis thaliana Chromosome II [GenBank ID#AC007019], region 9369406-9370082; mRNA: NCBI RefSeq ID#NM_127769 – 397 bp; Protein: NCBI RefSeq ID#NP_179791 – 104 aa)

(g) AtproPep7 (Unannotated AtproPep; Gene: No GenBank GeneID, located on Arabidopsis thaliana Chromosome V [GenBank ID#AB019236], region 3121350-3121577; mRNA: No mRNA predicted, therefore no ID#; a 228 bp open reading frame (ORF) is encoded in genomic DNA as shown; Protein: Not predicted, therefore no ID #, Translation of ORF encoded on chromosome V yields the 75 aa sequence shown.)

Orthologs:

(h) BnproPep1 from canola (*Brassica napus*) (GenBank ID#CD816645, Protein: 95 aa)

(i) StproPep1 from potato (*Solanum tuberosum*) (GenBank ID#CV505388, Protein: 116 aa)

(j) PbproPep1 from poplar (*Populus balsamifera*) (GenBank ID#CV230975, Protein: 121 aa)

(k) BeproPep1 from birch (GenBank ID#CD276952, Protein: 110 aa)

(l) GmproPep1 from soybean (*Glycine max*) (GenBank ID#CD401281, Protein: 115 aa)

(m) MsproPep1 from alfalfa (*Medicago sativa*) (GenBank ID#BI311441, Protein: 127 aa)

(n) VvproPep1 from grape (*Vitis vinifera*) (GenBank ID#CF604664, Protein: 83 aa)

(o) OsproPep1 from rice (*Oryza sativa*) (GenBank ID#CF333408; Locus tag:Os04g54590, Protein 154 aa)

(p) OsproPep2 from rice (*Oryza sativa*) (GenBank ID#AK111113; Locus tag:Os08g07600, Protein: 93 aa)

(q) TaproPep1 from wheat (*Triticum aestivum*) (GenBank ID#AL809059, Protein: 82 aa)

(r) TaproPep2 from wheat (*Triticum aestivum*) (GenBank ID#BF201609, Protein: 75 aa)

(s) ZmproPep1 from maize (*Zea mays*) (GenBank ID#DN214793, Protein:142 aa)

(t) HvproPep1 from barley (*Hordeum vulgare*) (GenBank ID#BQ763246, Protein: 93 aa)

Fig. 2 shows the sequence of an AtPep1 receptor gene (At1g73080) and its deduced amino acid sequence. Also noted are several features of the receptor polypeptide: a signal sequence (residues 1-24); cysteine pairs (residues 64 and 71; and residues 836 and 854); leucine-rich repeats (residues 76-827); transmembrane domain (residues 870-892); kinase domain (residues 927-1208); and an intron (between residues 1099 and 1100).

Fig. 3 shows the structure of a second AtPep1 receptor gene (At1g17750) and its deduced amino acid sequence. Also noted are several features of the receptor polypeptide: a signal sequence (residues 1-26); cysteine pairs (residues 62 and 71; and residues 709 and 727); leucine-rich repeats (residues 99-697); transmembrane domain (residues 738-760); kinase domain (residues 793-1079); and an intron (between residues 966 and 967).

Fig. 4 shows the concentration dependence of synthetic AtPep peptides deduced from the seven members of the AtproPep1 gene family in the

alkalinization assay. Peptide concentrations (left to right for each peptide): 0.25 nM, 2.5 nM, 25 nM, 250 nM.

Fig. 5 shows activity of synthetic AtPep1 peptides from the C-terminus of AtPep1 in the alkalinization assay, from a 9-mer (SSGRPGQHN) to a full-length 23-mer. Ten microliter aliquots of each peptide solution were tested for activity at 0.25 nM (gray), 2.5 nM (dotted), and 25 nM (black).

Fig. 6 shows the activity of single alanine amino acid substitutions at every position in the 15-mer peptide at the carboxy terminus of AtPep1 (RGKEKVSSGRPGQHN) in the alkalinization assay. The set of substituted 15-mer peptides was assayed using four-day-old Arabidopsis cells. Ten ml of each peptide (2.5 pmoles) was added to 1ml of cells to make a final concentration of 2.5nM. After 20 min, the pH of the media was recorded. The data is the average of three separate experiments.

The invention will be further described in the following example, which does not limit the scope of the invention described in the claims.

Detailed Description of the Invention

We have isolated novel defense peptides *At*Pep1 and *At*Pep2 from Arabidopsis and identified similar defense signal peptides in a variety of other plants, including many crop plants of commercial importance. This the first demonstration of the involvement of a plant-derived peptide signal in defense against pathogens. The *proAtPep1* and *2* precursor genes have been identified and belong to a seven-member gene family. The DNA sequence of the gene from *Arabidopsis thaliana* that encodes AtproPep1 and the deduced amino acid sequence of AtproPep1 are provided in Fig. 1. Orthologs have been identified in such dicots as canola, potato, poplar, alfalfa, soybean, grape and tomato, and in such monocots as rice, wheat, maize and barley, and are likely commonly found across the plant kingdom. The sequences or several paralogs and orthologs of AtPep1 are also provided in Fig. 1.

The rapid, sensitive alkalinization assay (described below) that was used to identify and purify the AtPep1 and 2 peptides is useful for identifying defense signal peptides from any plant whose signaling pathways result from peptide-receptor interactions that initiate intracellular signaling through MAP kinases and proton pumps in the plasma membrane. Within minutes after adding systemin to cells, an ATP-driven proton pump is inhibited, causing the extracellular medium of

the cells to become alkaline. When aliquots (e.g., 1-10 μL) from fractions from plant tissues that eluted from HPLC columns were added to 1 mL of suspension cultured plant cells grown at low pH (e.g., pH 5), some fractions caused the cell medium to increase in pH. The identification of a peptide as a defense signal peptide is confirmed by application of the peptide (e.g., as a plant fraction or isolated peptide) to plants or by expression of a transgene encoding the peptide, including, but not limited to, the gene encoding a pro-form of the peptide (such as, for example, AtproPep1, the pro-form of AtPep1), in plants and observation of detectable defense signal peptide activity, such as, for example, enhanced disease resistance.

We established a suspension cultured Arabidopsis cell line to be used in the alkalinization assay to seek novel peptides in Arabidopsis leaf extracts. Cells are grown unbuffered near pH 5 or less in order to record an alkalinization of about 1 pH unit in response to peptide signals. In order to establish an Arabidopsis suspension cell culture for use in the alkalinization assay, Arabidopsis cells were regularly transferred and maintained in the growth chamber room for several months, when they equilibrate at a pH of about 5.0 during exponential growth. Cell cultures that grow at low pH from several other plant species, including tomato (*Lycopersicon esculentum*), tobacco (*Nicotiana tabacum*), alfalfa (*Medicago sativa*), maize (*Zea mays*), petunia (*Petunia hybrida*), nightshade (*Solanum nigrum*), and sweet potato (*Ipomoea batatus*) have been developed for use in the alkalinization assay, and developing similar cultures from other plants is readily accomplished.

For the work described in Example 1, we used a typical purified peptide fraction from Arabidopsis leaves. Kilogram quantities of leaf material were extracted and peptides were separated on an HPLC column. A 10 microliter (μL) aliquot from each fraction eluting from the column was assayed with the alkallinization assay using suspension cultured Arabidopsis cells. Two novel peaks were identified that were called AtPep1 and AtPep2. The peptides were purified through several additional column separations until homogeneous, as verified by MALDI-MS and amino-terminal sequencing. The peptides were each 23 amino acids in length and the amino acids sequences of the two were identical at 10 residues. Neither peptide was post-translationally modified. The two peptides were chemically synthesized and, in alkalinization assays exhibited identical

activities as the native peptides, in the sub-nanomolar concentration range. Searches of protein data bases revealed that the peptides were derived from the C-terminus of two members of a seven-member gene family. The deduced precursors were from 92 to just over 100 amino acids in length and did not have leader sequences at their N-termini.

These properties are similar to tomato systemin (Pearce et al., Science 253:895-897, 1991). Other similarities between proAtPep1 prosystemin were the absence of a leader sequence in the precursors, the low nMolar concentrations needed to activate the alkalinization response, the processing of the peptides from the C-termini of their precursor proteins, and the presence of KEK motifs in the precursors that are commonly found in proteins that are involved in protein-protein interactions. The expression of the proAtPep1 gene in excised Arabidopsis leaves was induced by AtPep1, similar to the expression of the prosystemin gene in tomato plants being induced by systemin.

The tissue-specific expression of the *proAtPep1* gene was analyzed using RT-PCR analyses. All tissues of the plant expressed the gene at low levels, which did not reveal a clue as to its function. To assess whether stresses to Arabidposis plants might affect AtPep1 gene expression, the plants were subjected to cold and drought stresses and treatments with abscissic acid (ABA), methyl jasmonate (MeJA), methyl salicylate (MeSA), UV-B, and wounding. Only MeJA and wounding induced a strong expression of the gene. These results indicated that the gene was behaving in Arabidopsis in a similar manner as systemin in tomato plants (Ryan et al., Plant Cell. 14:251-264, 2002) and suggested that the peptide may be a defense signaling peptide.

Pathogen defense is well characterized in Arabidopsis, where two defense pathways have been identified in which jasmonate is a signaling component (Lorenzo and Solano, Curr. Opin. Plant Biol 8:532-540, 2005; Lorenzo et al., Plant Cell 16:1938-1950, 2004.). In one pathway, wounding and jasmonate activate defensive genes through the octadecanoid pathway, with COI1 and AMYC2 playing major roles in transcription of defensive genes that includes LOX2 and VSP2 (Lorenzo et al., Plant Cell 16:1938-1950, 2004). In the second pathway jasmonate, in concert with ethylene, activates PDF1.2, and several PR proteins, with active oxygen playing a key signaling role (Lorenzo et al., Plant Cell

15:165-178, 2003; Penninckx et al., Plant Cell 10:2103-2113, 1998; Penninckx et al., Plant Cell 8:2309-2323, 1996; Coego et al., Plant Cell 17:2123-2137, 2005; Mackerness et al., Plant Cell Environ. 22:1413-1423, 1999). To assess the possible involvement of AtPep1 with the known pathways, Arabidopsis plants were excised and supplied with the AtPep1 peptide at 10 nM concentrations, and several known wound-inducible genes and pathogen defense genes were assayed for expression levels two hours later. Only the pathogen defense genes were induced, with PDF2.2 and PR-1 being most strongly expressed, with PR-3, PR-4 and TAT expressed at lower levels. *LOX2* and *VSP2* genes were not induced by the peptide.

Arabidopsis plants were transformed with a 35S-AtPep1 fused gene, and many stable transformants that strongly expressed the gene were recovered. The overexpression of the gene did not visibly affect the growth of the transgenic plants compared to wild type plants. The progeny of a stable transformant that strongly expressed the gene was analyzed to determine if the overexpression of the gene would affect the expression of defense genes. Plants constitutively overexpressing the proAtPep1 gene also constitutively overexpressed the AtPep1-inducible genes.

To determine if the transgenic plants were more resistant to a pathogen, the soil of plants in which the transgenic line was grown was inoculated with the root pathogen *Pythium irregulare* and the plants were monitored with time to assess pathogenicity. The aerial parts of the transgenic plants infected with Pythium were visibly more robust than infected wild type plants. However, the roots of the infected transgenic plants were much denser and healthier that roots of infected wild type plants. The transgenic plants were growing almost as well as uninfected wild type plants. These experiments indicated that overexpression of the gene was enhancing resistance to the root pathogen

Using photoaffinity labeling of AtPep1, a high affinity binding protein for the peptide was identified on the surface of Arabidopsis suspension cultured cells by photoaffinity labeling. The protein was purified to homogeneity. The binding of the photoaffinity label to the receptor is strongly competed by unlabeled AtPep1, but not by tomato systemin. The binding of AtPep1 is powerfully competed by suramin, a potent inhibitor of many ligand-receptor interactions, including the binding of systemin with its receptor. Amino acid sequence analysis has shown the protein to be a leucine-rich repeat (LRR) receptor kinase. The gene (At1G73080) contains 27 LRR motifs, a 23 amino acid membrane-spanning

region, and an intracellular protein kinase domain. The protein is glycosylated, as evidenced by the loss of about 20% of its mass by enzymatic deglycosylation. An ortholog (At1G17750) is present in Arabidopsis, and LRR receptor kinase orthologs from rice and morning glory have been reported in GenBank that that share a high percentage of amino acid identity with the AtPep1 receptor. The paralog in Arabidopsis shares over 90% amino acid identity with AtPep1, but it has only 24 LRRs and does not have either a transmembrane domain or a kinase domain. The DNA and deduced amino acid sequences for the AtPep1 receptors are provided in Fig. 2 (Ag1g73080) and Fig. 3 (At1g17750).

The AtPep1 peptide, like systemin, is apparently a cytosol-derived peptide that involves the jasmonate/ethylene signaling pathway. How the peptide is processed from a precursor and arrives at the cell surface to activate a signaling pathway is unknown, but, like systemin, transport of either the precursor or the processed peptide to the cell surface may occur in order for the peptide to react with its receptor.

AtPep1 is a component of the defense signaling of Arabidopsis plants and is therefore the first plant peptide hormone to be associated with a known pathogen defense pathway in any plant.

Table 1 shows examples of the C-terminal sequences of paralogs and orthologs of *AtproPep1* that we have identified in a wide variety of plants. In addition to the examples listed in Table 1, for example, two orthologs of the AtPep1 precursor gene (called prepro*Le*Pep1) have been identified in tomato plants. The tomato cDNA has been isolated and shown to code for an AtPep1-like defense peptide. Other paralogs and orthologs of AtprpPep1 in these and other plant species may be found by amino acid sequence homology. Additional defense signal peptides may be identified by screening plants for peptides having defense peptide activity, as described herein.

Table 1: AtPep1, Paralogs, Orthologs and Dicot and Monocot Consensus Sequences

Peptide	Source	Alignment of C-terminal Sequences
AtPep1	*Arabidopsis thaliana*	10 20 | | -ATKVK AKQRG KEKVS SGRPG QHN

Paralogs		
AtPep2	*Arabidopsis thaliana*	-SLNVM RKGIR KQPVS SGKRG GVN
AtPep3	*Arabidopsis thaliana*	-DNKAK SKKRD KEKPS SGRPG QTN
AtPep4	*Arabidopsis thaliana*	-EIKAR GKNKT KPTPS SGKGG KHN
AtPep5	*Arabidopsis thaliana*	-GLPGK KNVLK KSRES SGKPG GTN
AtPep6	*Arabidopsis thaliana*	-ITAVL RRRPR PPPYS SGRPG QNN
AtPep7	*Arabidopsis thaliana*	-VSGNV AARKG KQQTS SGKGG GTN
Orthologs		
BnPep1	Canola (*Brassica napus*)	-VARLT RRRPR PP-YS SGQPG QIN
StPep1	Potato (*Solanum tuberosum*)	-PTERR GRPPS RPKVG SGPPP QNN
PbPep1	Poplar (*Populus balsamifera*)	-DAAVS ALARR TPPVS RGGGG QTN
BePep1	Birch (*Betula* spp.)	-DLVMA VNAPP RPSLT PGSGA QIN
GmPep1	Soybean (*Glycine max*)	-ASLMA TRGSR GSKIS DGSGP QHN
MsPep1	Alfalfa (*Medicago sativa*)	-LSSMG RGGPR RTPLT QGPPP QHN
VvPep1	Grape (*Vitis vinifera*)	-EKVRE KQKKG EDGES VGRPG KKN
Dicot consensus sequence		10 20 | | -XXXXX XXXXX XXXXS SGX(P/G)(P/G) QXN Residues 2-13 are rich in K and R
OsPep1	Rice (*Oryza sativa*)	-ARLRP KPPGN PREGS GGNGG HHH
OsPep2	Rice (*Oryza sativa*)	-DDSKP TRPGA PAEGS GGNGG AIH
TaPep1	Wheat (*Triticum aestivum*)	-AVRRP RPPTT PREGR GGGGG SHN
TaPep2	Wheat (*Triticum aestivum*)	-AAPAP QRPGA PAEGA GGQGG IIH
ZmPep1	Maize (*Zea mays*)	-VRRRP TTPGR PREGS GGNGG NHH

HvPep1	Barley (*Hordeum vulgare*)	-QLARP RPPGP PRQGH GGDGG AIH
Monocot consensus sequence		10 20 | | -XXXXX XXPGX P(R/A)EGX GGXGG X(H/I)H Residues 2-7 are rich in K and R

Among the native defense signal peptides identified so far, the shortest deduced peptide sequence is 23 amino acids in length (for example, AtPep1) and the longest is 36 amino acids (AtPep3).

The dicot and monocot disease signal peptides in particular have a substantial degree of homology in the C-terminal region of the peptides, allowing us to define consensus sequences as shown in Table 1, which shows C-terminal alignments for AtPep paralogs and orthologs that correspond to the sequence of AtPep1 and dicot and monocot consensus sequences for defense signal peptides.

According to one aspect of the invention, defense signal peptides are provided that comprise a dicot or monocot consensus sequence or have a high degree of amino acid sequence homology to such a consensus sequence, e.g., 75 percent homology, or 80 percent homology, or 90 percent homology. In one embodiment, such defense signal peptides are 10 amino acid residues in length or longer and have defense signal peptide activity.

Analogs of AtPep1 were synthesized and assayed in the alkalinization assay. An analog of AtPep1 missing the carboxy-terminal amino acid was completely inactive, whereas deletions from the amino-terminus of the peptide resulted in a sequential reduction in activity, until peptides with 9 carboxy-terminal amino acids remaining (SSGRPGQHN) were inactive. A peptide consisting of only the 15 C-terminal amino acids was nearly as active as the native peptide (at approximately 2.5 nM), but analogs of 14 amino acids and smaller were progressively less active.

AtPep1 is 23 amino acid residues in length. However, truncated forms of AtPep1 of 10 amino acid residues from the C-terminus of AtPep1 retain activity, and peptides having 15 residues retain substantially full defense signal peptide activity (Example 3). As a result of the foregoing studies of shorter defense signal peptides, according to another embodiment of the invention, defense signal peptides comprising at least 10 amino acid residues, particularly those including

sequences from the C-terminus of native defense signal peptides or including consensus dicot or monocot sequences are provided. Such shorter peptides may be 11 or more, 12 or more, or 15 or more amino acid residues in length, provided that they retain substantial defense signal peptide activity.

The 15-mer was substituted with alanine in each position to assess which amino acids were necessary for the alkalinating activity. A Ser to Ala substitution at position 7, counting from the amino-terminus of the 15-mer, and a Gly to Ala substitution at position 9 exhibited little activity. Computer modeling predicted that these two amino acids would be involved in a hairpin-turn within the peptide region of -SSGR- (compare with residues 15-18 of the sequence of AtPep1 shown in Table 1). Substituting Ala for Ser (-ASGR-) abolished the predicted turn and severely reduced activity (half-maximal activity at approximately 25 nM), while substituting Ala for Gly was even less active (half-maximal activity of > 250 nM). However, neither of these analogs were able to compete with the non-substituted 15-mer for receptor binding, indicating that the structural changes in this region may have severely modified the conformation without competing for the receptor binding site. Other Ala substitutions had no effect on activity. These results will guide the skilled artisan in making desired substitutions in other defense signal peptides.

Definitions and Methods

The following definitions and methods are provided to better define the present invention and to guide those of ordinary skill in the art in the practice of the present invention. Unless otherwise noted, terms are to be understood according to conventional usage by those of ordinary skill in the relevant art. Definitions of common terms in molecular biology may also be found in Rieger et al., Glossary of Genetics: Classical and Molecular, 5th edition, Springer-Verlag: New York, 1991; and Lewin, Genes V, Oxford University Press: New York, 1994. The nomenclature for DNA bases as set forth at 37 CFR 1.822 is used. The standard one- and three-letter nomenclature for amino acid residues is used.

Polynucleotides

"Polynucleotide." The term "polynucleotide" refers to a polymer of nucleotide monomers, including but not limited to ribonucleotides or

deoxyribonucleotides or nucleotide analogues. Polynucleotides include, for example, DNA and RNA molecules, including cDNA, genomic DNA, primers, probes, vectors, and so on, and include single- and double-stranded forms thereof. Polynucleotides according to the invention may be chemically modified by well known methods by labeling, coupling to solid supports, etc.

"Defense signal peptide (or polypeptide) polynucleotide". The term "defense signal peptide polynucleotide" refers to a polynucleotide that encodes a defense signal peptide, and a "defense signal polypeptide polynucleotide" refers to a polynucleotide that encodes a defense signal polypeptide (i.e., a polypeptide that, when processed in a plant cell, produces a defense signal peptide), whether a cDNA or genomic sequence or synthetic form thereof. Such polynucleotides may comprise wild-type polynucleotides sequences encoding defense signal polypeptides, such as those listed in Table 1, operably linked to a heterologous promoter, i.e., a promoter not associated in nature with such native, or wild-type, polynucleotide sequences. Alternatively, such polynucleotides may comprise non-naturally occurring recombinant polynucleotides that comprise a sequence that encodes a defense signal peptide operably linked to a suitable promoter. For expression of defense signal peptides for exogenous application to plants, a defense signal peptide polynucleotide may be operably linked to a promoter suitable for expression in a bacterial, fungal, insect, or other suitable cell. For transformation of plants or plant cells or tissues, a defense signal peptide may be operably linked to a promoter suitable for expression in a plant cell, i.e., a plant promoter. According to another embodiment, a heterologous promoter may be introduced into a plant or a plant cell or tissue for insertion into the genome, thereby producing an insertion of the promoter upstream of a sequence that encodes a defense signal peptide, operably linking the sequence encoding the defense signal peptide to the heterologous promoter. Such an expression unit, including the heterologous promoter and the sequence encoding the defense control peptide, is another embodiment of a defense signal peptide (or polypeptide) polynucleotide.

"Native". The term "native" refers to a naturally-occurring ("wild-type") polynucleotide, polypeptide or peptide.

"Isolated". An "isolated" polynucleotide is one that has been substantially separated or purified away from other polynucleotide sequences in the cell of the

organism in which the polynucleotide naturally occurs, i.e., other chromosomal and extrachromosomal DNA and RNA, by conventional purification methods. The term also embraces recombinant polynucleotides (including promoter insertions operably linked to a defense signal peptide gene) and chemically synthesized polynucleotides.

Fragments, Probes, and Primers. A fragment of a polynucleotide is a portion of a polynucleotide that is less than full-length and comprises at least a minimum length capable of hybridizing specifically with a native polynucleotide sequence under stringent hybridization conditions. The length of such a fragment is preferably at least 15 nucleotides, more preferably at least 20 nucleotides, and most preferably at least 30 nucleotides of a native polynucleotide sequence.

A "probe" is an isolated polynucleotide to which is attached a conventional detectable label or reporter molecule, e.g., a radioactive isotope, ligand, chemiluminescent agent, or enzyme. "Primers" are isolated polynucleotides that are annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, then extended along the target DNA strand by a polymerase, e.g., a DNA polymerase. Primer pairs can be used for amplification of a polynucleotide sequence, e.g., by the polymerase chain reaction (PCR) or other conventional nucleic-acid amplification methods.

Probes and primers are generally 15 nucleotides or more in length, preferably 20 nucleotides or more, more preferably 25 nucleotides, and most preferably 30 nucleotides or more. Such probes and primers hybridize specifically to the target polynucleotide sequence under high stringency hybridization conditions under at least moderately stringent conditions.

Methods for preparing and using probes and primers are described, for example, in Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, ed. Sambrook et al., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989 (hereinafter, "Sambrook et al., 1989"); Current Protocols in Molecular Biology, ed. Ausubel et al., Greene Publishing and Wiley-Interscience, New York, 1992 (with periodic updates) (hereinafter, "Ausubel et al., 1992); and Innis et al., PCR Protocols: A Guide to Methods and Applications, Academic Press: San Diego, 1990. PCR-primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such-as Primer

(Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge, Mass.).

Primers and probes based on the native defense signal polypeptide polynucleotide sequences that are disclosed herein can be used to confirm (and, if necessary, to correct) the disclosed polynucleotide sequences by conventional methods, e.g., by re-cloning and re-sequencing.

"Substantial similarity". A first polynucleotide is "substantially similar" to a second polynucleotide if, when optimally aligned (with appropriate nucleotide insertions or deletions) with the other polynucleotide (or its complementary strand), there is at least about 75% nucleotide sequence identity, preferably at least about 80% identity, more preferably at least about 85% identity, and most preferably at least about 90% identity. Sequence similarity can be determined by comparing the nucleotide sequences of two polynucleotides using sequence analysis software such as the Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, Madison, Wis.

Alternatively, two polynucleotides are substantially similar if they hybridize under stringent conditions.

"Operably Linked". A first nucleic-acid sequence is "operably linked" with a second nucleic-acid sequence when the first nucleic-acid sequence is placed in a functional relationship with the second nucleic-acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein coding regions, in reading frame.

"Recombinant". A "recombinant" polynucleotide is made by an artificial combination of two otherwise separated segments of sequence, e.g., by chemical synthesis or by the manipulation of isolated segments of polynucleotides by genetic engineering techniques. Techniques for nucleic-acid manipulation are well-known (see, e.g., Sambrook et al., 1989, and Ausubel et al., 1992). Methods for chemical synthesis of polynucleotides are discussed, for example, in Beaucage and Carruthers, Tetra. Letts. 22:1859-1862, 1981, and Matteucci et al., J. Am. Chem. Soc. 103:3185, 1981. Chemical synthesis of polynucleotides can be performed, for example, on commercial automated oligonucleotide synthesizers.

Preparation of Recombinant or Chemically Synthesized Polynucleotides; Vectors, Transformation, Host cells. Natural or synthetic polynucleotides according to the present invention can be incorporated into recombinant nucleic-acid constructs, typically DNA constructs, capable of introduction into and replication in a host cell. Such a construct preferably is a vector that includes a replication system and sequences that are capable of transcription and translation of a polypeptide-Oncoding sequence in a given host cell.

For the practice of the present invention, conventional compositions and methods for preparing and using vectors and host cells are employed, as discussed, inter alia, in Sambrook et al., 1989, or Ausubel et al., 1992.

A cell, tissue, organ, or organism into which has been introduced a foreign polynucleotide, such as a recombinant vector, is considered "transformed", "transfected", or "transgenic." A "transgenic" or "transformed" cell or organism also includes progeny of the cell or organism and progeny produced from a breeding program employing such a "transgenic" plant as a parent in a cross and exhibiting an altered phenotype resulting from the presence of a recombinant polynucleotide construct.

A number of vectors suitable for stable transfection of plant cells or for the establishment of transgenic plants have been described in, e.g., Pouwels et al., Cloning Vectors: A Laboratory Manual, 1985, supp. 1987); Weissbach and Weissbach, Methods for Plant Molecular Biology, Academic Press, 1989; and Gelvin et al., Plant Molecular Biology Manual, Kluwer Academic Publishers, 1990. Typically, plant expression vectors include, for example, one or more cloned plant genes under the transcriptional control of 5' and 3' regulatory sequences and a dominant selectable marker. Such plant expression vectors also can contain a promoter regulatory region (e.g., a regulatory region controlling inducible or constitutive, environmentally- or developmentally-regulated, or cell- or tissue-specific expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal.

Examples of constitutive plant promoters include, but are not limited to, the cauliflower mosaic virus (CaMV) 35S promoter, which confers constitutive, high-level expression in most plant tissues (see, e.g., Odel et al., Nature 313:810, 1985), including monocots (see, e.g., Dekeyser et al., Plant Cell 2:591, 1990; Terada and

Shimamoto, Mol. Gen. Genet. 220:389, 1990); the nopaline synthase promoter (An et al., Plant Physiol. 88:547, 1988) and the octopine synthase promoter (Fromm et al., Plant Cell 1:977, 1989).

A variety of plant gene promoters that are regulated in response to environmental, hormonal, chemical, and/or developmental signals, also can be used for expression of defense signal peptides in plant cells, including promoters regulated by (1) heat (Callis et al., Plant Physiol. 88:965, 1988), (2) light (e.g., pea rbcS-3A promoter, Kuhlemeier et al., Plant Cell 1:471, 1989; maize rbcS promoter, Schaffner and Sheen, Plant Cell 3:997, 1991; or chlorophyll a/b-binding protein promoter, Simpson et al., EMBO J. 4:2723, 1985), (3) hormones, such as abscisic acid (Marcotte et al., Plant Cell 1:969, 1989), (4) wounding (e.g., wunl, Siebertz et al., Plant Cell 1:961, 1989); or (5) chemicals such as methyl jasminate, salicylic acid, or Safener. It may also be advantageous to employ (6) organ-specific promoters (e.g., Roshal et al., EMBO J. 6:1155, 1987; Schernthaner et al., EMBO J. 7:1249, 1988; Bustos et al., Plant Cell 1:839, 1989), including promoters that express specifically in the root, leaf, seed, etc.

Plant expression vectors optionally include RNA processing signals, e.g., introns, which may be positioned upstream or downstream of a polypeptide-encoding sequence in the transgene. In addition, the expression vectors may also include additional regulatory sequences from the 3'-untranslated region of plant genes (Thornburg et al., Proc. Natl. Acad. Sci. USA 84:744 (1987); An et al., Plant Cell 1:115 (1989), e.g., a 3' terminator region to increase mRNA stability of the mRNA, such as the PI-II terminator region of potato or the octopine or nopaline synthase 3' terminator regions.

Useful dominant selectable marker genes include genes encoding antibiotic resistance genes (e.g., resistance to hygromycin, kanamycin, bleomycin, G418, streptomycin or spectinomycin); and herbicide resistance genes (e.g., phosphinothricin acetyltransferase). A useful strategy for selection of transformants for herbicide resistance is described, e.g., in Vasil, Cell Culture and Somatic Cell Genetics of Plants, Vols. I-III, Laboratory Procedures and Their Applications Academic Press, New York, 1984.

An expression vector for expression of a defense signal peptide or polypeptide in a plant may also comprise a gene encoding another polypeptide, including a herbicide-tolerance gene (e.g., tolerance to glyphosate, glufosinate,

etc.); a polypeptide conferring insect resistance (e.g., a *Bacillus thuringensis* insecticidal protein or a Xenorhabdus insecticidal protein); a pathogen protein (e.g., virus coat protein); a trait for improving yield, drought resistance, cold tolerance, etc.; a trait for modifying the oil, protein or starch composition of seeds; or another gene that has a desirable activity when expressed in a plant. For example, U.S. Pat. No. 5,571,706 describes the introduction of the N gene into tobacco to confer resistance to tobacco mosaic virus; WO 95/28423 describes the expression of the Rps2 gene from *Arabidopsis thaliana* in plants as a means of creating resistance to bacterial pathogens including *Pseudomonas syringae*; WO 98/02545 describes the introduction of the Prf gene into plants to obtain broad-spectrum pathogen resistance; and U.S. Patent No. 6,762,285 describes the expression of the Bs2 resistance proteins in plants to confer resistance to *Xanthomonas campestris*. Such plant defense genes may also be co-expressed on the same or a different expression vector with a defense signal polypeptide or peptide.

Nucleic-Acid Hybridization; "Stringent Conditions"; "Specific". The term "stringent conditions" is functionally defined with regard to the hybridization of a nucleic-acid probe to a target polynucleotide (i.e., to a particular nucleic-acid sequence of interest) by the specific hybridization procedure discussed in Sambrook et al., 1989, at 9.52-9.55. See also, Sambrook et al., 1989 at 9.47-9.52, 9.56-9.58; Kanehisa, Nucl. Acids Res. 12:203-213, 1984; and Wetmur and Davidson, J. Mol. Biol. 31:349-370, 1968.

Regarding the amplification of a target nucleic-acid sequence (e.g., by PCR) using a particular amplification primer pair, "stringent conditions" are conditions that permit the primer pair to hybridize only to the target nucleic-acid sequence to which a primer having the corresponding wild-type sequence (or its complement) would bind and preferably to produce a unique amplification product.

The term "specific for (a target sequence)" indicates that a probe or primer hybridizes under given hybridization conditions only to the target sequence in a sample comprising the target sequence.

Nucleic-Acid Amplification. As used herein, "amplified DNA" refers to the product of nucleic-acid amplification of a target nucleic-acid sequence. Nucleic-acid amplification can be accomplished by any of the various nucleic-acid amplification methods known in the art, including the polymerase chain reaction

(PCR). A variety of amplification methods are known in the art and are described, inter alia, in U.S. Pat. Nos. 4,683,195 and 4,683,202 and in PCR Protocols: A Guide to Methods and Applications, ed. Innis et al., Academic Press, San Diego, 1990.

See also the Examples below regarding RT-PCR, for example.

Nucleotide-Sequence Variants of Native Defense Signal Polypeptide Polynucleotides and Amino Acid Sequence Variants of Native Defense Signal Proteins and Peptides. Using the nucleotide and the amino-acid sequences disclosed herein, those skilled in the art can create DNA molecules, polypeptides and peptides that have minor variations in their nucleotide or amino acid sequence, respectively.

"Variant" DNA molecules are DNA molecules containing minor changes in a native sequence, i.e., changes in which one or more nucleotides of a native sequence is deleted, added, and/or substituted, preferably while substantially maintaining a desired biological activity. Variant DNA molecules can be produced, for example, by standard DNA mutagenesis techniques or by chemically synthesizing the variant DNA molecule or a portion thereof. Such variants preferably do not change the reading frame of the protein-coding region of the polynucleotide and preferably encode a protein having no change, only a minor reduction, or an increase in a desired biological activity.

Amino-acid substitutions are preferably substitutions of single amino-acid residues. DNA insertions are preferably of about 1 to 10 contiguous nucleotides and deletions are preferably of about 1 to 30 contiguous nucleotides. Insertions and deletions are preferably insertions or deletions from an end of the protein-coding or non-coding sequence and are preferably made in adjacent base pairs. Substitutions, deletions, insertions or any combination thereof can be combined to arrive at a final construct.

Preferably, variant polynucleotides according to the present invention are "silent" or "conservative" variants. "Silent" variants are variants of a native sequence or a homolog thereof in which there has been a substitution of one or more base pairs but no change in the amino-acid sequence of the polypeptide or peptide encoded by the sequence. "Conservative" variants are variants of a native (or consensus) sequence in which at least one codon in the protein-coding region of the gene has been changed, resulting in a conservative change in one or more

amino acid residues of the encoded polypeptide encoded, i.e., an amino acid substitution. A number of conservative amino acid substitutions are listed below. In addition, one or more codons encoding cysteine residues can be substituted for, resulting in a loss of a cysteine residue and affecting disulfide linkages in the polypeptide.

Table 2: Conservative Amino-Acid Substitutions

Original Residue	Conservative Substitutions
Ala	Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn; Gln
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg; Gln; Glu
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

Substantial changes in function are made by selecting substitutions that are less conservative than those listed above, e.g., causing changes in: (a) the structure of the polypeptide backbone in the area of the substitution; (b) the charge or hydrophobicity of the polypeptide at the target site; or (c) the bulk of an amino acid side chain. Substitutions generally expected to produce the greatest changes in protein properties are those in which: (a) a hydrophilic residue, e.g., seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g., leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histadyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is

substituted for (or by) one not having a side chain, e.g., glycine.

Polypeptides and Peptides

Three and one-letter code for amino acids. For the polypeptide and peptide sequences presented herein, either the three-letter code or the one-letter code may be used for representing amino acid residues, as provided in Table 3 below.

Table 3: Three-letter Code and One-letter Code for Amino Acids

Amino Acid	Three-Letter Code	One-Letter Code
Alanine	Ala	A
Cysteine	Cys	C
Aspartic acid	Asp	D
Glutamic acid	Glu	E
Phenylalanine	Phe	F
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Lysine	Lys	K
Leucine	Leu	L
Methionine	Met	M
Asparagine	Asn	N
Proline	Pro	P
Glutamine	Gln	Q
Arginine	Arg	R
Serine	Ser	S
Threonine	Thr	T
Valine	Val	V
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Unknown or Unspecified	Xaa	X

"Defense Signal Polypeptide"; "Defense Signal Peptide". In general, a peptide is considered a short polypeptide. The term "defense signal polypeptide" (or protein) refers to a polypeptide encoded by a defense signal protein polynucleotide, including, but not limited to, the polynucleotides listed in Table 1, and other polynucleotides that encode orthologs, paralogs, homologs, and variants of a native defense signal polypeptide. Defense signal peptides result from the processing of a native defense signal polypeptide, such as AtproPep1, in a plant cell. As a result, a native defense signal polypeptide includes sequences in addition to defense signal peptide sequences. Recombinant polypeptides that are not processed intracellularly, but that have defense signal peptide activity, are also considered defense signal polypeptides or peptides.

Recombinant fusion polypeptides may be made that, when processed in a plant cell, result in the production of more than one defense signal peptide, or in the production of a defense signal peptide and another biologically active polypeptide or peptide.

The term "defense signal peptide" refers to a peptide about ten or more amino acids in length that has substantial defense signal peptide activity. Such defense signal peptides may have a length be 11, 12, 13, 14, 15, or more amino acids. AtPep1 is a native defense signal peptide from Arabidopsis that is 23 amino acids in length, although sequences from the C-terminal end of AtPep1 as short as 10 amino acids retain substantial defense signal peptide activity, and such truncated peptides increase in activity with increasing length. Defense signal peptides longer than 23 amino acids retain defense signal peptide activity. The native defense signal polypeptides that we have identified range encodes propeptides of 75 to 154 amino acids that are processed intracellularly to produce the shorter defense signal peptides. Defense signal peptides up to about 160 amino acid residues, or 100, or 90, or 80, or 70, or 60, or 50, or 40, or 30, or 23, or 20 amino acid residues are included among the defense signal peptides of the invention.

Defense signal peptides may be produced by expression of a polynucleotide that encodes such a peptide intracellularly, e.g., in a plant cell, or in a non-plant cell, e.g., a bacterial, fungal, insect, or other cell used in recombinant production of polypeptides. Alternatively, defense signal peptides may be produced by chemical synthesis. Techniques for chemical synthesis of polypeptides are described, for example, in Merrifield, J. Amer. Chem. Soc. 85:2149-2156, 1963, and peptide synthesizers are commercially available. For chemical synthesis, shorter forms of the defense signal peptides are preferable to longer forms, including but not limited to, defense signal peptides between about 10 and about 30 amino acids in length.

Polypeptide Sequence Homology. Ordinarily, defense signal peptides encompassed by the present invention are at least about 75 percent homologous to a native defense signal peptide, including but not limited to any of the defense signal peptides listed in Table 1 or a dicot or monocot consensus defense signal peptide sequence, or at least about 80 percent, 85 percent, 90 percent, or 100 percent (complete) homology, and that has substantial defense signal peptide activity. Such homology is considered to be "substantial homology," although

more important than shared amino-acid sequence homology is the possession of characteristic structural features and highly conserved amino acid residues from the C-terminal region of native defense signal peptides or consensus sequences.

Polypeptide homology is typically analyzed using sequence analysis software such as the Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, Madison, Wis.). Polypeptide sequence analysis software matches homologous sequences using measures of homology assigned to various substitutions, deletions, substitutions, and other modifications.

"Isolated," "Purified," "Homogeneous" Polypeptides and Peptides. An "isolated" polypeptide or peptide has been separated from the cellular components (polynucleotides, lipids, carbohydrates, and other polypeptides) that naturally accompany it. Such a polypeptide or peptide can also be referred to as "pure" or "homogeneous" or "substantially" pure or homogeneous. Thus, a polypeptide which is chemically synthesized is isolated. A defense signal peptide or polypeptide is also considered "isolated" if it is the product of the expression of a recombinant polynucleotide (even if expressed in a homologous cell type). Thus, if AtPep1, for example, is recombinantly expressed in an Arabidopsis plant, it is considered "isolated" if the polynucleotide that encodes it is under the control of a promoter that is different from the native AtproPep1 promoter, or if the polynucleotide encodes a polypeptide other than the wild-type, or native, AtproPep1 polypeptide but, when processed in a plant cell produces a native AtPep1 peptide, or the AtPep1 peptide produced by expression of the polynucleotide and processing of the encoded polypeptide differs from that of the native AtPep1 peptide in any way, for example in length or sequence.

A monomeric polypeptide or peptide is isolated when at least 60% by weight of a sample is composed of the polypeptide or peptide, or 90% or more, or 95% or more, or more than 99%. Protein purity or homogeneity is indicated, for example, by polyacrylamide gel electrophoresis of a protein sample, followed by visualization of a single polypeptide band upon staining the polyacrylamide gel; high pressure liquid chromatography; or other conventional methods.

Protein Purification. The polypeptides and peptides of the present invention can be purified by any of the means known in the art. Various methods of protein purification are described, e.g., in Guide to Protein Purification, ed. Deutscher,

Meth. Enzymol. 185, Academic Press, San Diego, 1990; and Scopes, Protein Purification: Principles and Practice, Springer Verlag, New York, 1982.

Variant and Modified Forms of Defense Signal Peptides and Polypeptides. Encompassed by the defense signal peptides and polypeptides of the present invention are variant peptides and polypeptides in which there have been substitutions, deletions, insertions or other modifications of a native (i.e., wild-type) peptide or polypeptide. The variants substantially retain structural characteristics and biological activities of a corresponding native peptide or polypeptide and are preferably silent or conservative substitutions of one or a small number of contiguous amino acid residues.

Regarding the terms "paralog" and "ortholog", homologous polynucleotide sequences and homologous polypeptide sequences may be paralogs or orthologs of the claimed polynucleotide or polypeptide sequence. Orthologs and paralogs are evolutionarily related genes that have similar sequence and similar functions. Orthologs are structurally related genes in different species that are derived by a speciation event. Paralogs are structurally related genes within a single species that are derived by a duplication event. Sequences that are sufficiently similar to one another will be appreciated by those of skill in the art and may be based upon percentage identity of the complete sequences, percentage identity of a conserved domain or sequence within the complete sequence, percentage similarity to the complete sequence, percentage similarity to a conserved domain or sequence within the complete sequence, and/or an arrangement of contiguous nucleotides or peptides particular to a conserved domain or complete sequence. Sequences that are sufficiently similar to one another will also bind in a similar manner to the same DNA binding sites of transcriptional regulatory elements using methods well known to those of skill in the art.

The term "equivalog" describes members of a set of homologous proteins that are conserved with respect to function since their last common ancestor. Related proteins are grouped into equivalog families, and otherwise into protein families with other hierarchically defined homology types. This definition is provided at the Institute for Genomic Research (TIGR) world wide web (www) website, "tigr.org" under the heading "Terms associated with TIGRFAMs".

"Allelic variant" or "polynucleotide allelic variant" refers to any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic

variation arises naturally through mutation, and may result in phenotypic polymorphism within populations. Gene mutations may be "silent" or may encode polypeptides having altered amino acid sequence. "Allelic variant" and "polypeptide allelic variant" may also be used with respect to polypeptides, and in this case the terms refer to a polypeptide encoded by an allelic variant of a gene.

"Splice variant" or "polynucleotide splice variant" as used herein refers to alternative forms of RNA transcribed from a gene. Splice variation naturally occurs as a result of alternative sites being spliced within a single transcribed RNA molecule or between separately transcribed RNA molecules, and may result in several different forms of mRNA transcribed from the same gene. This, splice variants may encode polypeptides having different amino acid sequences, which may or may not have similar functions in the organism. "Splice variant" or "polypeptide splice variant" may also refer to a polypeptide encoded by a splice variant of a transcribed mRNA.

As used herein, "polynucleotide variants" may also refer to polynucleotide sequences that encode paralogs and orthologs of the presently disclosed polypeptide sequences. "Polypeptide variants" may refer to polypeptide sequences that are paralogs and orthologs of the presently disclosed polypeptide sequences.

A native defense signal peptide or polypeptide sequence can be modified by conventional methods, e.g., by acetylation, carboxylation, phosphorylation, glycosylation, ubiquitination, and labeling, whether accomplished by in vivo or in vitro enzymatic treatment or by the synthesis of a defense signal peptide or polypeptide using modified amino acids.

Labeling. There are a variety of conventional methods and reagents for labeling polypeptides and fragments thereof. Typical labels include radioactive isotopes, ligands or ligand receptors, fluorophores, chemiluminescent agents, and enzymes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed, e.g., in Sambrook et al., 1989 and Ausubel et al., 1992.

Peptide Fragments. The present invention also encompasses fragments of a defense signal peptide that lacks at least one residue of a native full-length defense signal peptide. Preferably, such a fragment retains substantial defense signal peptide activity, including but not limited to substantial activity in an alkalinization assay and/or the ability to enhance disease resistance in a plant.

"Defense Signal Peptide Activity"; Biological activity of Polypeptides or Peptides. The terms "biological activity", "biologically active", "activity" and "active" refer primarily to the characteristic biological activity or activities of a native defense signal peptide or polypeptide. Defense signal peptide activity includes activity in an alkalinization assay, as described herein. Substantial defense signal peptide activity in an alkalinization assay includes a change of at least 0.2 pH units when a ten microliter aliquot of a solution having a concentration of 25 nM of the peptide is added to one mL of plant cells in the assay. More substantial defense signal peptide activity in the assay is the observation of a change in pH of at least 0.2 pH units using a solution having concentrations of 2.5 nM or 0.25 nM of the peptide, or when a change of at least 0.5 pH units are observed at a given peptide concentration, or when the activity is at least 25 percent, or 50 percent, or 75 percent that of a native defense signal polypeptide.

Alternatively, a substantial defense signal peptide activity is the ability to enhance plant disease resistance and substantially improve yield of plant product, with enhancement of plant disease resistance evidenced by reduced disease symptoms, enhanced resistance to disease infestation, etc., when a defense signal peptide is applied to a plant exogenously or recombinantly expressed within a plant. A defense signal peptide substantially enhances disease resistance of a plant if it increases the resistance of a plant to a pathogen of at least 10 percent as compared to a control plant under similar conditions, or more substantially, of at least 25, or 50, or 75, or 100 percent, as measured by standard quantitative measures of plant disease resistance to a given pathogen, e.g., the number or size of lesions, increased growth, survival rate, rate of disease progression, higher root mass, better seed viability, seed quantity and quality, etc. Alternatively, a substantial defense signal peptide activity is present where the peptide, when applied to a plant exogenously or recombinantly expressed within a plant, confers a substantial change in any resistance to a biotic or abiotic stress that involves the jasmonate/ethylene or salicylic acid pathways, as measured by standard methods.

Fusion Polypeptides. The present invention also provides fusion polypeptides including, for example, heterologous fusion polypeptides in which a defense signal polypeptide coding sequence is joined to a heterologous promoter (i.e., a promoter from gene other than the promoter that is operably linked to that coding sequence in nature), or in which the coding sequence for the defense signal

peptide is joined to a fusion partner, i.e., a protein-coding sequence other than sequences with which the coding sequence for the defense signal peptide is joined in nature). Such fusion polypeptides can exhibit biological properties (such as substrate or ligand binding, enzymatic activity, antigenic determinants, etc.) derived from each of the fused sequences.

Polypeptide Sequence Determination. The sequence of a polypeptide of the present invention can be determined by any of the various methods known in the art.

Polypeptide Coupling to a Solid-Phase Support. The polypeptides of the present invention can be free in solution or coupled to a solid-phase support, e.g., nitrocellulose, nylon, column packing materials (e.g., Sepharose beads), magnetic beads, or glass wool, by conventional methods.

Antibodies

The present invention also encompasses polyclonal and/or monoclonal antibodies capable of specifically binding to a particular defense signal peptide and/or fragments thereof. Such antibodies are raised against a defense signal peptide or fragment thereof and are capable of distinguishing a defense signal peptide from other polypeptides, i.e., are specific for the particular defense signal peptide.

For the preparation and use of antibodies according to the present invention, including various immunoassay techniques and applications, see, e.g., Goding, Monoclonal Antibodies: Principles and Practice, 2d ed, Academic Press, New York, 1986; and Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1988. Defense signal peptide-specific antibodies are useful, for example in: purifying a defense signal peptide polypeptide from a biological sample, such as a host cell expressing a recombinant defense signal peptide; in cloning a paralog, ortholog, or homolog from an expression library; as antibody probes for protein blots and immunoassays; etc.

Antibodies can be labeled by any of a variety of conventional methods. Suitable labels include, but are not limited to, radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent agents, chemiluminescent agents, magnetic particles, etc.

Obtaining Paralogs, Orthologs, and Homologs of Defense Signal Peptides

As discussed in the Examples below, defense signal peptides homologous to AtPep1 and other defense signal peptides exist in many plant species. Based upon the availability of the defense signal peptide and polypeptide sequences and their corresponding gene sequences disclosed herein, paralogs and orthologs can be obtained by conventional methods, e.g., by screening a cDNA or genomic library with a probe that specifically hybridizes to a native defense signal peptide sequence under at least moderately stringent conditions, by PCR or another amplification method using a primer or primers that specifically hybridize to a native defense signal peptide or polypeptide sequence under at least moderately stringent conditions, or by screening an expression library using defense signal peptide-specific antibodies.

Plant Transformation and Regeneration; Transformed Plant Cells, Plants, and Parts and Products of Transformed Plants

Various polynucleotide constructs that include a sequence that encodes a defense signal polypeptide or a defense signal peptide are useful for producing plants having having enhanced disease resistance or enhanced resistance to another biotic and abiotic stress that involves the jasmonate/ethylene or salicylic acid pathways.

Polynucleotides that comprise a sequence that encodes a defense related polypeptide or a defense related peptide can be expressed in plants or plant cells under the control of an operably linked promoter that is capable of expression in the plant or plant cell. Any well-known method can be employed for plant cell transformation, culture, and regeneration in the practice of the present invention with regard to a particular plant species. Conventional methods for introduction of foreign DNA into plant cells include, but are not limited to: (1) Agrobacterium-mediated transformation (Lichtenstein and Fuller In: Genetic Engineering, Vol 6, Rigby, ed., London, Academic Press, 1987; and Lichtenstein and Draper, in: DNA Cloning, Vol II, Glover, ed., Oxford, IRI Press, 1985); (2) particle delivery (see, e.g., Gordon-Kamm et al., Plant Cell 2:603, 1990; or BioRad Technical Bulletin 1687), (3) microinjection (see, e.g., Green et al., Plant Tissue and Cell Culture, Academic Press, New York, 1987), (4) polyethylene glycol (PEG) procedures (see, e.g., Draper et al., Plant Cell Physiol. 23:451, 1982); Zhang and Wu, Theor. Appl.

Genet. 76:835, 1988), (5) liposome-mediated DNA uptake (see, e.g., Freeman et al., Plant Cell Physiol. 25:1353, 1984), (6) electroporation (see, e.g., Fromm et al., Nature 319:791, 1986); and (7) vortexing method (see, e.g., Kindle, Proc. Natl. Acad. Sci. USA 87:1228, 1990).

Once a transformed plant cell or tissue has been obtained, it is possible to regenerate a full-grown plant from it. Means for regeneration vary from species to species. In one approach a suspension of transformed protoplasts or a petri plate containing transformed explants is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced in the callus tissue. These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is usually reproducible and repeatable. Plant regeneration is described, for example, in Evans, et al., *Handbook of Plant Cell Cultures, Vol.* 1: (MacMillan Publishing Co., New York, 1983); and Vasil I. R. (ed.), *Cell Culture and Somatic Cell Genetics of Plants,* Acad. Press, Orlando, Vol. I, 1984, and Vol. III, 1986). Practically all plants can be regenerated from cultured cells or tissues, including monocots, dicots, gymnosperms, etc.

After the DNA construct is stably incorporated in transgenic plants, it can be transferred to other plants by sexual crosses or by asexual propagation. With respect to sexual crossing, any of a number of standard breeding techniques can be used depending upon the species to be crossed. Cultivars can be propagated in accord with common agricultural procedures known to those in the field.

The term "plant" encompasses any higher plant and progeny thereof, including monocots, dicots, gymnosperms, and other plants and includes parts of plants, including reproductive units of a plant (e.g., seeds), fruit, flowers, etc.

A "reproductive unit" of a plant is any totipotent part or tissue of the plant from which one can obtain a progeny of the plant, including, for example, seeds, cuttings, tubers, buds, bulbs, somatic embryos, cultured cells (e.g., callus or suspension cultures), etc.

According to one aspect of the invention, plant cells are provided that comprise a polynucleotide sequence that comprises a sequence that encodes a

defense signal peptide or polypeptide operably linked to a plant promoter. Another aspect of the invention is directed to plants comprising such cells, i.e., transformed or transgenic plants. Another aspect of the invention is a part or product of such plants.

Agronomically and commercially important products and/or compositions of matter derived from transgenic plants according to the invention include, but are not limited to, animal feed, commodities, products and by-products that are intended for use as food for human consumption or for use in compositions and commodities that are intended for human consumption, including but not limited to plant parts, including but not limited to seeds, seed pods, flowers (including flower buds), fruit, tubers, stems, cuttings, pollen, and products derived from processing such plant parts, including but not limited to flour, meal, syrup, oil, starch, cakes, cereals, and the like. Such compositions may be defined as containing detectable amounts of a polynucleotide sequence of the invention as set forth herein, and thus are also diagnostic for any transgenic event containing such nucleotide sequences. These products are more likely to be derived from crops propagated with fewer pesticides and organophosphates as a result of their incorporation of the nucleotides of the present invention for controlling plant disease. For example, such commodities and commodity products can be produced from seed produced from a transgenic plant, wherein the transgenic plant comprises cells that express a defense signal protein of the present invention.

Identifying Transgenic Plants According to the Invention and Parts and Products Thereof

Transgenic plants according to the present invention, parts of such plants, and products derived from the processing of such plants, can be readily identified by using probes and primers to specifically identify the presence of a transgene that encodes a defense signal peptide or the presence of a specific defense signal peptide. In order to perform such an identification, a biological sample thought to contain such a plant, part or product is contacted with a probe that binds specifically to the transgene containing a defense signal peptide- or polypeptide-encoding polynucleotide (such as one or more PCR primers, cDNA probe, etc.), and detecting such binding (e.g., by identifying the production of an amplification product of a diagnostic size after gel electrophoresis, or by autoradiography). Alternatively, one may use a proble that binds specifically to the defense signal

peptide or polypeptide itself, such as an antibody probe, wherein binding can be detected by an enzyme-linked immunosorbent assay (ELISA), etc.).

Conferring Resistance to Biotic and Abiotic Stresses to Plants and Enhancing Plant Growth

As one aspect of the invention, resistance to disease resistance, or to another biotic and abiotic stress that involves the jasmonate/ethylene pathway or the salicylic acid pathway, is conferred on a plant, or resistance may be enhanced in the plant, or growth of the plant is enhanced, by expression of polynucleotides that encode one or more defense peptides in cells of the plant.

As another aspect of the invention, methods are provided that comprise growing a seed into a plant, wherein the plant comprises cells comprising a polynucleotide sequence comprising a sequence that encodes a defense signal protein or polypeptide, wherein the plant exhibits one or more of the following: improved yield of plant product, reduced disease symptoms, or enhanced resistance to disease infestation; compared to a control plant lacking the recombinant polynucleotide.

According to another aspect of the invention, improved yield, reduced disease symptoms, or enhanced resistance to disease infestation is conferred or enhanced, by application of compositions comprising one or more defense signal peptides to a plant.

Where absolute immunity against infection by a pathogen or detrimental affects or other stresses is not be conferred, the severity of the disease is reduced and symptom development is delayed. This method of imparting resistance has the potential for enhancing plant resistance to a variety of diseases for which other approaches were ineffective in providing effective control.

The methods of the present invention are useful in imparting resistance to a wide variety of pathogens including viruses, bacteria, and fungi.

With regard to the use of the compositions and methods of the present invention to enhance plant growth, various forms of plant growth enhancement or promotion can be achieved. This can occur as early as when plant growth begins from seeds or later in the life of a plant. For example, plant growth according to the present invention encompasses greater yield, increased percentage of seeds germinated, increased plant size, greater biomass, more and bigger fruit, earlier fruit coloration, earlier flower opening, improved flower longevity (i.e., shelf-life), and earlier fruit

and plant maturation. As a result, the present invention provides significant economic benefit to growers. For example, early germination and early maturation permit crops to be grown in areas where short growing seasons would otherwise preclude their growth in that locale. Increased percentage of seed germination results in improved crop stands and more efficient seed use. Greater yield, increased size, and enhanced biomass production allow greater revenue generation from a given plot of land.

To confer such enhanced resistance, one may express a single gene copy, or in order to express a defense signal peptide at high levels, e.g., expression of multiple copies of a transgene encoding such a defense signal peptide and/or the use of strong promoters to drive expression may be employed. Expression of a transgene encoding a defense signal peptide in plant cells at a sufficiently high level may initiate the plant defense response constitutively in the absence of signals from the pathogen. A constitutive plant promoter can be used. Alternatively, an inducible promoter, or an organ- or tissue-specific promoter, for example, can be used.

If a plant cell is selected to be transformed, it may be of any type capable of being transformed, preferably one with an agronomic, horticultural, ornamental, economic, or commercial value. Examples of such plant cells include, but are not limited to: Acacia, alfalfa, aneth, apple, apricot, artichoke, arugula, asparagus, avocado, banana, barley, beans, beet, blackberry, blueberry, broccoli, brussels sprouts, cabbage, canola, cantaloupe, carrot, cassava, castorbean, cauliflower, celery, cherry, chicory, cilantro, citrus, clementines, clover, coconut, coffee, corn, cotton, cucumber, Douglas fir, eggplant, endive, escarole, eucalyptus, fennel, figs, garlic, gourd, grape, grapefruit, honey dew, jicama, kiwifruit, lettuce, leeks, lemon, lime, Loblolly pine, linseed, mango, melon, mushroom, nectarine, nut, oat, oil palm, oil seed rape, okra, olive, onion, orange, an ornamental plant, palm, papaya, parsley, parsnip, pea, peach, peanut, pear, pepper, persimmon, pine, pineapple, plantain, plum, pomegranate, poplar, potato, pumpkin, quince, radiata pine, radicchio, radish, rapeseed, raspberry, rice, rye, sorghum, Southern pine, soybean, spinach, squash, strawberry, sugarbeet, sugarcane, sunflower, sweet potato, sweetgum, tangerine, tea, tobacco, tomato, triticale, turf grass, turnip, a vine, watermelon, wheat, yams, and zucchini.

Compositions Comprising Defense Signal Peptides for Application to Plants

According to one embodiment of the invention, compositions for application to plants comprise an oil flowable suspension, comprising purified defense signal peptides or unpurified forms of the peptides, including lysed or unlysed bacterial cells or fractions thereof that contain one or more of the defense signal peptides disclosed herein. Any such bacterial host cell expressing the novel polynucleotides disclosed herein and producing a defense signal peptide is contemplated to be useful, such as Bacillus spp., including B. thuringiensis, B. megaterium, B. subtilis; B. cereus, Escherichia spp., including E. coli, and/or Pseudomonas spp., including P. cepacia, P. aeruginosa, and P. fluorescens.

In another embodiment, compositions for application to plants comprise a water dispersible granule or powder comprising purified or unpurified defense signal peptides.

In another embodiment, compositions for application to plants comprise a wettable powder, spray, emulsion, colloid, aqueous or organic solution, dust, pellet, or collodial concentrate comprising purified or unpurified defense signal peptides. Such dry forms of the insecticidal compositions may be formulated to dissolve immediately upon wetting, or alternatively, dissolve in a controlled-release, sustained-release, or other time-dependent manner. Alternatively, such a composition may consist of a combination of one or more of the following compositions: lysed or unlysed bacterial cells, spores, crystals, and/or purified crystal proteins.

In another embodiment, compositions for application to plants comprise an aqueous solution or suspension comprising purified or unpurified defense signal peptides. Such aqueous solutions or suspensions may be provided as a concentrated stock solution which is diluted prior to application, or alternatively, as a diluted solution ready-to-apply.

Such compositions may be formulated in a variety of ways. They may be employed as wettable powders, granules or dusts, by mixing with various inert materials, such as inorganic minerals (phyllosilicates, carbonates, sulfates, phosphates, and the like) or botanical materials (powdered corncobs, rice hulls, walnut shells, and the like). The formulations may include spreader-sticker adjuvants, stabilizing agents, other pesticidal additives, or surfactants. Liquid formulations may be aqueous-based or non-aqueous and employed as foams,

suspensions, emulsifiable concentrates, or the like. The ingredients may include rheological agents, surfactants, emulsifiers, dispersants, or polymers. Detergents may be included to facilitate uptake of the defense signal peptides by plant tissues and cells.

Regardless of the method of application, the amount of the active component(s) are applied at an amount that is effective to confer enhanced disease resistance to plants, which will vary depending on such factors as, for example, the specific disease to be controlled, the specific plant or crop to be treated, the environmental conditions, and the method, rate, and quantity of application of the composition.

Such compositions may be made by formulating purified or unpurified defense signal peptides with the desired biologically-acceptable carrier. The compositions may be formulated prior to administration in an appropriate means such as lyophilized, freeze-dried, dessicated, or in an aqueous carrier, medium or suitable diluent, such as saline or other buffer, for example. The formulated compositions may be in the form of a dust or granular material, or a suspension in oil (vegetable or mineral), or water or oil/water emulsions, or as a wettable powder, or in combination with any other carrier material suitable for agricultural application. Suitable agricultural carriers can be solid or liquid and are well known in the art.

The term "biologically-acceptable carrier" refers to all carriers that are compatible with the growth and development of a cultured cell or tissue, an excised plant part, a seed, a plant grown under greenhouse or field conditions, or other biological entity, e.g., aqueous solutions, buffers, adjuvants, etc. that are ordinarily used in connection with the biological entity, including but not limited to any carrier used in bacterial or plant cell or tissue culture and agriculturally-acceptable carriers. The term "agriculturally-acceptable carrier" covers all adjuvants, e.g., inert components, dispersants, surfactants, tackifiers, binders, etc. that are ordinarily used in formulation technology for compositions used in agriculture to be applied to plants, soils, etc. The formulations may be mixed with one or more solid or liquid adjuvants and prepared by various means, e.g., by homogeneously mixing, blending and/or grinding.

Such compositions of this invention are applied to the environment of the plant for uptake into plant tissues and cells, typically onto the foliage of the plant

or crop to be protected, by conventional methods, such as by spraying. The strength and duration of application will be set with regard to conditions specific to the particular pest(s), crop(s) to be treated and particular environmental conditions. The proportional ratio of active ingredient to carrier will naturally depend on the chemical nature, solubility, and stability of the defense signal protein(s), as well as the particular formulation contemplated.

Other application techniques, e.g., dusting, sprinkling, soaking, soil injection, seed coating, seedling coating, spraying, aerating, misting, atomizing, and the like, are also feasible and may be required under certain circumstances.

The defense signal peptides of the invention may be employed in such compositions singly, in a mixture of defense signal peptides, or in combination with other compounds, including and not limited to other proteins or chemical compounds used for treatment of plants, including but not limited to proteins or chemical compounds used to treat plants for pathogens, insect pests, etc. The method of the invention may also be used in conjunction with other treatments such as surfactants, detergents, polymers or time-release formulations.

The compositions of the present invention may be formulated for either systemic or topical use. The concentration of insecticidal composition which is used for environmental, systemic, or foliar application will vary widely depending upon the nature of the particular formulation, means of application, environmental conditions, and degree of biocidal activity. Typically, the composition will be present in the applied formulation at a concentration of at least about 0.1% by weight and may be up to and including about 99% by weight. Dry formulations of the compositions may be from about 0.1% to about 99% or more by weight of the composition, while liquid formulations may generally comprise from about 0.1% to about 99% or more of the active ingredient by weight.

The formulation may be administered to a particular plant or target area in one or more applications as needed, with a typical field application rate per hectare ranging on the order of from about 0.1 g to about 1 kg, 2 kg, 5, kg, or more of active ingredient.

Identifying Defense Signal Proteins

According to one aspect of the invention, methods are provided for identifying native defense signal peptides from plants and also for screening synthetic peptides for defense signal peptide activity.

A sensitive, rapid "alkalinization assay" (see Examples) is useful for isolate native defense signal peptides from plants or synthetic defense signal peptides produced by chemical synthesis or other means. Cultured plant cells, for example suspension cell cultures that grow at about pH 5 are used. Several laboratories have developed such cell cultures for Arabidopsis, tomato (*Lycopersicon esculentum*), tobacco (*Nicotiana tabacum*), alfalfa (*Medicago sativa*), maize (*Zea mays*), petunia (*Petunia hybrida*), nightshade (*Solanum nigrum*), and sweet potato (*Ipomoea batatus*), for example. Within minutes after adding systemin to cells, an ATP-driven proton pump is inhibited, causing the extracellular medium of the cells to become alkaline. When 1-10 μL aliquots from fractions from plant tissues, e.g., leaves, that have eluted from HPLC columns were added to 1 mL of suspension cultured cells, some fractions caused the cell medium to increase in pH. In order to confirm that a candidate peptide is a defense signal peptide, the peptide is applied to a plant as described in the Examples below or a polynucleotide sequence encoding the candidate peptide is expressed in a plant in order to observe whether disease resistance is enhanced in the plant. Confirmation of the identity of a candidate peptide may be obtained by determining whether the peptide induces defense gene expression (for example, of *PDF1.2* and *PR-1*), e.g., by supplying a solution of the peptide to excised leaves through their cut petioles then analyzing transcript levels, e.g., by semi-quantitative RT-PCR.

Identifying Compounds that Interact with Receptors for Defense Signal Proteins and That Enhance Plant Disease Resistance

According to another aspect of the invention, substances other than peptides and polypeptides, for example, chemical compounds, are screened for their ability to enhance plant defense against diseases. In one approach, the alkalinization assay is used to screen such substances. Candidate substances are added to cultured plant cells and a rise in pH indicates that a candidate substance interacts with a receptor. In a second approach, candidate substances are assayed for binding by an AtPep1 receptor or another receptor for a defense signal peptide. A composition comprising one or more candidate substances that are selected after being screened in an alkalinization assay or receptor binding assay may then be administered to plants in a greenhouse or field trial to assess whether the candidate substance(s) confer enhanced plant defense against a disease. Substances that have

activity in conferring enhanced plant defense may be formulated according to standard formulation approaches for application to plants, to seeds, to the soil, etc. The present invention also includes compositions comprising an amount of such substances that is effective to enhance plant defense against a disease and a biologically (including agriculturally) compatible carrier. Such compositions may also include other ingredients that are used in formulations for application to plants as detailed above.

The invention will be better understood by reference to the following Examples, which are intended to merely illustrate the best mode now known for practicing the invention. The scope of the invention is not to be considered limited thereto.

EXAMPLE 1: Isolation and analysis of *At*Pep1 and paralogs and orthologs thereof

Innate immunity is initiated in animals and plants through the recognition of a variety of pathogen associated molecules that in animals are called "pathogen-associated molecular patterns," or PAMPS, and in plants are called elicitors. Peptides derived from pathogens can be powerful elicitors of plant defense responses (Hahlbrock et al., Proc. Natl. Acad. Sci. USA 92:4150-4157, 1995; van den Askerveken et al., Plant Physiol. 103:91-96, 1993; Kammpren, Curr. Opin. Plant Biol. 4:295-300, 2001; Kunze et al., Plant Cell, 16:3496-3507, 2004; Navarro et al., Plant Physiol. 135:1113-1128, 2004; Fellbrich et al., Plant J. 32:375-390, 2002; He et al., Cell 73:1255-1266, 1993), but plant-derived peptides have not been identified previously that are elicitors of immune responses directed against pathogens.

We have isolated and characterized a 23 amino acid peptide, called *At*Pep1, that is a signaling component of the innate immune system of Arabidopsis. The peptide precursor gene is transcribed in response to elicitors generated by pathogens, and *At*Pep1 is produced to amplify the signaling pathways. Seven paralogs of the *AtproPep1* gene are present in the Arabidopsis genome, and orthologs have been identified in species of several agriculturally important families including Solanaceae, Poaceae, Salicaceae, Vitaceae, and Fabaceae. *At*Pep1 and its paralogs and orthologs play important roles as endogenous signals to amplify innate immunity.

Materials and Methods

Plant growth conditions. *Arabidopsis thaliana* ecotype Columbia seeds were grown in soil in four-inch square pots for six days under low light at approximately 18°C. Germinated seedlings were then grown under day lengths of 16 hours at 21°C. Mutant plants were grown in autoclaved soil.

Alkalinization Assay. Arabidopsis suspension cells were grown with shaking in the dark in 125 mL flasks, using 40 mL NT media as previously described (Pearce et al., Proc. Natl. Acad. Sci. USA 98:12843-12847, 2001). The cells were transferred weekly (2.5 mL) and used for assays 3-5 days after transfer. One mL aliquots of cells were transferred to wells of 24-well culture plates and allowed to equilibrate for one hour while agitated on a rotary shaker at 160 rpm. Aliquots of 1-10 µL from extracts or fractions eluted from HPLC columns were added to cells and the pH of the media was monitored after 20 min.

Purification of AtPep1. *Arabidopsis thaliana* (Columbia ecotype), 28 days after planting, consisting of rosettes, flowers, stems and seed pods, were harvested, frozen in liquid nitrogen, ground to a powder, and stored at –20 °C. Peptides were extracted from 600 g of powder as previously described (Pearce et al., Nature 411:817-820, 2001; Pearce and Ryan, J. Biol. Chem. 278:30044-30050, 2003), using 1200 mL 1% trifluoroacetic acid (TFA). The clear extract was applied to a reversed-phase C18 flash column (Bondesil, Varian Analytical Instruments, Walnut Creek, CA) that was equilibrated with 0.1% TFA/H_2O. After washing with equilibration buffer, the column was eluted with 50% methanol/0.1%TFA. The eluate was vacuum-evaporated and lyophilized to dryness. This material was dissolved in 0.1% TFA and chromatographed on a G-25 Sephadex column (2.5 X 33 cm), and the fractions were monitored with the alkalinization assay. The broad peak of activity that eluted between 1-1.5 void volumes was collected and lyophilized. The yield of dry powder was 109 mgs.

Two hundred forty mg of the powder was dissolved in 9 mL of 0.1%TFA/H_2O, centrifuged, and the clarified solution was applied to a 5 micron, 10 X 250 mm semi-preparative C18 column (#218TP510, Vydac, Hesperia, CA) with a flow rate of 2 ml/min and monitored at 225 nm. After 2 min, a gradient from 0-40% acetonitrile/0.1%TFA was applied to the column and 1 min fractions were collected and assayed as above. A defined activity peak was identified in fractions 36-37 and designated Arabidopsis Peptide 1 (*At*Pep1). *At*Pep1 was

further purified by strong cation exchange chromatography on a 5 μm, 4.6 X 200 mm PolySulfoethyl Aspartamide™ column (The Nest Group, Southborough, MA) equilibrated with 5 mM potassium phosphate, pH 3, in 25% acetonitrile. Two min after applying *At*Pep1 to the column, a gradient of 0-100% elution buffer consisting of 5 mM potassium phosphate, 1 M potassium chloride, pH 3, in 25% acetonitrile was applied for 60 min. Absorbance was monitored at 214 nm and the flow rate was 1 mL/min. Fractions were collected at minute intervals and 10 μL aliquots were assayed for alkalinization activity. The fractions with activity, 58 and 59, were pooled and lyophilized. Further purification of *At*Pep1 was performed on a narrow-bore reversed-phase 218TP52, 5 μm, 2.1 X 250 mm C18 column (Vydac, Hesperia, CA) that had been equilibrated with 0.1% TFA/$H_2$0. The lyophilized material was dissolved in the equilibration buffer and applied to the column. After two min, a gradient of 0-50% acetonitrile in 0.1% TFA was applied over 90 min with a flow rate of 0.25 mL/min, and monitored at 214 nm. Fractions were collected at 1 min intervals and assayed as above. The activity was present in fractions 48-50, which were pooled and lyophilized. Further purification was obtained on the same narrow-bore column but using a 0- 50% methanol/0.05% TFA gradient over 90 min for elution. The activity was found exclusively in fractions 63-64. These fractions were pooled and subjected to amino acid sequence analysis and MALDI-mass spectroscopy.

Peptide sequence analysis and synthesis. N-terminal sequence analysis was performed using Edman chemistry on an Applied Biosystems Procise Model 492 protein sequencer. MALDI-mass spectroscopy was performed on a PerSeptive Biosystems Voyager time-of-flight mass spectrometer equipped with a nitrogen laser (337 nm) with α-cyano-4-hydroxycinnamic acid as the matrix. Peptide synthesis was performed using Fmoc (N-(9-fluorenyl)methoxycarbonyl) chemistry by solid phase techniques using an Applied Biosystems Model 431 synthesizer. Synthetic peptides were purified by reversed-phase C18 HPLC. Peptide stocks (250 μM) were assayed for purity and the mass verified with a Finnigan LC/Q mass spectrometer using direct injection.

Plant stress and hormone treatments. To examine effects of cold stress, plants were placed in a refrigerated growth chamber set to 2°C. To simulate drought stress conditions, plants grown under standard growth chamber conditions

were grown without watering. Methyl jasmonate (Bedoukian Research Inc., Danbury CT) was applied as a 625 μM solution in 0.1% Triton X-100 to the upper surface of leaves and the plants were incubated in plexiglass boxes. Methyl salicylate (Sigma-Aldrich, St. Louis MO), was applied to leaf surfaces at 2 mM in a 0.1% Triton X-100 solution. Ethephon (Phytotechnology Laboratories, Shawnee Mission KS) was sprayed on plants as a 7 mM solution in 0.1% Triton X-100 (Sigma-Aldrich, St. Louis MO). ABA effects were analyzed by spraying plants with a 100 μM solution (mixed isomer, Sigma-Aldrich, St. Louis MO) in 0.1% Triton X-100 (Denekamp and Smeekens, Plant Physiol. 132, 1415-1423, 2003).

Excised-leaf assays. *At*Pep1 peptide dissolved in double distilled water was supplied to excised leaves of 3 to 4 week old *Arabidopsis* plants. Leaves were excised and the petioles were immersed in 800 μL centrifuge tubes containing either the peptide solution or distilled water, and placed in a closed clear plexiglass box containing a thin layer of water for humidity, and a small opening to allow air to enter. Boxes were incubated in a growth chamber under the plant growth conditions described above and sprayed with a fine mist of distilled water every half hour to ensure humidity and prevent wilting. To determine variations in basal levels of the *AtproPep1* transcript among assays, four different leaves from four different plants were used for each treatment, and leaves supplied with either water or *At*Pep1 were taken from the same plants. Assays were terminated by immersing the leaves in liquid nitrogen.

Hydrogen peroxide accumulation was visualized using diaminobenzidine (DAB) (Thordal-Christensen et al., Plant J. 11:1187-1194, 1997).

Semi-quantitative RT-PCR analysis of relative gene expression levels. RNA was isolated using Trizol reagent and manufacturer's instructions (Invitrogen, Carlsbad CA), and 2 μg of RNA template was reverse transcribed with a RETROscript kit (Ambion, Austin TX). PCR reactions were carried out with ExTaq Hot Start polymerase and reagents (Fisher Scientific, Pittsburgh PA). The *AtproPep1* forward and reverse primers with the respective sequences of 5' CTT ATC AGA TCT CAA TGG AGA AAT C 3' and 5' CAA TGT AAC TTA AAG TGC CTA ATT ATG 3' generated a 310 bp intron-spanning product. Primers to ß-tubulin (At5g62690) of 5' CAA CGC TAC TCT GTC TGT CC 3' and 5' TCT GTG AAT TCC ATC TCG TC 3' generated a 681 base pair intron-spanning product. An initial denaturing/polymerase activating step of 5 minutes at 94°C was followed by 31 repetitions of the following

three steps: a thirty second denaturation phase at 94°C, a thirty second annealing period at 55.5°C, and a one minute elongation step at 72°C. The amplification program was terminated with a ten minute final 72°C elongation phase.

The products of each reaction were separated by electrophoresis and were visualized on a Bio Imaging System (SynGene, Frederick MD) using GeneSnap version 6.00.26 software (SynGene, Frederick MD). A high resolution image of the gel was was analyzed using GeneTools analysis software version 3.02.00 (SynGene, Frederick MD). Relative band intensities for each band were normalized to the β-tubulin band. A numerical ratio of amplified *AtproPep1* cDNA to amplified tubulin cDNA was obtained for every sample. To calculate average values, semi-quantitative RT-PCR assays were performed in duplicate, and RNA extractions were performed in triplicate.

Transformation of Arabidopsis with a CaMV 35S:proAtPep1 gene. Genomic DNA was isolated from Arabidopsis leaves using the DNAzol reagent (Invitrogen, Carlsbad CA). The genomic sequence encoding AtproPep1 was amplified using a forward primer 5' ATA AAG AGT CAC ACC CAA TAC CG 3' and a reverse primer 5' TGA TAC TGG TTA TGA ACT TAT GAT GG 3' to generate a 1078 base pair product. A 5' Xho I recognition site and a 3' BamH I site were amplified onto the genomic fragment for ligation into the pART-7 vector (Gleave, Plant Mol. Biol. 20:1203-1207, 1992). Both the proAtPep1 genomic product and the pART-7 vector were digested with BamH I and Xho I enzymes (Promega Biosciences Inc., San Luis Obispo CA), and ligated using the LigaFast rapid DNA ligation system (Promega Biosciences Inc., San Luis Obispo CA). The construct was transformed into chemically-competent *E. coli* TOP10F' cells (Invitrogen, Carlsbad CA) that were plated out on LB-ampicillin (50 μg/mL). A plasmid clone containing the full *AtproPep1* genomic DNA insert with no nucleotide errors was used to generate an *AtproPep1*/pBART construct. Both pBART and *AprotPep1*/PART-7 plasmid were digested with Not I (Promega Biosciences Inc., San Luis Obispo CA) to enable ligation of the CaMV 35S/*AtproPep1* expression cassette into the digested pBART plasmid using the Promega LigaFast kit (Promega Biosciences Inc., San Luis Obispo CA). An empty pART-7 vector was digested with Not I to generate a control pBART construct. TOP10F' chemically competent cells were transformed with the constructs and grown in Luria-Berftani media containing 100 μg/mL spectinomycin (Sigma-

Aldrich, St. Louis MO), 40 µL of a 40 mg/mL solution of X-Gal (Sigma-Aldrich, St. Louis MO) and 40 µL of a 100 mM IPTG (Sigma-Aldrich, St. Louis MO) stock. A pBART clone containing the *CaMV 35S/proAtPep1* construct, and a second clone containing the empty CaMV 35S construct, were transformed into *Agrobacterium tumefaciens* strain AGLO cells (Lazo and Ludwig, Biotechnology (N Y) 9:963-967, 1991) by electroporation using a BioRad electroporator (BioRad Laboratories, Hercules, CA). The transformed cells were grown on 2XYT media (Lazo et al., Biotechnology (N Y) 9:963-967, 1991) containing 100 µg/mL spectinomycin, and viable colonies were screened using RT-PCR with the pART F and pART R primers.

Liquid cultures of *Agrobacterium* carrying the CaMV 35S:*AtproPep1* or empty CaMV 35S constructs were grown in 2XYT media and used for floral dip transformation of *Arabidopsis* plants (Clough and Bent, Plant J. 16:735-743, 1998). Transformed plants were grown to maturity, and the seed was collected and planted. Newly germinated seedlings were treated with a 350 µM solution of the herbicide BASTA (glufosinate ammonium, brand name Finale; Farnam Companies Inc., Phoenix AZ) four times at three day intervals, and healthy plants were screened for the *proAtPep1* transgene via PCR. Plants that were both glufosinate-resistant and that amplified products of the appropriate size were grown to maturity and the seeds planted to recover T2 progeny.

Growth and inoculation of plants with *Pythium irregulare*. Two strains of the oomycete root pathogen *Pythium irregulare*, strain 110305) were grown on water-agar (1%) plates for maintenance of stock cultures and, after growing at room temperature in the dark for one week, were stored at 4°C. *Pythium* stocks for infection assays were grown on 1 x potato dextrose agar (Sigma-Aldrich, St. Louis MO) in the dark for one week at room temperature.

Week-old *P. irregulare* cultures were scraped from the plates into 20 mL of sterile distilled water, the mixture was lightly ground with a mortar and pestle to produce a uniform suspension. Aliquots (250 µL) of the suspension or water were pipetted into the soil of plants having a rosette diameter of 2-3 cm. Plants were grown for 25 days as described above and assayed. The experiments were repeated five times. After two and a half weeks, the plants were photographed to show rosette morphology, and at three and a half weeks were harvested and the roots examined. The day prior to harvest, plants were not watered, so that the soil would easily separate

from the roots. Soil was gently rinsed from the roots of each plant with water, taking care to minimize damage, and each plant was trimmed at the base of the rosette to fully expose the root structure, and photographed.

Identification the *At*proPep1 gene and homologous genes. The gene locus encoding the *At*Pep1 peptide precursor was identified using the National Center for Biotechnology Information (NCBI) TBLASTN version 2.2.7 algorithm (Altschul et al., Nucleic Acids Res. 25:3389-3402, 1997) to search genomic sequences from *Arabidopsis thaliana*. To determine possible localization of the protein in the cell, several predictive programs were employed, including pSORT (Nakai and Kanehisa, Proteins 11:95-110, 1991), ChloroP41 and MitoProt (Emanuelsson et al., J. Mol. Biol. 300:1005-1016, 2000; Emanuelsson et al., Prot. Sci. 8:978-984, 1999). Orthologs to the *AtproPep1* gene were identified using the NCBI TBLASTN version 2.2.7 and Institute of Genomic Research (TIGR) TBLASTN 2.0MP algorithms (Gish, TBLASTN 2.0MP-WashU [27-Aug-2000] [linux-i686 21:46:47 28-Aug-2000] Copyright1996-2000 Washington University, Saint Louis, Missouri USA. [http://blast.wustl.edu]). The predicted protein sequence for each was aligned using the program Clustal W version 1.8, available at the Baylor College of Medicine Search Launcher website.

Analysis of gene expression using RT-PCR. Gene expression was analyzed using RT-PCR (Nishimura et al., Plant Cell 16:1365-1377, 2004). Forward and reverse primers used for RT-PCR analysis are shown in Table 4.

Results

We have isolated a 23 amino acid peptide called *At*Pep1 from extracts of Arabidopsis leaves that exhibits characteristics of an elicitor of the innate immune response. Endogenous peptide elicitors of innate immunity have not been previously known. The identification and isolation of the peptide from soluble extracts of Arabidopsis leaves was facilitated by its ability, at sub-nanomolar concentrations, to cause an alkalinization response that is typical of elicitors (Moyen and Johannes, Plant Cell Environ. 19:464-470, 1996; Felix and Boller, Plant J. 7:381-389, 1999; Pearce et al., Proc. Natl. Acad. Sci. USA 98:12843-12847, 2001; Pearce et al., Nature 411:817-820, 2001; Pearce and Ryan, J. Biol. Chem. 278:30044-30050, 2003).

A bioactive component, *At*Pep1, was identified and purified to homogeneity. Peptides present in a 1% TFA/water extract of Arabidopsis tissues

were passed through a reverse phase semi-preparative C18 flash chromatography column and separated on a G-25 Sepharose column. The breakthrough peak was applied to a C18 HPLC column and 10 μL from 2 mL fractions from the column were assayed for alkalinization activity. The peak identified as *At*Pep1 was further purified through two additional chromatography steps and finally purified by narrow bore HPLC. Fractions were assayed for alkalinization activity and the active peak was analyzed by MALDI mass spectroscopy. The amino acid sequence of the purified peptide was determined by Edman degradation. Its identity as a peptide was established by its molecular mass (M/Z, 2492.65) and amino acid sequence (from amino terminus to carboxy terminus, ATKVKAKQRGKEKVSSGRPGQHN (see Table 1) with a calculated molecular mass of 2491.8). The kD determined by mass spectroscopy matched the kD calculated from the amino acid sequence, indicating that the peptide was not post-translationally modified. The chemically synthesized peptide was found to be as active as the native *At*Pep1, with a 1/2 maximal activity of 0.25 nM.

The sequence of *At*Pep1 was identified in GenBank as being derived from the accession At5g64900, which encodes a small protein of 92 amino acids, with its C-terminal 23 amino acids comprising *At*Pep1. Fig. 1 shows the amino acid sequence of the *At*Pep1 precursor protein, *AtproPep1,* deduced from the protein encoded by the gene At5g64900. The *At*Pep1 sequence at the carboxyl terminus of the precursor protein, is underlined. The amino acid sequence of the precursor protein is highly charged and lacks a leader sequence, indicating that it is not synthesized through the secretory pathway, but rather on cytoplasmic ribosomes.

Expression analysis of *AtproPep1* in response to abiotic and biotic cues. As a first step in seeking a possible function for *AtproPep1* and its encoded peptide, the basal expression level of the gene was assessed in leaves, stems, roots and flowers of Arabidopsis plants was studied using semi-quantitative RT-PCR analysis of *AtproPep1* gene expression in response to treatment of leaves with MeJA, ethephon, MeSA, and *At*Pep1. The relative abundance of the *proAtPep1* transcript was estimated from the expression of the β-tubulin gene as a control. Leaves were wounded by crushing once across the mid-vein with a hemostat. Plants were sprayed with a 250 μM solution of MeJA in 0.1% Triton X-100; with a 2 μM solution of MeSA in 0.1% Triton X-100 or with a 7 mM solution of

ethephon in 0.1% Triton X-100. *At*Pep1 peptide (10 nM in water) was supplied through cut petioles of excised leaves. Total RNA was extracted and analyzed. The forward and reverse primers for real-time PCR (RT-PCR) analysis are shown in Table 4 below.

Table 4: RT-PCR primers

Gene	Primers	Product size
AtproPep1 (At5g64900)	5' CTTATCAGATCTCAATGGAGAAATC 3' (F*) 5' CAATGTAACTTAAAGTGCCTAATTATG 3' (R)	310 bp
PDF1.2 (At5g44420)	5' ATGGCTAAGTTTGCTTCCA 3' (F) 5' TTAACATGGGACGTAACAGATAC 3' (R)	243 bp
PR-1 (At2g14610)	5' GGAGCTACGCAGAACAACTA 3' (F) 5' AGTATGGCTTCTCGTTCACA 3' (R)	306 bp
TAT3 (At2g24850)	5' TACAGGGGTAGTTCAAGCAA 3' (F) 5' CCTAGAGCCACTCCTGGTAT 3' (R)	330 bp
LOX2 (At3g45140)	5' ACGGTAGAAGACTACGCACA 3' (F) 5' TAAGGTCTCGAGCTCCTCTT 3' (R)	312 bp
VSP2 (At5g24770)	5' CAAAATATGGATACGGGACA 3' (F) 5' ATTGCCAACGATGTTGTATC 3' (R)	317 bp
ATTI3 (At2g43530)	5' TGGCAATGAAGTCAGTTTCT 3' (F) 5'AGAAGTCGCAGAAGCACTTA 3' (R)	231 bp
β-tubulin (At5g62690)	5' CAACGCTACTCTGTCTGTCC 3' (F) 5' TCTGTGAATTCCATCTCGTC 3' (R)	681 bp

*F=Forward primers; R=Reverse primers.

The *AtproPep1* gene was expressed at low levels in all tissues, giving no clues as to its possible function. Monitoring the expression of *AtproPep1* in intact plants exposed to different environmental conditions and chemicals, including drought and cold stress, UV-B irradiation, wounding, methyl jasmonate (MeJA), methyl salicylate (MeSA), abscissic acid (ABA) and ethephon®, provided more definitive clues. Whereas most treatments did not cause changes in expression of *AtproPep,* wounding, MeJA, ethephon, and *At*Pep1 all induced expression of *AtproPep1*, indicating a possible relationship of the gene and its encoded peptide in

plant defense. Transcription of the gene in response to wounding was detected within about 8 h, whereas spraying the plants with a 250 μM solution of MeJA or 7 mM ethephon induced a strong expression of the gene within an hour. Supplying 10 nM *At*Pep1 through cut petioles of excised leaves induced expression of the *AtproPep1* gene within two hours.

*At*Pep1 regulates transcription of pathogen defense genes. The expression of *AtproPep1* in response to MeJA and ethylene (Et) suggested that the encoded peptide might have a role in activating innate immunity in Arabidopsis, although an endogenous peptide had not previously been reported in the innate immune system of any plant. The jasmonic acid (JA)/Et signaling pathway in Arabidopsis activates the expression of defensive genes including *PDF1.2* (*defensin*), while the salicylic acid (SA) pathway activates several pathogen-related (PR) genes (Penninckx et al., Plant Cell 8:2309-2323, 1996; Lorenzo et al., Plant Cell 15:165-178, 2003; Zimmerli et al., Plant J. 40:633-646, 2004; Penninckx et al., Plant Cell 10:2103-2113, 1998; Hammond-Kosack and Parker, Curr. Opin. Biotechnol. 14:177-193, 2003; Mauch-Mass and Matreau, Ann. Bot. 82:535-540, 1998).

In order to determine whether *At*Pep1 regulates defense gene expression, we determined the fold induction of defense related genes in excised Arabidopsis leaves in response to 10 nM *At*Pep1 supplied through their cut petioles. After 2 hr, transcript levels were analyzed for expression levels of *PDF1.2* (*defensin*), *PR-1* (*pathogenesis-related* 1) *LOX2* (*lipoxygenase 2), VSP2* (*vegetative storage protein2*) and ATTI3 (*Arabidopsis thaliana trypsin inhibitor 3*), relative to levels in untreated excised leaves. Expression was determined by semi-quantitative RT-PCR using a β-tubulin gene as a control. Supplying excised Arabidopsis leaves with solutions of *At*Pep1 through their cut petioles induced a strong expression of *PDF1.2* and *PR-1*.

We also performed similar assays with an Arabidopsis triple mutant (*fad3-2, fad7-2, fad8*; McConn and Browse, Plant Cell 8:403-416, 1996) that is incapable of synthesizing jasmonic acid, and a mutant (*ein2-1*; Guzman and Ecker, Plant Cell 2:513-523, 1990) that is incapable of perceiving ethylene. *At*Pep1 was supplied at 10 nM for 2 hr, and RNA isolated and assayed by semi-quantitative RT-PCR for gene expression levels. *At*Pep1 did not induce the expression of *AtproPep1*,

PDF1.2 or *PR-1*. These experiments suggested that *At*Pep1 acts upstream from the JA/Et and SA pathways to activate *PDF1.2, PR-1* and *AtproPep1*.

We also studied the accumulation of H_2O_2 in leaves supplied for 2 hr with water, 10 nM *At*Pep1, or with 10 nM *At*Pep1, all containing 1 mg/mL of DAB to visualize H_2O_2 accumulation. Leaves treated with *At*Pep1 and DAB were also co-supplied with 100 μM DPI, an inhibitor of NADPH oxidase. We also studied the transcription of *PDF1.2* and *PR-1* in leaves of wild-type plants in response to supplying with 10 nM *At*Pep1 in the presence or absence of DPI (diphenylene iodonium chloride), an inhibitor of NADPH oxidase in both plants and animals (O'Donnell et al., Biochem. J. 290:41-49, 1993). The expression of each gene was analyzed by RT-PCR and compared to expression in excised plants treated only with water. The results indicated that reactive oxygen species (ROS) generated in both the JA/Et and SA pathways is required for *PDF1.2* and *PR-1* transcription (Penninckx et al., Plant Cell 8:2309-2323, 1996; Hammond-Kosack and Parker, Curr. Opin. Biotechnol. 14:177-193, 2003; Mackerness et al., Plant Cell and Environ. 22:1413-1423, 1999).

AtproPep1 over-expression in Arabidopsis enhances innate immunity. Arabidopsis plants were transformed with a *CaMV-35S-AtproPep1* transgene in order to assess the effects of the constitutive synthesis of *At*Pep1 on the expression of defense genes using semi-quantitative RT-PCR. In previous studies, overexpression of the tomato prosystemin precursor gene (McGurl et al., Proc. Natl. Acad. Sci. USA 91:9799-9802, 1994) caused a constitutive over-expression of defense genes. This is apparently due to the constitutive synthesis of prosystemin in the cytoplasm of cells (prosystemin, like *At*proPep1, lacks a leader sequence) where it is processed to systemin. Analysis of transgenic Arabidopsis plants overexpressing *AtproPep1* behaved in a similar manner as the prosystemin gene in that it caused an over-expression of defense genes, in this case of *PDF1.2* and *PR-1*. The fold expression of the various genes (compared to wild-type) was found to be: *AtproPep1*, 12.7 ± 6.4; PDF1.2, 4.4 ± 0.5; PR-1, 2.1 ± 0.4; LOX2, 1.0 ± 0.2; VSP-2, 0.9 ± 0.1; and ATTI3, 1.2 ± 0.1. These results indicated that plants over-expressing *AtproPep1* were synthesizing *At*Pep1 in the absence of pathogen attacks or elicitors, constitutively signaling the defense response.

Transgenic Arabidopsis plants constitutively over-expressing *AtproPep1* were assayed for enhanced resistance against a root pathogen, *Pythium irregulare*, an oomycete that has been employed previously to demonstrate the effects of signaling mutants of Arabidopsis on disease resistance (Staswick et al., Plant J. 15:747-754, 1998; Vijayan et al., Proc. Natl. Acad. Sci. USA 95:7209-7214, 1998). The soils of young wild type (Columbia) and transgenic plants overexpressing a 35S:*AtproPep1* gene (having rosette diameters of 2-3 cm) were inoculated with either a suspension of *Pythium irregulare* strain 110305 propagules, or with sterile water, and the plants were grown for 25 days post-inoculation. Five repetitions were performed with 16 plants of each genotype in each experiment. The aerial parts of the wild-type plants inoculated with *Pythium* were slightly smaller than uninoculated wild-type or transgenic plants. However, the roots of plants from duplicate experiments in which the root masses of uninoculated wild-type and transgenic plants are compared to the roots of inoculated wild-type and transgenic plants clearly showed that the over-expression of *AtproPep1* had enhanced the resistance of the plants toward the root pathogen.

AtProPep1 paralogs. *AtproPep1* belongs to a seven-member gene family in Arabidopsis of which one gene is unannotated. Three paralogs, At5g64890, At5g64900 (*AtproPep1*), and At5g64905, are sequentially encoded in a 5.5 kilobase region of chromosome V (NCBI Arabidopsis Genome Database). Paralogs At5g09980 and At5g09990 and the unannotated gene are also found on chromosome V, but in a 3.8 kb region at a distal region on the second arm of the chromosome. At2g22000 is found on chromosome II. In comparing the amino acid sequences of the open reading frames of the paralogs, a low overall amino acid sequence identity was found, but within the C-terminal region of each gene where the putative *At*Pep1 sequences reside, the amino acid identities ranged from 35% to 65%. Four genes, At5g64905, At5g64900, At5g64890 and At5g09980, are expressed relatively strongly in excised Arabidopsis leaves in response to supplying either *At*Pep1 through cut petioles, or by spraying with MeJA. However, spraying plants with MeSA strongly induced only two of the genes that are induced by *At*Pep1 and MeJA, i.e. At5g64890 and At5g64905. This differential regulation of *At*proPep1 paralogs suggests that a complex signaling network is at play in the leaves, and that cross-talk occurrs between the JA/Et and SA pathways, regulating the expression levels of the paralogs. This differential

expression of *AtproPep1* paralogs was also found in the results from several recent microarray analyses of Arabidopsis genes transcribed in response to pathogens and elicitors. In these analyses the *AtproPep1* paralogs were included without any knowledge of their possible signaling roles.

We performed a transcript analysis in order to determine the increases in transcription of *AtproPep1* paralogs in Arabidopsis leaves in response to the pathogens *P. infestans* (oomycete), *B. cineria* (fungus), and *Ps. syringae* DC 3000 (bacteria) (Toufighi et al., Plant J. 43:153-163, 2005; Craigon et al., Nucleic Acids Res. 32:D575-577, 2004), and to the elicitors NPP1, HrpZ, flg22, and elf18, derived from oomycetes, and bacteria, respectively (Kammpren, Curr. Opin. Plant Biol. 4:295-300, 2001; Kunze et al., Plant Cell, 16:3496-3507, 2004; Navarro et al., Plant Physiol. 135:1113-1128, 2004; Fellbrich et al., Plant J. 32:375-390, 2002); He et al., Cell 73:1255-1266, 1993). All of these treatments strongly induce the transcription of At5g64890 and At5g64905, the two paralogs that are strongly induced by treating leaves with MeJA, MeSA and *At*Pep1.. However, the lack of induction of At5g64900 by the pathogens, a gene induced by MeJA and *At*Pep1, and the induction of At5g64900, At5g64890 and At5g64905 by elicitors as well as by the pathogens, indicates that differential induction of the *AtproPep* family of genes may be governed by the types of elicitors related to individual pathogens. The data presented herein supports a model in which the paralogs At5g64900, At5g64905, and At5g64890 are transcribed in response to elicitors of the JA/Et signaling pathway, while the genes At5g64905, and At5g64890 are transcribed in response to elicitors of SA. The nascent proproteins or processed peptides are transported to the apoplast, where they interact with a cell surface receptor(s) to amplify the immune response. Thus, the *AtproPep1* paralogs are components of an amplification system for a broad spectrum of elicitors that activate the innate immune response.

Discussion

Some fundamental similarities are found among signaling components of animal and plant innate immune systems, including the recognition of PAMPS and/or elicitors from pathogens, the involvement of LRR receptor kinases that monitor the signals, and the resulting activation of defense gene transcription of genes involved in early steps of innate immunity. Several peptides originating from plant pathogens can activate the plant innate immune response, including

Pep13, AVR9, and elicitins derived from fungi (Hahlbrock et al., Proc. Natl. Acad. Sci. USA 92:4150-4157, 1995; van den Askerveken et al., Plant Physiol. 103:91-96, 1993; Kammpren, Curr. Opin. Plant Biol. 4:295-300, 2001), and the peptides hrpZ, NPP1, flg22 and elf13 from bacteria (Kunze et al., Plant Cell, 16:3496-3507, 2004; Navarro et al., Plant Physiol. 135:1113-1128, 2004; Fellbrich et al., Plant J. 32:375-390, 2002; He et al., Cell 73:1255-1266, 1993), as examples. However until this report endogenous plant peptides have not been reported that are involved with signaling roles directed against pathogen attacks. We have discovered a family of genes that encode small peptides are rapidly and are strongly transcribed along with defense genes in response to pathogens and their elicitors, appearing to assure a rapid, strong amplification of the innate immune response. The low expression or lack of expression of some of the *AtproPep* genes demonstrates that the paralogs are differentially expressed in response to pathogen infections and elicitors. Fusions of all paralogs with green fluorescent protein (GFP) and beta-glucuronidase (GUS) are used to investigate expression of the paralogs to determine their tissue-specific expression and their roles in defense responses.

The induction of the defense responses by *At*Pep1 is mediated by a binding protein on the cell surface of Arabidopsis suspension cultured cells that interacts with *At*Pep1 with the characteristics of a receptor, further supporting the fundamental concept proposed in the model described above in which the *AtproPep1* paralogs serve as components of an amplification system for a broad spectrum of elicitors that activate the innate immune response..

Searches of plant genomic databases and EST collections identified *AtproPep* orthologs in species from several plant families. Table 1, *supra*, shows the C-terminal sequences of paralogs and orthologs of *AtproPep1* aligned with the *At*Pep1 peptide sequence. Paralogs are grouped above and dicot and monocot orthologs are grouped below.

The regions with the highest amino acid identities among the deduced proteins occur within the C-terminal residues of each where the *At*Pep homologs are found. The deduced canola peptide exhibited the highest identity with the *At*Pep peptides, being a Brassicaceae species. All of the putative *At*Pep homologs have a conserved glycine at residue #17 (numbers aligned with *At*Pep1), and all but paralogs from the Poeceae family contain an asparagine at residue #23. Each

peptide contains several proline, glycine, and serine residues within a ten amino acid C-terminal region that may be important for receptor recognition.

The chemical and physiological properties of the *At*Pep1 family members, their precursor proteins, and their genes, are strikingly similar to the properties of the 18 amino acid peptide signal systemin, its precursor prosystemin, and its gene, that are components of the signaling pathway for defense against herbivorous pests of the Solanaceae family (Ryan and Pearce, "Peptide hormones/systemins," in Encyclopedia of Biological Chemistry, ed. Lennarz and Lane, vol 3, pp. 381-384, Elsevier, 2004). Both *At*Pep1 homologs and tomato systemin homologs are cleaved from the carboxy (C)-termini of precursor proteins that lack leader peptides, both precursors are small, highly positively charged proteins, and each activates defense genes. The mechanism for processing *At*Pep1 is not known, nor is it known if it is the *At*Pep1 or its peptide precursor that is transported to the apoplast, or if more than one receptor is involved in recognizing the different peptides.

These results indicate that the major role for receptor-mediated defense-signaling peptides in plants is to amplify signaling that is initiated by wounding and elicitors to mount a rapid, strong defense against herbivores and pathogens. If *AtproPep* orthologs behave in the same manner when over-expressed in other plant species by constitutively expressing defense genes, they may provide an important new approach to enhance innate immunity in a broad spectrum of agriculturally important crops.

EXAMPLE 2: An LRR receptor kinase is a component of AtPep1 amplification of innate immunity in Arabidopsis

Over 200 LRR receptor kinases are present in the Arabidopsis genome. Only a few LRR receptor kinases have been identified that interact with the peptide signals in plants. This includes the receptor for the defense peptide signal, systemin (Scheer and Ryan, Proc. Natl. Acad. Sci. 99:9585-9590, 2002) and receptors for the developmental peptide signals CLAVATA1 (Clark and Meyerowitz, Cell 89:575-585, 1997), phytosulfokine-alpha (Matsubayashi et al., Science 296:1470-1472, 2002) and the pathogen-derived peptide flg22 (Gomez-Gomez and Boller, Trends Plant Sci. 7:251-256, 2000).

The discovery that the *At*Pep family of endogenous peptides from Arabidopsis leaf extracts cause an alkalinization of the medium of Arabidopsis suspension cultured cells indicates that the peptide plays a role in plant cells by interacting with a cell surface receptor. Investigations of the biological role of the peptide in Arabidopsis plants indicated that it activates the innate immune system of the plants and functions through an interaction with a cell surface receptor.

We have isolated from Arabidopsis a cell surface LRR Thr/Ser kinase receptor for *At*Pep1, a 23 amino acid signal that that amplifies defense genes for innate immunity. The interaction of *At*Pep1 with the receptor is saturable and exhibits a Kd of 0.25 nM. Two SALK mutant lines with T-DNA insertions in exons of At1g73080 do not express the receptor gene and are not labeled by an *At*Pep1 photoaffinity analog that was used to identify and isolate the receptor protein. However, in contrast to wild-type plants, the SALK insertional lines constitutively expressed high levels of *PDF1.2* and *PR-I,* indicting that the receptor was negatively regulating defense gene expression in the absence of *At*Pep1. The AtPep1 receptor plays a central role in amplifying innate immunity activating defense gene expression when interacting with AtPep peptides that are synthesized in response to elicitors.

Methods

Synthesis of AtPep1 analogs. AtPep1, Cys-AtPep1 and Tyr-AtPep1 were synthesized using solid-phase instrumentation, (peptide synthesizer Model 431A; Applied Biosystems, Foster City, CA). After synthesis, the polypeptides were purified using C18 reverse-phase, high-performance liquid chromatography (HPLC), as previously described (Pearce and Ryan, J. Biol. Chem. 278:30044-30050, 2003; Pearce et al., Proc. Natl. Acad. Sci. USA 98:12843-12847, 2001; Pearce et al., Nature 411:817-820, 2001; Scheer and Ryan, The Plant Cell 11:1525-1535, 1999; Shevchenko et al., Anal. Chem. 68:850-858, 1996). Stock solutions of the peptides (2.5 mM) were prepared in water and stored at -20 °C. Iodination of Tyr-AtPep1 was performed using IODO-GEN Pre-Coated Iodination Tubes (Pierce, Rockford, IL). Two hundred µl of NaI (20 mM in 0.1 M phosphate buffer, pH 8.0) solution was oxidized in the IODO-GEN Pre-Coated Iodination Tube for 6 min. Oxidized NaI solution was transferred into a 1.5 ml tube containing 100 nmol Tyr-AtPep1, and maintained at room temperature for 6 min in the dark gently

agitating the tube every 30 sec. Iodinated Tyr-AtPep1 was purified using HPLC as described previously (Scheer and Ryan, The Plant Cell 11:1525-1535, 1999), and quantified with bicinchoninic acid (Pierce, Rockford, IL). All analogs were anlyzed by LCQ ion trap mass spectrometry (Finnigan, San Jose, CA). Cys-AtPep1 was coupled through a disulfide bond to the photoaffinity cross-linker, N-(4-[*p*-azidosalicylamido]butyl)-3'-(2'-pyridyldithio)propionamide (APDP) (Pierce, Rockford, IL) to produce azido-Cys-AtPep1, which was purified by HPLC by methods previously described (Scheer and Ryan, Proc. Natl. Acad. Sci. 99:9585-9590, 2002).

Radioactive iodinations of Tyr-AtPep1 and azido-Cys-AtPep1 were performed using 2 mCi of $Na^{125}I$ and 12.5 nmol and the products were purified by HPLC (Scheer and Ryan, Proc. Natl. Acad. Sci. 99:9585-9590, 2002). The specific activity of the purified mono-iodinated Tyr-AtPep1 and azido-Cys-AtPep1 were 2.58 mCi/nmol, while the The specific activity of the diiodinated forms was 5.15 mCi/nmol. The mono-iodinated analog of azido-Cys-AtPep1 was found to comprise 90% of the iodinated analog and was employed for photoaffinity labeling.

Alkalinization assay. Medium alkalinization activity of Arabidopsis suspension cultured cells by AtPep1 analogs was analyzed as previously described for systemin (Pearce et al., Science 253:895-897, 1991), RALF (Pearce et al., Proc. Natl. Acad. Sci. USA 98:12843-12847, 2001), HypSys peptides (Pearce and Ryan, J. Biol. Chem. 278:30044-30050, 2003; Pearce et al., Nature 411:817-820, 2001), and AtPep1 (*supra*). The alkalinization activity of azido-Cys-AtPep1 were carried out after incubation the analog with cells in darkness for 10 min, when the pH was recorded.

Binding assays. Binding assays of Arabidopsis cells with $^{125}I_1$-Tyr-AtPep1 were performed by the methods of Scheer and Ryan (Pearce et al., Nature 411:817-820, 2001) modified as follows: Arabidopsis suspension cultured cells were subcultured and grown for 4-5 days, washed with culture medium, and diluted with medium to a level of 0.2 mg fresh weight/ml. Two mL of cells were aliquoted into each well of 12-well culture plates and allowed to equilibrate on an orbital shaker (160 rpm) at room temperature for 1 hr. $^{125}I_1$-Tyr-AtPep1 was added to the medium, and 500 μL of cells were removed at selected times and filtered through a

2.5 cm Type A/E Glass Fiber Filter (Pall Corporation, Ann Arbor, MC) using a 12-well vacuum filtration manifold (Millipore, Bedford, MA). The filtered cells were washed three times with 5 mL of cold MS medium containing 3 % sucrose, suspended in 1 mL MS medium containing 3 % sucrose in a glass test tube, and analyzed for total radioactivity in a gamma-ray counter (Isodata 2020; Isodata Inc., Palatine, IL). Specific binding was calculated by subtracting nonspecific binding (binding in the presence of 250-fold native AtPep1) from total binding. $^{125}I_1$-Tyr-AtPep1 bound to the cell surface within a minute, and equilibrated within 4 min.

Photoaffinity labeling. To 1 mL of cultured cells in darkness was added ^{125}I-azido-Cys-AtPep1 (0.25 nM final concentration as described above) and the cells were shaken for 10 min on an orbital shaker (160 rpm) at room temperature. The cells were transferred to a 1.5 mL glass tube, centrifuged at 10,000 x g and the sedimented cells were dispersed in 1 mL cold MS medium containing 3 % sucrose and centrifuged as above. This wash was repeated twice. The cells were resuspended in 1 mL MS medium and irradiated with UV-B for 10 min on ice to photoactivate the azido group to effect the crosslinking. The cells were washed with 1 mL MS medium, centrifuged as above, and resuspended in 400 µL of 5 % SDS. The cells were disrupted by boiling for 30 min, and the insoluble debris was removed by centrifugation at 10,000 x g. Proteins in the clear supernatant were precipitated by adding 1.25 volumes of methanol/chloroform (vol/vol). After centrifugation at 10,000 x g, the pellet was recovered and dissolved in 100 µL of Laemmli sample buffer containing 5 % SDS, boiled for 10 min, and separated using 8 % SDS-PAGE. The gels were dried and exposed to X-ray film for 50 hr to visualize labeled proteins. In competition assays, unlabeled AtPep1 and suramin were added to the cells and incubated for 10 min before adding ^{125}I-azido-Cys-AtPep1. The same procedures described above were employed to detect labeled proteins.

Purification of AtPep1-binding protein. ^{125}I-azido-Cys-AtPep1 (0.25 nM) was added to 1 L of Arabidopsis suspension cultured cells and incubated for 10 min in the dark as described above, and collected on Miracloth (Calbiochem, San Diego, CA). After washing the cells with 1 L of cold water, the cells were suspended in 500 ml of MS medium containing 3 % sucrose, and irradiated with UV-B for 15 min while being mixed on an orbital shaker at 160 rpm. The cells were again collected on Miracloth and washed with 1 L cold water. Microsomal

fractions were prepared by differential centrifugation as previously described (Pearce et al., Proc. Natl. Acad. Sci. USA 98:12843-12847, 2001), and stored at -80 °C. This process was repeated three times and the microsomal proteins were pooled. Protein was measured by Bio-Rad Protein Assay reagent (BIO-RAD, Hercules, CA) using bovine serum albumin as a standard.

Purification of AtPep1-binding protein from the membranes was performed as described previously with modifications (Pearce et al., Proc. Natl. Acad. Sci. USA 98:12843-12847, 2001). Briefly, 25 and 55 mg of radiolabeled and unlabeled microsomal proteins, respectively, were mixed and separated by 2 sets of 7.5 % SDS-PAGE (0.6 X 14 X 9 cm) for 16 with 37 V at room temperature. The gels were sliced horizontally into 5 mm width and the radioactivity measured with a gamma-counter (Isodata 2020; Isodata Inc., Palatine, IL). The gel slices near 180 kD, containing the highest radioactivity were pooled. The gel slices were mascerated and the proteins were recovered by incubating the mascerate three times for 30 min with 25 mL of 20 mM Tris-HCl, 0.5 M NaCl, pH 7.5. The eluted proteins were pooled and incubated with 75 μL of Conconavalin A-Sepharose 4B (Amersham Bioscience, Piscataway, NJ) for 6 h at room temperature to trap the Con A-protein complexes. The Con A-Sepharose was washed with 50 mL of 0.5 M alpha-methyl-D-glucoside three times, followed by 50 mL of H_2O three times, to elute loosely bound proteins. Bound proteins were eluted by boiling the Con A-Sepharose in 500 μL of 5 % SDS, and the eluted proteins were precipitated by adding 1.25 volumes of methanol/chloroform (2:1, vol/vol). Half of the eluted proteins were separated using 7.5 % SDS-PAGE (0.15 X 14 X 8.5 cm) at100 V for 8 h at room temperature. The gel was sliced horizontally into 1 mm widths, and the gel slice containing highest radioactivity was digested with trypsin (Promega, Madison, WI) according to Shevchenko et al. (Anal. Chem. 68:850-858, 1996). The other half of the eluted proteins were digested with Peptide-N-Glycosidase F (PNGase F) (Prozyme, San Leandro, CA), separated by 7.5 % SDS-PAGE, recovered as above, and digested with trypsin. Peptides generated in the trypsin digests were analyzed by matrix-assisted laser desorption ionization time-of-flight mass spectrometer (MALDI-TOF MS), Voyager-DE RP Biospectrometry Workstation (Applied Biosystems, Framingham, MA). Protein was identified by searching in the National Center for Biotechnology Information database using Mascot (www.matrixscience.com).

Analysis of T-DNA insertional lines. *Arabidopsis thaliana* T-DNA insertional lines, SALK 014538, SALK 059281 and SALK 064539, were obtained from ABRC (Ohio State University) through The Arabidopsis Information Resource. The plants were screened by RT-PCR using gene-specific primer pairs and a primer specific for the T-DNA left border. Total RNA was purified from rosette leaves.

Microsomal fractions were prepared from one-month-old plants by differential centrifugation as previously described (Scheer and Ryan, Proc. Natl. Acad. Sci. 99:9585-9590, 2002). The membranes were photoaffinity labeled (Takayama et al, Nature 413:534-538, 2001) using the radiolabeled azido analog used with suspension cultured cells. The membranes were incubated for 60 min at approximately 4 ^{0}C, irradiated with UV-B for 15 min, separated by SDS-PAGE, dried and analyzed by radioautography to identify labeled proteins.

Results

An analog of AtPep1 was synthesized with a Tyr residue attached to its N-terminus and radiolabeled with ^{125}I to quantify binding. The Tyr analog and mono- and di-iodo-Tyr-AtPep1 analogs were separated by HPLC, and the concentration-dependent activities of AtPep1 and these analogs where tested in the alkalinization assay. All were found to be as fully active as AtPep1 in the alkalinization assay. The mono-iodinated Tyr-AtPep1 comprised about 95% of the iodinated proteins and was employed for binding studies. Saturation kinetics of the binding of mono-^{125}I-Tyr-AtPep1 with Arabidopsis suspension cultured cells (six repetitions using 10^{6} cells/assay) showed that the ^{125}I-labeled peptide was maximally bound to Arabidopsis cells within about 10 min, saturating the sites at about 0.1 nM peptide. A Kd of 0.25 nM was estimated from Scatchard analysis of the saturation data, which is typical of ligand-receptor interactions and indicates that AtPep1 was interacting with a cell surface binding protein with the characteristics of a receptor.

A photoaffinity labeled AtPep1 was prepared by synthesizing an analog with a Cys residue at its N-terminus so that an ^{125}I-labeled azido adduct, APDP, could be crosslinked to the peptide through a disulfide bond. All procedures were in darkness unless otherwise specified. The azido-Cys-AtPep1was purified by

HPLC and assayed for its alkalinization response under red light to avoid photoactivating the azido group. The analog was as fully active as the native peptide in the alkalinization assay. The azido analog was iodinated with ^{125}I and incubated with Arabidopsis suspension cultured cells for 10 min and subjected to UV-B irradiation to activate the azido group for crosslinking with binding proteins. SDS-PAGE analyses of the radiolabeled membrane proteins revealed a single labeled protein band of Mr approximately 180 kD. Pre-incubation of 2.5 nM AtPep1 to cells totally abolished labeling. Tomato systemin (LeSys), a nonhomologous 18 amino acid peptide signal from tomato plants (Pearce et al., Science 253:895-897, 1991) did not compete for binding with the labeled analog and had not effect on photoaffinity labelling. Suramin, a polycyclic non-specific inhibitor of peptide hormone-receptor interactions in both animals and plants (Stratmann et al., Proc. Natl. Acad. Sci. USA, 97:8862-8867, 2000), inhibited the labeling of the 180 kD protein by the photoaffinity AtPep1 analog at 100 nM, supporting a membrane association for the 180 kD labeled protein. The labeled protein band from SDS-PAGE gels was eluted and incubated with the carbohydrase PNGaseF to enzymatically remove covalently bound carbohydrates. This enzyme caused a decrease in the kD of the photoafinity labeled protein from about 180 kD to about 150 kD, indicating that the binding protein was glycosylated.

Purification of the radiolabeled protein from 1 L of Arabidopsis cells in late log phase was achieved using final steps of ConA-Sepharose affinity chromatography followed by SDS-PAGE. After electrophoresis, the labeled 180 kD protein band was excised from the gel and the protein recovered. Half of the protein was digested with trypsin and the fragments were analyzed by MALDI-TOF mass spectroscopy. The other half of the eluted protein was treated with the enzyme PNGase F to remove carbohydrate and was subjected to gel electrophoresis. The protein in the 150 kD band was recovered, digested with trypsin, and was also analyzed by mass spectroscopy. The amino acid sequences of 18 tryptic fragments from the 180 kD peptide exactly matched sequences of the Arabidopsis LRR receptor kinase gene, At1g73080. The deglycosylated protein yielded three large fragments of from 12 to 18 amino acids in length that were also exact matches to sequences within the At1g73080 gene. The nucleotide sequence

of the AtPep1 receptor gene (At1g73080) the deduced receptor polypeptide is provided in Fig. 2 shows the structure of the At1g73080 gene, which is comprised of an 646 amino acid extracellular domain containing 27 LRR motifs; a 22 amino acid transmembrane domain; and a 280 amino acid Ser/Thr receptor kinase domain.

The leaves of two SALK T-DNA insertional lines having insertions in the exons of the gene At1g73080, SALK 014538 and SALK 059281, did not express the gene when analyzed by RT-PCR, using wild-type plants and a SALK 059281 insertional mutant of gene At5g55480 as controls. Microsomal membrane proteins from the two At1g73080 mutant lines, and from wild type plants and the control SALK 059281 line, were analyzed for proteins that were specifically photoaffinity labeled by ^{125}I-azido-Cys-AtPep1. A protein was labeled in the membranes from wild-type plants and the SALK 059281 plants, but not from the SALK mutants unable to express the receptor gene. The label was found in a 180 kD protein band and in a slightly lower doublet band that appears to be degradation products of the receptor protein, since they were labeled. The proteins labeled in the wild-type and SALK 059281 microsomal membranes were absent when the membranes were preincubated with 2.5 nM AtPep1 and then photoaffinity labeled, indicating that the proteins labeled in the membranes of wild-type and SALK 059281, and not in membranes from the lines with mutants in the At1g73080 gene, were the AtPep1 receptor and its biologically active fragments.

To further investigate the role of the 180 kD protein as the receptor of AtPep1, competition experiments were performed between the $^{125}I_1$-Tyr-AtPep1 analog and synthetic peptide homologs derived from sequences at the N-termini of all seven of the AtPep family members that corresponded to the 23 amino acid AtPep1 (gene At5g684900). The peptides were purified after synthesis on HPLC and each assayed for biological activities at increasing concentrations in the alkalinization assay to determine the concentration of each that caused maximal activity. Fig. 4 shows the concentration dependence of synthetic AtPep peptides deduced from the seven members of the AtproPep1 gene family in the alkalinization assay. All peptides except those derived from At1g09980 gene and the unannotated gene were fully active at about 2.5 nM concentrations. The two genes with diminished activity were from Subfamily II, which reside relatively

close together on Chromosome V. All seven synthetic peptides competed with the $^{125}I_1$-Tyr-AtPep1 analog for binding with Arabidopsis suspension cultured cells, with a pattern similar to their biological activities in the alkalinization response. As in the alkalinization assay, the peptides derived from the At5g09980 gene and the unannotated gene were much weaker competitors than peptides derived from the other genes.

EXAMPLE 3: Shorter peptides from the C-terminus of AtPep1 possess substantial defense signal peptide activity

Analogs of AtPep1 from the C-terminus of AtPep1 were synthesized and assayed in the alkalinization assay. One mL aliquots of 4 day old Arabidopsis cells were allowed to equilibrate on an orbital shaker at 180 rpm for one hour. A ten µL aliquot of each peptide solution was added to the cells. Peptide concentrations of 0.25 nM, 2.5 nM, and 25 nM were tested. After 20 min, the pH of the media was recorded. The results are shown in Fig. 5. An analog of AtPep1 missing the carboxy-terminal amino acid was completely inactive, whereas deletions from the amino-terminus of the peptide resulted in a sequential reduction in activity, until peptides with 9 carboxy-terminal amino acids remaining (SSGRPGQHN) were inactive. A peptide with 10 carboxy-terminal amino acids remaining had substantial defense signal peptide activity at the 25 nM peptide concentration, causing a change in pH of over 0.20 units, and longer analogs had progressively greater activity. A peptide consisting of only the 15 C-terminal amino acids was nearly as active as the native peptide at approximately 2.5 nM and had substantial activity even at the lowest concentration tested. It is expected that substantial defense signal peptide activity will be retained by analogs of other defense signal peptides that comprise sequences from the C-terminus of the peptides.

EXAMPLE 4: Alanine substitutions in residues of a 15-mer analogs from the C-terminus of AtPep1

The 15-mer from the carboxy-terminus of AtPep1 (RGKEKVSSGRPGQHN) was substituted with alanine at each position to assess which amino acids were necessary for the alkalinizing activity. Fig. 6 shows the effect of these single alanine amino acid substitutions on the activity of the 15-mer peptide in the alkalinization assay. The set of substituted 15-mer peptides was assayed using four-day-old Arabidopsis cells. Ten ml of each peptide (2.5 pmoles)

was added to 1ml of cells to make a final concentration of 2.5nM. After 20 min, the pH of the media was recorded. The data is the average of three separate experiments.

A Ser to Ala substitution at position 7, counting from the amino-terminus of the 15-mer, and a Gly to Ala substitution at position 9, exhibited little activity. Computer modeling predicted that these two amino acids would be involved in a beta-turn within the peptide region of -SSGR- (compare with residues 15-18 of the sequence of AtPep1 shown in Table 1). Substituting Ala for Ser (-ASGR-) abolished the predicted turn and severely abolished activity (half-maximal activity at ~250 nM), while substituting Ala for Gly was even less active (half-maximal activity of > 250 nM. However, neither of these analogs were able to compete with the non-substituted 15-mer for receptor binding, indicating that the structural changes in this region may have severely modified the conformation without competing for the receptor binding site. Peptides with alanine substitutions at all residues were synthesized and assayed, with most showing no differences in activity than the native AtPep1.

All publications, patents and patent applications are incorporated herein by reference. While in the foregoing specification, this invention has been described in relation to certain preferred embodiments thereof, and many details have been set forth for purposes of illustration, it will be apparent to those skilled in the art that the invention is susceptible to additional embodiments and that certain of the details herein may be varied considerably without departing from the basic principles of the invention.

WHAT IS CLAIMED IS:

1. A composition comprising an isolated defense signal peptide 10 or more amino acid residues in length that has substantial defense signal peptide activity.
2. The composition of claim 1 wherein the defense signal peptide is 15 or more amino acid residues in length.
3. The composition of claim 1 wherein the defense signal peptide is 20 or more amino acid residues in length.
4. The composition of claim 1 wherein the defense signal peptide is 23 or more amino acid residues in length.
5. The composition of claim 1 wherein the defense signal peptide is between about 10 and about 50 amino acid residues in length.
6. The composition of claim 1 wherein the defense signal peptide is between about 15 and about 50 amino acid residues in length.
7. The composition of claim 1 comprising a polypeptide that is processed in a plant cell to produce the defense signal peptide.
8. The composition of claim 1 wherein the defense signal peptide comprises a sequence having at least 75 percent homology with a polypeptide selected from the group consisting of AtPep1, AtPep2, AtPep3, AtPep4, AtPep5, AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2.
9. The composition of claim 1 wherein the defense signal peptide comprises a sequence having at least 80 percent homology with a polypeptide selected from the group consisting of AtPep1, AtPep2, AtPep3, AtPep4, AtPep5, AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2.
10. The composition of claim 1 wherein the defense signal peptide comprises a sequence having at least 85 percent homology with a polypeptide selected from the group consisting of AtPep1, AtPep2, AtPep3, AtPep4, AtPep5, AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2.
11. The composition of claim 1 wherein the defense signal peptide comprises a sequence having at least 90 percent homology with a polypeptide selected from the group consisting of AtPep1, AtPep2, AtPep3, AtPep4, AtPep5,

AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2.

12. The composition of claim 1 wherein the defense signal peptide comprises a sequence having complete homology with a member of the group consisting of AtPep1, AtPep2, AtPep3, AtPep4, AtPep5, AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2.
13. The composition of claim 1 wherein the defense signal peptide comprises a sequence having at least 90 percent homology with a dicot defense signal peptide consensus sequence.
14. The composition of claim 1 wherein the defense signal peptide comprises a sequence having complete homology with a dicot defense signal peptide consensus sequence.
15. The composition of claim 1 wherein the defense signal peptide comprises a sequence having at least 90 percent homology with a monocot defense signal peptide consensus sequence.
16. The composition of claim 1 wherein the defense signal peptide comprises a sequence having complete homology with a dicot defense signal peptide consensus sequence.
17. The composition of claim 1 comprising a biologically acceptable carrier.
18. A plant comprising the composition of claim 17 applied to a surface of said plant.
19. A seed comprising the composition of claim 17 applied to a surface of the seed.
20. An isolated polynucleotide comprising a sequence that encodes a defense signal peptide operably linked to a plant promoter, wherein expression of the polynucleotide in a cell of a plant causes the plant to exhibit an improvement compared to a control plant lacking the polynucleotide that is selected from the group consisting of improved yield of plant product, reduced disease symptoms, and enhanced resistance to disease infestation.
21. The isolated polynucleotide of claim 20 wherein the defense signal peptide is 10 or more amino acid residues in length.
22. The isolated polynucleotide of claim 20 wherein the defense signal peptide is 15 or more amino acid residues in length.

23. The isolated polynucleotide of claim 20 wherein the defense signal peptide is 20 or more amino acid residues in length.

24. The isolated polynucleotide of claim 20 wherein the defense signal peptide is 23 or more amino acid residues in length.

25. The isolated polynucleotide of claim 20 that comprises a sequence that encodes a polypeptide that is processed in a plant cell to produce the defense signal peptide.

26. The isolated polynucleotide of claim 20 wherein the defense signal peptide is between about 10 and about 50 amino acid residues in length.

27. The isolated polynucleotide of claim 20 wherein the defense signal peptide is between about 15 and about 50 amino acid residues in length.

28. The isolated polynucleotide of claim 20 that encodes a polypeptide that is processed in a plant cell to produce the defense signal peptide.

29. The isolated polynucleotide of claim 20 wherein the defense signal peptide comprises a sequence having at least 75 percent homology with a polypeptide selected from the group consisting of AtPep1, AtPep2, AtPep3, AtPep4, AtPep5, AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2.

30. The isolated polynucleotide of claim 20 wherein the defense signal peptide comprises a sequence having at least 80 percent homology with a polypeptide selected from the group consisting of AtPep1, AtPep2, AtPep3, AtPep4, AtPep5, AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2.

31. The isolated polynucleotide of claim 20 wherein the defense signal peptide comprises a sequence having at least 85 percent homology with a polypeptide selected from the group consisting of AtPep1, AtPep2, AtPep3, AtPep4, AtPep5, AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2.

32. The isolated polynucleotide of claim 20 wherein the defense signal peptide comprises a sequence having at least 90 percent homology with a polypeptide selected from the group consisting of AtPep1, AtPep2, AtPep3,

AtPep4, AtPep5, AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2.

33. The isolated polynucleotide of claim 20 wherein the defense signal peptide comprises a sequence having complete homology with a member of the group consisting of AtPep1, AtPep2, AtPep3, AtPep4, AtPep5, AtPep6, AtPep7, BnPep1, StPep1, PbPep1, GmPep1, MsPep1, VvPep1, OsPep1, OsPep2, TaPep1, TaPep2, ZmPep1, and HvPep2.

34. The isolated polynucleotide of claim 20 wherein the defense signal peptide comprises a sequence having at least 90 percent homology with a dicot defense signal peptide consensus sequence.

35. The isolated polynucleotide of claim 20 wherein the defense signal peptide comprises a sequence having complete homology with a dicot defense signal peptide consensus sequence.

36. The isolated polynucleotide of claim 20 wherein the defense signal peptide comprises a sequence having at least 90 percent homology with a monocot defense signal peptide consensus sequence.

37. The isolated polynucleotide of claim 20 wherein the defense signal peptide comprises a sequence having complete homology with a dicot defense signal peptide consensus sequence.

38. The polynucleotide of claim 20 wherein the sequence that encodes a defense signal peptide has at least 80 percent sequence similarity to a polynucleotide sequence selected from the group consisting of *AtproPep1*, *AtproPep2*, *AtproPep3*, *AtproPep4*, *AtproPep5*, *AtproPep6*, *AtproPep7*, *BnproPep1*, *StproPep1*, *PbproPep1*, *GmproPep1*, *MsproPep1*, *VvproPep1*, *OsproPep1*, *OsproPep2*, *TaproPep1*, *TaproPep2*, *ZmproPep1*, and *HvproPep2*.

39. The polynucleotide of claim 20 wherein the sequence that encodes a defense signal peptide has at least 90 percent sequence similarity to a polynucleotide sequence selected from the group consisting of *AtproPep1*, *AtproPep2*, *AtproPep3*, *AtproPep4*, *AtproPep5*, *AtproPep6*, *AtproPep7*, *BnproPep1*, *StproPep1*, *PbproPep1*, *GmproPep1*, *MsproPep1*, *VvproPep1*, *OsproPep1*, *OsproPep2*, *TaproPep1*, *TaproPep2*, *ZmproPep1*, and *HvproPe*.

40. The polynucleotide of claim 20 wherein the sequence that encodes a defense signal peptide has at least 95 percent sequence similarity to a polynucleotide sequence selected from the group consisting of *AtproPep1*, *AtproPep2*, *AtproPep3*, *AtproPep4*, *AtproPep5*, *AtproPep6*, *AtproPep7*, *BnproPep1*, *StproPep1*, *PbproPep1*, *GmproPep1*, *MsproPep1*, *VvproPep1*, *OsproPep1*, *OsproPep2*, *TaproPep1*, *TaproPep2*, *ZmproPep1*, and *HvproPe*.

41. The polynucleotide of claim 20 wherein the sequence that encodes a defense signal peptide has complete sequence similarity to a polynucleotide sequence selected from the group consisting of *AtproPep1*, *AtproPep2*, *AtproPep3*, *AtproPep4*, *AtproPep5*, *AtproPep6*, *AtproPep7*, *BnproPep1*, *StproPep1*, *PbproPep1*, *GmproPep1*, *MsproPep1*, *VvproPep1*, *OsproPep1*, *OsproPep2*, *TaproPep1*, *TaproPep2*, *ZmproPep1*, and *HvproPe*.

42. An isolated polynucleotide comprising a sequence that encodes a defense signal peptide operably linked to a heterologous promoter.

43. The isolated polynucleotide of claim 42 wherein expression of the polynucleotide in a cell of a plant causes the plant to exhibit an improvement compared to a control plant lacking the polynucleotide that is selected from the group consisting of improved yield of plant product, reduced disease symptoms, and enhanced resistance to disease infestation.

44. The isolated polynucleotide of claim 42 wherein the promoter is a constitutive promoter.

45. The polynucleotide of claim 42 wherein the promoter is a non-constitutive promoter.

46. The polynucleotide of claim 42 wherein the promoter is an organ- or tissue-specific promoter.

47. The polynucleotide of claim 42 wherein the promoter is an inducible promoter.

48. A cell comprising the isolated polynucleotide of claim 20 or claim 42.

49. The cell of claim 48 selected from the group consisting of a plant cell, a bacterial cell, a fungal cell and an insect cell.

50. A plant cell of claim 48.

51. A plant comprising the plant cell of claim 50.

52. A plant of claim 51 selected from the group consisting of Acacia, alfalfa, aneth, apple, apricot, artichoke, arugula, asparagus, avocado, banana, barley, beans, beet, blackberry, blueberry, broccoli, brussels sprouts, cabbage, cantaloupe, carrot, cassava, castorbean, cauliflower, celery, cherry, chicory, cilantro, citrus, clementines, clover, coconut, coffee, corn, cotton, cucumber, Douglas fir, eggplant, endive, escarole, eucalyptus, fennel, figs, garlic, gourd, grape, grapefruit, honey dew, jicama, kiwifruit, lettuce, leeks, lemon, lime, Loblolly pine, linseed, mango, melon, mushroom, nectarine, nut, oat, oil palm, oil seed rape, okra, olive, onion, orange, an ornamental plant, palm, papaya, parsley, parsnip, pea, peach, peanut, pear, pepper, persimmon, pine, pineapple, plantain, plum, pomegranate, poplar, potato, pumpkin, quince, radiata pine, radicchio, radish, rapeseed, raspberry, rice, rye, sorghum, Southern pine, soybean, spinach, squash, strawberry, sugarbeet, sugarcane, sunflower, sweet potato, sweetgum, tangerine, tea, tobacco, tomato, triticale, turf grass, turnip, a vine, watermelon, wheat, yams, and zucchini.

53. The plant of claim 51 that exhibits reduced symptoms from, or enhanced resistance to, a disease caused by an organism of a genus selected from the group consisting of Alternaria, Ascochyta, Aspergillus, Botrytis, Cercospora, Colletotrichum, Diplodia, Erwinia, Erysiphe, Fusarium, Gaeumanomyces, Helminthosporium, Macrophomina, Magnaporthe, Mycosphaerella, Nectria, Peronospora, Phoma, Phymatotrichum, Phytophthora, Plasmopara, Podosphaera, Pseudomonas, Puccinia, Puthium, Pyrenophora, Pyricularia, Pythium, Rhizoctonia, Scerotium, Sclerotinia, Septoria, Thielaviopsis, Uncinula, Venturia, Verticillium, and Xanthomonas.

54. A part of the plant of claim 51 selected from the group consisting of seeds, seed pods, flowers, fruit, tubers, stems, cuttings, and pollen.

55. A product resulting from processing a plant of claim 51 or a part thereof.

56. A composition comprising the polynucleotide of claims 20 or 42 and a biologically acceptable carrier.

57. A method of making a defense signal peptide comprising expressing in a cell a polynucleotide of claims 20 or 42.

58. The method of claim 57 wherein the cell is a bacterial cell, a fungal cell, or an insect cell.

59. The method of claim 57 further comprising purifying the defense signal peptide.

60. A method of making a transgenic plant comprising introducing into a cell of a plant a polynucleotide of claims 20 or 42, thereby producing a transformed cell, and regenerating a transgenic plant from the transformed cell, wherein, compared to a control plant lacking the polynucleotide, the transgenic plant exhibits a characteristic selected from the group consisting of substantially improved yield of plant product, substantially reduced disease symptoms, and substantially enhanced resistance to disease infestation.

61. A method of making a plant that comprises a transgene comprising a sequence that encodes a defense signal peptide operably linked to a plant promoter, the method comprising sexually crossing a plant that comprises the transgene with a plant that lacks the transgene, thereby producing a plurality of progeny plants, and selecting a progeny plant comprising the transgene.

62. A method of making a plant that comprises a transgene comprising a sequence that encodes a defense signal peptide operably linked to a plant promoter, the method comprising asexually reproducing a plant that comprises the transgene, thereby producing a plurality of progeny plants, and selecting a progeny plant comprising the transgene.

63. A method of growing a plant comprising planting a seed that comprises a polynucleotide sequence of claim 20 or claim 42, and growing the seed to produce a plant, wherein, compared to a control plant lacking said polynucleotide sequence, the plant grown from the seed exhibits a characteristic selected from the group consisting of substantially improved yield of plant product, substantially reduced disease symptoms, and substantially enhanced resistance to disease infestation.

64. A method for detecting a plant cell comprising a polynucleotide of claim 20 or 42 in a biological sample, the method comprising contacting the biological sample with a probe that binds specifically to the polynucleotide, and detecting said binding.

65. The method of claim 64 wherein the probe is a PCR primer, the method comprising performing PCR on the sample and detecting said binding by detecting an amplification product diagnostic of the presence of the polynucleotide in the sample.

66. A kit for detecting a plant cell comprising a polynucleotide of claim 20 or 42 in a biological sample, the kit comprising one or more probes that bind specifically to the polynucleotide, and instructions for use.

67. A method for detecting a plant cell comprising a polynucleotide of claim 20 or 42 in a biological sample, the method comprising contacting the biological sample with a probe that binds specifically to the defense signal peptide, and detecting said binding.

68. The method of claim 67 wherein the probe is an antibody.

69. A kit for detecting a plant cell comprising a polynucleotide of claim 20 or 42 in a biological sample, the kit comprising one or more probes that bind specifically to the defense signal peptide, and instructions for use.

70. A plant cell comprising an insertion of a foreign promoter upstream of a coding sequence for a defense signal protein, wherein the foreign promoter is operably linked to the coding sequence for the defense signal protein and the plant is characterized by a substantially enhanced resistance to a disease compared to a control plant lacking the insertion of the foreign promoter.

71. A method of making a transgenic plant comprising (a) introducing into cells of a plant a polynucleotide that comprises a heterologous promoter, thereby producing a cell comprising an insertion of the heterologous promoter upstream of a coding sequence for a defense signal protein, wherein expression of the defense signal protein is controlled by the foreign promoter, and (b) regenerating a transgenic plant from said cell comprising the insertion.

72. A method of identifying a defense signal peptide comprising: (a) providing a plurality of candidate peptides having a length of at least 10 amino acids; (b) assaying said plurality of candidate peptides for defense signal peptide activity in an alkalinization assay; and (c) selecting a candidate peptide that has substantial defense signal peptide activity.

73. The method of claim 72 comprising providing the candidate peptides by chemically synthesizing said plurality of candidate peptides.

74. The method of claim 72 comprising administering the candidate peptide to a plant by applying a composition comprising the candidate peptide to the plant.

75. The method of claim 72 comprising administering the candidate peptide to a plant by expressing within a cell of the plant a polynucleotide that comprises a sequence that encodes the candidate peptide, thereby producing the candidate peptide within the cell of the plant.

76. A method of identifying a substance that enhances defense of a plant against a disease comprising (a) contacting an isolated AtPep1 receptor with a plurality of candidate substances; (b) selecting a candidate substance that has a detectable interaction with the isolated AtPep1 receptor; and (c) applying the selected candidate substance to a plant to determine whether the selected candidate substance enhances defense of the plant against a disease.

77. The method of claim 76 wherein the candidate substances are peptides.

78. The method of claim 76 wherein the candidate substances are non-peptide chemical compounds.

Fig. 1

(a) AtproPep1

mRNA - 499 bp

```
  1 actcacatat aaaaaacagc ttcactcctc tcaccaaaac taatcagatt aataaaagtt
 61 ttcctctgtc ttatcagatc tcaatggaga aatcagatag acgaagcgaa gaaagtcacc
121 tatggattcc tcttcagtgc ctcgaccaaa ccctcagagc tatcttgaaa tgccttggtc
181 tttttcatca agattctccg acaacgtcct ctcccggaac ttcgaaacag ccgaaggagg
241 aaaaagaaga cgttaccatg gaaaaggagg aggtcgttgt gacgagtaga gccacaaagg
301 tcaaggcaaa gcaaaggggg aaggagaaag ttagctcagg ccgtcctggc caacataatt
361 aggcacttta agttacattg tttagtctaa ttatttgcag tcgaaatgtg ttaatttaat
421 atcactgttt tacttttttа ttatatcaac aatctacaga caaacaaaat ttcattaagt
481 tcttgttcac tatacgagt
```

Protein - 92 aa

```
 1 MEKSDRRSEE SHLWIPLQCL DQTLRAILKC LGLFHQDSPT TSSPGTSKQP KEEKEDVTME
61 KEEVVVTSRA TKVKAKQRGK EKVSSGRPGQ HN
```

(b) AtproPep3

mRNA - 568 bp

```
  1 acattgagag atacaaagtt gtctctctga catataccta gctgctcgat aactcaccaa
 61 actattggat ttcaatggag aaattagata aacggaggga agaagaaact tatctatgga
121 ttccagttca gtttctcgac caagctctca tagctgtctt gaaatgtatt ggtcttcttt
181 gtcagccagc gaagaaaact gcgccgtctc cggtaacttt taaccagccg gaggaacaag
241 aggaagacta tggtgttgct ctgaaagacg atgatgtcgt tgtgttgctt agggacaaca
301 aggccaaatc aaagaaaagg gataaagaaa agcctagttc aggtcgtcct ggccaaacta
361 atagtgtacc caacgcggca atacaagttt ataaggagga ttaagaagtc aaaaattgag
421 tcgaaaaatc caagaggcca atgagtcagt cattgctcct ttttttttt tactcaaact
481 tctatgaaaa actcgtacgt agtttatttt ggtttcctca tttttcaaga cagcaaaatt
541 gaccagaatg tatatacttt tgaatcgg
```

Protein - 109 aa

```
 1 MEKLDKRREE ETYLWIPVQF LDQALIAVLK CIGLLCQPAK KTAPSPVTFN QPEEQEEDYG
61 VALKDDDVVV LLRDNKAKSK KRDKEKPSSG RPGQTNSVPN AAIQVYKED
```

(c) AtproPep4

mRNA - 486 bp

```
  1 atcaacctaa taacacacaa cactaaatct ctttcccaaa aaaagattaa gaagtcaacg
 61 atggagaatc tcagaaatgg agaagataac ggttctttga tcccatttac gttctttgat
121 caatcttcag tgacgattcc tctcttgaag tgttccggtc tcgaaagttc atcatcatca
181 tcttcttctt gcgatctttc gtcatcacac agcgaggaag atgagagtat cgatataaag
241 gaggaggaag aagaagaaga agaagatggc atgaccattg aaatcaaagc gagagggaag
301 aacaagacta agcctacgcc aagttcagga aaaggaggca aacacaatta gagttcattc
361 atataccgag gaaattaaac aaataaatgc atttgtataa aatacttaga gctataatac
421 agtggagttt ttttatagtc atttgtttcg aatatgaatt ggattaataa agatcgagtt
481 ttattt
```

Protein - 96 aa

```
 1 MENLRNGEDN GSLIPFTFFD QSSVTIPLLK CSGLESSSSS SSSCDLSSSH SEEDESIDIK
61 EEEEEEEEDG MTIEIKARGK NKTKPTPSSG KGGKHN
```

(d) AtproPep5

mRNA - 460 bp

```
  1 acttagctct cacgaagcag aattgaagaa aaacatggag agaggagttt cttattatct
 61 atggattcct tttaagttca tccaccaaac tttcggatct cttttactca agcttctcgg
121 tttgcgatct ccatctgatc atagttttcc ggaggatggg gaggaggaag ttaaggttgt
181 ggaagtgtcg tcgaggggtc ttcccgggaa aaagaatgta ctaaagaagt cgagagaaag
241 ttccggcaag ccgggaggca ccaacaagaa gccgttttag tttttcactt caactaataa
301 tatttgacgg agaaattctt ccttacattt tcatctattt agtgtaagat ctaagagaat
361 agtttcatat ttgtatcgta taattcctga agattgcaac tcctacgagt cctttatttt
421 ttttctttaa gacaataact aaagagagac gtgaatcata
```

Protein - 81 aa

```
 1 MERGVSYYLW IPFKFIHQTF GSLLLKLLGL RSPSDHSFPE DGEEEVKVVE VSSRGLPGKK
61 NVLKKSRESS GKPGGTNKKP F
```

(e) AtproPep2

mRNA - 412 bp

```
  1 acaaagaaaa tttgaggaga aagtcacata tagaggaact tagaagatag cgaagatgca
 61 gcaagagaga gatcacaaaa gagattgttg caagctcatg cctcaaactg tcaaggcttt
121 cttcaagtgt ctgagattca gacgttcttc ttcttcttct tcagacatgg tgaaagctag
181 agcaagaaat gaagagaaag aagaaccttc atctatcgaa acttcaacta ggagtctcaa
241 cgtaatgagg aaagggataa ggaaacaacc agttagctcg ggaaaacgag gtggagttaa
301 cgactacgac atgtaactag aatcttgatg tagaattgga taatcttgtt tggtagttac
361 tctacaacat actttctttg catctcatga atcatcatga tatattgata tt
```

Protein - 86 aa

```
 1 MQQERDHKRD CCKLMPQTVK AFFKCLRFRR SSSSSSDMVK ARARNEEKEE PSSIETSTRS
61 LNVMRKGIRK QPVSSGKRGG VNDYDM
```

(f) AtproPep6

mRNA - 397 bp

```
  1 ggtcaaacta gacacaacac ttaatgcatt gagcagaaga agaagaagaa gaagaattaa
 61 gaagagaaag aaaacaaaaa acatggaagt taatggagaa gaagagagaa gaagtagaag
121 agaagatgaa gaaaaagaag attactacta ctctcttctc aactctccat gttctgtttg
181 taacaaattt gttcaagcca tattgaagtg tcttggtctt gagtcatcat caataccacc
241 atcttcatca tcatcatcac catccttagt agaagaagaa gattcaggaa ctgaaactgt
301 tgaagaaaca ggatttatgg cgaggataac agcagtgtta agaaggagac caagaccacc
361 accttatagc tcaggacgac ctggtcaaaa caattga
```

Protein - 104 aa

```
 1 MEVNGEEERR SRREDEEKED YYYSLLNSPC SVCNKFVQAI LKCLGLESSS IPPSSSSSSP
61 SLVEEEDSGT ETVEETGFMA RITAVLRRRP RPPPYSSGRP GQNN
```

(g) AtproPep7

mRNA - 228 bp open reading frame (ORF)

```
  1 atgaatgttt ttttttttgt ttctgaatta ttgtcacatc tttcttttca atatgaaatt
 61 tctaatggaa aatgtgtata tgtaataatg ttggtgacga agatatcaca agaagtagag
121 gaagagacag aggtagtgaa tataccgagg agtgtggtgt cggggaacgt tgcagcgcga
181 aagggtaagc agcaaacgag ttccgggaag ggtggaggta ccaactag
```

Protein - 75 aa

```
 1 MNVFFFVSEL LSHLSFQYEI SNGKCVYVIM LVTKISQEVE EETEVVNIPR SVVSGNVAAR
61 KGKQQTSSGK GGGTN
```

(h) BnproPep1

Protein - 95 aa

```
 1 MEVNGEEKRS YRREDEEKEV YYPLLNSPCS AFHKTVQAIL KCLGLESSSI
51 SPSSSSSQDP GTETVQETGF MAMVARLTRR RPRPPYSSGQ PGQIN
```

(i) StproPep1

Protein - 116 aa

```
  1 MFYLQEGIKA ILKCLGFESS KLVHQASSSS SSSSMSDINK NEEEESEKQE
 51 QECVLFQEDG NKQGSDSTND NYKNDPPVEN DDEDPPQSET LILPTERRGR
101 PPSRPKVGSG PPPQNN
```

(j) PbproPep1

Protein - 121 aa

```
  1 MDKGSSTKEE IQGDVLQISH SPSIFVEAFN ALLRCLGLGT VDHQRITQES
 51 SSTSSSKQED DEKASEESPQ YPPPTRTSDP QADPPTDTSE DPSTDAAVSA
101 LARRTPPVSR GGGGQTNTTT S
```

(k) BeproPep1

Protein - 110 aa

```
  1 MEESSANDQA TTAHTKVVYF LEEALRAIFK CLGLETKPQD DPPSSQLEDA
 51 SSTTKQAVAD NSSTADPELA DPPSTTETSE VAATASIDLV MAVNAPPRPS
101 LTPGSGAQIN
```

(l) GmproPep1

Protein - 115 aa

```
  1 MEGSSPSIEE ERTATFYVYH PCYFLQQALR ALLKCVGIDE SENTMCSQAN
 51 KQEKSSLPQT PSADDPITNS PTHKSSPDAA DPPSTTNQTI IIASLMATRG
101 SRGSKISDGS GPQHN
```

(m) MsproPep1

Protein - 127 aa

```
  1 MEETTERLST KKEEKTMTFY VYHPCYCLEE IFKTFLRCFG IESTQTKEEE
 51 DSSTSLLKPH ACACASDSNV ALKDRYYSSS SNKKSSQEEG VADPPPSTST
101 QTINLSSMGR GGPRRTPLTQ GPPPQHN
```

(n) VvproPep1

Protein - 83 aa

```
 1 MNDDAEQRQR SHAGDDGQEG LDLGRLPPNP CGHGVDRSSW RPHGGGPFVF
51 CFCPCLAGEK VREKQKKGED GESVGRPGKK NEIL
```

(o) OsproPep1

Protein - 154 aa

```
  1 MDRVEEKEGN RFQEPASDRC EDNEDKEQDN SEESSSVDQR KEEEEEEKEG
 51 CEEATPAAAA AAAAPSFFAH PCSLLQYIAR VCACCLGLSD SFCDPKASSV
101 LVPEPEPAAA DPSQEGEEDM KSSEATTRVR AARLRPKPPG NPREGSGGNG
151 GHHH
```

(p) OsproPep2

Protein - 93 aa

```
 1 MAMSSSPASP PPSFLIGGAQ AQLLRHREEM LLVLPSPPSG RQLPSEEEEA
51 APCAVNGRST ILAAADDSKP TRPGAPAEGS GGNGGAIHTA ASS
```

(q) TaproPep1

Protein - 82 aa

```
 1 MGMADWFGGG GTRPSAAPAA SLNSSREEAG EAADIGTREI SKTTTGRGFY
51 MREVIMRVRA VRRPRPPTTP REGRGGGGGS HN
```

(r) TaproPep2

Protein - 75 aa

```
 1 MASPSPSFLL QLVRYVWSLP SQFMGATARA LPASREGAGG AIRPSFAAPA
51 PQRPGAPAEG AGGQGGIIHE ASPVP
```

(s) ZmproPep1

Protein - 142 aa

```
  1 MDERGEKEEE HGVVEEETAA VVLKEVEVEM EMVGGSEEAS AAPLLLAHPC
 51 SLLQLLLRAC AGCLVRLLHG HCSDGANDDP KAAADDDDAA PEAAAAAAAA
101 AGDGGDKAAT YLYMQEVWAV RRRPTTPGRP REGSGGNGGN HH
```

(t) HvproPep1

Protein - 93 aa

```
 1 MASSAPPAFL PQLVQPVSVL PDQPPSAPAE GTGGQVMVLN DASSLPLQLM
51 RTPPGEGAGG RIHRQLARPR PPGPPRQGHG GDGGAIHAIL LEL
```

Fig. 2

AtPep1 Receptor (At1g73080)

```
      1 tgaaagaccc aaacctaatg aatgttaacc actaattgac cattcaccaa
ccaattattt

     61 aatgaaatat ctttgttagt ttcgttattt agtattgtta acggtttctt
actcttttg

    121 actacatcag acggacgtaa aacgacatcg ttgtcgaata ttcaaaagat
tcacaatttg

    181 acaaagagaa acagagacga cttgtttcta aaaaaaccac gtgtgtctga
aaacggaaaa

    241 aaagaagact gaatgagaaa cggcgtgtaa aaagaaaacg cgttgaaggt
taggctctca

    301 caatcgttgg tatacagaga gaccaaacat ctcgtcataa aaaacggcaa
gaatcatcag

    361 ttactttata cccatcaatc aagtcttgtc ctttttctcc ttctctctct
catacgagct

    421 tctttcctgc tgatgaggct tgagctttaa attttcaatc ttgattgaga
ttctgcatgt

    481 ttctcgatct ttaaactcag ATGAAGAATC TTGGGGGGTT GTTCAAAATT
CTTCTGCTTT
                            M  K  N  L  G  G  L  F  K  I  L  L
L  F

    541 TCTTCTGTCT CTTTCTATCG ACCCACATAA TTTCCGTTTC TTGTTTAAAC
TCAGATGGGC
          F  C  L  F  L  S  T  H  I  I  S  V  S  C  L  N  S  D
G  L

    601 TAACTCTACT CTCTCTTCTG AAGCATTTGG ATAGAGTACC ACCACAAGTT
ACTTCGACAT
          T  L  L  S  L  L  K  H  L  D  R  V  P  P  Q  V  T  S
T  W

    661 GGAAAATAAA CGCATCTGAA GCAACTCCAT GTAACTGGTT CGGTATCACT
TGTGACGATT
          K  I  N  A  S  E  A  T  P  C  N  W  F  G  I  T  C  D
D  S

    721 CTAAGAATGT TGCGTCTCTC AACTTCACTC GTTCTAGGGT TTCAGGTCAA
TTGGGTCCGG
          K  N  V  A  S  L  N  F  T  R  S  R  V  S  G  Q  L  G
P  E
```

781 AAATTGGGGA GCTCAAAAGC TTGCAGATTT TGGATCTGAG TACTAACAAT
TTCTCCGGGA
I G E L K S L Q I L D L S T N N F S
G T

841 CTATACCTTC CACTTTAGGA AACTGTACCA AACTCGCTAC TCTAGATTTG
TCTGAAAATG
I P S T L G N C T K L A T L D L S E
N G

901 GATTCTCTGA TAAGATCCCA GATACTCTCG ATAGCTTGAA GAGGTTGGAG
GTGCTTTATC
F S D K I P D T L D S L K R L E V L
Y L

961 TTTACATAAA CTTCCTCACT GGTGAGTTAC CTGAATCCTT GTTTCGAATT
CCGAAGCTGC
Y I N F L T G E L P E S L F R I P K
L Q

1021 AGGTTTTATA TCTTGACTAT AACAATCTCA CCGGTCCGAT TCCTCAAAGT
ATTGGTGATG
V L Y L D Y N N L T G P I P Q S I G
D A

1081 CTAAGGAGCT TGTGGAGCTG AGTATGTATG CGAATCAGTT CTCTGGTAAC
ATCCCTGAGT
K E L V E L S M Y A N Q F S G N I P
E S

1141 CGATTGGGAA TAGCAGTAGT CTGCAGATTC TTTATTTGCA CAGGAACAAG
TTAGTTGGTT
I G N S S S L Q I L Y L H R N K L V
G S

```
    1201 CATTACCTGA AAGTCTCAAT CTTTTGGGGA ATCTCACTAC TCTGTTTGTT
GGTAACAACA
           L  P  E   S  L  N   L  L  G  N   L  T  T   L  F  V   G  N
N  S

    1261 GTCTACAAGG GCCGGTTCGT TTCGGATCAC CTAATTGCAA GAATTTGTTG
ACTTTAGATT
           L  Q  G   P  V  R   F  G  S  P   N  C  K   N  L  L   T  L
D  L

    1321 TGTCATACAA TGAATTCGAA GGCGGTGTTC CACCTGCATT GGGAAATTGC
AGTAGCCTTG
           S  Y  N   E  F  E   G  G  V  P   P  A  L   G  N  C   S  S
L  D

    1381 ACGCTTTAGT CATTGTGAGT GGTAACTTGT CAGGTACAAT CCCTTCCTCA
TTGGGTATGT
           A  L  V   I  V  S   G  N  L  S   G  T  I   P  S  S   L  G
M  L

    1441 TGAAGAATCT CACAATTCTT AACCTTTCCG AGAATCGTCT CTCTGGGAGT
ATCCCCGCAG
           K  N  L   T  I  L   N  L  S  E   N  R  L   S  G  S   I  P
A  E

    1501 AGCTCGGGAA CTGCAGTAGC TTGAACTTGT TGAAGCTGAA CGATAACCAG
CTTGTAGGCG
           L  G  N   C  S  S   L  N  L  L   K  L  N   D  N  Q   L  V
G  G

    1561 GAATACCGAG TGCATTAGGT AAGCTGAGGA AGCTAGAAAG TCTGGAGCTT
TTCGAAAACC
           I  P  S   A  L  G   K  L  R  K   L  E  S   L  E  L   F  E
N  R

    1621 GGTTTTCGGG TGAGATTCCT ATTGAGATAT GGAAGAGTCA GAGTCTTACG
CAGTTGCTAG
           F  S  G   E  I  P   I  E  I  W   K  S  Q   S  L  T   Q  L
L  V

    1681 TTTATCAAAA CAATCTCACT GGTGAACTAC CTGTGGAAAT GACTGAGATG
AAGAAGCTAA
           Y  Q  N   N  L  T   G  E  L  P   V  E  M   T  E  M   K  K
L  K

    1741 AGATCGCTAC GCTGTTCAAC AACAGCTTTT ATGGAGCGAT ACCACCGGGT
TTAGGTGTGA
           I  A  T   L  F  N   N  S  F  Y   G  A  I   P  P  G   L  G
V  N

    1801 ACAGCAGCTT AGAAGAGGTT GACTTTATTG GTAACAAACT TACAGGAGAG
ATACCGCCAA
           S  S  L   E  E  V   D  F  I  G   N  K  L   T  G  E   I  P
P  N
```

```
    1861 ATCTATGCCA TGGAAGGAAG TTGAGAATAC TCAACTTGGG TTCTAATCTG
CTTCATGGTA
           L  C  H  G  R  K  L  R  I  L  N  L  G  S  N  L  L  H
G  T

    1921 CAATACCAGC TTCTATTGGT CACTGTAAGA CCATCAGGAG ATTCATCCTT
AGAGAAAATA
           I  P  A  S  I  G  H  C  K  T  I  R  R  F  I  L  R  E
N  N

    1981 ACCTTTCAGG TCTTCTTCCT GAGTTTTCTC AGGATCATAG TCTTTCTTTT
CTTGATTTCA
           L  S  G  L  L  P  E  F  S  Q  D  H  S  L  S  F  L  D
F  N

    2041 ATAGCAACAA CTTCGAAGGA CCAATCCCGG GCAGCCTCGG AAGCTGTAAG
AATCTCTCGA
           S  N  N  F  E  G  P  I  P  G  S  L  G  S  C  K  N  L
S  S

    2101 GTATTAACCT ATCTCGAAAC AGATTCACGG GGCAGATACC TCCACAACTT
GGGAATCTAC
           I  N  L  S  R  N  R  F  T  G  Q  I  P  P  Q  L  G  N
L  Q

    2161 AAAACCTTGG TTACATGAAT CTTTCTCGTA ATCTTCTTGA AGGGTCTCTA
CCAGCTCAGC
           N  L  G  Y  M  N  L  S  R  N  L  L  E  G  S  L  P  A
Q  L

    2221 TATCTAACTG TGTGAGTTTA GAGCGTTTTG ATGTTGGCTT CAACTCATTA
AACGGTTCAG
           S  N  C  V  S  L  E  R  F  D  V  G  F  N  S  L  N  G
S  V

    2281 TTCCTTCAAA CTTTAGTAAC TGGAAAGGCT TGACGACTTT AGTTCTCAGC
GAGAACCGGT
           P  S  N  F  S  N  W  K  G  L  T  T  L  V  L  S  E  N
R  F

    2341 TTTCAGGAGG TATTCCACAG TTCTTGCCTG AGCTTAAGAA GCTGTCAACT
CTGCAGATTG
           S  G  G  I  P  Q  F  L  P  E  L  K  K  L  S  T  L  Q
I  A
```

2401 CTAGAAATGC TTTTGGTGGT GAGATTCCTT CGTCGATTGG GTTGATAGAG
GATCTGATCT
R N A F G G E I P S S I G L I E D L
I Y

2461 ATGACTTGGA CCTTAGTGGA AACGGATTGA CAGGTGAAAT TCCAGCCAAG
TTGGGAGATC
D L D L S G N G L T G E I P A K L G
D L

2521 TCATCAAGTT AACAAGACTC AACATATCTA ACAACAATTT GACAGGATCT
TTATCGGTTC
I K L T R L N I S N N N L T G S L S
V L

2581 TCAAAGGTCT TACCTCATTG CTACATGTTG ATGTCTCCAA CAATCAGTTC
ACAGGTCCAA
K G L T S L L H V D V S N N Q F T G
P I

2641 TACCAGATAA CTTGGAGGGT CAGTTGTTAT CTGAGCCGTC GTCGTTTTCA
GGAAATCCAA
P D N L E G Q L L S E P S S F S G N
P N

2701 ACCTCTGCAT TCCACATTCC TTCTCTGCTA GCAACAATAG CCGCAGCGCG
TTAAAGTACT
L C I P H S F S A S N N S R S A L K
Y C

2761 GTAAAGATCA ATCTAAAAGC AGGAAGAGTG GCCTTAGCAC CTGGCAAATC
GTGCTAATAG
K D Q S K S R K S G L S T W Q I V L
I A

2821 CGGTCTTATC GTCTTTATTA GTCTTGGTTG TGGTCCTTGC TCTTGTTTTC
ATTTGCCTAC
V L S S L L V L V V V L A L V F I C
L R

2881 GTCGTCGCAA AGGAAGACCA GAGAAAGATG CTTATGTCTT CACTCAGGAG
GAAGGCCCAT
R R K G R P E K D A Y V F T Q E E G
P S

2941 CTTTGTTGTT GAACAAAGTT CTTGCAGCAA CTGACAATCT AAATGAAAAG
TACACCATTG
L L L N K V L A A T D N L N E K Y T
I G

3001 GAAGAGGAGC TCATGGAATT GTGTACAGAG CTTCTTTAGG CTCCGGAAAG
GTCTACGCTG
R G A H G I V Y R A S L G S G K V Y
A V

3061 TGAAGAGACT TGTATTCGCG TCTCACATCC GCGCTAACCA GAGTATGATG
AGGGAGATTG
K R L V F A S H I R A N Q S M M R E
I D

3121 ATACAATCGG TAAAGTCAGG CACAGGAATC TGATTAAGTT AGAAGGGTTT
TGGCTGAGGA
T I G K V R H R N L I K L E G F W L
R K

3181 AAGACGACGG TTTAATGCTG TATAGATACA TGCCAAAAGG AAGTCTTTAC
GACGTTCTCC
D D G L M L Y R Y M P K G S L Y D V
L H

3241 ACGGTGTTAG CCCGAAAGAA AATGTGCTAG ACTGGTCTGC ACGGTACAAT
GTAGCACTTG
G V S P K E N V L D W S A R Y N V A
L G

3301 GTGTCGCTCA TGGACTAGCC TATCTACACT ATGACTGCCA TCCCCCGATT
GTTCACCGTG
V A H G L A Y L H Y D C H P P I V H
R D

3361 ACATCAAACC AGAGAACATA CTCATGGACT CAGATTTGGA GCCTCACATT
GGGGATTTCG
I K P E N I L M D S D L E P H I G D
F G

3421 GTTTGGCTCG CCTTCTTGAT GACTCAACGG TTTCAACTGC AACTGTTACA
GGCACCACCG
L A R L L D D S T V S T A T V T G T
T G

3481 GCTACATTGC ACCAGgtaat gcatcttctc attatacata gtggacttgg
tataatctgg
Y I A P E

3541 tttagtgttc aaaccgagtt agttaccggt taaaaaagtc tgttaggaag
atactctgtt

3601 tcttattagc taatttcaca attaaactgc agAAAACGCT TTCAAAACCG
TGAGGGGAAG
N A F K T V R
G R

3661 AGAATCAGAC GTTTACAGTT ATGGAGTCGT GTTACTTGAG CTGGTTACGA
GGAAGAGAGC
E S D V Y S Y G V V L L E L V T R K
R A

3721 GGTGGACAAA TCTTTCCCGG AAAGTACAGA TATAGTAAGC TGGGTGAGAT
CTGCCTTGAG
V D K S F P E S T D I V S W V R S A
L S

3781 CAGCAGCAAC AACAATGTGG AGGATATGGT AACAACAATC GTCGATCCGA
TTCTCGTGGA
S S N N N V E D M V T T I V D P I L
V D

3841 CGAGCTTCTG GATTCGAGTC TTAGGGAGCA GGTGATGCAA GTGACGGAAC
TGGCACTGAG
E L L D S S L R E Q V M Q V T E L A
L S

3901 TTGTACACAG CAAGATCCGG CAATGAGACC AACGATGAGA GATGCGGTGA
AACTGTTGGA
C T Q Q D P A M R P T M R D A V K L
L E

3961 AGATGTGAAA CATCTGGCAA GAAGCTGCTC CTCTGATTCA GTTCGGTAAt
ctcgttactt
D V K H L A R S C S S D S V R *

4021 tgtgcagagc agaaggagga aactaaagga ctgttatcag tggtaacgta
actgggctta

4081 ccggtaatgt aactgggcca ataatgtaaa atatggctta ttgaaggccc
aaatatgacg

4141 gccctttaat tgtaaccgtg tttgtttgtg aaataaaatc tcgtttatca
aatttctgtt

4201 tcctatttta tttttaaaaa aagtgttggg aaaattttcg tttggccgga
gatgaagata

4261 ggggcggttc aggagagttg gcttatacgg actccggtac ttaggccggc
ggttcaaggc

4321 tcaattcgcc ggaatggaaa gccgcagctt gagtttcctc tctcgaacag
tgagcttaag

4381 ctcttctttc ttccattgaa ttttttttt gggaacgcat tatcttgtgc
acacttgtag

4441 taacttggtc tatatcaaat tgaggttgaa gatgaaagtt cagttatttc
ctgtgatttg

4501 caatttcct

_____ Signal sequense Residues 1-24

C C Cysteine paire Residues 64 and 71; and Residues 836 and 854

Leucine-rich repeat Residues 76-827

_____ Transmembrane domain Residues 870-892

Kinase domain Residures 927-1208

Intron between 1099 and 1100.

Fig. 3

AtPep1 Receptor (At1g17750)

```
      1 attctagtgt agacgacaga taccagagat cttgattaaa ttccaatata
taatgtttat

     61 gaaagattta aatctaacaa agtaacagta ggatgaaatg tcaatagaaa
attagcgtcc

    121 aaagaagctt tctcttgaat aagctaagaa aacaaaacgt ggaaaatgga
atatttaaaa

    181 ccacgaaaca gtctccgagt agtaatggaa atagcgagaa agaaacacga
aacagcgttg

    241 aaggtcgcca attcgtaaag ttaggttcac aatctctgac gaagagttaa
ccaaaaagcg

    301 cgctcttttt ctctctcacc aaactcatcg atcgtttctt aaataatgca
atctgttcct

    361 gtcactaaat ccagataccc tttcaaatcc aaaagctctc tctttttttt
ttccgcctct

    421 cattctgggt tcaagggttg ttgagtgagg ttactacgta cgagtgtttc
atatttcagt

    481 ctcttgagct ctaatctcaa ATGAGGAATC TTGGGTTACT CGAAATTACT
CTGCTTTGCT
                            M  R  N  L  G  L  L  E  I  T  L  L
C  S

    541 CTCTCTTTGT CTATTTCCGT ATAGATTCTG TCTCTAGTTT AAACTCAGAT
GGTTTGGCTT
          L  F  V  Y  F  R  I  D  S  V  S  S  L  N  S  D  G  L
A  L

    601 TACTCTCGCT TCTCAAGCAC TTTGATAAAG TCCCACTTGA AGTAGCTTCG
ACGTGGAAGG
          L  S  L  L  K  H  F  D  K  V  P  L  E  V  A  S  T  W
K  E

    661 AGAACACATC TGAAACCACT CCATGTAATA ATAACTGGTT TGGTGTCATT
TGTGATCTTT
          N  T  S  E  T  T  P  C  N  N  N  W  F  G  V  I  C  D
L  S

    721 CTGGTAATGT CGTCGAGACC CTTAATTTGT CTGCTTCTGG GCTTTCAGGC
CAATTAGGTT
          G  N  V  V  E  T  L  N  L  S  A  S  G  L  S  G  Q  L
G  S
```

781 CTGAAATTGG GGAGCTTAAG AGCTTGGTCA CATTGGATCT CAGTCTTAAC
AGTTTCTCTG
E I G E L K S L V T L D L S L N S F
S G

841 GTTTATTGCC TTCCACTTTA GGAAACTGTA CTTCACTTGA GTATTTGGAT
TTGTCTAACA
L L P S T L G N C T S L E Y L D L S
N N

901 ATGATTTTTC TGGAGAAGTT CCTGATATTT TTGGTAGCTT GCAGAATTTG
ACGTTTCTGT
D F S G E V P D I F G S L Q N L T F
L Y

961 ATCTTGATCG CAATAATCTT AGTGGTCTCA TTCCTGCAAG TGTTGGTGGG
TTGATAGAGC
L D R N N L S G L I P A S V G G L I
E L

1021 TCGTAGATCT GAGGATGTCA TATAATAACT TGTCTGGTAC CATTCCAGAG
TTGCTTGGGA
V D L R M S Y N N L S G T I P E L L
G N

1081 ACTGTAGTAA GCTGGAATAT CTGGCTTTGA ACAACAACAA GTTAAATGGT
TCTTTGCCAG
C S K L E Y L A L N N N K L N G S L
P A

1141 CAAGTCTCTA TCTACTCGAG AATCTTGGTG AGCTATTTGT CAGTAACAAC
AGCCTTGGAG
S L Y L L E N L G E L F V S N N S L
G G

1201 GGAGGCTTCA TTTTGGTTCT AGCAACTGCA AGAAATTGGT TTCTTTAGAT
CTCTCGTTCA
R L H F G S S N C K K L V S L D L S
F N

1261 ATGATTTCCA AGGCGGTGTT CCACCTGAGA TAGGCAACTG CAGTAGCCTT
CACTCTTTAG
D F Q G G V P P E I G N C S S L H S
L V

1321 TCATGGTGAA ATGCAACTTG ACAGGTACAA TCCCATCATC AATGGGTATG
TTGAGAAAGG
M V K C N L T G T I P S S M G M L R
K V

1381 TTTCGGTTAT TGACCTTTCC GATAATCGTC TCTCGGGGAA TATCCCTCAA
GAGCTTGGGA
S V I D L S D N R L S G N I P Q E L
G N

1441 ACTGCAGCAG CTTGGAAACC TTGAAGCTGA ACGACAACCA GCTCCAAGGC
GAGATACCAC
C S S L E T L K L N D N Q L Q G E I
P P

1501 CTGCATTGAG TAAGCTAAAG AAGCTACAAA GCCTGGAGCT TTTTTTTAAT
AAGCTGTCCG
A L S K L K K L Q S L E L F F N K L
S G

1561 GTGAGATTCC TATTGGCATA TGGAAGATTC AGAGTCTGAC ACAGATGCTC
GTTTATAACA
E I P I G I W K I Q S L T Q M L V Y
N N

1621 ACACTCTCAC CGGGGAACTA CCAGTTGAAG TAACTCAGCT GAAGCACCTT
AAGAAGCTTA
T L T G E L P V E V T Q L K H L K K
L T

1681 CACTGTTTAA CAACGGCTTT TATGGAGATA TACCAATGAG TTTAGGCCTG
AATCGAAGCT
L F N N G F Y G D I P M S L G L N R
S L

1741 TAGAGGAGGT GGACCTTCTT GGTAACCGTT TTACAGGGGA GATACCACCC
CATCTCTGCC
E E V D L L G N R F T G E I P P H L
C H

1801 ATGGACAGAA GTTGAGATTG TTCATCTTGG GTTCTAATCA GCTTCATGGT
AAGATACCAG
G Q K L R L F I L G S N Q L H G K I
P A

1861 CGTCTATTCG TCAGTGTAAG ACCCTTGAGC GAGTCAGACT TGAAGATAAC
AAACTTTCAG
S I R Q C K T L E R V R L E D N K L
S G

1921 GTGTTCTTCC GGAATTCCCT GAGAGTCTTA GTCTTTCCTA TGTGAACCTC
GGAAGCAATA
V L P E F P E S L S L S Y V N L G S
N S

1981 GCTTTGAAGG ATCCATCCCG CGCAGCTTGG GAAGCTGTAA AAATCTCTTG
ACTATTGACC
F E G S I P R S L G S C K N L L T I
D L

2041 TTTCTCAAAA CAAACTCACG GGTCTGATAC CTCCAGAACT GGGAAATCTG
CAAAGCCTTG
S Q N K L T G L I P P E L G N L Q S
L G

2101 GACTGTTGAA CCTTTCACAT AATTATCTGG AAGGTCCTCT GCCATCCCAG
CTATCAGGCT
L L N L S H N Y L E G P L P S Q L S
G C

2161 GTGCGAGACT TCTGTATTTT GATGTTGGAT CCAACTCATT GAACGGTTCT
ATTCCATCAA
A R L L Y F D V G S N S L N G S I P
S S

2221 GCTTCAGAAG CTGGAAAAGC TTGTCCACTT TAGTTCTCAG TGACAATAAT
TTTCTAGGAG
F R S W K S L S T L V L S D N N F L
G A

2281 CTATTCCACA GTTCTTGGCA GAGCTTGACC GACTCTCAGA TCTGCGGATA
GCTCGAAATG
I P Q F L A E L D R L S D L R I A R
N A

2341 CTTTTGGAGG TAAGATTCCT TCCTCGGTTG GCTTGTTGAA GAGTCTACGC
TATGGCTTAG
F G G K I P S S V G L L K S L R Y G
L D

2401 ACCTCAGTGC GAACGTATTT ACGGGTGAGA TTCCAACCAC ACTGGGGGCT
CTTATCAATC
L S A N V F T G E I P T T L G A L I
N L

2461 TTGAACGTCT CAACATATCC AACAACAAGT TGACAGGGCC TTTATCGGTT
CTTCAAAGTC
E R L N I S N N K L T G P L S V L Q
S L

2521 TTAAGTCATT GAATCAAGTT GACGTCTCGT ATAATCAGTT CACGGGTCCA
ATACCCGTAA
K S L N Q V D V S Y N Q F T G P I P
V N

2581 ATCTGTTATC AAATTCTTCA AAGTTTTCTG GAAATCCAGA CCTCTGCATT
CAAGCTTCTT
L L S N S S K F S G N P D L C I Q A
S Y

2641 ACTCAGTGAG TGCCATAATC CGCAAAGAGT TTAAATCTTG CAAAGGTCAA
GTCAAACTTA
S V S A I I R K E F K S C K G Q V K
L S

2701 GCACGTGGAA GATCGCCCTT ATAGCAGCTG GGTCCTCACT ATCCGTATTG
GCTTTGCTTT
T W K I A L I A A G S S L S V L A L
L F

2761 TTGCTCTCTT TTTGGTTTTA TGCCGGTGCA AAAGAGGAAC CAAGACAGAA
GATGCTAATA
A L F L V L C R C K R G T K T E D A
N I

2821 TCCTCGCAGA GGAAGGTCTG TCCTTGTTGC TGAACAAAGT TCTAGCAGCC
ACTGACAATC
L A E E G L S L L L N K V L A A T D
N L

2881 TAGATGACAA GTACATCATT GGAAGAGGAG CTCATGGAGT TGTTTACAGA
GCTTCTTTAG
D D K Y I I G R G A H G V V Y R A S
L G

2941 GATCAGGCGA AGAATACGCC GTGAAGAAAC TCATCTTTGC GGAACACATT
CGCGCAAACC
S G E E Y A V K K L I F A E H I R A
N Q

3001 AAAATATGAA GCGGGAGATC GAAACAATCG GGCTAGTCAG GCACAGAAAT
CTCATTCGGT
N M K R E I E T I G L V R H R N L I
R L

3061 TAGAAAGATT TTGGATGAGG AAAGAAGATG GCTTAATGCT GTATCAGTAC
ATGCCCAATG
E R F W M R K E D G L M L Y Q Y M P
N G

3121 GAAGCCTACA CGACGTTTTG CACAGAGGTA ATCAAGGAGA AGCAGTTCTT
GACTGGTCTG
S L H D V L H R G N Q G E A V L D W
S A

3181 CACGGTTCAA CATAGCCCTT GGGATTTCAC ATGGACTGGC GTATTTACAT
CATGATTGTC
R F N I A L G I S H G L A Y L H H D
C H

3241 ATCCACCAAT AATTCACCGC GACATCAAAC CAGAGAACAT ACTCATGGAC
TCGGATATGG
P P I I H R D I K P E N I L M D S D
M E

3301 AGCCTCACAT TGGAGATTTC GGATTGGCTC GGATTCTAGA TGACTCAACA
GTTTCAACGG
P H I G D F G L A R I L D D S T V S
T A

3361 CCACTGTTAC TGGCACAACT GGGTACATTG CACCAGgtat atatacttct
caacataaca
T V T G T T G Y I A P E

3421 cgttcgtatt ttgttcaccg ttaccttatt catgctctga tgaccatatt tctatcaaac

3481 agAAAATGCG TACAAGACGG TGAGGAGCAA GGAATCAGAT GTTTACAGTT ATGGAGTTGT

N A Y K T V R S K E S D V Y S Y G V V

3541 TTTGCTCGAG CTGGTAACAG GAAAGAGAGC ACTGGACAGA TCTTTCCCGG AAGATATCAA

L L E L V T G K R A L D R S F P E D I N

3601 CATTGTGAGC TGGGTCAGAT CTGTATTAAG CAGCTACGAG GATGAAGACG ATACTGCTGG

I V S W V R S V L S S Y E D E D D T A G

3661 TCCAATCGTT GATCCAAAAC TTGTGGATGA GCTTCTGGAT ACGAAGCTCA GGGAACAAGC

P I V D P K L V D E L L D T K L R E Q A

3721 AATCCAAGTC ACAGACTTGG CTCTTAGATG TACAGACAAG AGGCCGGAGA ACAGACCATC

I Q V T D L A L R C T D K R P E N R P S

3781 GATGAGAGAT GTGGTGAAAG ATTTGACTGA TTTGGAAAGT TTTGTAAGAA GCACTTCGGG

M R D V V K D L T D L E S F V R S T S G

3841 TTCAGTTCAC TAGtttcata agttgcaggt ttatatagtg tactgttctt tgaaaccact

S V H *

3901 aatataattg taactaccat tgaataccgt gaagatttga gaaacatata tatccaaaga

3961 gacttaattt tattgcataa gttggaattg ttgggggaaa taaattacaa gtattaagag

4021 aggcttcaca aatttgccaa tctcagtata ttttagccag tgctaccaga aagctcgcca

4081 gagaaggcac taccatacga tgtggaaccg gttggtgtct gtgatgatgt tgcgatccct

4141 cgagctttcc gcatctgcct tagtctgaaa acatctgtgg gcaaaggctg gctttgtaca

4201 gttgctgtat catcttcttc ttgccctgtt aacagagaga tagaataaga taagtaacta

4261 tcatttaacg gttgttctaa gggatagtag aagttatatg catctgcagc taagaagtgc

4321 caaacatttg agtgatgaca tagtcatgag cca

_____ Signal sequense Residues 1-26

C C Cysteine paire Residues 62 and 71; and residues 709 and 727

Leucine-rich repeat Residues 99-697

_____ Transmembrane domain Residues 738-760

Kinase domain Residues 793-1079

Intron between residues 966 and 967.

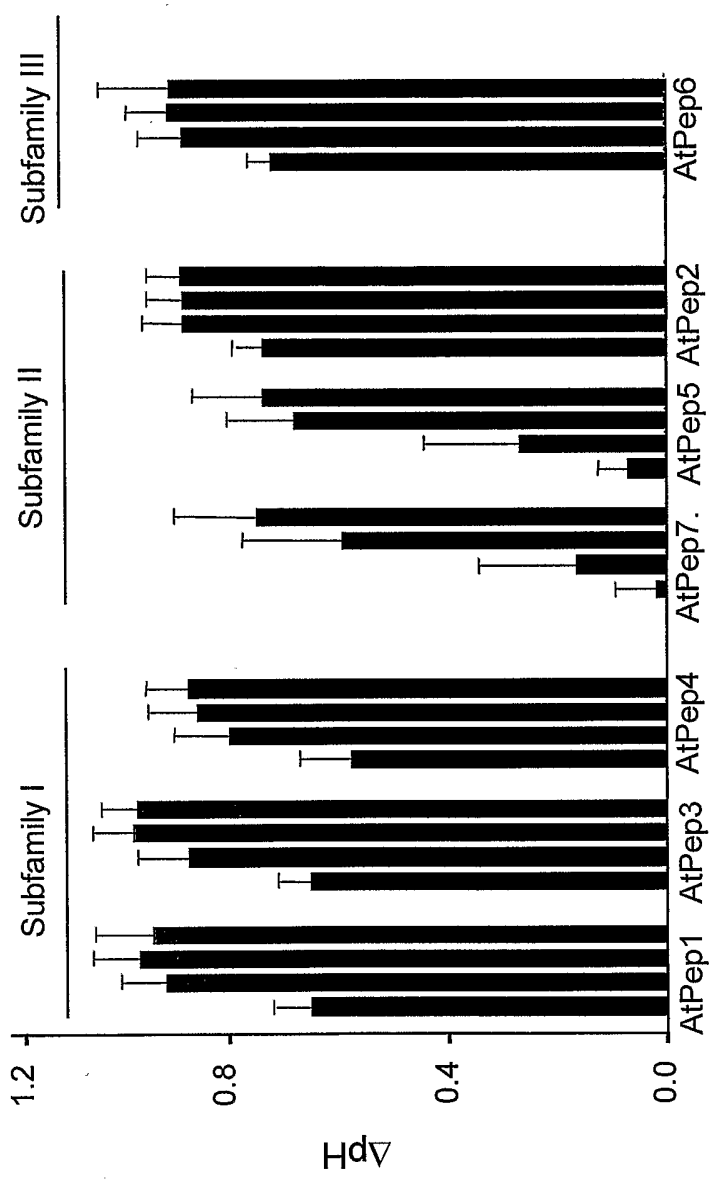
Subfamily I
Subfamily II
Subfamily III
AtPep1
AtPep3
AtPep4
AtPep7.
AtPep5
AtPep2
AtPep6
1.2
0.8
0.4
0.0
ΔpH

Fig. 4

Fig. 5

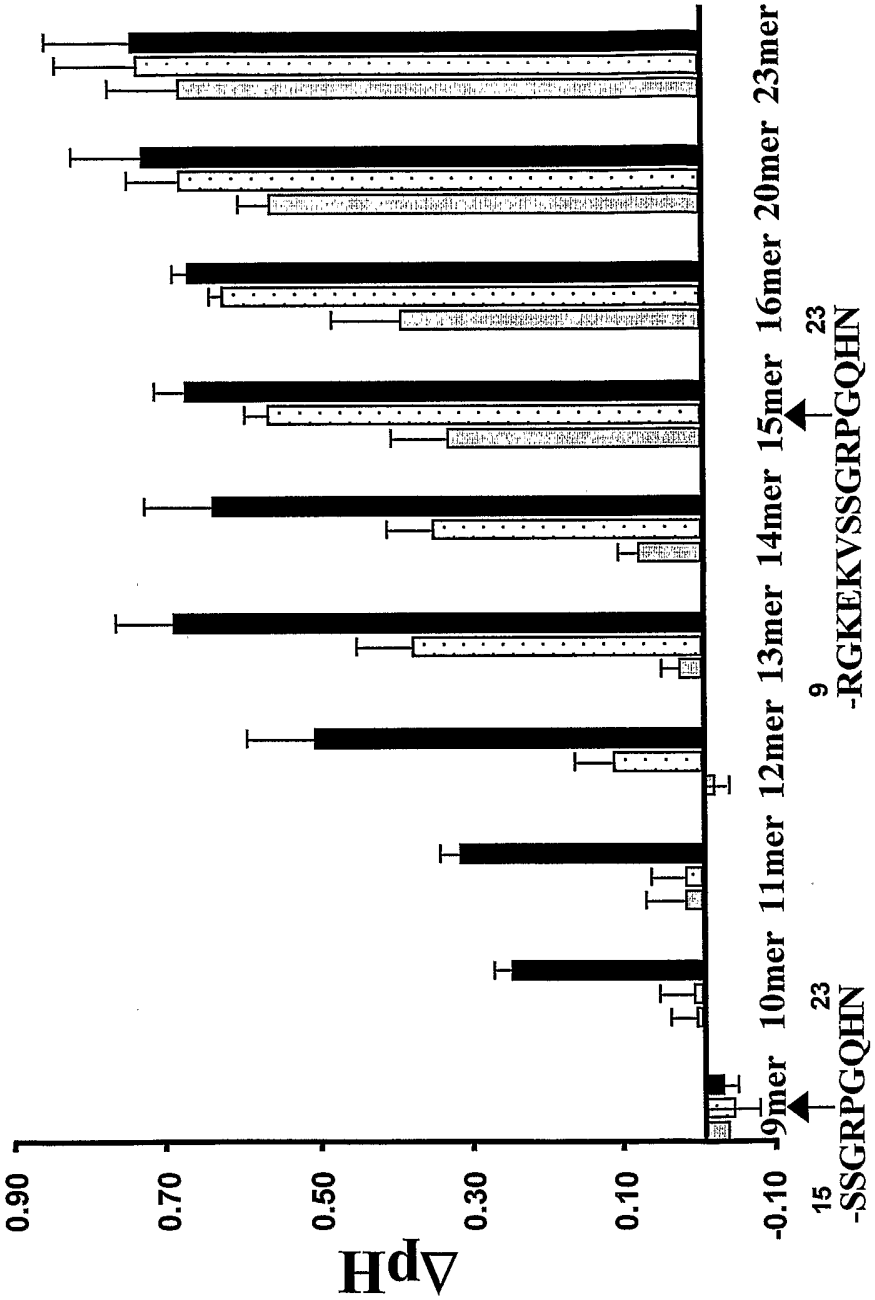
ΔpH
0.90
0.70
0.50
0.30
0.10
-0.10
9mer
10mer
11mer
12mer
13mer
14mer
15mer
16mer
20mer
23mer
15
-SSGRPGQHN
23
9
-RGKEKVSSGRPGQHN
23

Fig. 6