

US 20040096934A1

# (19) United States (12) Patent Application Publication (10) Pub. No.: US 2004/0096934 A1 Freyssinet et al.

# May 20, 2004 (43) **Pub. Date:**

## (54) PEPSIN-SENSITIVE MODIFIED BACILLUS THURINGIENSIS INSECTICIDAL TOXIN

(76) Inventors: Georges Freyssinet, St. Cyr au Mont d'Or (FR); Cecile Rang, St. Georges d'Orques (FR); Roger Frutos, St. Georges d'Orques (FR)

> Correspondence Address: **BAKER & BOTTS 30 ROCKEFELLER PLAZA** NEW YORK, NY 10112

- (21) Appl. No.: 10/665,460
- (22) Filed: Sep. 19, 2003

#### **Related U.S. Application Data**

(63) Continuation of application No. PCT/FR02/00772, filed on Mar. 4, 2002.

#### (30) **Foreign Application Priority Data**

Mar. 19, 2001 (FR)...... 01/03691

## **Publication Classification**

(51) Int. Cl.<sup>7</sup> ..... C07K 14/325; C12P 21/06 (52) 

#### (57) ABSTRACT

The invention relates to the degradation of Bacillus thuringiensis Cry proteins in the digestive tracts of mammals and concerns Bacillus thuringiensis Cry proteins having a peptide sequence that has been modified in such a way as to make said proteins sensitive to the specific enzymes in the digestive tracts of mammals, in particular pepsins. According to the invention, the Cry proteins are modified by inserting pepsin cleavage sites in the peptide sequence thereof. The invention also relates to transformed plants expressing said modified Cry proteins.

atgaatcgaa ataatcaaaa tgaatatgaa attattgaag cccccattg tgggtqtcca tcagaagaag aattaaggta tcctttggca aqtgaaccaa atgcagcgtt acaaaatatg aactataaag aatacttaca aatgacagaa gaggaataca ctgaatctta tataaatcct agtttatcta ttagtggtag agaagcatta cagactgcgc ttactgttat taggagaata ctcggggctt taggt**tta**cc gttttctgga caaata**tta**a gtttttatca attcctttta aatacactgt ttccattaaa tgaaacagct atatttgaag ctttcatgcg acagttagag gaacttttaa atcaacaaat aacagaattt gcaagaaatc aggcacttgc aagattgcaa ggattaggag aatcttttaa t**tta**tatcaa cqttcccttc aaaatttttt ggctgaacga aatgaaacac gaaatttaag tttattacqt gctcaattta tagctttaga acttgaattt ttaaatgcta ttecattgtt tgcattaaat ggacagcagt taccattact gtcattatat gcacaagctt taaatttaca tttgttatta ttaaaagaag catctctttt tggagaagga tttggattca cacaggggga aatttccaca tattatgaac gtcaattgga actaaccgct aagtacacta attactgtga aact**ttt**tat aatacaggtt tagaacgttt aagaggaaca aatactgaaa gttttttaag atatcatcaa ttccgtagag aaatgacttt attattatta gaattattag cgctatttcc atattatgaa ttacgacttt atccaacggg atcaaaccca cagettacae gtgagttata tacagaaceg attttatta atecaecage taatttagga ctttgccgac gttttggtac taatccctat aatacttttt ctgagetega aaatgeette attegeeeae cacatettt tgaaaggetg aatagettaa caateageag taategattt eeattateat etaattttat ggaatatttt tcaggacata cgttacgccg tagttatctg aacgaatcag cattacaaga agaaagttat ggcctaatta caaccacaag agcaacaatt aatcooggat tagaaggaac aaacogcata gagtcaacgg cattagaatt

tcgttctgca ttgataggta tatatggctt aaatagagct tctttttac caggaggett gtttaatggt acgaettete etgetaatgg aggatgtaga gaactetatg aaacaaatga agaattacca ccagaagaaa gtaccggaag ttcaacccat agactatete atttaacett ttttagettt caaactaate aggetggate tatagetaat geaggaagtt tacetaetta tttatttace cgtcgtgaat tagaacttaa taatacgatt accccaaata gaattacaca attaccattg ttaaaggcat ctgcaccttt atcgggtact acgttattaa aaggtccagg atttacagga gggggtatac tecgaagaac aactaatgge acatttggaa cgttaagatt aacgttaaat tcaccattaa cacaacaata tcgcctaaga ttacgttttg cctcaacagg aaatttcagt ataaggttac tccgtggagg gttatctatc ggtgaattaa gattagggag cacaatgaac agagggcagg aactaactta cgaateettt tteacaagag agtttaetae tactggtccg ttcaatccgc cttttacatt tacacaaget caagagatte taacattaaa tgcagaaggt ttaagcaccg gtggtgaata ttatatagaa agaattgaaa tt**tta**cct**tt a**aatccggca cgagaagcgg aagaggaatt agaagcggcg aagaaagcg

MNRNNQNEYEIIEAPHCGCPSEEELRYPLASEPNAALQNMNYKEYLQMTEEEYTESYINPSLSISGREALQTALTLLGRILGALGLPFSGQILSFYQFLLNTLFPLNETAIFEAFMRQLEELLNQQITEFARNQALARLQGLGESFNLYQRSLQNFLAERNETRNLSLLRAQFIALELEFLNAIPLFALNGQQLPLLSLYAQALNLHLLLLKEASLFGEGFGFTQGEISTYYERQLELTAKYTNYCETFYNTGLERLRGTNTESFLRYHQFRREMTLLLLELLALFPYYELRLYPTGSNPQLTRELYTEPILFNPPANLGLCRRFGTNPYNTFSELENAFIRPPHLFERLNSLTISSNRFPLSSNFMEYFSGHTLRRSYLNESALQEESYGLITTRATINPGLEGTNRIESTALEFRSALIGIYGLNRASFLPGGLFNGTTSPANGGCRELYETNEELPPEESTGSSTHRLSHLTFSFQTNQAGSIANAGSLPTYLFTRRELELNNTITPNRITQLPLLKASAPLSGTTLLKGPGFTGGGILRRTTNGTFGTLRLTLNSPLTQQYRLRLRFASTGNFSIRLLRGGLSIGELRLGSTMNRGQELTYESFFTREFTTGPFNPFTFTQAQEILTLNAEGLSTGGEYYIERIEILPLNPAREAEELEAAKKAKKAKKA

US 2004/0096934 A1

Oligonucléotide n°53 : tgaatatgaaattattgaagccccccattg Oligonucléotide n°54 : tgggtgtccatcagaagaagaagaattaaggtatcctttggca Oligonucléotide n°55 : tcctttggcaagtgaaccaaatgcagc Oligonucléotide n°56 : gaactataaagaatacttacaaatg Oligonucléotide n°57 : caaatgacagaagaggaatacactga Oligonucléotide n°58 : tacactgaatcttatataaa Oligonucléotide n°59 : tattagtggtagagaagcattacagactgcgcttac Oligonucléotide n°60 : cagactgcgcttactgttattaggagaatactcggg Oligonucléotide n°61 : gggctttaggtttaccgttttctgg Oligonucléotide n°62 : ttctggacaaatattaagtttttatcaa Oligonucléotide n°63 : cttttaaatacactgtttccattaaatgaaacagctatat Oligonucléotide n°64 : acagctatatttgaagctttcatg Oligonucléotide n°65 : ctttcatgcgacagttagaggaactt Oligonucléotide n°66 : gaggaacttttaaatcaacaaataac Oligonucléotide n°67 : ggattaggagaatcttttaat Oligonucléotide n°68 : tcttttaatttatatcaacgttc Oligonucléotide n°69 : ccttcaaaatt**tt**ttggctga Oligonucléotide n°70 : ttggctgaacgaaatga Oligonucléotide n°71 : cgaaatgaaacacgaaatttaag Oligonucléotide n°72 : acacgaaatttaagtttattacgtgctcaatttatag Oligonucléotide n°73 : gctcaatttatagctttagaacttgaatttttaaatgctattccattg Oligonucléotide n°74 : ccattgtttgcattaaatggacagcag Oligonucléotide n°75 : aatggacagcagttaccattactgtca Oligonucléotide n°76 : ccattactgtcattatatgcacaaget Oligonucléotide n°77 : tatgcacaagctttaaatttacattt Oligonucléotide n°78 : ttattaaaagaagcatctctttt Oligonucléotide n°79 : tggagaaggatttggattcacacag Oligonucléotide n°80 : cacatattatgaacgtcaattgga

US 2004/0096934 A1

Oligonucléotide n°81 : tactgtgaaactttttataatacaggtt Oligonucléotide n°82 : tacaggtttagaacgtttaagagga Oligonucléotide n°83 : aatactgaaagttttttaagatatcatc Oligonucléotide n°84 : gtagagaaatgactttattattattagaattattagcgctatttccatatt Oligonucléotide n°85 : tttccatattatgaattacgactttatccaac Oligonucléotide n°86 : cttacacgtgagttatatacaga Oligonucléotide n°87 : tatacagaaccgattttatttaatccacc Oligonucléotide n°88 : ccaccagctaatttaggactttgccgac Oligonucléotide n°89 : ctttgccgacgttttggtactaatccc Oligonucléotide n°90 : catctttttgaaaggctgaatag Oligonucléotide n°91 : taatcgatttccattatcatctaattttat Oligonucléotide n°92 : ctaattttatggaatatt**tt**tcaggacatacgttac Oligonucléotide n°93 : tagttatctgaacgaatcagcattacaagaaga Oligonucléotide n°94 : caagaagaaagttatggcct Oligonucléotide n°95 : caattaatcccggattagaaggaacaaaccgcata Oligonucléotide n°96 : gagtcaacggcattagaatttcgttctgca Oligonucléotide n°97 : ggtatatatggcttaaatagagcttc Oligonucléotide n°98 : tagagcttcttttttaccaggaggcttgtt Oligonucléotide n°99 : ctgctaatggaggatgtagagaactctatga Oligonucléotide n°100 : ctctatga**a**acaaatga Oligonucléotide n°101 : acaaatgaagaattaccacc Oligonucléotide n°102 : attaccaccagaagaaagtaccggaag Oligonucléotide n°103: agactatctcatttaaccttttttagcttt Oligonucléotide n°104 : gctaatgcaggaagtttacctacttat Oligonucléotide n°105 : cctacttatttatttacccgtcgtga Oligonucléotide n°106 : acccgtcgtgaattagaacttaataatacgatt Oligonucléotide n°107 : attaccattgttaaaggcatctgc Oligonucléotide n°108 : aaggcatctgcacctttatcgggtactacg

# Patent Application Publication May 20, 2004 Sheet 6 of 9 US 2004/0096934 A1

Oligonucléotide	n°109	:	tcgggtactacgttattaaaaggtccagg	
Oligonucléotide	n°110	:	acatttggaacgttaagattaacgttaaattcaccattaa	
Oligonucléotide	n°111	:	cacaacaatatcgcctaaga <b>tta</b> cgttttgcctcaac	
Oligonucléotide	n°112	:	aaatttcagtataaggttactccgtggaggg	
Oligonucléotide	n°113	:	ctccgtggagggttatctatcggtga	
Oligonucléotide	n°114	:	tctatcggtga <b>atta</b> agattagggagcac	
Oligonucléotide	n°115	:	caagagattctaacattaaatgcagaaggt	
Oligonucléotide	n°116	:	aatgcagaaggt <b>tta</b> agcaccggtggtgaata	
Oligonucléotide	n°117	:	gtggtgaatattatatagaaagaattgaaatt	
Oligonucléotide	n°118	:	agaattgaaatt <b>tta</b> cct <b>t</b> t <b>a</b> aatccggcacgagaag	
Oligonucléotide	n°119	:	cgagaagcggaagagga <b>a</b> ttagaagcggcg	

atgaatcgaa ataatcaaaa tgaatatgaa attattgatg cccccattg tgggtgtcca tcagatgacg atgtgaggta tcctttggca agtgacccaa atgcagcgtt acaaaatatg aactataaag attacttaca aatgacagat gaggactaca ctgattetta tataaateet agtttateta ttagtggtag agatgcagtt cagactgcgc ttactgttat taggagaata ctcggggctt taggtgttcc gttttctgga caaatattaa gtttttatca attcctttta aatacactgt ggccagttaa tgatacagct atatgggaag ctttcatgcg acaggtggag gaacttgtca atcaacaaat aacagaattt gcaagaaatc aggcacttgc aagattgcaa ggattaggag aatcttttaa tgtatatcaa cgttcccttc aaaattggtt ggctgatcga aatgatacac gaaatttaag tttattacgt gctcaattta tagctttaga ccttgatttt gttaatgcta ttccattgtt tgcagtaaat ggacagcagg ttccattact gtcagtatat gcacaagctt taaatttaca tttgttatta ttaaaagaag catctctttt tggagaagga tggggattca cacaggggga aatttccaca tattatgaac gtcaattgga actaaccgct aagtacacta attactgtga aacttggtat aatacaggtt tagaacgttt aagaqqaaca aatactgaaa gtttttaag atatcatcaa ttccgtagag aaatgacttt agtggtatta gatgttgtgg cgctatttcc atattatgat gtacgacttt atccaacggg atcaaaccca cagettacae gtgaggtata tacagateeg attgtattta atecaecage taatttagga ctttgccgac gttggggtac taatccctat aatacttttt ctgagetega aaatgeette attegeeeae cacatetttt t**gaa**aggetg aatagettaa caateageag taategattt ceagttteat etaattttat ggaatatttt tcaggacata cgttacgccg tagttatctg aacgattcag cagtacaaga agatagttat ggcctaatta caaccacaag agcaacaatt aatcccggag ttgatggaac aaaccgcata gagtcaacgg cattagaatt tcgttctgca ttgataggta tatatggctt aaatagagct tcttttgtcc caggaggett gtttaatggt acgaettete etgetaatgg aggatgtaga gatetetatg atacaaatga tgaattacea eeagatgaaa gtaceggaag

US 2004/0096934 A1

ttcaacccat agactatete atttaacett ttttagettt caaactaate aggetggate tatagetaat geaggaagtg taeetaatta tgtttggaee egtegtgatg tggaeettaa taataegatt aceecaaata gaattaeeaa attaeeattg gtaaaggeat etgeaeetgt ttegggtaet aeggtettaa aaggteeagg atttaeagga gggggtatae teegaagaae aaetaatgge acatttggaa egttaagagt aaeggttaat teaeeattaa caeaaaata tegeetaaga ttaeegtttg eeteaaeagg aaattteagt ataagggtae teegtggagg ggtteetate ggtgatgtta gattagggag eaeaatgaae agagggeagg aaetaaetta egaateett tteeaaagag agttaeetae taetggteeg tteaateege etttaeatt taeeaaaget eaagagtte taaeagtgaa tgeagaaggt gttageaeeg gtggtgaata ttatatagat agaattgaaa ttgteeetgt gaateeggea egagaagegg aagaagggttt agaageggeg aagaaageg

MNRNNQNEYE IIDAPHCGCP SDDDVRYPLA SDPNAALQNM NYKDYLQMTD EDYTDSYINP SLSISGRDAV QTALTLLGRI LGALGVPFSG QILSFYQFLL NTLWPVNDTA IWEAFMROVE ELVNQQITEF ARNQALARLQ GLGESFNVYQ RSLQNWLADR NDTRNLSLLR AQFIALDLDF VNAIPLFAVN GQQVPLLSVY AOALNLHLLL LKEASLFGEG WGFTQGEIST YYERQLELTA KYTNYCETWY NTGLERLRGT NTESFLRYHQ FRREMTLVVL DVVALFPYYD VRLYPTGSNP OLTREVYTDP IVFNPPANLG LCRRWGTNPY NTFSELENAF IRPPHLFERL NSLTISSNRF PVSSNFMEYF SGHTLRRSYL NDSAVQEDSY GLITTRATI NPGVDGTNRI ESTALEFRSA LIGIYGLNRA SFVPGGLFNG TTSPANGGCR DLYDTNDELP PDESTGSSTH RLSHLTFFSF QTNQAGSIAN AGSVPTYVWT RRDVDLNNTI TPNRITQLPL VKASAPVSGT TVLKGPGFTG GGILRRTTNG TFGTLRVTVN SPLTQQYRLR LRFASTGNFS IRVLRGGVSI GDVRLGSTMN RGOELTYESF FTREFTTTGP FNPPFTFTQA QEILTVNAEG VSTGGEYYID RIEIVPVNPA REAEEDLEAA KKA

### PEPSIN-SENSITIVE MODIFIED BACILLUS THURINGIENSIS INSECTICIDAL TOXIN

### FIELD OF THE INVENTION

**[0001]** The present invention relates to the degradation of *Bacillus thuringiensis* Cry proteins in the mammalian digestive tract. It relates to *Bacillus thuringiensis* Cry proteins, the peptide sequence of which has been modified so as to make them sensitive to the specific enzymes in the mammalian digestive tract, in particular to pepsins. According to this invention, the Cry proteins are modified by insertion of pepsin-cleavage sites into their peptide sequence. The invention also relates to transformed plants expressing these modified Cry proteins.

[0002] Bacteria of the species Bacillus thuringiensis (hereinafter referred to as Bt) are well known for the insecticidal toxins which they produce. These Gram-positive bacteria form a parasporal crystal protein during their stationary phase, which is greatly responsible for their insecticidal activity. The crystal protein of Bt consists of an insecticidal toxin which is protein in nature, referred to as Cry protein, and encoded by a cry gene. By virtue of its insecticidal properties, this Cry protein was used in protecting crops against pest insects, as an alternative solution to synthetic insecticides. Currently, this agronomic use is essentially implemented by two methods, direct spreading of the product as a biopesticide, and genetic transformation of the plants grown, with a gene encoding a Cry protein. Depending on the strains of Bt from which they are derived, the Cry proteins have insecticidal activities with respect to different insect spectra. The main orders of insects against which the Cry toxins are active are Lepidoptera, Coleoptera and Diptera, but some toxins are effective with respect to other insect orders. All the Cry proteins isolated from the various strains of Bt are grouped together in a classification as a function of their sequence homologies, and they are assigned a code in order to distinguish them (Crickmore et al., 1998, Microbiol. Molec. Biol. Review (62(3), 807-813). The advantage of using these toxins in agriculture therefore lies in their specificity of action with respect to one or more given insect orders, but also in their lack of toxicity with respect to mammals, to birds, to amphibians and to reptiles.

[0003] This lack of toxicity with respect to mammals has made it impossible to develop the culturing of transgenic plants expressing a Cry protein, and to use the seeds from these plants for human and animal foodstuffs. However, although they are non-toxic with respect to mammals, some of these proteins are relatively undegraded in the mammalian digestive tract, and this lack of degradation leads to a relatively long persistence of the toxin in the digestive tract of said mammals. In addition, a lack of persistence of Cry proteins in the mammalian digestive tract is one of the criteria taken into account by the administrative authorities (for example the United States Environmental Protection Agency—EPA) which grant marketing authorizations in the food sector for seeds containing these proteins or for products derived from these seeds.

#### BACKGROUND OF THE INVENTION

**[0004]** The present invention makes it possible to overcome the drawback mentioned above. This invention is based on the principle according to which the stability of certain Cry proteins in the mammalian digestive tract is thought to be due to a lack of sensitivity of these proteins to the specific enzymes in said digestive tract, in particular to the proteases. The solution to this problem therefore lies in the artificial integration of specific sites, specific to the enzymes of the mammalian digestive tract, into the Cry protein. A subject of the present invention is therefore modified Cry proteins sensitive to the specific enzymes in the mammalian digestive tract, in particular the specific proteases in the mammalian stomach, and more particularly the pepsins. Pepsin is a particular enzyme of the protease family, and it is the major protease present in the mammalian stomach (95% of stomach proteases). It is an aspartic protease which acts at an optimum pH of 2. Pepsin is an enzyme of choice as a source of degradation of Cry proteins since it is not present in the digestive tube of insects, in particular of the Lepidoptera, in which the pH of the digestive tube is between 10 and 11 (Terra, W. B. and C. Ferreira, 1994, Insect digestive enzymes: properties, compartmentalization and function. Comp. Biochem. Physiol. 109B: 1-62). This lack of pepsin in insects therefore guarantees that introducing pepsin-specific sites into the Cry proteins does not present a risk of increasing their degradation in the insect digestive tube. The present invention is therefore a solution to the technical problem set out above, namely an increase in the sensitivity of the Cry proteins to enzymes of the mammalian digestive tract, without alteration of the insecticidal properties of said Cry proteins.

**[0005]** However, the Cry protein is a very organized protein, the activated form of which is made up of three domains, and in which the structure-function relationships are very strong within and between the domains. This high level of organization of the Cry proteins does not permit the random insertion of mutations into the protein. Specifically, the insertion of cleavage sites specific to mammalian stomach enzymes must not alter the insecticidal properties of the toxins.

**[0006]** The Cry proteins are naturally produced by the bacterium *Bacillus thuringiensis* in the form of inactive protoxins. The natural method of action of these proteins involves solubilization of the crystal protein in the insect intestine, proteolytic degradation of the released protoxin, attachment of the activated toxin to the receptors in the insect intestine, and insertion of the toxin into the apical membrane of the intestinal cells so as to create ion channels or pores. The proteolytic degradation of the protoxin in the insect intestine takes place under the joint action of the alkaline pH and of the serine proteases (essentially trypsin) of the digestive juice (Schnepf et al., 1998).

**[0007]** The Cry toxins consist of three structural domains, domain I, domain II and domain III. Domain I occupies approximately the N-terminal half of the activated toxin. Domains II and III each occupy approximately a quarter of the activated toxin. Domain III is located at the C-terminal end of the activated toxin. Each domain of the Cry protein has its own structure and its own function.

**[0008]** Domain I consists of seven  $\alpha$ -helices, 6 amphiphilic helices and a hydrophobic helix, connected to one another via inter-helix loops consisting of a few amino acids. This domain is the transmembrane domain, responsible for the formation of the ion channel or pore (Aronson et al., 1995; Chen et al., 1993; Manoj-Kumar and Aronson,

1999; Masson et al., 1999; Rang et al., 1999; Coux et al., 1999). The formation of the transmembrane pore by the a-helices of domain I in fact involves four Cry proteins which form a complete pore with their four respective α4-helices (Masson et al., 1999). A cylindrical pore of four  $\alpha$ 4-helices therefore forms. The inside of this pore consists of the hydrophilic faces of the amphiphilic helices; since the negatively charged residues are present on the hydrophilic faces, they are in the lumen of the pore, in aqueous medium, and perform their ion transport function. The outside of the pore consists of the hydrophobic faces which anchor the pore in the lipid membrane. The formation of the pore by the a-helices of domain I therefore involves very strong structure-function relationships and conformational changes over time. The introduction of mutations into the  $\alpha$ -helices of domain I therefore has a high probability of disturbing the function of this domain and therefore the activity of the toxin.

[0009] Domains II and III of the activated toxin consist of  $\beta$ -sheets, which are themselves also in a very compacted form. These two domains are involved in receptor site recognition (specificity) and in toxin stability (Abdul-Rauf and Ellar, 1999; Dean et al., 1996; Hussain et al., 1996; Lee et al., 1999; Rajamohan et al., 1996, 1998; Wu and Dean, 1996). Domain III exchanges induce changes in specificity (de Maagd et al., 1999). This region is much less conserved, and therefore more variable, than domain I. It is involved in the specificity of each toxin. This variability and these interactions specific to each toxin are involved in the nature of the very specific host spectrum of each toxin and are involved in the recognition of different receptor sites. Recognition of the receptor takes place via loops in domain II and in domain III and the conformation of these loops varies subtly from one toxin to the other as a function of the arrangement and of the interactions between domains II and III. Domain I also interferes with the other two domains and influences the general conformation (Rang et al., 1999, 2001). In addition, very little is known about the structurefunction relationships within these two domains, and no information is actually available regarding the conformation required for recognition of a receptor site. It is therefore very difficult to predict consequences of introducing modifications into domains II and III on the specificity, the ability to recognize the receptor sites and the toxicity of the Cry proteins. Moreover, it is known that mutations generated in domains II and III very often induce destabilization of the toxin in insects, leading to a loss of toxicity.

**[0010]** Salt bridges also exist between domains I and II of the Cry proteins. These bridges play an important role in the stability of the toxin and in the functioning thereof. Artificial elimination of these bridges in Cry1Aa1 shows that the protoxins and activated toxins are less stable than the parental protein (Vachon et al., 2000). These salt bridges are present between domain II and the 7-helix of domain I. The acknowledged importance of these bridges implies that mutations in domain II and the  $\alpha$ 7-helix of domain I have a high risk of disturbing the function of the Cry proteins.

#### DETAILED DESCRIPTION OF THE INVENTION

**[0011]** The present invention relates to a pepsin-sensitive modified Cry protein, characterized in that it has at least one additional pepsin cleavage site.

[0012] The term "Cry protein" is intended to mean the insecticidal protein produced by a strain of bacterium Bacillus thuringiensis (hereinafter referred to as Bt), the various holotypes of which, which exist and which are to come, are referenced by the Bt classification committee (Crickmore, 2001) and accessible on the Internet site at "www.biols.susx.ac.uk/Home/NeilCrickmore/Bt/index.html." In particular, this Cry protein is encoded by a cry gene, either naturally by the Bt bacterium, or in a recombinant manner in a host organism transformed with a cry gene or with a gene comprising at least the coding sequence of a Cry protein. The Cry proteins according to the invention also comprise Cry proteins the sequence of which has been artificially modified so as to increase their insecticidal activity or their resistance to treatment conditions. This definition also includes Cry protein fragments which conserve the insecticidal activity, such as the truncated Cry proteins comprising only the N-terminal portion of a complete Cry protein, in particular domain I of this protein (WO 94/05771). Also included are the fused Cry proteins, as described in international patent application WO 94/24264. Preferably, the Cry protein according to the invention is selected from the Cry1, Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 proteins. Preferably, it is the Cry9C protein, and preferably the Cry9Ca1 protein (Lambert et al., Appl. Environm. Microbiol. 62, 80-86; WO 94/05771). In particular, the present invention also fits any Cry protein, the toxicity of which has been improved, such as, for example, those described in patent applications WO 97/49814 or WO 99/00407.

**[0013]** According to the present invention, the Cry protein is modified. The term "modified Cry protein" is intended to mean a Cry protein, the peptide sequence of which is different from the sequence of the native Cry protein from which it is derived. This sequence difference is the result of artificial modifications introduced by genetic engineering, in particular the insertion or the substitution of specific amino acid residues into or in said peptide sequence. In particular, the modified Cry protein is produced by modification of the nucleotide sequence encoding it, in particular by the technique of site-directed mutagenesis well known to those skilled in the art (Hutchinson C. A. et al., 1978, J. Biol. Chem. 253: 6551). Preferably, the modification of the Cry protein consists of an amino acid residue substitution.

[0014] The modified Cry protein according to the invention is pepsin-sensitive. The pepsin focuses its proteolytic action on specific cleavage sites consisting of the amino acids leucine, phenylalanine and glutamic acid. The proteolysis takes place on the C-terminal side of the residue concerned. According to the invention, the term "pepsinsensitive" is intended to mean the property, for the modified Cry protein, of undergoing proteolysis by pepsin. Proteolysis of the Cry protein leads to partial or total loss of the insecticidal activity of said protein. The pepsin-sensitivity can therefore be measured by bringing a modified Cry protein according to the invention into contact, preferably in vitro, with a pepsin, and then measuring the loss of insecticidal activity of said modified Cry protein in comparison with a native Cry protein which has not been modified according to the invention. By way of example, the tests described in Examples 7 and 8 can be used to measure the pepsin sensitivity of a Cry protein according to the invention. Alternatively, the Western blotting technique can also be used to measure said pepsin sensitivity. Using this

technique, the sensitivity is measured by observing the structural degradation of the modified Cry protein after contact with a pepsin. This observation consists of the disappearance or the decrease in intensity of a band corresponding to the Cry protein on a gel electrophoresis transfer membrane, compared to a native Cry protein which has not been modified according to the invention. The use of these techniques is part of the general knowledge of those skilled in the art.

[0015] The modified Cry protein according to the invention is characterized in that it has at least one additional pepsin cleavage site. The term "pepsin cleavage site" is intended to mean a site consisting of at least one amino acid residue recognized as a site of proteolysis by pepsin. The amino acid residues recognized by pepsin are leucine, phenylalanine or glutamic acid. The expression "additional pepsin cleavage site" is intended to mean an additional cleavage site compared to the native Cry protein as produced by the Bt bacterium.

**[0016]** Preferably, the additional pepsin cleavage site is represented by an amino acid residue selected from leucine, phenylalanine or glutamic acid residues. According to a particular embodiment of the invention, the modified Cry protein has several additional pepsin cleavage sites represented by the same amino acid residue. According to another embodiment of the invention, the modified Cry protein has several additional pepsin cleavage sites represented by different amino acid residues.

[0017] According to a particular embodiment of the invention, the modified Cry protein according to the invention is characterized in that it has at least one additional pepsin cleavage site in at least one of the inter- $\alpha$ -helix loops of domain I. The expression "inter- $\alpha$ -helix loops of domain I" is intended to mean the peptide chains linking the seven  $\alpha$ -helices of domain I of the Cry proteins as described in Grochulski et al. (1995) and Li et al. (1991). According to the invention, the Cry protein should have at least one additional pepsin cleavage site. In addition, said additional cleavage site is in at least one of the inter- $\alpha$ -helix loops of domain I. The term "additional" is therefore understood to be supplementary compared to the number of pepsin cleavage sites naturally present in the inter- $\alpha$ -helix loops of domain I of the native Cry protein as produced by the Bt bacterium. This definition means that the modified Cry protein according to the invention is characterized in that it has a number of pepsin cleavage sites in its inter- $\alpha$ -helix loops of domain I which is greater than the number of these sites in the same native Cry protein as produced by the Bt bacterium, the difference between said numbers being at least equal to 1.

**[0018]** According to a particular embodiment of the invention, the modified Cry protein according to the invention has at least one pepsin cleavage site in the inter- $\alpha$ -helix loop linking the  $\alpha$ 3 and  $\alpha$ 4 helices of domain I.

**[0019]** According to a preferred embodiment of the invention, the modified Cry protein is a modified Cry9C protein. Preferably, the modified Cry protein is a modified Cry9Ca1 protein having a pepsin cleavage site positioned on amino acid residue 164. In particular, the arginine residue naturally present at position 164 on the Cry9Ca1 protein is replaced with an amino acid residue chosen from leucine, phenylalanine and glutamic acid residues, on the Cry9Ca1 protein

modified according to the invention. Preferably, the Cry9Ca1 protein modified according to the invention is selected from the Cry proteins the sequences of which are represented by the identifiers SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:8.

**[0020]** The present invention also relates to a pepsinsensitive modified Cry protein, characterized in that the additional pepsin cleavage sites which it possesses are introduced by substituting aspartic acid residues with glutamic acid residues, substituting tryptophan residues with phenylalanine residues, and substituting valine or isoleucine residues with leucine residues. Preferably, the degree of substitution which said modified Cry protein has is 25%. The expression "degree of substitution" is intended to mean the percentage of amino acid residues of the native Cry protein which are replaced with amino acid residues corresponding to pepsin cleavage sites in the modified Cry protein of the invention.

**[0021]** A subject of the present invention is also a method for increasing the pepsin sensitivity of the Cry proteins, characterized in that at least one additional pepsin cleavage site is introduced into said Cry proteins. The expression "increasing the pepsin sensitivity of the Cry proteins" is intended to mean an increase in the pepsin sensitivity of the Cry proteins obtained by said method compared to the corresponding native Cry proteins, this increase resulting in proteolytic destruction and a loss of insecticidal activity of the Cry proteins, these effects possibly being partial or total.

**[0022]** The introduction of at least one pepsin cleavage site is carried out artificially by genetic engineering. In particular, it involves an insertion or a substitution of amino acid residues. Preferably, it involves a substitution. Such a substitution can be readily carried out by the site-directed mutagenesis technique well known to those skilled in the art.

**[0023]** Preferably, the Cry protein to which the method according to the invention applies is selected from the Cry1, Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 proteins. Preferably, it is the Cry9C protein, and preferably the Cry9Ca1 protein.

**[0024]** In particular, the additional pepsin cleavage site is represented by an amino acid residue chosen from leucine, phenylalanine and glutamic acid residues.

**[0025]** According to a particular embodiment of the invention, the method according to the invention is characterized in that at least one additional pepsin cleavage site is introduced into at least one of the inter- $\alpha$ -helix loops of domain I of said Cry protein.

**[0026]** According to another particular embodiment of the invention, the method according to the invention is characterized in that at least one additional pepsin cleavage site is introduced into the inter- $\alpha$ -helix loop linking the  $\alpha$ 3 and  $\alpha$ 4 helices of domain I.

**[0027]** According to a preferred embodiment of the invention, the present method applies to a Cry9C protein. Preferably, it applies to a Cry9Ca1 protein, and the additional pepsin cleavage site is introduced by substitution of amino acid residue 164. In particular, the arginine residue naturally present at position 164 on the Cry9Ca1 protein is replaced with an amino acid residue chosen from leucine, phenylalanine and glutamic acid residues.

**[0028]** The present invention also relates to a method for increasing the pepsin sensitivity of the Cry proteins, characterized in that the additional pepsin cleavage sites are introduced by substituting aspartic acid residues with glutamic acid residues, substituting tryptophan residues with phenylalanine residues, and substituting value or isoleucine residues with leucine residues.

**[0029]** Preferably, the degree of substitution introduced into said Cry protein is 25%.

**[0030]** The present invention also relates to a polynucleotide encoding a modified Cry protein according to the invention. According to the present invention, the term "polynucleotide" is intended to mean a natural or artificial nucleotide sequence which may be of the DNA or RNA type, preferably of the DNA type, in particular double-stranded.

[0031] The present invention also relates to a chimeric gene comprising, functionally linked to one another, at least one promoter which is functional in a host organism, a polynucleotide encoding a modified Cry protein according to the invention, and a terminator element which is functional in this same host organism. The various elements which a chimeric gene can contain are, firstly, regulatory elements for the transcription, the translation and the maturation of proteins, such as a promoter, a sequence encoding a signal peptide or a transit peptide, or a terminator element constituting a polyadenylation signal and, secondly, a polynucleotide encoding a protein. The expression "functionally linked to one another" means that said elements of the chimeric gene are linked to one another in such a way that the functioning of one of these elements is affected by that of another. By way of example, a promoter is functionally linked to a coding sequence when it is capable of affecting the expression of said coding sequence. The construction of the chimeric gene according to the invention and the assembly of its various elements can be carried out using techniques well known to those skilled in the art, in particular those described in Sambrook et al. (1989, Molecular Cloning: A Laboratory Manual, Nolan C. ed., New York: Cold Spring Harbor Laboratory Press). The choice of the regulatory elements constituting the chimeric gene depends essentially on the host species in which they must function, and those skilled in the art are capable of selecting regulatory elements which are functional in a given host organism. The term "functional" is intended to mean capable of functioning in a given host organism.

[0032] According to a particular embodiment of the invention, the chimeric gene contains a "constitutive" promoter. A constitutive promoter according to the present invention is a promoter which induces the expression of a coding sequence in all the tissues of a host organism and continuously, i.e. during the entire duration of the life cycle of said host organism. Some of these promoters can be tissue-specific, i.e. can express the coding sequence continuously, but only in a particular tissue of the host organism. Constitutive promoters can originate from any type of organism. Among the constitutive promoters which may be used in the chimeric gene of the present invention, mention may be made, by way of example, of bacterial promoters, such as that of the octopine synthase gene or that of the nopaline synthase gene, viral promoters, such as that of the gene controlling transcription of the 19S or 35S RNAs of the cauliflower mosaic virus (Odell et al., 1985, Nature, 313, 810-812), or the promoters of the cassava vein mosaic virus (as described in patent application WO 97/48819). Among the promoters of plant origin, mention will be made of the promoter of the ribulose-biscarboxylase/oxygenase (RuBisCO) small subunit gene, the promoter of a histone gene as described in application EP 0 507 698, the promoter of the EF1- $\alpha$  gene (WO 90/02172), the promoter of an actin gene (U.S. Pat. No. 5,641,876), or the promoter of a ubiquitin gene (EP 0342926).

[0033] According to another particular embodiment of the invention, the chimeric gene contains an inducible promoter. An inducible promoter is a promoter which only functions, i.e. which only induces expression of a coding sequence, when it is itself induced by an inducing agent. This inducing agent is generally a substance which can be synthesized in the host organism subsequent to a stimulus external to said organism, this external stimulus possibly being physical or chemical, biotic or abiotic in nature. Such promoters are known, such as, for example, the promoter of the plant O-methyltransferase class II (COMT II) gene described in patent application WO 00/56897, the Arabidopsis PR-1 promoter (Lebel et al., 1998, Plant J. 16(2): 223-233), the EAS4 promoter of the tobacco sesquiterpene synthase gene (Yin et al., 1997, Plant Physiol. 115(2), 437-451), or the promoter of the gene encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase (Nelson et al., 1994, Plant Mol. Biol. 25(3): 401-412).

**[0034]** Among the terminator elements which can be used in the chimeric gene of the present invention, mention may, for example, be made of the nos terminator element of the gene encoding *Agrobacterium tumefaciens* nopaline synthase (Bevan et al., 1983, Nucleic Acids Res. 11(2), 369-385), or the terminator element of a histone gene as described in application EP 0 633 317.

**[0035]** According to a particular embodiment of the invention, the promoter and the terminator element of the chimeric gene according to the invention are both functional in plants.

**[0036]** It also appears to be important for the chimeric gene to additionally comprise a signal peptide or a transit peptide which makes it possible to control and orient the production of the Cry protein specifically in a cellular compartment of the host organism, such as, for example, the cytoplasm, in a particular compartment of the cytoplasm, or the cell membrane or, in the case of plants, in a particular type of cellular compartment, for example the chloroplasts, or in the extracellular matrix.

**[0037]** The transit peptides can be either single or double. The double transit peptides are optionally separated by an intermediate sequence, i.e. they comprise, in the direction of transcription, a sequence encoding a transit peptide of a plant gene encoding an enzyme located in plastids, a portion of sequence of the mature N-terminal portion of a plant gene encoding an enzyme located in plastids, and then a sequence encoding a second transit peptide of a plant gene encoding an enzyme located in plastids, and then a sequence encoding a second transit peptide of a plant gene encoding an enzyme located in plastids. Such double transit peptides are, for example, described in patent application EP 0 508 909.

**[0038]** Signal peptides of use according to the invention which may be mentioned include in particular the signal peptide of the tobacco PR-1 $\alpha$  gene described by Cornelissen

et al. (1987, Nucleic Acid Res. 15, 6799-6811), in particular when the chimeric gene according to the invention is introduced into plant cells or plants.

[0039] The present invention also relates to a vector containing a chimeric gene according to the invention. Such a vector is of use for transforming a host organism and expressing a modified Cry protein according to the invention in said organism. This vector may be a plasmid, a cosmid, a bacteriophage or a virus. In general, the main qualities of this vector should be an ability to maintain itself and to self-replicate in the host organism's cells, in particular by virtue of the presence of an origin of replication, and to express therein a modified Cry protein. The choice of such a vector and also the techniques for inserting the chimeric gene according to the invention therein are widely described in Sambrook et al. (1989, Molecular Cloning: A Laboratory Manual, Nolan C. ed., New York: Cold Spring Harbor Laboratory Press) and are part of the general knowledge of those skilled in the art. The vector used in the present invention may also contain, in addition to the chimeric gene of the invention, a chimeric gene containing a selectable marker. This selectable marker makes it possible to select the host organisms effectively transformed, i.e. those having incorporated the vector. Among the selectable markers which can be used in many host organisms, mention may be made of markers containing genes for resistance to antibiotics, such as that of the hygromycin phosphotransferase gene (Gritz et al., 1983, Gene 25: 179-188). Preferably, the host organism to be transformed is a plant. Among the selectable markers which can be used in plants, mention may be made of markers containing genes for tolerance to herbicides, such as the bar gene (White et al., NAR 18: 1062, 1990) for tolerance to bialaphos, the EPSPS gene (U.S. Pat. No. 5,188,642) for tolerance to glyphosate or else the HPPD gene (WO 96/38567) for tolerance to isoxazoles. Mention may also be made of genes encoding readily identifiable enzymes such as the GUS enzyme, or genes encoding pigments or enzymes which regulate the production of pigments in the transformed cells. Such selectable marker genes are in particular described in patent applications WO 91/02071 and WO 95/06128.

[0040] The present invention also relates to host organisms transformed with a vector as described above. The term "host organisms" is intended to mean any type of organism, in particular plants or microorganisms such as bacteria, viruses, fungi or yeast. The term "transformed host organism" is intended to mean a host organism which has incorporated into its genome the chimeric gene of the invention and, consequently, produces a modified Cry protein according to the invention in its tissues. To obtain the host organisms according to the invention, those skilled in the art can use one of the many known methods of transformation. One of these methods consists in bringing the cells to be transformed into contact with polyethylene glycol (PEG) and the vectors of the invention (Chang and Cohen, 1979, Mol. Gen. Genet. 168(1), 111-115; Mercenier and Chassy, 1988, Biochimie 70(4), 503-517). Electroporation is another method, which consists in subjecting the cells or tissues to be transformed and the vectors of the invention to an electric field (Andreason and Evans, 1988, Biotechniques 6(7), 650-660; Shigckawa and Dower, 1989, Aust. J. Biotechnol. 3(1), 56-62). Another method consists in directly injecting the vectors into the host cells or tissues by microinjection (Gordon and Ruddle, 1985, Gene (33(2), 121-136). Advantageously, the "biolistic" method may be used. It consists in bombarding cells or tissues with particles onto which the vectors of the invention are adsorbed (Bruce et al., 1989, Proc. Natl. Acad. Sci. US 86(24), 9692-9696; Klein et al., 1992, Biotechnology 10(3), 286-291; U.S. Pat. No. 4,945, 050). Preferably, the transformation of plants will be carried out using bacteria of the Agrobacterium genus, preferably by infecting the cells or tissues of said plants with A. tumefaciens (Knopf, 1979, Subcell. Biochem. 6, 143-173; Shaw et al., 1983, Gene 23(3): 315-330) or A. rhizogenes (Bevan and Chilton, 1982, Annu. Rev. Genet. 16: 357-384; Tepfer and Casse-Delbart, 1987, Microbiol. Sci. 4(1), 24-28). Preferably, the transformation of plant cells with Agrobacterium tumefaciens is carried out according to the protocol described by Ishida et al. (1996, Nat. Biotechnol. 14(6), 745-750).

[**0041**] These various techniques are in particular described in the following patents and patent applications: U.S. Pat. Nos. 4,459,355, 4,536,475, 5,464,763, 5,177,010, 5,187,073, EP 267,159, EP 604 662, EP 672 752, U.S. Pat. Nos. 4,945,050, 5,036,006, 5,100,792, 5,371,014, 5,478, 744, 5,179,022, 5,565,346, 5,484,956, 5,508,468, 5,538,877, 5,554,798, 5,489,520, 5,510,318, 5,204,253, 5,405,765, EP 270 615, EP 442 174, EP 486 233, EP 486 234, EP 539 563, EP 674 725, WO 91/02071 and WO 95/06128.

**[0042]** The present invention also relates to a method for producing the modified Cry proteins according to the invention. This method comprises at least the steps of:

- [0043] a) culturing a transformed host organism according to the invention in a culture medium suitable for the growth and for the multiplication of said organism,
- [0044] b) extracting the Cry proteins produced by the transformed organism cultured in step (a).

**[0045]** Depending on the host organism chosen to carry out this method and depending on the chimeric gene which it contains, the Cry proteins produced are either produced in the host organism, or are secreted into the culture medium. It ensues that the extraction provided for in step (b) may require a step for destroying the microorganisms, or at least the cells of which they are composed, in order to release the Cry proteins if said proteins are not secreted into the culture medium. The extraction step common to the two possibilities (proteins secreted or not secreted) consists of removal of the host organisms or debris from these organisms by filtration or centrifugation of the culture medium.

**[0046]** According to a particular embodiment, this method for producing the modified Cry proteins can also comprise an additional step (c) of purification of the Cry proteins produced, from the culture medium.

**[0047]** According to a preferred embodiment, the host organism is a microorganism. Preferably, the host organism is a *Bacillus thuringiensis* bacterium and the culturing performed in step (a) is continued until the sporulation phase of said bacteria.

**[0048]** The present invention also comprises plants transformed with a vector according to the invention, characterized in that they contain a chimeric gene according to the invention stably integrated into their genome, and express a modified Cry protein in their tissues. The invention also

extends to the parts of these plants, and the descendants of these plants. The expression "part of these plants" is intended to mean any organ of these plants, whether it is aerial or subterranean. The aerial organs are the stems, the leaves and the flowers. The subterranean organs are mainly the roots, but they can also be tubers. The term "descendants" is intended to mean mainly the seeds containing the embryos derived from the reproduction of these plants with one another. By extension, the term "descendants" applies to all the plants and seeds formed in each new generation derived from crosses between a plant, in particular a plant variety, and a transformed plant according to the invention.

**[0049]** The transformed plants according to the invention may be monocotyledones or dicotyledones. Preferably, these plants are plants of agronomic value. Advantageously, the monocotyledonous plants are wheat, maize and rice. Advantageously, the dicotyledonous plants are rapeseed, soybean, tobacco and cotton.

[0050] According to a particular embodiment of the invention, the transformed plants according to the invention contain, in addition to a chimeric gene according to the invention, at least one other gene containing a polynucleotide encoding a protein of interest. Among the polynucleotides encoding a protein of interest, mention may be made of polynucleotides encoding an enzyme for resistance to a herbicide, for example the polynucleotide encoding the bar enzyme (White et al., NAR 18: 1062, 1990) for tolerance to bialaphos, the polynucleotide encoding the EPSPS enzyme (U.S. Pat. No. 5,188,642; WO 97/04103) for tolerance to glyphosate or else the polynucleotide encoding the HPPD enzyme (WO 96/38567) for tolerance to isoxazoles. Also contained in these plants may be polynucleotides for resistance to diseases, for example a polynucleotide encoding the oxalate oxidase enzyme as described in patent application EP 0 531 498 or U.S. Pat. No. 5,866,788, or a polynucleotide encoding an antibacterial and/or antifungal peptide such as those described in patent applications WO 97/30082, WO 99/24594, WO 99/02717, WO 99/53053 and WO 99/91089. Mention may also be made of polynucleotides encoding agronomic characteristics of the plant, in particular a polynucleotide encoding a  $\Delta$ -6 desaturase enzyme as described in U.S. Pat. Nos. 5,552,306 and 5,614,313, and patent applications WO 98/46763 and WO 98/46764, or a polynucleotide encoding a serine acetyltransferase (SAT) enzyme as described in patent applications WO 00/01833 and PCT/FR 99/03179. The transformed plants according to the invention can also contain a polynucleotide encoding another insecticidal toxin, for example a polynucleotide encoding another Bacillus thuringiensis Cry protein (for example, see international patent application WO 98/40490).

**[0051]** A subject of the present invention is also monoclonal or polyclonal antibodies directed against a modified Cry protein according to the invention, or a fragment thereof. The techniques for producing antibodies are widely described in the general literature and in reference works such as Immunological Techniques Made Easy (1998, 0. Cochet, J. -L. Teillaud, C. Sautès eds., John Wiley & Sons, Chichester). Preferably, the antibodies according to the invention are used in tests, or kits, for detecting the Cry proteins according to the invention. **[0052]** The examples below make it possible to illustrate the present invention without, however, limiting the scope thereof.

#### EXAMPLES

#### Example 1

#### Creation of a Pepsin Cleavage Site at Amino Acid 164 of the Cry9Ca1 Toxin

**[0053]** A pepsin-specific site is introduced into the *Bacillus thuringiensis* Cry9Ca1 toxin by substituting the arginine naturally present at position 164 in this toxin with one of the three amino acids recognized by pepsin: leucine, phenylalanine or glutamic acid. Amino acid 164 is present in the inter- $\alpha$ -helix loop linking the  $\alpha$ 3 and  $\alpha$ 4 helices of domain I (hereinafter referred to as  $\alpha$ 3- $\alpha$ 4 inter-helix loop)

**[0054]** The native sequence of the  $\alpha 3$ - $\alpha 4$  inter-helix loop is between aspartic acid 159 and value 168. The sequence of this loop is as follows: DRNDTRNLSV. This amino acid sequence corresponds to the following DNA sequence extending from base 475 to base 504:

GAT CGA AAT GAT ACA CGA AAT TTA AGT GTT Asp Arg Asn Asp Thr Arg Asn Leu Ser Val

**[0055]** Codon 164 (CGA) encoding arginine is modified to a codon encoding either leucine or phenylalanine or glutamic acid. The codon possibilities are as follows:

Leucine:	TTA,	TTG,	CTT,	CTC,	CTA	or	CTG
Phenylalanine:	TTT	or TT	2				
Glutamic acid:	GAA	or GAG	3				

**[0056]** The choice of preferential codons in the sitedirected mutagenesis depends on the organism in which the modified cry gene must be expressed and therefore varies accordingly. This choice is part of the general knowledge of those skilled in the art, who will adapt the preferential codons as a function of the chosen organism for production. In this example, the chosen organism for expression is the *B. thuringiensis* bacterium. The codons preferentially used by *B. thuringiensis* to encode leucine, phenylalanine or glutamic acid are, respectively, TTA (leucine), TTT (phenylalanine) and GAA (glutamic acid).

**[0057]** The modification for expression in Bt can therefore be carried out using the following mutagenesis oligonucleotides (in the oligonucleotides described in the examples below, the codon in upper case letters corresponds to the mutated codon, and the bases and amino acids in bold characters correspond to the bases and amino acids specifically mutated):

Oligonucleotide No. 1: 5'-gat cga aat gat aca **TTA** aat tta agt gtt gtt-3'

Asp Arg Asn Asp Thr Leu Asn Leu Ser Val Val

**[0058]** Oligonucleotide No. 1 allows the replacement of arginine 164 with a leucine.

Oligonucleotide No. 2: 5'-gat cga aat gat aca **TTT** aat tta agt gtt gtt-3'

Asp Arg Asn Asp Thr Phe Asn Leu Ser Val Val

**[0059]** Oligonucleotide No. 2 allows replacement of arginine 164 with a phenylalanine.

Oligonuclectide No. 3: 5'-gat cga aat gat aca **GAA** aat tta agt gtt gtt-3'

Asp Arg Asn Asp Thr **Glu** Asn Leu Ser Val Val

**[0060]** Oligonucleotide No. 3 allows replacement of arginine 164 with a glutamic acid.

**[0061]** The characteristics of the bacterial strains of *Escherichia coli* used to modify the sequence of the cry9Ca1 gene are as follows:

- [0062] JM 109 of genotype recA1 supE44 endA1 hsdR17 gyrA96 relA1 thiD (lac-proAb) F' (traD36 proAB+lacIq lacZ DM15)
- [0063] BMH 71-18 mut S of genotype thi, supE, (lac-proAB), (mutS::Tn10)(F', proAB, lacIqZ $\Delta$ M15).

[0064] The plasmid DNA is prepared by minipreparation according to the alkaline lysis technique (Birboim and Doly, 1979). Each bacterial colony is grown in 2 ml of LB medium supplemented with the appropriate antibiotic, overnight at 37° C. with shaking (200 rpm). The culture is then transferred into a microtube and then centrifuged at 13 500 g for 5 min. After removal of the supernatant, the bacteria are resuspended in 100  $\mu$ l of a solution of 25 mM Tris-HCl, pH 8, and 10 mM EDTA containing RNase A at a final concentration of 100  $\mu g/ml.$  200  $\mu l$  of a 0.2 M NaOH solution containing 1% SDS are added and the suspension is mixed twice by inverting the microtube. 150  $\mu$ l of a 2.55 M potassium acetate solution, pH 4.5, are added and the suspension is incubated for 5 min in ice. After centrifugation for 15 min at 13 500 g, the supernatant is transferred into a microtube containing 1 ml of cold ethanol. After centrifugation for 30 min at 13 500 g, the supernatant is removed and the pellet is washed with 1 ml of 70% ethanol. The pellet containing the DNA is dried for a few minutes under vacuum and then taken up in 50  $\mu$ l of sterile distilled water. The samples are then placed at 65° C. for 30 min.

**[0065]** The digestions with restriction endonucleases are carried out for 1  $\mu$ g of DNA in a final volume of 20  $\mu$ l in the presence of one tenth of the final volume of 10× buffer recommended by the supplier for each enzyme, and using 5 units of enzyme. The reaction is incubated for 2 to 3 h at the optimal temperature for the enzyme.

**[0066]** Dephosphorylation of the 5' ends engendered by restriction enzyme is carried out with calf intestine alkaline phosphatase. The reaction is carried out using 5  $\mu$ l of 10x dephosphorylation buffer (500 mM Tris-HCl, pH 9.3, 10 mM MgCl2, 1 mM ZnCl<sub>2</sub> and 10 mM spermidine) and one unit of enzyme per  $\mu$ g of DNA in a final volume of 50  $\mu$ l. The reaction is incubated for one hour at 37° C. in the case of

overhanging 5' ends or at 55° C. in the case of blunt ends or 3' overhanging ends. After dephosphorylation, the enzyme is then inactivated for 30 min at 65° C. and then removed with two volume for volume extractions with a phenol-chloroform-isoamyl alcohol (25-24-1) mixture. The ligations are formed using T4 phage DNA ligase. They are carried out with an amount of vector equal to 100 ng and an insert/ vector molar ratio of between 5 and 10. The final volume of the reaction is 30  $\mu$ l and comprises 3  $\mu$ l of 10× ligation buffer (300 mM Tris-HCl, pH 7.8, 100 mM MgCl<sub>2</sub>, 100 mM DTT and 10 mM ATP) and 3 units of enzyme. The reaction is incubated overnight at 14° C.

**[0067]** The mutagenesis oligonucleotide (oligonucleotide No. 1, oligonucleotide No. 2 and oligonucleotide No. 3) are phosphorylated in the 5' position in order to allow ligation. 100 pmol of oligonucleotide are incubated for 30 min at 37° C. with 5 units of T4 polynucleotide kinase in a final volume of 25  $\mu$ l in the presence of 2.5  $\mu$ l of 10× phosphorylation buffer (700 mM TrisHCl, pH 7.6, 100 mM MgC12 and 50 mM DTT) in the presence of ATP at a final concentration of 1 mM. The enzyme is then inactivated at 70° C. for 10 min.

[0068] The site-directed mutagenesis is carried out according to a conventional method described below. Other procedures known to those skilled in the art are described in the literature and give identical results. The site-directed mutagenesis method used is that described by the manufacturer for the use of the Altered Sites II system marketed by the company Promega. A detailed description of the mutagenesis system and of the protocol can be found on the internet site of the company Promega at the address http:// www.promega.com. The cry9Ca1 gene is pre-cloned into a phagemide pAlter-1 (Promega) carrying the tetracycline resistance gene and the ampicillin resistance gene containing a point mutation. The DNA fragment to be mutated is pre-cloned into the plasmid pAlter-1. 0.5 pmol of plasmid DNA are denatured by adding 2 µl of 2 M NaOH, 2 mM EDTA in a final volume of 20  $\mu$ l and incubating for 5 min at ambient temperature.  $2 \mu l$  of 2 M ammonium acetate, pH 4.6, and 75  $\mu$ l of ethanol are added and the mixture is incubated at -70° C. for 30 min. After centrifugation at 14 000 g for 15 min at 4° C., the pellet is then rinsed with 200  $\mu$ l of 70% ethanol and recentrifuged at 14 000 g for 15 min at 4° C. The denatured DNA pellet is then dried under vacuum and resuspended in  $100 \,\mu$ l of sterile distilled water. 10 µl of denatured DNA, i.e. 0.05 pmol, are mixed with 0.25 pmol of phosphorylated ampicillin-resistance gene repair oligonucleotide, 0.25 pmol of tetracycline-resistance gene destruction oligonucleotide and 1.25 pmol of phosphorylated mutagenesis oligonucleotide (oligonucleotide No. 1, No. 2 or No 3) in the presence of hybridization buffer (20 mM Tris-HCl, pH 7.5, 10 mM MgCl2, 50 mM NaCl) and incubated at 75° C. for 5 min, and then slowly cooled to ambient temperature. 5  $\mu$ l of sterile distilled water, 3  $\mu$ l of 10× synthesis buffer (100 mM Tris-HCl, pH 7.5, 20 mM DTT, 10 mM ATP, 5 mM dNTP), 10 units of T4 DNA polymerase and 3 units of T4 DNA ligase are added and the reaction is incubated for 90 min at 37° C. 200 µl of competent E. coli BMH 71-18 bacteria are then incubated in the presence of 1.5  $\mu$ l of the preceding reaction, in ice for 30 min. A heat shock is then performed by placing the bacteria at 42° C. for 50 sec and then in ice for 2 min. 900 µl of LB medium are then added and the suspension is incubated at 37° C. for one hour with shaking. 4 ml of LB medium supplemented with ampicillin at the final concentration of

100  $\mu$ g/ml are then added and the culture is incubated overnight at 37° C. with shaking. A minipreparation of plasmid DNA is prepared from the 4 ml of culture according to the plasmid DNA extraction protocol described above. 200 µl of competent E. coli JM109 bacteria are then incubated in the presence of 1 ng of plasmid DNA, in ice for 30 min. A heat shock is then performed by placing the bacteria at 42° C. for 50 sec, and then in ice for 2 min. 900  $\mu$ l of LB medium are then added and the suspension is incubated at 37° C. overnight with shaking. 100  $\mu$ l of bacterial suspension are then plated out on a Petri dish containing solid LB medium supplemented with ampicillin at the final concentration of 100 µg/ml. The recombinants obtained are screened in order to find the clone of interest. This search is carried out by isolating the plasmid DNA of several colonies by the minipreparation technique described above, and then by sequencing this DNA. The recombinants are then selected using medium supplemented with tetracycline at the final concentration of 12.5  $\mu$ g/ml. The correctness of the desired mutation and the verification of the lack of undesirable mutations are controlled by sequencing the DNA after site-directed mutagenesis. DNA samples for the sequencing are purified with the Wizard Plus SV Minipreps DNA Purification System (Promega) according to the procedure recommended by the supplier, and the sequencing is carried out on an ABI 377 automatic sequencer (Perkin-Elmer) using sequencing reactions carried out according to the chain termination method (Sanger et al., 1977), by PCR using the ABI PRISM BigDye terminator Cycle Sequencing Kit system. For carrying out the sequencing reactions and the automatic analysis of the samples, the procedures used are those recommended by the supplier (Applied Biosystems).

## Example 2

## Creation of Pepsin Cleavage Sites in the α3-α4 Inter-helix Loop of the Cry9Ca1 Toxin

**[0069]** Pepsin-specific sites are introduced into the  $\alpha$ 3- $\alpha$ 4 inter-helix loop of the Cry9Ca1 toxin by substituting at least one amino acid of this inter-helix loop with an amino acid recognized by pepsin, namely leucine, phenylalanine and glutamic acid. Codons encoding these three amino acids will therefore be created in place of the codons naturally present in the region extending from base 475 to base 504. The codon possibilities for these three amino acids are described in Example 1.

**[0070]** As in Example 1, the selected organism for producing the modified Cry protein is the *B. thuringiensis* bacterium, and the choice of the replacement codons is therefore identical to that of Example 1. In addition, if another organism for production is selected, those skilled in the art will be able to adjust the preferential codons as a function of the organism for production selected.

**[0071]** Various alternative sequences for the  $\alpha 3 \cdot \alpha 4$  interhelix loop are possible, each having a variable number of leucine, phenylalanine or glutamic acid residues. Some of these possibilities are given in Table 1. The possibilities for modification of the  $\alpha 3 \cdot \alpha - 4$  inter-helix loop are not limited to those given in Table 1 below. The aim of the list given in Table 1 is to illustrate some of the possibilities for modifications. Those skilled in the art, being aware of the codons specific for each amino acid according to the organisms, will be able to adapt the teaching described in this example to all the possibilities for modifying the  $\alpha 3 \cdot \alpha 4$  inter-helix loop, in particular to those which are not described in Table 1.

TABLE 1

		essible modifications of the $\alpha$ 3— $\alpha$ 4 ix loop of the Cry9Cal toxin
Protein	Amino acid sequence	Nucleotide sequence
CryCal	DRNDTRNLSV	gat cga aat gat aca cga aat tta agt gtt Asp Arg Asn Asp Thr Arg Asn Leu Ser Val
Mutant No.	L ELNEFLNSV	gaA TTA aat gaA TTT TTa aat tta agt gtt Glu Leu Asn Glu Phe Leu Asn Leu Ser Val
Mutant No.	2 ELNELLNLSV	gaA TTA aat gaA TTa TTa aat tta agt gtt Glu Leu Asn Glu Leu Leu Asn Leu Ser Val
Mutant No.	3 ELLEFLLLSV	gaA TTa TTA gaA TTT TTa TTA tta agt gtt Glu Leu Leu Glu Phe Leu Leu Leu Ser Val
Mutant No.	4 ELLELLLLSV	gaA TTa TTA gaA TTa TTa TTA tta agt gtt Glu Leu Leu Glu Leu Leu Leu Leu Ser Val
Mutant No.	5 ELLEELLLSV	gaA TTa TTA gaA GAa TTa TTa tta agt gtt Glu Leu Leu Glu Glu Leu Leu Leu Ser Val

TABLE	1-continu	ed
тарыы	T-COUCTUR	eu

Examples of possible modifications of the $\alpha 3-\alpha 4$ inter-helix loop of the Cry9Cal toxin         Amino acid sequence         Nucleotide sequence         Mutant No. 6 ERLEFLLLSV       gaA cga TTA gaA TTT TTA TTA tta agt gtt												
Protein			Nucleotide sequence									
Mutant No.	6	ERLEFLLLSV	gaA cga TTA gaA TTT TTa TTA tta agt gtt Glu Arg Leu Glu Phe Leu Leu Leu Ser Val									
Mutant No.	7	ERLELLLLSV	gaA cga TTA gaA TTa TTa TTA tta agt gtt Glu Arg Leu Glu Leu Leu Leu Leu Ser Val									
Mutant No.	8	ERLEELLLSV	gaA TTa GAA gaA TTa TTa TTA tta agt gtt Glu Leu Glu Glu Leu Leu Leu Leu Ser Val									
Mutant No.	9	ELLEEEELSV	gaA TTa TTA gaA GAa GAa GAA tta agt gtt Glu Leu Leu Glu Glu Glu Glu Leu Ser Val									

[0072] The substitution of several amino acids within the  $\alpha 3$ - $\alpha 4$  inter-helix loop requires, for each of the mutants, the successive use of several mutagenesis oligonucleotides. The

mutagenesis oligonucleotides required to create the examples of mutants given in Table 1 are presented below (numbered from 4 to 20).

Oligonucleotide	No.	4:				aca Thr									
Oligonucleotide	No.	5:				aca Thr									
Oligonucleotide	No.	6:				cga Arg									
Oligonucleotide	No.	7:		-	-	cga Arg							-		2
Oligonucleotide	No.	8:				cga Arg									
Oligonucleotide	No.	9:		-	-	cga Arg							_	-	2
Oligonucleotide	No.	10:	-			cga Arg		-					-		-
Oligonucleotide	No.	11:				cga Arg									
Oligonucleotide	No.	12:				ttg Leu									
Oligonucieotide	No.	13:				ttg Leu					-				
Oligonucleotide	No.	14:				ttg Leu					-				
Oligonucleotide	No.	15:				ttg Leu									
Oligonucleotide	No.	16:				ttg Leu									
Oligonucleotide	No.	17:				ttg Leu	-	-	-		-				
Oligonucleotide	No.	18:				ttg Leu									
Oligonucleotide	No.	19:				ttg Leu	-	-		-	-				
Oligonucleotide	No.	20:	caa	aat	tgg	ttg	gct	gaA	TTa	TTA	gaa	gaa	gaa	gaa	tta

**[0073]** The successive site-directed mutagenesis procedure is similar to the procedure described in Example 1. The difference lies in the combination of oligonucleotides. For each of the examples of mutants described in Table 1, the successive combinations of oligonucleotides are described below.

**[0074]** Mutant No. 1: The creation of mutant No. 1 requires two successive series of site-directed mutagenesis according to the protocol described in Example 1, using oligonucleotide No. 6 in the first mutagenesis and oligonucleotide No. 13 in the second. Oligonucleotide No. 13 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 6.

**[0075]** Mutant No. 2: The creation of mutant No. 2 requires two successive series of site-directed mutagenesis according to the protocol described in Example 1, using oligonucleotide No. 8 in the first mutagenesis and oligonucleotide No. 12 in the second. Oligonucleotide No. 12 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 8.

**[0076]** Mutant No. 3: The creation of mutant No. 3 requires three successive series of site-directed mutagenesis according to the protocol described in Example 1, using oligonucleotide No. 4 in the first mutagenesis, oligonucleotide No. 7 in the second and oligonucleotide No. 14 in the third. Oligonucleotide No. 7 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 4 and oligonucleotide No. 14 is defined to recognize the modifications introduced during the first mutagenesis with outgonucleotide No. 4 and oligonucleotide No. 14 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 4 and No. 7.

[0077] Mutant No. 4: The creation of mutant No. 4 requires three successive series of site-directed mutagenesis according to the protocol described in Example 1, using oligonucleotide No. 4 in the first mutagenesis, oligonucleotide No. 9 in the second and oligonucleotide No. 15 in the third. Oligonucleotide No. 9 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 4 and oligonucleotide No. 15 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 4 and No. 9.

[0078] Mutant No. 5: The creation of mutant No. 5 requires three successive series of site-directed mutagenesis according to the protocol described in Example 1, using oligonucleotide No. 4 in the first mutagenesis, oligonucleotide No. 11 in the second and oligonucleotide No. 16 in the third. Oligonucleotide No. 11 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 4 and oligonucleotide No. 16 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 4 and No. 11.

**[0079]** Mutant No. 6: The creation of mutant No. 6 requires three successive series of site-directed mutagenesis according to the protocol described in Example 1, using oligonucleotide No. 4 in the first mutagenesis, oligonucleotide No. 7 in the second and oligonucleotide No. 17 in the third. Oligonucleotide No. 7 is defined to recognize the

modifications introduced during the first mutagenesis with oligonucleotide No. 4 and oligonucleotide No. 17 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 4 and No. 7.

**[0080]** Mutant No. 7: The creation of mutant No. 7 requires three successive series of site-directed mutagenesis according to the protocol described in Example 1, using oligonucleotide No. 4 in the first mutagenesis, oligonucleotide No. 9 in the second and oligonucleotide No. 18 in the third. Oligonucleotide No. 9 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 4 and oligonucleotide No. 18 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 4 and oligonucleotide No. 18 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 4 and No. 9.

**[0081]** Mutant No. 8: The creation of mutant No. 8 requires three successive series of site-directed mutagenesis according to the protocol described in Example 1, using oligonucleotide No. 4 in the first mutagenesis, oligonucleotide No. 9 in the second and oligonucleotide No. 19 in the third. Oligonucleotide No. 9 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 4 and oligonucleotide No. 19 is defined to recognize the modifications introduced during the first mutagenesis with one of the modifications introduced during the first two mutageneses with oligonucleotide No. 9.

**[0082]** Mutant No. 9: The creation of mutant No. 9 requires three successive series of site-directed mutagenesis according to the protocol described in Example 1, using oligonucleotide No. 5 in the first mutagenesis, oligonucleotide No. 10 in the second and oligonucleotide No. 20 in the third. Oligonucleotide No. 10 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 5 and oligonucleotide No. 20 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 5 and No. 10.

**[0083]** According to this protocol, the oligonucleotides are divided up into three categories, 1st series oligonucleotides, 2nd series oligonucleotides and 3rd series oligonucleotides. This division is as follows:

	oligonucleotides No. 4, 5, 6 and 8 oligonucleotides No. 7, 9, 10, 11, 12 and 13
3rd series oligonucleotides:	oligonucleotides No. 14, 15, 16, 17, 18, 19 and 20.

**[0084]** The complete protocol for producing these mutants is identical to that described in Example 1. This protocol is common to each of the series of mutageneses, only the mutagenesis oligonucleotide and the oligonucleotide for inhibition/restoration of the resistance to the antibiotic change. The passing onto the following mutation takes place after screening of the clone of interest which has integrated the preceding mutation. If this step is the final step of the first series or of the second series of mutagenesis, the material derived from this series of experiments is re-used as initial material for, respectively, the 2nd or 3rd series of mutagenesis.

esis using, respectively, the 2nd or 3rd series oligonucleotides. A second cycle of mutagenesis can then be carried out using the plasmid DNA obtained as DNA matrix and also the oligonucleotide for repair of the tetracycline resistance gene and the oligonucleotide for destruction of the ampicillin resistance gene and a 2nd series mutagenesis oligonucleotide. The recombinants are then selected using medium supplemented with tetracycline at the final concentration of 12.5  $\mu$ g/ml. A third cycle of mutagenesis can be carried out using the plasmid DNA obtained at the end of the second cycle of mutagenesis as DNA matrix and also the oligonucleotide for repair of the ampicillin resistance gene and the oligonucleotide for destruction of the tetracycline resistance gene and a 3rd series mutagenesis oligonucleotide. The recombinants are then selected using medium supplemented with ampicillin at the final concentration of 100  $\mu$ g/ml. After all the series of mutagenesis required to produce a mutant have been carried out, the steps for controlling the mutations are carried out as described in Example 1.

#### Example 3

## Creation of Pepsin Cleavage Sites in the $\alpha 4-\alpha 5$ , $\alpha 5-\alpha 6$ and $\alpha 6-\alpha 7$ Inter-helix Loops of the Cry9Ca1 Toxin

**[0085]** The positions of the native sequences of the  $\alpha 4$ - $\alpha 5$ , ( $\alpha 5$ - $\alpha 6$  and  $\alpha 6$ - $\alpha 7$  inter-helix loops of the Cry9Ca1 toxin are given in Table 2 below. The nucleotide sequences and the corresponding positions in the cry9Ca1 gene are given in Table 3.

TABLE 2

Positi		he α4–α5, α5–α6 and α6–α7 f the Cry9Ca1 toxin
Loop	Sequence	Position
Loop a4–a5 Loop a5–a6 Loop a6–a7	FAVNGQQVPLL LFGEGWGF LRGTN	Phenylalanine 187 to leucine 197 Leucine 216 to phenylalanine 223 Leucine 257 to asparagine 261

## [0086]

TABLE	3

Position and sequences of $cry9Ca1$ gene encoding the $a4-a5$ , $a5-a6$ and $a6-a7$ inter-helix loopsLoopSequencePosition													
Loop		Seq	lence	Э									Position
Loop	α4—α5	TTT	GCA	GTA	AAT	GGA	CAG	CAG	GTT	CCA	TTA	CTG	559—591
Loop	α5—α6	CTT	TTT	GGA	GAA	GGA	TGG	GGA	TTC				646—669
Loop	α6—α7	TTA	AGA	GGA	ACA	AAT							769—783

**[0087]** The superposition of the nucleotide and amino acid sequences are as follows:

Loop	$\mathbf{T}\mathbf{T}\mathbf{T}$	GCA	GTA	AAT	GGA	CAG	CAG	$\mathbf{GTT}$	CCA	TTA	CTG	
$\alpha 4-\alpha 5$ :	Phe	Ala	Val	Asn	Gly	Gln	Gln	Val	Pro	Leu	Leu	
Loop	$\mathbf{CTT}$	$\mathbf{T}\mathbf{T}\mathbf{T}$	GGA	GAA	GGA	TGG	GGA	TTC				
α5—α6	Leu	Phe	Gly	Glu	Gly	Trp	Gly	Phe				
Loop	TTA	AGA	GGA	ACA	AAT							
α6–α7	Leu	Arg	Gly	$\mathbf{Thr}$	Asn							

**[0088]** Pepsin-specific sites are introduced into the  $\alpha 4$ - $\alpha 5$ ,  $\alpha 5$ - $\alpha 6$  or  $\alpha 6$ - $\alpha 7$  inter-helix loops of the Cry9Ca1 toxin by substituting at least one amino acid of these inter-helix loops with an amino acid recognized by pepsin, namely leucine, phenylalanine and glutamic acid. Codons encoding these three amino acids will therefore be created in place of the codons naturally present in the region extending from bases 559 to 591 ( $\alpha 4$ - $\alpha 5$  inter-helix loop), 646 to 669 ( $\alpha 5$ - $\alpha 6$  inter-helix loop), and 769 to 783 ( $\alpha 6$ - $\alpha 7$  inter-helix loop). The codon possibilities for these three amino acids are described in Example 1.

**[0089]** As in Example 1, the chosen organism for producing the modified Cry protein is the *B. thuringiensis* bacterium, and the choice of the replacement codons is therefore identical to that of Example 1. In addition, if another organism for production is chosen, those skilled in the art will be able to adjust the preferential codons as a function of the chosen organism for production.

[0090] Various alternative sequences for the  $\alpha 4 \cdot \alpha 5$ ,  $\alpha$ 5- $\alpha$ 6 and  $\alpha$ 6- $\alpha$ 7 inter-helix loops are possible, each one having a variable number of leucine, phenylalanine or glutamic acid residues. Several of these various possibilities are given in Tables 4, 5 and 6. The possibilities for modification of the  $\alpha 4$ - $\alpha 5$ ,  $\alpha 5$ - $\alpha 6$  and  $\alpha 6$ - $\alpha 7$  inter-helix loops are not limited to those given in Tables 4, 5 and 6 below. The aim of the list given in Tables 4, 5 and 6 is to illustrate some of the possibilities for modification without limiting the scope of the invention to these illustrations. Those skilled in the art, being aware of the codons specific for each amino acid according to the organisms, will be able to adapt the teaching described in this example to all the possibilities for modifying the  $\alpha 4$ - $\alpha 5$ ,  $\alpha 5$ - $\alpha 6$  and  $\alpha 6$ - $\alpha 7$  inter-helix loops, in particular to those which are not described in Tables 4, 5 and 6.

	Examples of possible modifications of the $\alpha4-\alpha5$ inter-helix loop of the Cry9Cal toxin											
Protein	Amino acid sequence	Nucl	leot:	ide s	seque	ence						
CryCal	FAVNGQVPLL										tta Leu	
Mutant No. 10	FLLNLFFLPLL										tta Leu	
Mutant No. 11	FLLNLEELPLL										tta Leu	
Mutant No. 12	FEENLEELPLL										tta Leu	_
Mutant No. 13	FEENFLLFPLL										tta Leu	
Mutant No. 14	FEENFEEFPLL										tta Leu	
Mutant No. 15	FLLNFEEFPLL										tta Leu	_
Mutant No. 16	FLLNEFFEPLL								-		tta Leu	

TABLE 4

# [0091]

TABLE 5

Examples of possible modifications of the $\alpha 5-\alpha 6$ inter-helix loop of the Cry9Cal toxin									
Protein	Amino acid sequence Nucleotide sequence								
Cry9Cal	LFGEGWGF		ttt Phe						
Mutant No. 17	LFLELFLF		ttt Phe		-				
Mutant No. 18	LFLLLFLF		ttt Phe						
Mutant No. 19	LFLEEFEL		ttt Phe		-				
Mutant No. 20	LFEEEFEL		ttt Phe						
Mutant No. 21	LFEEEFEE		ttt Phe						

## [0092]

## TABLE 6

Exam	ples of poss	ible mod	ificatio	ns	
of the $\alpha 6-\alpha 7$	inter-helix	loop of	the Cr	y9Cal to	oxin

Protein	Amino acid sequence	Nucleotide sequence
Cry9Ca1	LRGTN	tta aga gga aca aat Leu Arg Gly thr Asn
Mutant No. 22	LLELN	tta TTa gAa TTa aat Leu Leu Glu Leu Asn

TABLE 6-continued

	ples of possible inter-helix loop	modifications o of the Cry9Cal toxin
Protein	Amino acid sequence	Nucleotide sequence
Mutant No. 1	23 LLFLN	tta TTa TTT TTa aat Leu Leu Phe Leu Asn
Mutant No. 3	24 LELLN	tta GAa TTa TTa aat Leu Glu Leu Leu Asn
Mutant No.	25 LLFFN	tta TTa TTT TTT aat Leu Leu Phe Phe Asn
Mutant No. 1	26 LEELN	tta GAa GAa TTa aat Leu Glu Glu Leu Asn
Mutant No. 1	27 LEFLN	tta GAa TTT TTa aat Leu Glu Phe Leu Asn
Mutant No. 3	28 LEFEN	tta GAa TTT GAa aat Leu Glu Phe Glu Asn
Mutant No. 3	29 LEEEN	tta GAa gAa GAa aat Leu Glu Glu Glu Asn

## 3-1-Creation of Pepsin Cleavage Sites in the $\alpha$ 4- $\alpha$ 5 Inter-helix Loop

**[0093]** The substitution of several amino acids within the  $\alpha$ 4- $\alpha$ 5 inter-helix loop requires, for each of the mutants, the successive use of several mutagenesis oligonucleotides. The mutagenesis oligonucleotides required to create the examples of mutants given in Table 4 are presented below (numbered from 21 to 34).

Oligonucleotide No. 21: gct att cca ttg ttt TTa Tta aat gga cag cag gtt Ala Ile Pro Leu Phe Leu Ile Asn Gly Gln Gln Val Oligonucleotide No. 22: gct att cca ttg ttt GAa GAa aat gga cag cag gtt Ala Ile Pro Leu Phe Glu Glu Asn Gly Gln Gln Val Oligonucleotide No. 23: tta tta aat gga cag cag TtA cca tta ctg tca gta Leu leu Ann Gly Gln Gln Leu Pro Leu Leu Ser Val Oligonucleotide No. 24: tta tta aat gga cag cag Ttt cca tta ctg tca gta Leu leu Asn Gly Gln Gln Phe Pro Leu Leu Ser Val Oiigonucleotide No. 25: tta tta aat gga cag cag gAA cca tta ctg tca gta Leu leu Asn Gly Gln Gln Glu Pro Leu Leu Ser Val Oligonucleotide No. 26: gaa gaa aat gga cag cag TtA caa tta ctg tca gta Glu Glu Asn Gly Gln Gln Leu Pro Lou Leu Ser Val Oligonucleotide No. 27: gaa gaa aat gga cag cag Ttt cca tta ctg tca gta Glu Glu Asn Gly Gln Gln Phe Pro Leu Leu Ser Val Oligonucleotide No. 28: cca ttg ttt tta tta aat TTa TTT TTT tta cca tta ctg tca gta Pro Lou Phe Lou Leu Asn Leu Phe Phe Leu Pro Leu Leu Ser Val Oligonucleotide No. 29: cca ttg ttt tta tta aat TTa GaA GaA tta cca tta ctg tca gta Pro Leu Phe Leu Leu Asn Leu Glu Glu Leu Pro Leu Leu Ser Val Oligonucleotide No. 30: cca ttg ttt gaa gaa aat TTa GaA GaA tta cca tta ctg tca gta Pro Leu Phe Glu Glu Asn Leu Glu Glu Leu Pro Leu Leu Ser Val Oligonucleotide No. 31: cca ttg ttt gaa gaa aat TTT TTA TTA ttt cca tta ctg tca gta Pro Leu Phe Glu Glu Asn Phe Leu Leu Phe Pro Leu Leu Ser Val Oligonucleotide No. 32: cca ttg ttt gaa gaa aat TTT GaA GaA ttt cca tta ctg tca gta Pro Leu Phe Glu Glu Asn Phe Glu Glu Phe Pro Leu Leu Ser Val Oligonucleotide No. 33: cca ttg ttt tta tta aat TTT GaA GaA ttt cca tta ctg tca gta Pro Leu Phe Leu Leu Asn Phe Glu Glu Phe Pro Leu Leu Ser Val Oligonucleotide No. 34: cca ttg ttt tta tta sat GAa TTT TTT gaa cca tta ctg tca gta Pro Leu Phe Leu Leu Asn Glu Phe Phe Glu Pro Leu Leu Ser Val

**[0094]** The successive site-directed mutagenesis procedure is similar to the procedure described in Example 2. The only difference lies in the combination of oligonucleotides. For each of the mutants described in Table 4, the successive combinations of oligonucleotides are described below.

[0095] Mutant No. 10: The creation of mutant No. 10 requires three successive series of site-directed mutageneses according to the protocol described below, using oligonucleotide No. 21 in the first mutagenesis, oligonucleotide No. 23 in the second and oligonucleotide No 28 in the third. Oligonucleotide No. 23 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 21 and oligonucleotide No. 28 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 21 and oligonucleotide No. 28 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 21 and 23. **[0096]** Mutant No. 11: The creation of mutant No. 11 requires three successive series of site-directed mutageneses according to the protocol described below, using oligonucleotide No. 21 in the first mutagenesis, oligonucleotide No. 23 in the second and oligonucleotide No 29 in the third. Oligonucleotide No. 23 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 21 and oligonucleotide No. 29 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 21 and 23.

**[0097]** Mutant No. 12: The creation of mutant No. 12 requires three successive series of site-directed mutageneses according to the protocol described below, using oligonucleotide No. 22 in the first mutagenesis, oligonucleotide No. 26 in the second and oligonucleotide No 30 in the third. Oligonucleotide No. 26 is defined to recognize the modifi-

13

cations introduced during the first mutagenesis with oligonucleotide No. 22 and oligonucleotide No. 30 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 22 and 26.

**[0098]** Mutant No. 13: The creation of mutant No. 13 requires three successive series of site-directed mutageneses according to the protocol described below, using oligonucleotide No. 22 in the first mutagenesis, oligonucleotide No. 27 in the second and oligonucleotide No 31 in the third. Oligonucleotide No. 27 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 22 and oligonucleotide No. 31 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 22 and 27.

**[0099]** Mutant No. 14: The creation of mutant No. 14 requires three successive series of site-directed mutageneses according to the protocol described below, using oligonucleotide No. 22 in the first mutagenesis, oligonucleotide No. 27 in the second and oligonucleotide No 32 in the third. Oligonucleotide No. 27 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 22 and oligonucleotide No. 32 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 22 and 27.

**[0100]** Mutant No. 15: The creation of mutant No. 15 requires three successive series of site-directed mutageneses according to the protocol described below, using oligonucleotide No. 21 in the first mutagenesis, oligonucleotide No. 24 in the second and oligonucleotide No 33 in the third. Oligonucleotide No. 24 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 21 and oligonucleotide No. 33 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 21 and 24. **[0101]** Mutant No. 16: The creation of mutant No. 16 requires three successive series of site-directed mutageneses according to the protocol described below, using oligonucleotide No. 21 in the first mutagenesis, oligonucleotide No. 25 in the second and oligonucleotide No 34 in the third. Oligonucleotide No. 25 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 21 and oligonucleotide No. 34 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 21 and 25.

**[0102]** According to this protocol, the oligonucleotides intended to create the mutants No. 10 to No. 16 described in Table 4 are divided up into three categories, 1st series oligonucleotides, 2nd series oligonucleotides and 3rd series oligonucleotides. This division is as follows:

1st series oligonucleotides:oligonucleotides No. 21 and 222nd series oligonucleotides:oligonucleotides No. 23, 24, 25, 26 and 273rd series oligonucleotides:oligonucleotides No. 28, 29, 30, 31, 32, 33and 34.

## 3-2-Creation of Pepsin Cleavage Sites in the $\alpha$ 5- $\alpha$ 6 Inter-helix Loop

**[0103]** The substitution of several amino acids within the  $\alpha$ 5- $\alpha$ 6 inter-helix loop requires, for each of the mutants, the successive use of several mutagenesis oligonucleotides. The mutagenesis oligonucleotides required to create the examples of mutants given in Table 5 are presented below (numbered from 35 to 44).

Oligonucleotide No. 35:	gat gca	tct ctt	ttt TTa	gaa gga	tgg gga	ttc	
	Asp Ala	Ser Leu	Phe Leu	Glu Gly	Trp Gly	Phe	
Ollgonucleotide No. 36:	gat gca	tct ctt	ttt TTa	TTa gga	tgg gga	ttc ac	a
	Asp Ala	Ser Leu	Phe Leu	Leu Gly	Trp Gly	Phe Th	c
Oligonucleotide No. 37:	gat gca	tct ctt	ttt gAa	gaa gga	tgg gga	ttc	
	Asp Ala	Ser Leu	Phe Glu	Glu Gly	Trp Gly	Phe	
Oligonucleotide No. 38:	tta gaa	gga tgg	gga TTa	aca sag	ggg gaa	att	
	Leu Glu	Gly Trp	Gly Leu	Thr Gln	Gly Glu	Ile	
Oligonucleotide No. 39:	gga gaa	gga tgg	gga GAA	. aca sag	ggg gaa	att	
	Gl <b>y</b> Glu	Gly Trp	Gay Glu	Thr Gln	Gay Glu	Ile	
Oligonucleotide No. 40:	gca tct	ctt ttt	tta gaa	TTa tTT	TTa ttc	aca ca	g gqg gaa att
	Ala Ser	Leu Phe	Leu Glu	Leu Phe	Leu Phe	Thr Gl	n Gly Glu Ile
Oligonucleotide No. 41:	gca tot	ctt ttt	tta tta	TTa tTT	TTa ttc	aca ca	g ggg gaa att
	Ala Ser	Leu Phe	Leu Leu	Leu Phe	Leu Phe	Thr Gl	n Gly Glu Ile
Oligonucleotide No. 42:	gca tct	ctt ttt	tta gaa	TTa tTT	TTa ttc	aca ca	g ggg gaa att
	Ala Ser	Leu Phe	Leu Glu	Glu Phe	Gln Leu	Thr Gl	n Gl <b>y</b> Glu Ile

#### -continued

Qligonucleotide No. 43: gca tct ctt ttt gaa gaa TTa tTT TTa ttc aca cag ggg gaa att Ala Ser Leu Phe Glu Glu Glu Glu Leu Thr Gln Gly Glu Ile Oligonucleotide No. 44: gca tct ctt ttt gaa gaa TTA tTT TTa gaa aca cag ggg gaa att Ala Ser Leu Phe Glu Glu Glu Glu Glu Glu Thr Gln Gly Glu Ile

**[0104]** The successive site-directed mutagenesis procedure is similar to the procedure described in Example 2. The only difference lies in the combination of oligonucleotides. For each of the mutants described in Table 5, the successive combination of oligonucleotides are described below.

**[0105]** Mutant No. 17: The creation of mutant No. 17 requires two successive series of site-directed mutageneses according to the protocol described below, using oligonucleotide No. 35 in the first mutagenesis and oligonucleotide No. 40 in the second. Oligonucleotide No. 40 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 35.

**[0106]** Mutant No. 18: The creation of mutant No. 18 requires two successive series of site-directed mutageneses according to the protocol described below, using oligonucleotide No. 36 in the first mutagenesis and oligonucleotide No. 41 in the second. Oligonucleotide No. 41 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 36.

**[0107]** Mutant No. 19: The creation of mutant No. 19 requires three successive series of site-directed mutageneses according to the protocol below, using oligonucleotide No. 35 in the first mutagenesis, oligonucleotide No. 38 in the second and oligonucleotide No. 42 in the third. Oligonucleotide No. 38 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 35 and oligonucleotide No. 42 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 35 and 38.

**[0108]** Mutant No. 20: The creation of mutant No. 20 requires three successive series of site-directed mutageneses according to the protocol below, using oligonucleotide No. 37 in the first mutagenesis, oligonucleotide No. 38 in the second and oligonucleotide No 43 in the third. Oligonucleotide No. 38 is defined to recognize the modifications

introduced during the first mutagenesis with oligonucleotide No. 37 and oligonucleotide No. 43 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 37 and 38.

**[0109]** Mutant No. 21: The creation of mutant No. 21 requires three successive series of site-directed mutageneses according to the protocol below, using oligonucleotide No. 37 in the first mutagenesis, oligonucleotide No. 39 in the second and oligonucleotide No 44 in the third. Oligonucleotide No. 39 is defined to recognize the modifications introduced during the first mutagenesis with oligonucleotide No. 37 and oligonucleotide No. 44 is defined to recognize the modifications introduced during the first two mutageneses with oligonucleotides No. 37 and 39.

**[0110]** According to this protocol, the oligonucleotides intended to create mutants No. 17 to No. 21 described in Table 5 are divided up into three categories, 1st series oligonucleotides, 2nd series oligonucleotides and 3rd series oligonucleotides. This division is as follows:

1st series oligonucleotides:	oligonucleotides No. 35, 36 and 37
2nd series oligonucleotides:	oligonucleotides No. 38, 39, 40 and 41
3rd series oligonucleotides:	oligonucleotides No. 42, 43 and 44.

#### 3-3-Creation of Pepsin Cleavage Sites in the α6-α7 Inter-helix Loop

**[0111]** The substitution of several amino acids within the  $\alpha 6-\alpha 7$  inter-helix loop requires, for each of the mutants, only one mutagenesis. The mutagenesis oligonucleotides required to create the examples of mutants given in Table 6 are presented below (numbered from 45 to 52).

Oligonucleotide No.	45:	ggt	tta	gat	cgt	tta	TTa	gAa	TTa	aat	act	gaa	agt	tgg
		Gly	Leu	Asp	Arg	Leu	Leu	Glu	Leu	Asn	Thr	Glu	Ser	Trp
Oligonucleotide No.	46:	ggt	tta	gat	cgt	tta	TTa	TTT	TTa	aat	act	gaa	agt	tgg
		Gly	Leu	Asp	Arg	Leu	Leu	Phe	Leu	Asn	Thr	Glu	Ser	Trp
Oligonucleotide No.	47 <b>:</b>	ggt	tta	gat	cgt	tta	GAa	TTa	TTa	aat	act	gaa	agt	tgg
		Gly	Leu	Asp	Arg	Leu	Glu	Leu	Leu	Asn	Thr	Glu	Ser	Trp
Oligonucleotide No.	48:	ggt	tta	gat	cgt	tta	TTa	TTT	TTT	aat	act	gaa	agt	tgg
		Gly	Leu	Asp	Arg	Leu	Leu	Phe	Phe	Asn	Thr	Glu	Ser	Trp
Oligonucleotide No.	49:	ggt	tta	gat	cgt	tta	GAa	GAa	TTa	aat	act	gaa	agt	tgg

-continued Cly Leu Asp Arg Leu Glu Glu Leu Asn Thr Glu Ser Trp Oligonucleotide No. 50: ggt tta gat cgt tta GAa TTT TTa aat act gaa agt tgg Cly Leu Asp Arg Leu Glu Phe Leu Asn Thr Glu Ser Trp Oligonucleotide No. 51: ggt tta gat cgt tta GAa TTT GAa aat act gaa agt tgg Cly Leu Asp Arg Leu Glu Phe Glu Asn Thr Glu Ser Trp Oligonucleotide No. 52: ggt tta gat cgt tta GAa gAa GAa aat act gaa agt tgg Cly Leu Asp Arg Leu Glu Glu Glu Gln Asn Thr Glu Ser Trp Oligonucleotide No. 45 is used to create mutant No. 22.

Oligonucleotide No. 46 is used to create mutant No. 23. Oligonucleotide No. 47 is used to create mutant No. 24. Oligonucleotide No. 48 is used to create mutant No. 25. Oligonucleotide No. 49 is used to create mutant No. 26. Oligonucleotide No. 50 is used to create mutant No. 27. Oligonucleotide No. 51 is used to create mutant No. 28. Oligonucleotide No. 52 is used to create mutant No. 29.

**[0112]** The complete protocol for producing these mutants is identical to that described in Example 2. This protocol is common to each of the series of mutageneses, only the mutagenesis oligonucleotide and the oligonucleotide for inhibition/restoration of the resistance to the antibiotic change.

#### Example 4:

## Creation of Pepsin Cleavage Sites in the α3-α4, α4-α5, α5-α6 and α6-α7 Inter-helix Loops of Various Cry Toxins

**[0113]** Several groups of Cry proteins exhibit structural similarities. They are in particular the proteins belonging to the Cry1, Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 or Cry20 families. These similarities are demonstrated in the literature (Schnepf et al., 1998). Other Cry proteins not cited in the literature can also exhibit structural and sequence similarities with these families. The aim of Example 4 is to demonstrate the applicability of the teaching of the present invention, as exemplified on the Cry9Ca1 protein in Examples 2 and 3, to all these structurally similar families.

[0114] The modifications in the inter-helix loops described in Examples 2 and 3 can be carried out in an equivalent manner for all the Cry proteins in which it is possible to identify inter-helix loops similar to those present in domain I of the Cry9Ca1 toxin. If the location and the sequence of these inter-helix loops are defined for these various Cry toxins, it is very easy for those skilled in the art to form modifications similar to those given in Examples 2 and 3 using the technical details provided in these same Examples 2 and 3. In the present example, the elements for creating specific sites for degradation by pepsin in the Cry toxins other than the Cry9Ca1 toxin, and in particular the Cry1, Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 proteins, are given. The modification of these inter-helix loops to create sites for degradation by pepsin in the Cry1, Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 or Cry20 toxins requires the following steps to be followed:

**[0115]** 1) Establish, according to the sequences and the locations of the inter-helix loops given in Tables

6-13 below, lists of possible mutants having one or more leucine, phenylalanine or glutamic acid residues as given in Tables 1, 4, 5 and 6 and in Examples 2 and 3.

- **[0116]** 2) Establish the sequences of the mutant genes taking into account the codon preference of the host organism and, if this organism is *B. thuringiensis*, preferentially using the codons TTA, TTT and GAA for leucine, phenylalanine and glutamic acid, respectively.
- **[0117]** 3) Synthesizing mutagenesis oligonucleotides for modifying the sequence of the genes encoding the toxins selected based on the model for those given in Examples 2 and 3.
- **[0118]** 4) Use single or multiple mutagenesis strategies as described in Examples 2 and 3 and according to the experimental protocols described in detail in Examples 2 and 3.

**[0119]** The location of the  $\alpha 3$ - $\alpha 4$ ,  $\alpha 4$ - $\alpha 5$ ,  $\alpha 5$ - $\alpha 6$  and  $\alpha 6$ - $\alpha 7$  inter-helix loops of domain I and their sequences are given, for the Cry1, Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 toxins, in Tables 7, 8, 9, 10, 11, 12 and 13 below. These sequences are given for each of the holotype proteins as defined by the *Bacillus thuringiensis* classification committee (Crickmore et al., 2001). However, since the intra-holotype sequence homologies, i.e. the sequence homologies between the various subtypes of the same holotype, are very high, those skilled in the art will be able to adapt the teaching of the present Example 4 to all the Cry protein subtypes.

TABLE 7

Location and sequence of the $\alpha$ 3- $\alpha$ 4 inter-helix loop in the Cry1 proteins								
Protein	Amino acid sequence	Position in protein	Nucleotide sequence	Position in gene				
CrylAa CrylAb CrylAc	DPTN DPTN DPTN	120 to 123	gatectactaat gatectactaat gatectactaat	358 to 369 358 to 369 358 to 369				

Location and sequence of the $\alpha$ 3– $\alpha$ 4 inter-helix loop in the Cry1 proteins								
Protein	Amino acid sequence	Position in protein	Nucleotide sequence	Position in gene				
CrylAd	DPTN	120 to 123	gatectactaat	358 to 369				
CrylAe	DPTN	120 to 123	gatectactaat	358 to 369				
CrylAf	DPTN	120 to 123	gatectactaat	358 to 369				
CrylAg	DPTN	120 to 123	gatectactaat	358 to 369				
CrylBa	NRDD	139 to 142	aaccgtgatgat	415 to 426				
CrylBb	NRND	144 to 147	aaccgaaatgat	430 to 441				
CrylBc	NRND	144 to 147	aaccgaaatgat	430 to 441				
CrylBd	NRND	144 to 147	aaccgaaatgat	430 to 441				
CrylCa	DPNN	119 to 122	gatectaataat	355 to 366				
CrylCb	DPDN	119 to 122	gateetgataat	355 to 366				
CrylDa	DPTN	119 to 122	gatectactaat	355 to 366				
CrylDb	DPSN	119 to 122	gateegtetaat	355 to 366				
CrylEa	DPTN	118 to 121	gatectactaat	352 to 363				
CrylEb	DPTN	117 to 120	gatectactaat	349 to 360				
CrylFa	NPNN	118 to 121	aatcctaataat	352 to 363				
CrylFb	NPNN	118 to 121	aateetaataat	352 to 363				
CrylGa	DPNN	118 to 121	gatectaataat	352 to 363				
CrylGb	DPDN	118 to 121	gatectgataac	352 to 363				
CrylHa	SPNN	122 to 125	tetectaataat	364 to 375				

TABLE 7-continued

Location and sequence of the  $\alpha 3-\alpha 4$ 

Protein	Amino acid sequence	Position in protein	Nucleotide sequence	Position in gene
CrylHb	SPNN	121 to 124	tctcctaataat	361 to 372
CrylIa	NRNN	148 to 151	aatcgtaataac	442 to 453
CrylIb	NRNN	148 to 151	aatcgtaataac	442 to 453
CrylIc	NRNN	148 to 151	aatcgtaataac	442 to 453
CrylId	NRNN	148 to 151	aatcgcaataac	442 to 453
CrylIe	NRNN	148 to 151	aatcgcaacaac	442 to 453
CrylJa	DPDN	119 to 122	gatcctgataac	355 to 366
CrylJb	TPDN	119 to 122	actecagataac	355 to 366
CrylKa	NRND	145 to 148	aaccgaaatgat	433 to 444

# [0120]

## TABLE 8

Location and sequence of the  $\alpha4-\alpha5$  inter-helix loop in the Cry1 proteins

Protein	Amino acid sequence	Position in protein	n Nucleotide sequence	Position in gene
Cry1Aa	LAVQNYQVPLL	148 to 158	ttggcagttcaaaattatcaagttcctcttta	442 to 474
	FLAVQNYQVPLL	148 to 158	ttgcagttcaaaattatcaagttcctctttta	442 to 474
Cry1Ab	FAVQNYQVPLL	148 to 158	ttgcagttcaaaattatcaagttcctcttta	442 to 474
Cry1Ac	FAVQNYQVPLL	148 to 158	tttgcagttcaaaattatcaagttcctcttta	442 to 474
	LAVQNYQVPLL	148 to 158	ttggcagtcaaaattatcaagttcctcttta	442 to 474
Cry1Ad	FTVQNYQVPLL	148 to 158	cctacagttcaaaattatcaagtacctcttcta	442 to 474
CrylAe	FTVQNYQVPLL	148 to 158	tttacagttcaaaattatcaagtacctcttcta	442 to 474
Cry1Af	FAVQNYQVPLL	148 to 158	tttgcagttcaaaattatcaagttcctcttta	442 to 474
Cry1Ag	LAVQNYQVPLL	148 to 158	ttggcagttcaaaattatcaagttcctcttta	442 to 474
Cry1Ba	FAIRNQEVPLL	167 to 177	ttcgcaattagaaaccaagaagttccattattg	499 to 531
Cry1Bb	FRIRNEEVPLL	172 to 182	ttcagaatacgaaatgaagaagttccattatta	514 to 546
Cry1Bc	FRIRNEEVPLL	172 to 182	ttcagaatacgaaatgaagaagttccattatta	514 to 546
Cry1Bd	FRIRNEEVPLL	172 to 182	ttcagaatacgaaatgaagaagttccattatta	514 to 546
Cry1Ca	FRISGFEVPLL	147 to 157	tttcgaatttctggatttgaagtacccctttta	439 to 471
Cry1Cb	FRIAGFEVPLL	147 to 157	tttcgaattgctggatttgaagtacccctttta	439 to 471
Cry1Da	FRVQNYEVALL	147 to 157	tttagagttcaaaattatgaagttgctctttta	439 to 471
Cry1Db	LRVRNYEVALL	147 to 157	ttaagagttcgtaattatgaagttgctctttta	439 to 471
CrylEa	LFSVQNYQVPFL	145 to 156	aattttacacttacaagttttgaaatccctctttta	433 to 468
Cry1Fb	NFTLTSFEIPLL	145 to 156	aattttacacttacaagttttgaaatccctctttta	433 to 468
Cry1Ga	RLAIRNLEVVNL	145 to 156	actttggcaattcggaatcttgaggtagtgaattta	433 to 468

-	Location and s	equence of the $\alpha4-\alpha5$ inter-helix loop in the Cry1 proteins	5_
Protein	Amino acid sequence	Position in protein Nucleotide sequence Position	in gene
Cry1Gb	LMAIPGPELATL	145 to 156 cttatggcaattccaggtttgaattagctactttta 433 to	5 468
Cry1Gb	LMAIPGPELATL	145 to 156 cttatggcaattccaggttttgaattagctacttta 433 to	o 468
Cry1Ha	LREQGFEIPLL	150 to 160 ctgagagaacaaggctttgaaattcctcttta 448 to	5 480
Cry1Hb	LREQGFEIPLL	149 to 159 ctgagagaacagggctttgaaattcctcttta 445 to	o 477
Cry1Ia	FAVSGEEVPLL	176 to 186 tttgcagtgtctggagaggaggtaccattatta 526 to	558
Cry1Ib	FAVSGEEVPLL	176 to 186 tttgcagtatctggtgaggaagtaccattatta 526 to	558
Cry1Ic	FAVSGEEVPLL	176 to 186 tttgcagtatctggtgaggaagtaccattatta 526 to	558
Cry1Id	FAVSGEEVPLL	176 to 186 tttgcagtttctggagaagaggtgccgctatta 526 to	558
Crylle	FAVSGEEVPLL	176 to 186 tttgcagtatcaggtgaggaagtaccattattg 526 to	558
Cry1Ja	FRIIGFEVPLL	147 to 157 tttcggataattggatttgaagtgccactttta 439 to	o 471
Cry1Jb	FRIPGFEVPLL	147 to 157 tttcggattcccggatttgaagtgccacttcta 439 to	o 471
Cry1Ka	FSIRNEEVPLL	173 to 183 ttcagcatacgaacgaagaggttccattattta 517 to	549

## TABLE 8-continued

# [0121]

# TABLE 9

Locatio	on and seque	nce of t	he (	$\alpha 5-\alpha 6$ inter-helix loop in t	he Cryl proteins
Protein	Amino acid sequence	Position prote		Nucleotide sequence	Position in gene
Cry1Aa	FGQRWGFD	178 to	185	tttggacaaaaggtggggatttgat	532 to 555
Cry1Ab	FGQRWGFD	178 to	185	tttggacaaaggtggggatttgat	532 to 555
Cry1Ac	FOQRWGPD	178 to	185	tttggacaaaggtggggatttgat	532 to 555
Cry1Ad	FGQRWGFD	178 to	185	ttggacaacgttgggggatttgat	532 to 555
Cry1Ae	FGQRWGLD	178 to	185	tttggacaacgttggggacttgat	532 to 555
Cry1Af	CGQRSGFD	175 to	182	tgtggacaaaggtcgggatttgat	523 to 546
Czy1Ag	FGQRWGFD	178 to	185	tttggacaaaggtggggatttgat	532 to 555
Cry1Ba	FGSEFGLT	197 to	204	ttggtagtgaatttgggcttaca	589 to 612
Cry1Bb	FGSEWGMA	202 to	209	tttggtagtgaatgggggatggca	604 to 627
Cey1Bc	FGSEWGMA	202 to	209	tttggtagtgaatgggggatggca	604 to 627
Cry1Bd	FGSEWGMA	202 to	209	tttggtagtgaatgggggatggca	604 to 627
Cry1Ca	FGERWGLT	177 to	184	ttggagaaagatggggattgaca	529 to 552
	FGERWGVT	177 to	184	ttggagaaagatgggggggggggagtgaca	529 to 552
Cry1Cb	FGARWGLT	177 to	184	tttggagcaagatggggattgaca	529 to 552
Cry1Da	FGERWGYD	177 to	184	ttcggagaaagatggggatatgat	529 to 552
Cry1Db	YGQRWGFD	177 to	184	tacggtcagagatggggctttgac	529 to 552
Cry1Ea	FGQAWGFD	176 to	183	tttgggcaggcttggggatttgat	526 to 549

Locatio	Location and sequence of the $\alpha 5-\alpha 6$ inter-helix loop in the Cry1 proteins					the Cryl proteins
Amino acid Position in Protein sequence protein Nucleotide sequence Position in gene						Position in gene
Cry1Eb	FGQRWGFD	175	to	182	tttggacaacgttgggggatttgat	523 to 546
Cry1Fa	FGQGWGLD	176	to	183	tttgggcagggttggggactggat	526 to 549
Cry1Fb	FGQGWGLD	176	to	183	tttgggcagggttggggctggat	526 to 549
Cr <b>y</b> 1Ga	FGERWGLT	176	to	183	tttggagaaagatggggattaaca	526 to 549
Cry1Gb	FGERWGLT	176	to	183	tttggggagagatggggattgaca	526 to 549
Cr <b>y</b> 1Ha	FGQRWGLD	180	to	187	tttgggcaaagatggggacttgac	538 to 561
Cr <b>y</b> 1Hb	FGQRWGLD	179	to	186	tttggacagagatggggacttgat	535 to 558
Cry1Ia	FGKEWGLS	206	to	213	tttggaaaagagtggggattatca	616 to 639
Cr <b>y</b> 1Ib	FGKEWGLS	206	to	213	tttgaaagaatggggattatca	616 to 639
Cr <b>y</b> 1Ic	FEKNGGLS	206	to	213	ttgaaaagaatggggggattatca	616 to 639
Cr <b>y</b> 1Id	FGKEWGLS	206	to	213	tttggaaaagaatggggggattgtca	616 to 639
Cr <b>y</b> 1Ie	FGKEWGLS	206	to	213	tttggaaaagagtggggattatct	616 to 639
Cr <b>y</b> 1Ja	FGERWGLT	177	to	184	ttggagagagatggggattgacg	529 to 552
Cr <b>y</b> 1Jb	FGERWGLT	177	to	184	ttcggagagagatggggattgacg	529 to 552
Cry1Ka	FGSEWGMS	203	to	210	tttggtagtgaatgggggatgtca	607 to 630

TABLE 9-continued

[0122]

TABLE 10

	Location			e of a6-a7 inter-hel: Cry1 proteins	ix loop
Protein	Amino acid sequence	Positic prote		n Nucleotide sequence	Position in gene
Cry1Aa	VWGPD	218 to	222	gtatggggaccggat	652 to 666
Cry1Ab	VWGPD	218 to	222	gtatggggaccggat	652 to 666
Cry1Ac	VWGPD	218 to	222	gtatggggaccggat	652 to 666
Cry1Ad	VWGPD	218 to	222	gtatggggaccggaa	652 to 666
Cry1Ae	VWGPD	218 to	222	gtatggggaccggat	652 to 666
Cry1Af	VWGPD	215 to	219	gtatggggaccggat	643 to 657
Cry1Ag	VWGPD	218 to	222	gtatggggacecgac	652 to 666
Cr <b>y</b> 1Ba	LRGTN	237 to	241	ttgagagggacaaaa	709 to 723
Cry1Bb	LRGTN	242 to	246	ttaagagggacaaat	724 to 738
Cry1Bc	LRGTN	242 to	246	ttaagagggacaaat	724 to 738
CrylBd	LRGTN	242 to	246	ttaagagggacaaat	724 to 738
Cry1Ca	LPKST	217 to	221	ttaccgaaatctacg	649 to 663
Cry1Cb	LPKST	217 to	221	ttaccaaaatcracg	649 to 663
Cr <b>y</b> 1Da	LEGRF	217 to	221	ttaggaaggtcgtttt	649 to 663

	Location		ence of a6-a7 inter-hel he Cry1 proteins	ix loop
Protein	Amino acid sequence	Position protei		Position in gene
Cr <b>y</b> 1Db	LEGSR	217 to 2	221 <b>ttagagggatctcga</b>	649 to 663
Cry1Ea	LPRTGG	216 to 2	221 ttaccacgaactggtggg	646 to 663
Cry1Eb	LPRNEG	215 to 2	220 ttaccacgtaatgaaggg	643 to 660
CrylFa	LRGTNT	216 to 2	221 ataagaggtaataatact	646 to 663
Cry1Fb	LRGTNT	216 to 2	221 ttaagaggtactaatact	646 to 663
Cr <b>y</b> 1Ga	IGGIS	216 to 2	220 <b>attggagggataagt</b>	646 to 660
Cr <b>y</b> 1Gb	LNVIR	216 to 2	220 <b>ttaaatgttataaga</b>	646 to 660
Cry1Ha	FGGVS	220 to 2	224 tttggtggtgtgtca	658 to 672
Cr <b>y</b> 1Hb	FGVVT	219 to 2	223 <b>tttggtgttgtaaca</b>	655 to 669
Cr <b>y</b> 1Ia	LRGTN	246 to 2	250 <b>ttgaggggtacaaat</b>	736 to 750
Cr <b>y</b> 1Ib	LRGTN	246 to 2	250 <b>ttgaggggtacaaat</b>	736 to 750
CrylIc	LRATN	246 to 2	250 <b>ttgagggctacaaat</b>	736 to 750
Cr <b>y</b> 1Id	LRGTN	246 to 2	250 <b>ttgaggggaacaaat</b>	736 to 750
CrylIe	LRGTN	246 to 2	250 <b>ttgagaggtacaaat</b>	736 to 750
Cr <b>y</b> 1Ja	LGPRS	217 to 2	221 cagggtttagatct	649 to 663
Cry1Jb	LGFTS	217 to 2	221 ctagggtttacttct	649 to 663
Cry1Ka	LRGTT	243 to 2	247 ttaagagggacaact	727 to 741

TABLE 10-continued

# [0123]

# TABLE 11

Location and	sequence of the $\alpha_3-\alpha_4$ inter-helix loop
in the Cry3, Cry4,	Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19
	and Cry20 proteins

Protein	Amino acid sequence	Position in protein Nucleotide sequence	Position in gene
Cry3Aa	NPVSSRN	153 to 159 <b>aatcctgtgagulcacgaaat</b>	457 to 477
Cry3Ba	APVNLRS	154 to 160 gcgcctgtaaatttacgaagt	460 to 480
Cry3Eb	TPLSLRS	154 to 160 acacctttaagtttgcgaagt	440 to 480
Cry3Ca	TPLTLRD	151 to 157 actcctttgactttcgagat	451 to 471
Cry4Aa	NNPNPQNTQD	160 to 169 aaataatcaaacccacaaaatactcaggat	478 to 501
Cry4Ba	EPNNQSYRTA	136 to 145 gagcctaataaaccagtcctatagaacagca	406 to 435
Cry7Aa	KQDDPEAILS	147 to 156 aaacaagatgatccagaagctatactttct	439 to 468
Cry7Ab	NPDDPATITR	147 to 156 aatcctgatgaccagcaactataacacga	439 to 468
Cry8Aa	NRNDARTRSV	158 to 167 <b>aatcgcaatgatgcaagaactaagtgtt</b>	472 to 501
Czy8Ba	NPNGSRALRD	159 to 168 <b>aatccaaatggttcaagagccttacgagat</b>	415 to 504

	Location and sequence of the a3-a4 inter-helix loop in the Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 proteins				
Protein	Amino acid sequence	Position in protein Nucleotide sequence	Position in gene		
Cry8Ca	NPHSTRSAAL	159 to 168 aacccacagtacacgaagcgcagcactt	475 to 504		
Cry9Aa	NPNSASAEEL	146 to 155 <b>aatcctaattctgcttctgctgaagaactc</b>	436 to 465		
Cry9Ba	RPNGVRANLV	134 to 143 agaccaacggcgtaagagcaaacttagtt	400 to 429		
Cry9Ca	DRNDTRNLSV	159 to 168 gatcgaaacgatacacgaaatttaagtgtt	475 to 504		
Cry9Da	RPNGARASLV	159 to 168 agaccaaatggcgcaagggcatccttagtt	475 to 504		
Cry9Ea	RPNGARANLV	159 to 168 <b>agaccgaacggagcaagagctaacttagtt</b>	475 to 504		
Cry10Aa	ARTHANAKAV	162 to 171 gcacgtacacacgctaatgctaaagcagta	484 to 513		
Cry16Aa	NYNPTSIDDV	109 to 118 <b>aattataatccaacttctatagacgatgta</b>	325 to 354		
Cry17Aa	NKDDPLAIAEL	127 to 131 aataaagatgaccccttggctatagctgaatta	379 to 411		
Cry19Aa	DPKSTGNLSTL	159 to 169 gatccaaaatacaggtaatttaagcacctta	475 to 507		
Cr <b>y</b> 19Ba	NKNNFASGEL	151 to 160 <b>aataaaataatttcgcaagtggtgaactt</b>	451 to 480		
Cr <b>y</b> 20Aa	ERNRTRENGQ	141 to 150 gaacgtaatagaactcgtgaaaacggacaa	421 to 450		

# [0124]

TABLE 12

		ration and sequence of the $\alpha4-\alpha5$ inter-helix loop Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 protein	<u>8</u>
Protein	Amino acid sequence	Position in protein Nucleotide sequence	Position in gene
Cry3Aa	ISGYEVL	186 to 192 atttctggatacgaggttcta	556 to 576
CRy3Ba	VSKFEVL	157 to 193 gtttccaaattcgaagttctg	559 to 579
CRY3Bb	VSKFEVL	187 to 193 gtttccaaattcgaagtgctg	559 to 579
Cry3Ca	VSGYEVL	184 to 190 gtctctggatacgaagttcta	550 to 570
Cry4Aa	LVNSCPPNPSDCDYYNILVL	188 to 207 cttgtaaactcttgtcctcctaatcctagtgattgcgattactataacat actagtatta	562 to 621 a
Cry4Ba	FSNLVGYELLLL	164 to 175 <b>tttagcaacttagtaggttatgaattattgttatta</b>	490 to 525
Cry7Aa	FKVTGYEIPLL	175 to 185 <b>tttaaggttactggatatgaaataccattacta</b>	523 to 555
Cry7Ab	FRVAGYEIPLL	175 to 185 <b>tttagggttgctggatatgaaataccattacta</b>	523 to 555
Cry8Aa	FAVSGHEVLLL	136 to 196 tttgcagtatccggacacgaagtactattatta	556 to 588
Cry8Ba	FRVTNFEVPFL	187 to 197 <b>tttcgagtgacaaauttgaagtaccatttcctt</b>	559 to 591
Cry8Ca	FSQTNYETPLL	187 to 197 <b>ttttacaaacgaattatgagactccactctta</b>	559 to 591
Cry9Aa	LTNGGSLARQNAQILLL	175 to 191 <b>ttaacgaatggtggtctgttagctagacaaaatgcccaaatattattat</b>	523 to 571 a
Cry9Ba	FGSGPGSQRFQAQLL	161 to 175 <b>tttggtagtggccctggaagtcaaaggtttcaggcacaatgttg</b>	481 to 525
Cry9Ca	FAVNGQQVPLL	187 to 197 tttgcagtaaatggcacagcaggttccattactg	559 to 591
Cry9Da	FGSGPGSQRJYATILL	116 to 200 <b>tttggctctggtcctggaagtcaaaattatgcaactatattactt</b>	556 to 600

TABLE	12-continue	d
-------	-------------	---

	in the Cry3,	Location and sequence of the $\alpha4-\alpha5$ inter-helix loop Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 proteins	_	
Protein	Amino acid sequence	Position in protein Nucleotide sequence	Positio in gene	
Cry9Ea	FGTOPOSQIWAVALL	186 to 200 tttggtacgggtcctggtagtcaaagagatgcggtagcgttgttg	556 to 6	00
Cry10Aa	LKNNASYRIPTL	189 to 200 ttaaaaaataatgctagctatcgfaataccaacactc	565 to 6	00
Cry16Aa	FKVKNYEVTVL	136 to 146 <b>tttaaggttaaaaattatgaagtaacagtgtta</b>	406 to 4	38
Cry17Aa	FKRANYEVLLL	155 to 165 <b>tttaaaagggcgaattatgaagtcttactatta</b>	463 to 4	95
Cry19Aa	VNNQOSPOYELLLL	187 to 200 gttaaataatcaggggagtccaggttatgagttacttttattg	559 to 6	00
Cry19Ba	FSLGGYETVLL	180 to 190 <b>ttctcattaggaggttatgaaacagtattatta</b>	538 to 5	70
Cry20Aa	LSRRGFETLLL	173 to 183 ctttctcgcagaggattcgaaactcttttatta	517 to 5	49

# [0125]

TABLE 13

	Location and sequence of the $\alpha$ 5- $\alpha$ 6 inter-helix loop in the Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 proteins							
Protein	Amino acid sequence	Position in protein	Nucleotide sequence	Position in gene				
Cry3Aa	GEEWGYE	215 to 222	ggagaagantggggatacgaa	643 to 663				
Cry3Ba	GEEWGYS	216 to 222	ggagaagaatggggatattct	646 to 666				
Cry3Bb	GEEWGYS	226 to 222	ggagaagaatggggatattct	646 to 666				
Cry3Ca	GTDWGYS	213 to 219	ggaacggattggggatattct	637 to 657				
Cry4Aa	FEAYLKNNRQFDYEL	227 to 241	tttgaagcgtatttaaaaaacaattcgattattttagag	679 to 723				
Cry4Ba	LINAQEWSL	193 to 201	ctcataaatgcacaagaatggtcttta	577 to 603				
	PHKCTRMVY	193 201	cctcataaatgcacaagaatggtctat	577 to 603				
Cry7Aa	GDKWGF	206 to 211	ggagataaatggggattc	616 to 633				
	GDKWEF	206 to 221	ggagataaatgggaattc	616 to 633				
Cry7Ab	GDKWGF	206 to 211	ggagataaatgggguttc	616 to 633				
Cry8Aa	GEEWGF	217 to 222	ggagaagagtggggattt	649 to 666				
Cry8Ba	GEEWGL	218 to 223	ggagaagaatggggattg	652 to 669				
Cry8Ca	GKEWGY	218 to 223	gggaaggaatggggatat	652 to 669				
Cry9Aa	RYGTNWGL	210 to 217	agatatggcactaattgggggcta	628 to 651				
Cry9Bal	KYGARWGL	194 to 201	aagtatggggcaagatggggatc	580 to 603				
Cry9Ca	LFGEGWGF	216 to 223	ctttttggagaaggatggggattc	646 to 669				
Cry9Da	IYGARWGL	219 to 226	atttatggagcagatgggggctg	655 to 678				
Cry9Ea	IYGARWOL	219 to 226	atctaggggcaagatggggactt	655 to 678				
Cry10Aa		219 to 226	acttatggggcaagatggggactt	655 to 673				
	IYGDAWNLYRELGP	265 to 178	${\tt atttatggagatgcatggaatttatatagagaattaggattt}$					
Cry17Aa		184 to 192	cttttaaataagttatagataatttt	550 to 576				
Cry19Aa		219 to 227	atttatggagataaatggtggagcgca	655 to 681				
Cry19Ba		209 to 215	atttacggaaaagaattagga	625 to 645				
Cry20Aa	LYRNQWL	202 to 208	ctttatagaaatcaatggtta	604 to 624				

# [0126]

TABLE 14

Location and sequence of the $\alpha 6-\alpha 7$ inter-helix loop in the Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 proteins							
Protein		Position in protein	Nucleotide sequence	Position in gene			
Cry3Aa Cry3Ba	RGSS RGST		agaggttcatct agaggttcaact	763 to 774 766 to 777			

TABLE 14-continued

Location and sequence of the $\alpha 6-\alpha 7$ inter-helix loop in the Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 proteins								
Protein	Amino acid sequence	Pos: prof	itid	on in		Pos	itio	on in gene
Cry3Bb	RGST	256	to	259	agaggttcaat	766	to	777
Cry3Ca	RGST	253	to	256	agaggttcgact	757	to	768
Cry4Aa	LIKTTPD	274	to	280	ttaattaaaacgacgcctgat	820	to	840
Cry4Ba	LRNKS	235	to	239	cttagaaataaatct	703	to	717
Cry7Aa	LNGST	245	to	249	ttgaacggttccact	733	to	747
Cry7Ab	LNGST	245	to	249	ttgaacggttccact	733	to	747
Cry8Aa	LKGTT	256	to	260	ttgaaggtaccact	766	to	780
Cry8Ba	LKGSS	257	to	261	ttaaaaggctcgagc	769	to	783
Cry8Ca	LRGTG	257	to	261	ttaagaggaacgggt	769	to	783
Cry9Aa	LRQRGTS	252	to	258	ctaagacaacgaggcactagt	754	to	774
Cry9Ba1	LRGTS	236	to	240	ttacgaggaacgagc	706	to	720
Cry9Ca	LRGTN	257	to	261	ttaagaggaacaaat	769	to	783
Cry9Da	LRGTT	260	to	264	ttaagaggcacaacc	778	to	792
Cry9Ea	VRGTN	260	to	264	gtagaggaacaaat	778	to	792
Cry10Aa	1rtnt	267	to	271	attagaactaatact	799	to	813
Cry16Aa	LKLDPN	210	to	215	ttaaaactagatccgaat	628	to	645
Cry17Aa	IKNKTRDF	224	to	231	ataaaaaataaaactagg-	670	to	693
					gatttt			
Cry19Aa				265	ttagaacagcaggt			795
				254	aaaaacaaatagga			162
Cry20Aa	DRSS	245	to	248	gatcgttcaagt	733	to	744

**[0127]** Mutants can be prepared for each of the cry genes mentioned in this example, based on the models of Examples 1, 2 and 3. The technical procedures which can be used to carry out the mutagenesis are similar to those given in Examples 1, 2 and 3.

#### Example 5

### Overall Increase in the Leucine, Phenylalanine and Glutamic Acid Content of the Cry Proteins

**[0128]** The overall increase in the leucine, phenylalanine and glutamic acid content of the Cry proteins is described below for the Cry9Ca1 toxin. Although this example is carried out on the Cry9Ca1 protein and the cry9Ca1 gene, its teaching is applicable to all the Cry toxins and all the cry genes. This teaching applies in particular to all the Cry toxins the sequence of which is known and filed in the Genbank database:

www.ncbi.nlm.bih.gov/Genbank/index.html.

**[0129]** The Genbank accession numbers for the cry genes are available on the following site:

www.biols.susx.ac.uk/Home/Neil\_Crickmore/Bt/index.html.

**[0130]** This teaching also applies to all the Cry toxins and cry genes, the sequences of which are not disclosed on Genbank.

**[0131]** Unlike the strategies described in Examples 1 to 4, the aim is not to modify a precise region of the toxin so as to integrate amino acids recognized by pepsin, but to increase, overall, the number of these sites by increasing the amount of leucine, of phenylalanine and of glutamic acid in said toxin. This strategy makes it possible to make the Cry toxin more sensitive to pepsin by increasing the percentage of residues recognized by pepsin. Glutamic acid (E; Glu) preferentially substitutes for aspartic acid (D; Asp), pheny-

lalanine (F; Phe) preferentially replaces tryptophan (W; Trp) and leucine (L; Leu) preferably replaces valine (V; Val) or isoleucine (I; Ile). This strategy require the creation of a three-dimensional model for the activated Cry9Ca1 toxin, created from the primary sequence of the protein by comparison with the three-dimensional structures of Cry1Aa1 and Cry3Aa1. The model was created using the Swiss-Model Protein Modelling Server (Peitsch, 1995; Peitsch, 1996; Guex and Peitsch, 1997). The server address is as follows:

#### www.expasy.ch/swissmod/swiss-model.html.

**[0132]** Preferably, the substitutions should reach a maximum level of 25%. The activated Cry9Ca1 toxin contains 31 aspartic acids, 9 tryptophans and 47 valines. There are naturally 26 glutamic acids, 35 phenylalanines and 62 leucines. Taking into account a maximum substitution of 25% for each of the amino acids, the relative ratios are as follows:

Amino acid	Number of residues in native Cry9Ca1	Number of residues in modified Cry9Ca1
Asp (D)	31	24
Glu (E)	26	33
Trp (W)	9	7
Phe (F)	35	37
Val (V)	47	36
Leu (L)	61	72

**[0133]** The substitution of isoleucine (I; Ile) with leucine can also be envisioned instead of or in addition to the substitution of valine with leucine. There are naturally 27 isoleucines in the Cry9Ca1 toxin. Taking into account a preferential degree of substitution of 25%, it is sufficient to replace 6 isoleucine residues with leucines.

**[0134]** It is possible to modify the sequence of the cry9Ca1 gene as shown below. The only aim of the dem-

onstration below is to illustrate the example, and it does not in any way limit the scope of the invention. This demonstration relates to aspartic acid, tryptophan and valine residue replacement. Those skilled in the art can very easily adapt this approach to any other cry gene, the sequence of which would be known, and in particular from the sequences available on Genbank and the accession numbers of which are mentioned on the following site:

www.biols.susx.ac.uk/Home/Neil\_Crickmore/Bt/index.html.

**[0135]** The cry genes generally expressed in transgenic plants are truncated genes, i.e. only the gene sequence encoding the activated toxin is introduced into these plants. The sequences given in this example correspond to this truncated version and extend, depending on whether it is a case of the gene or the protein, from the initiation codon or from the first methionine to 15 codons or amino acids downstream of the conserved block 5 which limits the activated toxin.

**[0136]** The sequence of the native and truncated cry9Ca1 gene is given in SEQ ID NO:1.

**[0137]** The sequence of the native and truncated Cry9Ca1 protein is given in SEQ ID NO:2.

**[0138]** The sequence of a modified cry9Ca1 gene in which all the codons encoding the valine, aspartic acid and tryptophan residues have been modified is given in **FIG. 1** (SEQ ID NO:9). This modified sequence can be used as a basis for defining the various mutagenesis oligonucleotides which may be used. The modified bases are represented in bold characters.

**[0139]** The sequence of a modified Cry9Ca1 protein in which all the value, aspartic acid and tryptophan residues have been modified is given in **FIG. 2** (SEQ ID NO:10) and the modified amino acids are represented in bold characters.

**[0140]** All the mutagenesis oligonucleotides which may make it possible to perform the valine, phenylalanine and glutamic acid residue replacements are given in **FIG. 3** (SEQ ID NOS:94 to 160). The modified bases are represented in bold characters.

**[0141]** A possibility for the use of certain oligonucleotides to create a modified cry9Ca1 gene in which the replacement with respect to codons encoding the valine, aspartic acid and tryptophan residues has been carried out at up to 25% is shown below by way of illustration. The aim of this illustration is to exemplify the strategy developed without limiting the scope of the invention. On the basis of the teaching of this example and of FIGS. 1 to 3 (SEQ ID NOS:9 and 10), those skilled in the art will be able to adapt other combinations of the oligonucleotides given in **FIG. 5** (SEQ ID NOS:94 to 160) or other oligonucleotides prepared along the same principle, in particular for replacing isoleucine residues.

**[0142]** The sequence of a cry9Ca1 gene modified by replacement with respect to the codons encoding the valine, aspartic acid and tryptophan residues up to 25% is given in **FIG. 4** (SEQ ID NO:11). The modified bases are in bold.

**[0143]** The sequence of a Cry9Ca1 protein modified by valine, aspartic acid and tryptophan residue replacement up to 25% is given in **FIG. 5** (SEQ ID NO:12). The modified amino acids are in bold.

[0144] The creation of a modified cry9Ca1 gene in which 25% of the value, aspartic acid and tryptophan codons have been modified, and the sequence of which is given in FIG. 4 (SEQ ID NO:11), can be carried out using, among those given in FIG. 5 (SEQ ID NOS:94 to 160), the following oligonucleotides:

[0145]	Oligonucleotide No. 60
[0146]	Oligonucleotide No. 62
[0147]	Oligonucleotide No. 67
[0148]	Oligonucleotide No. 72
[0149]	Oligonucleotide No. 77
[0150]	Oligonucleotide No. 78
[0151]	Oligonucleotide No. 80
[0152]	Oligonucleotide No. 82
[0153]	Oligonucleotide No. 83
[0154]	Oligonucleotide No. 88
[0155]	Oligonucleotide No. 90
[0156]	Oligonucleotide No. 92
[0157]	Oligonucleotide No. 96
[0158]	Oligonucleotide No. 97
[0159]	Oligonucleotide No. 103
[0160]	Oligonucleotide No. 111

**[0161]** The method preferably used is a multiple mutagenesis with a mixture of the oligonucleotides mentioned immediately above. The site-directed mutagenesis procedure is similar to that described in Example 1, the only difference being that a mixture of mutagenesis oligonucleotides is used in this example, whereas a single mutagenesis oligonucleotide is used in Example 1. The protocol used is that described in Examples 1 to 4. It is common to each of the mutagenesis series, only the mutagenesis oligonucleotide and the oligonucleotide for inhibition/restoration of the resistance to the antibiotic change.

#### Example 6

#### Production of Modified Cry Proteins in *B. Thuringiensis* and Purification

**[0162]** The native and modified genes are inserted, with their promoter and terminator sequences, into the *E. coli-B. thuringiensis* pHT3101 shuttle vector (Lereclus et al., 1989).

**[0163]** The plasmid DNA is prepared by minipreparation according to the alkaline lysis technique (Birboim and Doly, 1979). Each bacterial colony is grown in 2 ml of LB medium supplemented with the appropriate antibiotic, overnight at 37° C. with shaking (200 rpm). The culture is then transferred into a microtube and then centrifuged at 13 500 g for 5 min. After removal of the supernatant, the bacteria are resuspended in 100  $\mu$ l of a solution of 25 mM Tris-HCl, pH 8, 10 mM EDTA containing RNase A at the final concentration of 100  $\mu$ g/ml. 200  $\mu$ l of a solution of 0.2 M NaOH, 1% SDS are added and the suspension is mixed twice by inverting the tube. 150  $\mu$ l of a 2.55 M potassium acetate solution, pH 4.5, are added and the suspension is incubated for 5 min in ice. After centrifugation for 15 min at 13 500

g, the supernatant is transferred into a microtube containing 1 ml of cold ethanol. After centrifugation for 30 min at 13 500 g, the supernatant is removed and the pellet is washed with 1 ml of 70% ethanol. The pellet containing the DNA is dried for a few minutes under vacuum and then taken up in 50  $\mu$ l of sterile distilled water. The samples are then placed at 65° C. for 30 min.

**[0164]** The digestions with restriction endonucleases are carried out per 1  $\mu$ g of DNA in a final volume of 20  $\mu$ l in the presence of one tenth of the final volume of 10× buffer recommended by the supplier for each enzyme and using 5 units of enzyme. The reaction is incubated for 2 to 3 h at the optimum temperature for the enzyme.

[0165] Dephosphorylation of the 5' ends engendered by restriction enzyme is carried out with calf intestine alkaline phosphatase. The reaction is carried out using 5  $\mu$ l of 10× dephosphorylation buffer (500 mM Tris-HCl, pH 9.3, 10 mM MgCl2, 1 mM ZnCl<sub>2</sub> and 10 mM spermidine) and one unit of enzyme per  $\mu g$  of DNA in a final volume of 50  $\mu$ l. The reaction is incubated for one hour at 37° C. in the case of overhanging 5' ends or at 55° C. in the case of blunt ends or 3' overhanging ends. After dephosphorylation, the enzyme is then inactivated for 30 min at 65° C. and then removed with two volume for volume extractions with a mixture of phenol-chloroform-isoamyl alcohol (25-24-1). The ligations are carried out using T4 phage DNA ligase. They are carried out with an amount of vector equal to 100 ng and an insert/vector molar ratio of between 5 and 10. The final volume of the reaction is 30  $\mu$ l and comprises 3  $\mu$ l of 10× ligase buffer (300 mM Tris-HCl, pH 7.8, 100 mM MgCl2, 100 mM DTT and 10 mM ATP) and 3 units of enzyme. The reaction is incubated overnight at 14° C.

[0166] The construct is inserted into an acrystalliferous strain of B. thuringiensis according to a method derived from that described in 1989 by Lereclus et al. and described elsewhere (Rang et al., 1999, 2000). A preculture of acrystalliferous Bacillus thuringiensis subsp. kurstaki HD-1 is incubated overnight at 37° C. with shaking in 10 ml of BHI medium (Difco). 250 ml of BHI medium are then inoculated with 5 ml of preculture and incubated at 37° C. with shaking until the OD at 600 nm of the culture reaches the value of 0.3. The culture is then centrifuged at 1 000 g at 4° C. for 10 min. The supernatant is removed and the bacterial pellet is rinsed with 50 ml of cold sterile distilled water. The bacteria are again centrifuged for 10 min at 1 000 g at 4° C. The pellet is taken up in 4 ml of a cold, sterile solution of 40% PEG-6000 and placed in ice. 200  $\mu$ l of bacteria are then mixed with 5  $\mu$ g of plasmid DNA and then placed in an electroporation cuvette 0.2 cm in diameter. The cuvette is then placed in the electroporation chamber and a current corresponding to the following characteristics: 2.5 kV, 1 000  $\Omega$ , 25  $\mu$ F, is supplied. The bacteria are then covered, placed in ice for 10 min before being added to 2 ml of BHI medium, and incubated at 37° C. with shaking for 90 min. 200  $\mu$ l of culture are then plated out onto Petri dishes containing usual solid medium (IEBC, 1994) supplemented with erythromycin at a final concentration of 25  $\mu$ g/ml, and incubated overnight at 28° C.

**[0167]** The recombinant strains of *Bacillus thuringiensis* expressing the native gene or the mutated genes are cultured in 250 ml of Usual medium containing 25  $\mu$ g/ml of erythromycin with shaking at 28° C. The bacterial growth is

verified by observation by phase-contrast light microscopy. The bacteria are grown until bacterial lysis after sporulation. The culture is then centrifuged at 5 000 g for 10 min. The pellet is washed with 25 ml of 1 M NaCl and the suspension is again centrifuged at 5 000 g for 10 min. The pellet is then taken up in 15 ml of sterile distilled water containing 1 mM of PMSF, incubated in ice, and treated with ultrasound (100 W) for 1 min in order to dissociate the aggregates between the spores and the crystals. The suspension is then loaded onto a discontinuous NaBr gradient made up of a layer of 4 ml of 38.5% concentration, of 4 layers of 6 ml of 41.9%, 45.3%, 48.9% and 52.7% and a layer of 3 ml of 56.3%. The gradient is then centrifuged at 20 000 g for 90 min at 20° C. The various components of the suspension (spores, cell debris, parasporal bodies) are positioned in the gradient at various levels depending on their density. Each band is recovered and washed three times with one volume of sterile distilled water. Each band is observed by phase-contrast light microscopy. The fraction containing inclusion bodies is stored at -20° C. in sterile distilled water containing 1 mM of PMSF, for subsequent analysis.

#### Example 7

#### Analysis of the Stability of the Proteins to Proteases

[0168] The first stability analysis performed is the verification of stability to trypsin. The proteins present in the parasporal inclusion body are solubilized for one hour at 37° C. in solubilizing buffer (50 mM Na<sub>2</sub>CO<sub>3</sub>, pH 10.8, 14.6 mM 2-mercaptoethanol). The suspension is then centrifuged at 14 000 g for 10 min in order to remove the insoluble material. One tenth of the total volume of 0.05% trypsin is then added to the supernatant and the mixture is incubated for 2 h at 37° C. The condition of the proteins after trypsin treatment is verified by SDS-polyacrylamide gel analysis according to the Laemmli method (1970). This technique allows the proteins to be separated according to their molecular mass by virtue of the presence of SDS, which confers an overall negative charge on all the proteins. The sample is first treated by adding one volume of 2× treatment solution (125 mM Tris-HCl, 20% glycerol, 4% SDS, 10% 2-mercaptoethanol, 0.01% bromophenol blue) and is then denatured for 5 min in boiling water. The sample is then loaded onto the gel and first passes through a first stacking gel made up of a 4% acrylamide-bisacrylamide mixture, 0.1% SDS, and 125 mM Tris-HCl, pH 6.8. The sample then passes through the separating gel made up of 12% acrylamide-bisacrylamide, 0.1% SDS and 375 mM Tris-HCl, pH 8.8, which makes it possible to separate the various proteins as a function of their size. The electrophoresis is carried out at 100 V in migration buffer (25 mM Tris-HCl, pH 8.3, 192 mM glycine, 0.1% SDS) until the bromophenol blue leaves the gel. The gel is then stained for one hour with a solution of 40% methanol-7% acetic acid containing 0.025% of Coomassie blue and then destained with a 50% methanol-10% acetic acid solution. The gel is ultimately fixed in a 5%methanol-7% acetic acid solution.

**[0169]** The second analysis is the verification of the stability to the digestive juices of insects. The trypsin-stable toxins are purified by FPLC (Pharmacia) using an anion exchange column (Q-Sepharose) equilibrated with a 40 mM  $Na_2CO_3$  solution, pH 10.7. The elution is carried out with a

gradient of 50 to 500 mM of NaCl. The OD at 280 nm of the fractions is measured and the fractions containing the proteins are analyzed by SDS-polyacrylamide gel electrophoresis. The fractions containing the toxin are pooled and dialyzed at 4° C. against distilled water for approximately 48 h until the proteins precipitate. The protein suspension is then centrifuged at 8 000 g and at 4° C. for 30 min. The toxins contained in the pellet are resuspended in distilled water and assayed according to Bradford (1976). They are then divided up into aliquot fractions of 100  $\mu$ g, lyophilized, and then stored at 4° C. Before they are used, the toxins are solubilized and brought to a concentration of 10 mg/ml with 25 mM Tris, pH 9.5, for the purpose of testing their stability to the digestive juices of Ostrinia nubilalis larvae. The digestive juice of the O. nubilalis larvae can be taken either by regurgitation induced by electric shock according to the procedure of Ogiwara et al. (1992), or by dissection of the larvae and collection of the intestinal juice with a pipette according to the method described by Baines et al. (1994). In both cases, between 100 and 200 individuals are required to collect the digestive juice. The juice collected is centrifuged at 15 000 g for 15 minutes at 4° C. before use. The protein concentration of the digestive juice is determined by the Bradford method (BioRad). The reaction is carried out for 15 minutes at 37° C. with a 1:1 ratio (based on the protein concentration of the digestive juice) of toxin to digestive juice. The reaction is stopped with a cocktail of protease inhibitors (Protease Inhibitors Set, Roche Diagnostics) mixed with an equivalent volume of 2× treatment solution (125 mM Tris-HCl, 20% glycerol, 4% SDS, 10% 2-mercaptoethanol, 0.01% bromo-phenol blue), and then incubated for 5 minutes in boiling water. The proteins are then analyzed by SDS-PAGE according to the procedure described above, in order to determine their resistance to the digestive juices of the larvae and their possible state of degradation.

**[0170]** The final type of stability analysis carried out is that of stability to pepsin. The lyophilized native and modified toxins are dissolved in a gastric buffer (0.5 mg NaCl, 1.75 ml 1M HCl in 250 ml  $H_2O$ , pH 2.0) simulating mammalian stomach fluid and containing 0.32% of pepsin. Samples are removed after 0, 5, 15, 60 and 240 minutes of incubation at 37° C. and then analyzed by SDS-polyacryla-mide gel electrophoresis as described above. These conditions are identical to those described in the EPA (United States Environmental Protection Agency) No. 4458108.

**[0171]** This series of analyses makes it possible to visualize the state of conservation of the native and mutated proteins, and therefore their stability, to various proteases present in insects (trypsin and digestive juices) and, consequently, to verify that the mutated proteins have effectively conserved their stability in insects. These analyses also make it possible to verify that the mutated proteins are effectively degraded by pepsin under the conditions similar to those present in the mammalian stomach.

## Example 8

## Analysis of the Insecticidal Properties

**[0172]** The analysis of insecticidal properties is carried out through two types of experiment for testing the two steps of the process of toxicity in insects: receptor site recognition and evaluation of the toxicity in vivo.

**[0173]** Analysis of the affinity of the toxins for the receptor site is carried out using toxin radiolabeled with iodine 125

(<sup>125</sup>I). The FPLC-purified and lyophilized activated toxins are taken up in storage buffer (20 mM Tris-HCl, pH 8.6) and analyzed by SDS-PAGE in order to verify their condition. An aliquot fraction is assayed according to the Bradford method (1976). The toxins are iodinated according to the chloramine-T method (Markwell, 1982). 25 µµg of toxins are incubated for 5 min at ambient temperature with 0.25 mCi of Na-<sup>125</sup>I and an "Iodo-bead" (Pierce) in 50  $\mu$ l of sodium carbonate buffer (50 mM Na<sub>2</sub>CO<sub>3</sub>, pH 10). The iodination reaction is then deposited at the surface of a dextran desalting column (Pierce) equilibrated with CBS buffer (50 mM Na<sub>2</sub>CO<sub>3</sub>, pH 10.8, 150 mM NaCl) in order to remove the free iodine. The labeling and the quality of the protein are verified by SDS-PAGE followed by autoradiography. The mean specific activity of a labeled toxin is 100 000 cpm/pmol.

[0174] In order to prepare the brush border membrane vesicles (BBMV) on which the study of the affinity of the toxins for the receptors is carried out, the insects are allowed to grow to the final larval stage. The insect used is Ostrinia *nubilalis*, but the methodology used is applicable to any other insect species. The use of another insect species requires the production conditions and the nutritive medium to be adapted to each of the species envisioned, which can be readily done by any individual skilled in the art. The Ostrinia nubilalis larvae are produced on meridic artificial nutritive medium (Lewis and Lynch, 1969; Reed et al., 1972; Ostlie et al., 1984). The method for producing the Ostrinia nubilalis larvae is that described by Huang et al. (1997). The larvae are produced individually in 128-well plates (Bio-Ba-128, C-D International). Each well contains 2 ml of artificial medium. After ten days, the larvae are transferred into larger dishes (18.4 cm in diameter and 7.6 cm high) containing 300 ml of artificial nutritive medium. Corrugated cardboard is placed inside by way of pupation site. During the larval phase, the temperature of the production cell is 25° C. with constant light (24 h). The pieces of cardboard containing the chrysalises are transferred into screened cages for the emergence and the production of the adults. Waxed paper is placed in the case to accept the eggs. The eggs are removed and kept on hold at 15° C. The production of the adults is carried out at 25° C. with 75% relative humidity and a photoperiod of 14 h.

[0175] To carry out the tests of affinity of the toxins for the receptor sites, the larvae are collected at the beginning of the 5th larval stage and placed under fasting conditions for 6 hours. They are then removed and placed on ice for 5 minutes. The larvae are dissected and the digestive tube is removed. The dissected digestive tubes are pooled in groups of 20, placed in a cryotube containing MET buffer (300 mM mannitol, 5 mM EGTA, 17 mM Tris-HCl, pH 7.5), frozen in liquid nitrogen and stored at  $-80^{\circ}$  C.

**[0176]** The BBMVs are prepared according to the differential magnesium precipitation method (Wolfersberger et al., 1987; Nielsen-LeRoux and Charles, 1992). The BBMVs are taken up in TBS buffer (20 mM Tris-HCl, pH 8.5, 150 mM NaCl) and the total protein concentration is determined by the Bradford method using the Biorad kit and bovine serum albumin (BSA) as standard (Bradford, 1976).

**[0177]** The in vitro receptor recognition assays are carried out in 1.5 ml polyethylene microtubes, in 20 mM sodium phosphate buffer, pH 7.4, containing 0.15 M of NaCl and

0.1% of bovine serum albumin (PBS/BSA). The assays are carried out, in duplicate, at ambient temperature in a total volume of 100  $\mu$ l, with 10  $\mu$ g of BBMV protein. The toxins attached to the BBMVs are separated from the free toxins by centrifugation at 14 000 g for 10 min at ambient temperature. The pellets of each sample, containing the toxin attached to the membrane, are rinsed twice with 200  $\mu$ l of cold PBS/BSA buffer (20 mM Tris/HCl, 150 mM NaCl, 0.1% BSA, pH 8.5) and then centrifuged. The pellets are finally resuspended in 200  $\mu$ l of PBS/BSA buffer and added to 3 ml of HiSafe 3 scintillant cocktail (Pharmacia) in a scintillation vial. The counting is performed in a liquid scintillation counter.

[0178] The direct binding assays are carried out according to the Nielsen-LeRoux and Charles protocol (1992). 30  $\mu$ g of BBMV per microtube are incubated with a series of concentrations of 1 to 100 mM of toxin labeled with <sup>125</sup>I-iodine in Tris/BSA buffer (20 mM Tris/HCl, 150 mM NaCl, 0.1% BSA, pH 8.5). The amount of nonspecific attachment is determined in parallel experiments in the presence of a 300-fold excess of unlabeled toxin. After incubation for 90 minutes at ambient temperature, the samples are centrifuged at 14 000 g for 10 minutes at 4° C. The pellets are rinsed twice with cold Tris/BSA buffer and resuspended in 150  $\mu$ l of the same buffer and added to 3 ml of HiSafe 3 scintillant cocktail (Pharmacia) in a scintillation vial. Each experiment is carried out in duplicate and each experimental point is counted twice in a liquid scintillation counter. The data are analyzed using the LIGAND software (Munson and Rodbard, 1980) marketed by the company Biosoft.

**[0179]** The homologous competition experiments are carried out as described above for the direct binding experiments, with 10  $\mu$ g of BBMVs in a total volume of 100  $\mu$ l for 90 min at ambient temperature. The BBMVs are incubated in a fixed concentration of 10 nM of toxin labeled with 125I-iodine in the presence of a series of concentrations (from 0.1 to 300 times the concentration of the labeled toxin) in Tris/BSA buffer. The value for the nonspecific binding (the binding always present in the presence of a 300-fold excess of the unlabeled toxin) is subtracted from the total value counted. Each experiment is carried out in duplicate and each experimental point is counted twice in a liquid scintillation counter. The data are analyzed using the LIGAND software (Munson and Rodbard, 1980) marketed by the company Biosoft.

[0180] The in vivo toxicity assays are carried out according to the procedure described by Lambert et al. (1996). The activated and solubilized toxin is incorporated into the nutritive medium at various concentrations either side of the 50% lethal dose (LD50) of Cry9Ca1 for Ostrinia nubilalis, which is 96.6 ng of toxin per  $cm^2$  of surface area of medium. Six doses, of 0.1 ng/cm<sup>2</sup>, 1 ng/cm<sup>2</sup>, 10 ng/cm<sup>2</sup>, 100 ng/cm<sup>2</sup>,  $1\ 000\ \text{ng/cm}^2$  and  $10\ 000\ \text{ng/cm}^2$  evaluate the LD50 values of the native and modified toxins. The toxicity assays are carried out on neonatal larvae in plates containing 24 wells of 2 cm<sup>2</sup> (Multiwell-24 plates, Coming Costar Corp.). 50 µl of each of the dilutions of toxin are plated out onto the medium and dried under a flow hood. One larva is placed in each well and a total of 24 larvae is used for each dose (one plate per dose). For each dose the assay is repeated at least three times. A control is carried out with distilled water. The plates are covered and placed at 25° C., 70% relative humidity and with a photoperiod of 16 h. The mortality is controlled after 7 days and the LD50 is calculated according to the probit method (Finney, 1971).

#### Bibliographical References

**[0181]** Abdul-Rauf, M. and Ellar, D. J. 1999. Mutations of loop 2 and loop 3 residues in domain II of *Bacillus thuringiensis* CryC  $\delta$ -endotoxin affect insecticidal specificity and initial binding to *Spodoptera littoralis* and *Aedes aegypti* midgut membranes. Current Microbiology. 39: 94-98.

**[0182]** Aronson, A. I., D. Wu, and C. Zhang. 1995. Mutagenesis of specificity and toxicity regions of a *Bacillus thuringiensis* protoxin gene. J. Bacteriol. 177: 4059-4065.

**[0183]** Baines, D., A. Brownright and J. L. Schwartz. 1994. Establishment of primary and continuous cultures of epithelial cells from larval lepidopteran midguts. J. Insect. Physiol. 40: 347-357.

**[0184]** Birnboim, H. C. and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7: 1513.

**[0185]** Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of micro-gram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72: 248-254.

**[0186]** Chen, X. J., Lee, M. K. and Dean, D. H. (1993). Site-directed mutations in a highly conserved region of *Bacillus thuringiensis*  $\delta$ -endotoxin affect inhibition of short circuit current across Bombyx mori midguts. Proc. Natl. Acad. Sci. USA. 90 9041-9045.

[0187] Coux, F., V. Vachon, C. Rang, K. Moozar, L. Masson, M. Royer, M. Bes, S. Rivest, J. L. Schwartz, R. Laprade and R. Frutos. 1999. Role of interdomain salt bridges in the pore-forming ability of insecticidal toxin Cry1Aa of *Bacillus thuringiensis*. 32nd Annual Meeting of the Society for Invertebrate Pathology. 22-27 Aug. 1999, Irvine, Calif., USA.

[0188] Crickmore, N. 2001. epunix.biols.susx.ac.uk/ Home/Neil Crickmore/Bt/list.html

**[0189]** Crickmore, N., Zeigler, D. R., Feitelson, J., Schnepf, E., van Rie, J., Lereclus, D., Baum, J., and Dean, D. H. 1998. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal proteins. Microbiol. Mol. Biol. Rev. 62: 807-813.

**[0190]** de Maagd, R. A., P. Bakker, N. Staykov, S. Dukiandjiev, W. Stiekema and D. Bosch. 1999. Identification of *Bacillus thuringiensis* delta-endotoxin Cry1C domain III amino acid residues involved in insect specificity. Appl. Environ. Microbiol. 65(10): 4369-4374.

**[0191]** Dean, D. H., Rajamohan, F., Lee, M. K., Wu, S. J., Chen, X. J., Alcantara, E., Hussain, S. R. 1996. Probing the mechanism of action of *Bacillus thuringiensis* insecticidal proteins by site-directed mutagenesis—a minireview. Gene. 179: 111-117.

**[0192]** Finney, D. J. 1971. Probit analysis. A statistical treatment of the sigmoid response curve. University Press, Cambridge.

**[0193]** Gazit, E. and Shai, Y. 1993. Structural and functional characterization of the  $\alpha$ 5 segment of *Bacillus thuringiensis*  $\delta$ -endotoxin. Biochemistry. 32: 3429-3436.

**[0194]** Grochulski, P., Masson, L., Borisova, S., Pusztai-Carey, M., Schwartz, J. -L., Brousseau, R. and Cygler, M. (1995) *Bacillus thuringiensis* CryIA(a) insecticidal toxin: crystal structure and channel formation. J. Mol. Biol. 254: 447-464.

**[0195]** Guex, N. and Peitsch, M. C. 1997. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modelling. Electrophoresis. 18: 2714-2723.

**[0196]** Huang, F. R., A. Higgins and L. L. Buschman. 1997. Baseline susceptibility and changes in susceptibility to *Bacillus thuringiensis* subsp. *kurstaki* under selection pressure in European corn borer (Lepidoptera: Pyralidae). J. *Econ. Entomol.* 90: 1137-1143.

**[0197]** Hussain, S. R. A., A. I. Aronson and D. H. Dean. 1996. Substitution of residues on the proximal side of Cry1A *Bacillus thuringiensis* delta-endotoxins affects irreversible binding to Manduca sexta midgut membrane. *Biochemical and Biophysical Research Communications*. 226(1) 8-14.

**[0198]** Hutchinson, C. A., S. Phillips, M. H. Edgell, S. Gillam, P. Jahnke and M. Smith. 1978. Mutagenesis at a specific position in a DNA sequence. J. Biol. Chem. 253: 6551.

**[0199]** IEBC. 1994. Collection of *Bacillus thuringiensis* and *Bacillus sphaericus* (classified by serotypes). Institut Pasteur Paris.

**[0200]** Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 227: 680-685.

**[0201]** Lambert, B., Buysse, L., Decock, C., Jansen, S., Piens, C., Saey, B., Seurinck, J., Van Audenhove K., van Rie, J., van Vliet, A. and Peferoen, M. 1996. A *Bacillus thuringiensis* insecticidal crystal protein with a high activity against members of the family Noctuidae. *Appl. Environ. Microbiol.* 62: 80-86.

**[0202]** Lee, M. K., You, T. H., Gould, F. L. and Dean, D. H. 1999. Identification of residues in domain III of *Bacillus thuringiensis* Cry1Ac toxin that affect binding and toxicity. *Appl. Environ. Microbiol.* 65: 4513-4520.

**[0203]** Lereclus, D., Arantes, S. O., Chauffaux, J. and Lecadet, M. M. 1989. Transformation and expression of a cloned delta-endotoxin gene in *Bacillus thuringiensis*. FEMS Microbiol. Lett. 60: 211-217.

**[0204]** Lewis, L. C. and R. E. Lynch. 1969. Rearing the European corn borer on diets containing corn leaf and wheat germ. *Iowa State J. Sci.* 44: 9-14.

**[0205]** Li, J., Carroll, J. and Ellar, D. J. (1991). Crystal structure of insecticidal  $\delta$ -endotoxin from *Bacillus thuring-iensis* at 2.5 Å resolution. *Nature* 353: 815-821.

**[0206]** Manoj Kumar, A. S. and Aronson, A. I. 1999. Analysis of mutations in the pore-forming region essential for insecticidal activity of a *Bacillus thuringiensis* δ-endotoxin. *J. Bacteriol.* 181: 6103-6107. **[0207]** Markwell, M. A. K. 1982. A new solid-state reagent to iodinate proteins. I-Conditions for the efficient labelling of antiserum. Anal. Biochem. 125: 427-432.

**[0208]** Masson, L., B. E. Tabashnik, Y. B. Liu, R. Brousseau and J. L. Schwartz. 1999. Helix 4 of the *Bacillus thuringiensis* Cry1Aa toxin lines the lumen of the ion channel. *J. Biol. Chem.* 274: 31996-32000.

**[0209]** Munson, P. J. & Rodbard, D. 1980. LIGAND: A versatile computerized approach for characterization of ligand-binding systems. Anal. Biochem. 107, 220-239.

**[0210]** Nielsen-LeRoux, C., Charles, J. F. 1992. Binding of *Bacillus sphaericus* binary toxin to a specific receptor on midgut brush border membranes from mosquito larvae. *Eur. J Biochem.* 210: 585-590.

**[0211]** Ogiwara, K., L. S. Indrasith, S. Asano and H. Hori. 1992. Processing of δ-endotoxin from *Bacillus thuringiensis* subsp. *kurstaki* HD-1 and HD-73 by gut-juices of various insect larvae. J. Invert. Pathol. 60: 121-126.

**[0212]** Ostlie, K. R., G. L. Hein, L. G. Higley, L. V. Kaster and Showers, W. B. 1984. European corn borer (Lepi-doptera: Pyralidae) development, larval survival, and adult vigor on meridic diets containing marker dies. *J. Econ. Entomol.* 77: 118-120.

**[0213]** Peitsch, M. C. 1995. Protein modeling by E-mail. *Bio/Technology*. 13: 658-660.

**[0214]** Peitsch, M. C. 1996. ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. *Biochem. Soc. Trans.* 24: 274-279.

**[0215]** Rajamohan, F., Alcantara, E., Lee, M. K., Chen, X. J., Curtiss, A. and Dean, D. H. (1996). Single amino acid changes in domain II of *Bacillus thuringiensis* CryIAb δ-endotoxin affect irreversible binding to Manduca sexta midgut membrane vesicles. J. Bacteriol. 177: 2276-2282.

**[0216]** Rajamohan, F., M. K. Lee and D. H. Dean. 1998. *Bacillus thuringiensis* insecticidal proteins: molecular mode of action. *Prog. Nucleic Acid Res. Mol. Biol.* 60: 1-27.

**[0217]** Rang, C., L. Lacey and R. Frutos. 2000. The crystal proteins from *Bacillus thuringiensis* subsp. *thompsoni* display a synergistic activity against the codling moth, *Cydia pomonella*. Current *Microbiology* 40: 200-204.

**[0218]** Rang, C., Vachon, V., Coux, F., Carret, C., Moar, W. J., Brousseau, R., Schwartz, J. L., Laprade, R. and Frutos, R. 2001. Exchange of domain I from *Bacillus thuringiensis* Cry1toxins influences protoxin stability and crystal formation. *Current Microbiology*. In Press.

**[0219]** Rang, C., Vachon, V., de Maagd, R. A., Villalon, M., Schwartz, J. -L., Bosch, D., Frutos, R. and Laprade, R. 1999. Interaction between functional domains of *Bacillus thuringiensis* insecticidal crystal proteins. *Appl. Environ. Microbiol.* 65: 2918-2925.

**[0220]** Reed, G. L., W. B. Showers, J. L. Huggans and S. W. Carter. 1972. Improved procedures for mass rearing the European corn borer. *J. Econ. Entomol.* 65: 1472-1476.

**[0221]** Sambrook, J., E. F. Fritch and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual, Nolan C. ed., New York: Cold Spring Harbor Laboratory Press. **[0222]** Sanger, F., Nicklen, S. & Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74: 5463-5467.

[0223] Schnepf, E., N. Crickmore, J. van Rie, D. Lereclus, J. Baum, J. Feitelson, D. R. Zeigler and D. H. Dean. 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* 62: 775-806.

**[0224]** Terra, W. B. and C. Ferreira. 1994. Insect digestive enzymes: properties, compartmentalization and function. *Comp. Biochem. Physiol.* 109B: 1-62.

**[0225]** Vachon, V. F., Coux, F., G. Prefontaine, C. Rang, L. Marceau, L. Masson, R. Brousseau, R. Frutos, J. L. Schwartz and R. Laprade, R. 2000. Role of  $\alpha$ -helix three charged residues in pore formation by the *Bacillus thuring*-

*iensis* insecticidal toxin Cry1Aa. 33rd Annual Meeting of the Society for Invertebrate Pathology. 13-18 Aug. 2000, Guanajuato, Mexico.

**[0226]** Wolfersberger, M. G., Luthy, P., Maurer, A., Parenti, P., Sacchi, V., Giordana B., Hanozet, G. 1987. Preparation and partial characterisation of amino acid transporting brush border membrane vesicles from the larval midgut of the cabbage butterfly (*Pieris brassicae*). *Comp. Biochem. Physiol.* 86: 301-308.

**[0227]** Wu, S. J. and Dean D. H. 1996. Functional significance of loops in the receptor binding domain of *Bacillus thuringiensis* CryIIIA δ-endotoxin. *J. Mol. Biol.* 255: 628-640.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 160 <210> SEQ ID NO 1 <211> LENGTH: 2019 <212> TYPE: DNA <213> ORGANISM: Bacillus thuringiensis <220> FEATURE: <221> NAME/KEY: CDS <222> LOCATION: (1)..(2019) <400> SEQUENCE: 1 48 atg aat cga aat aat caa aat gaa tat gaa att att gat gcc ccc cat Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His 10 tgt ggg tgt cca tca gat gac gat gtg agg tat cct ttg gca agt gac 96 Cys Gly Cys Pro Ser Asp Asp Asp Val Arg Tyr Pro Leu Ala Ser Asp 25 20 30 cca aat gca gcg tta caa aat atg aac tat aaa gat tac tta caa atg 144 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu Gln Met 35 40 45 aca gat gag gac tac act gat tct tat ata aat cct agt tta tct att Thr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu Ser Ile 192 55 60 agt ggt aga gat gca gtt cag act gcg ctt act gtt gtt ggg aga ata 240 Ser Gly Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Ile 65 70 75 80 288 ctc ggg gct tta ggt gtt ccg ttt tct gga caa ata gtg agt ttt tat Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Val Ser Phe Tyr 85 90 95 caa ttc ctt tta aat aca ctg tgg cca gtt aat gat aca gct ata tgg 336 Gln Phe Leu Leu Asn Thr Leu Trp Pro Val Asn Asp Thr Ala Ile Trp 100 105 110 gaa gct ttc atg cga cag gtg gag gaa ctt gtc aat caa caa ata aca 384 Glu Ala Phe Met Arg Gln Val Glu Glu Leu Val Asn Gln Gln Ile Thr 120 125 115 gaa ttt gca aga aat cag gca ctt gca aga ttg caa gga tta gga gac 432 Glu Phe Ala Arg Asn Gln Ala Leu Ala Arg Leu Gln Gly Leu Gly Asp 135 130 140 tct ttt aat gta tat caa cgt tcc ctt caa aat tgg ttg gct gat cga 480 Ser Phe Asn Val Tyr Gln Arg Ser Leu Gln Asn Trp Leu Ala Asp Arg 145 150 155 160 aat gat aca cga aat tta agt gtt gtt cgt gct caa ttt ata gct tta 528

-continued

													tin	ueu			
Asn	Asp	Thr	Arg	Asn 165	Leu	Ser	Val	Val	<b>A</b> rg 170	Ala	Gln	Phe	Ile	Ala 175	Leu		
-	ctt Leu	-		-		-			-		-	-			-		
-	gtt Val			-		-		-		-					-		
	tta Leu 210																
	GJÀ ddd																
	tac Tyr																
Leu	aga Arg	Gly	Thr 260	Asn	Thr	Glu	Ser	Trp 265	Leu	Arg	Tyr	His	Gln 270	Phe	Arg		
Arg	gaa Glu	Met 275	Thr	Leu	Val	Val	Leu 280	Åsp	Val	Val	Ala	Leu 285	Phe	Pro	Tyr		
Tyr	gat Asp 290	Val	Arg	Leu	Tyr	Pro 295	Thr	Ğİy	Ser	Asn	Pro 300	Gln	Leu	Thr	Arg		
Glu 305	gta Val	Tyr	Thr	Asp	Pro 310	Ile	Val	Phe	Asn	Pro 315	Pro	Ala	Asn	Val	Gly 320		
Leu	tgc Cys	Arg	Ārg	Trp 325	Gly	Thr	Asn	Pro	<b>Tyr</b> 330	Asn	Thr	Phe	Ser	Glu 335	Leu		
Ğlu	aat Asn	Ala	Phe 340	Ile	Arg	Pro	Pro	His 345	Leu	Phe	Asp	Arg	Leu 350	Asn	Ser		
Leu	aca Thr	Ile 355	Ser	Ser	Asn	Arg	Phe 360	Pro	Val	Ser	Ser	Asn 365	Phe	Met	Asp		
Tyr	tgg Trp 370	Ser	Ğİy	His	Thr	Leu 375	Arg	Arg	Ser	Tyr	Leu 380	Asn	Asp	Ser	Ála		
Val 385	caa Gln ccc	Glu	Asp	Ser	<b>Ty</b> r 390	Gly	Leu	Ile	Thr	Thr 395	Thr	Arg	Ala	Thr	Ile 400		
Asn	Pro	ĞÎy	Val	Asp 405	Gly	Thr	Asn	Arg	Ile 410	Glu	Ser	Thr	Āla	Val 415	Asp		
Phe	Arg	Ser	Ala 420	Leu	Ile	Gly	Ile	Tyr 425	Gly	Val	Asn	Arg	Ala 430	Ser	Phe		
Val	Pro	Gly 435	Gly	Leu	Phe	Asn	Gly 440	Thr	Thr	Ser	Pro	Ala 445	Asn	Gly	Gly		
Cys	Arg 450	Asp	Leu	Tyr	Asp	Thr 455	Asn	Asp	Glu	Leu	Pro 460	Pro	Asp	Glu	Ser		
acc	gga	ayt	ıca	acc	σατ	aya	ula	LUT	σατ	yιτ	acc	LLT	LLT	ayc	LLT	1440	

-continued

	-con	htinued
Thr Gl <b>y</b> Ser Ser Thr His 465 470	Arg Leu Ser His Val Thr Phe 475	e Phe Ser Phe 480
	tct ata gct aat gca gga agt Ser Ile Ala Asn Ala Gly Ser 490	
	gat gtg gac ctt aat aat acg Asp Val Asp Leu Asn Asn Thr 505	
	cca ttg gta aag gca tct gca Pro Leu Val Lys Ala Ser Ala 520 525	a Pro Val Ser
	ggt cca gga ttt aca gga ggg Gly Pro Gly Phe Thr Gly Gly 535 540	
	aca ttt gga acg tta aga gta Thr Phe Gly Thr Leu Arg Val 555	
	tat cgc cta aga gtt cgt ttt Tyr Arg Leu Arg Val Arg Phe 570	5
	gta ctc cgt gga ggg gtt tct Val Leu Arg Gly Gly Val Ser 585	
	atg aac aga ggg cag gaa cta Met Asn Arg Gly Gln Glu Leu 600 605	a Thr Tyr Glu
	ttt act act act ggt ccg ttc Phe Thr Thr Thr Gly Pro Phe 615 620	
	caa gag att cta aca gtg aat Gln Glu Ile Leu Thr Val Asn 635	
	tat tat ata gat aga att gaa Tyr Tyr Ile Asp Arg Ile Glu 650	
	gcg gaa gag gat tta gaa gcg Ala Glu Glu Asp Leu Glu Ala 665	
gcg Ala		2019
<pre>&lt;210&gt; SEQ ID NO 2 &lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Bacillu</pre>	s thuringiensis	
<400> SEQUENCE: 2		
Met Asn Arg Asn Asn Glr 1 5	Asn Glu Tyr Glu Ile Ile Asp 10	D Ala Pro His 15
Cys Gly Cys Pro Ser Asp 20	Asp Asp Val Arg Tyr Pro Leu 25	1 Ala Ser Asp 30
Pro Asn Ala Ala Leu Glr 35	Asn Met Asn Tyr Lys Asp Tyr 40 45	
Thr Asp Glu Asp Tyr Thr 50	Asp Ser Tyr Ile Asn Pro Ser 55 60	: Leu Ser Ile
Ser Gly Arg Asp Ala Val 65 7(	Gln Thr Ala Leu Thr Val Val 75	l Gly Arg Ile 80

Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Val Ser Phe Tyr Gln Phe Leu Leu Asn Thr Leu Trp Pro Val Asn Asp Thr Ala Ile Trp Glu Ala Phe Met Arg Gln Val Glu Glu Leu Val Asn Gln Gln Ile Thr Glu Phe Ala Arg Asn Gln Ala Leu Ala Arg Leu Gln Gly Leu Gly Asp Ser Phe Asn Val Tyr Gln Arg Ser Leu Gln Asn Trp Leu Ala Asp Arg Asn Asp Thr Arg Asn Leu Ser Val Val Arg Ala Gln Phe Ile Ala Leu Asp Leu Asp Phe Val Asn Ala Ile Pro Leu Phe Ala Val Asn Gly Gln 180 185 190 Gln Val Pro Leu Leu Ser Val Tyr Ala Gln Ala Val Asn Leu His Leu Leu Leu Lys Asp Ala Ser Leu Phe Gly Glu Gly Trp Gly Phe Thr 
 Gln Gly Glu Ile Ser Thr Tyr Tyr Asp Arg Gln Leu Glu Leu Thr Ala

 225
 230
 235
 240
 Lys Tyr Thr Asn Tyr Cys Glu Thr Trp Tyr Asn Thr Gly Leu Asp Arg 245 250 255 Leu Arg Gly Thr Asn Thr Glu Ser Trp Leu Arg Tyr His Gln Phe Arg 260 265 270 Arg Glu Met Thr Leu Val Val Leu Asp Val Val Ala Leu Phe Pro Tyr Tyr Asp Val Arg Leu Tyr Pro Thr Gly Ser Asn Pro Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Ile Val Phe Asn Pro Pro Ala Asn Val Gly Leu Cys Arg Arg Trp Gly Thr Asn Pro Tyr Asn Thr Phe Ser Glu Leu Glu Asn Ala Phe Ile Arg Pro Pro His Leu Phe Asp Arg Leu Asn Ser Leu Thr Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe Met Asp Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Asn Asp Ser Ala Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala Thr Ile 385 390 395 400 Asn Pro Gly Val Asp Gly Thr Asn Arg Ile Glu Ser Thr Ala Val Asp 405 410 415 Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val As<br/>n Arg Ala Ser Phe420 425 430Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly Cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp Glu Ser 450 455 460 Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe 

Lin Thr As G Gin Als Gly Ser 11e Als As An Als Gly Ser Val Pro Thr 495 Yr Val Trp Thr Arg Arg Arg Arg Val Asp Leu Ang Am Thr Ile Thr Pro 500 Ser Pro Leu Thr Gln Leu Pro Leu Val Lye Als Ser Als Pro Val Ser 510 Ser Pro Leu Thr Val Leu Lye Gly Fro Gly Phe Thr Gly Gly Gly Ile Leu 530 Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Thr Val Am 550 Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Als Ser Thr 550 Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Gly Gly Gly Er The Gly Arg 550 Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Gly Gly Gly Leu Gly Arg 550 Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Gly Gly Gly Cle Thr Val Arg 550 Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Gly Gln Glu Leu Thr Tyr Glu 550 Ser Pro Leu Thr Gln Als Glu Glu Flo Untr Thr Gly Pro Phe Asn Pro Pro 610 610 610 611 Ser Thr Gly Gly Glu Tyr Tyr Ile Arg Arg Ile Glu He Val Pro 660 Fra Arg Glu Arg Glu Als Glu Glu Asn Als Glu Glu Gly 617 Fra Phe Thr Gln Als Glu Glu He Uru Thr Val Ann Als Glu Gly 618 Fra Phe Thr Gln Als Glu Glu Glu Asn Leu Glu Als Als Glu Gly 619 Fra Arg Glu Das Glu Glu Glu Asn Leu Glu Als Als Glu Gly 610 Fra Arg Glu Als Glu Glu Glu Asn Leu Glu Als Als Glu Glu Her 710 711 712 712 712 712 712 712 712																	
500     505     510       stm: Arg Tie Thr Ohn Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 510     510       stm: Arg Tie Thr Ohn Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 510     515       stm: Arg Tie Thr Ohn Leu Lys Cily Pro Cily Phe Thr Cily Cily Cily Ile Leu 530     550       stm: Arg Thr Thr Ann Oily Thr Ene Cily Thr Leu Arg Val Thr Val Asn 550     550       stm: Front Cin Cil Tyr Arg Leu Arg Cily Cily Val Ser The Cily Ala Ser Thr 555     550       stm: Front Cil Cil Yer Cily Pro Cily Pro Phe Asn Pro Pro 610     555       stm: Front Cil Cil Cil Yer Thr Thr Thr Cily Pro Phe Asn Pro Pro 610     600       stm: Front Cil Cil Cil Yer Tyr Tie Asp Arg Tie Cilu Tie Val Pro 640     655       stal Ser Thr Cily Cil Vi Tyr Tyr Tie Asp Arg Tie Cilu Tie Val Pro 640     655       stal Ser Thr Cily Cil Vi Tyr Tyr Tie Asp Arg Tie Cilu Tie Val Pro 645     650       stal Ser Thr Cily Cilu Ang Cilu Ala Cilu Cilu Asp Leu Cilu Ala Ala Lye Lye 640     655       stal Ser Thr Cily Cilu Tyr Tyr Tie Asp Arg Tie Cilu Tie Val Pro 640     655       stal Ser Thr Cily Cilu Ang Cilu Asp Leu Cilu Ala Ala Lye Lye 640     650       stal Ser Thr Cily Cilu Ala Cilu Cilu Asp Leu Cilu Ala Ala Lye Lye 640     650       stal Ser Thr Cily Cilu Ala Cilu Cilu Asp Leu 640     650       stal Ser Thr Cilu Ala Cilu Cilu Asp Leu 650     650       stal Ser Thr Cily Cilu Ala Cilu Cilu Asp Leu 710     650       stal Ser Thr Cily Cilu Ala Cilu Cilu Asp Cilu Ala Cilu Ala Cilu Asp Cilu	Gln T	'hr	Asn	Gln		Gly	Ser	Ile	Ala		Ala	Gly	Ser	Val		Thr	
<pre>Man Arg lle Thr Gin Lew Pro Lew Val Lye Ala Ser Ala Pro Val Ser 515 Thr Gin Lew Pro Lew Val Lye Ala Ser Ala Pro Val Ser 530 Thr Thr Val Lew Lye Gly Pro Gly Phe Thr Gly Gly Gly Ile Lew 540 Thr Thr Val Lew Lye Gly Pro Gly Phe Thr Lew Arg Val Thr Val Am 555 Thr 550 Thr Thr Am Oly Thr Phe Gly Thr Lew Arg Val Thr Val Am 555 Thr 550 Thr 560 Thr Arg Glu Gly Lew Gly Gly Val Ser Ile Gly Ap 550 Thr 565 Thr Yar Glue Arg Val Lew Arg Gly Gly Gly Val Ser Ile Gly Ap 550 Thr 550 Thr Het Am Arg Gly Gln Glu Lew Thr Tyr Glu 605 The Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Am Pro Pro 610 The Thr Gln Ala Glu Glu Ile Um Thr Val Am Ala Glu Gly 651 Thr Gly Gly Glu Tyr Tyr Ile App Arg Ile Glu Ile Val Pro 660 Thr 615 Thr Thr Gly Glu Ala Clu Glu Ala Ala Glu Gly 650 Thr 616 Thr 300 Thr 616 Thr Thr Gly Pro Phe Am Pro Pro 660 Thr 615 Thr Glu Ala Glu Glu Glu Lew Thr Yar Us 660 Thr 617 Thr Gly Glu Tyr Tyr Ile App Arg Ile Glu Ile Val Pro 660 Thr 617 Thr 717 Thr 618 Thr 717 Thr 619 Thr 619 Thr 610 Thr 617 Thr 610 Thr 617 Thr 717 Thr 618 Thr 717 Thr 619 Thr 610 Thr 617 Thr 611 Thr 717 Thr 71 Thr 612 Thr 717 Thr 612 Thr 717 Thr 613 Thr 717 Thr 614 Thr 617 Thr 717 Thr 71 Thr 617 Thr 617 Thr 717 Thr 712 Thr 717 Thr 712 Thr 717 Thr 712 Thr 717 Thr 713 Thr 717 Thr 714 Thr 71 Thr 715 Th</pre>	Tyr V	7al	Trp		Arg	Arg	Asp	Val	-	Leu	Asn	Asn	Thr		Thr	Pro	
bit Thr Thr Val Leu Lye Gly Pro Gly Pho Thr Cly Gly Gly Ile Leu         530         try Avg of Thr Thr Aan Gly Thr Phe Gly Thr Leu Arg Val Thr Val Aan         555         see Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Pho Ala Ser Thr         570         bit Aan Pho Sen Tile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp         580         see Pro Leu Thr Gln Gln Tyr Arg Leu Arg Gly Gly Gln Glu Leu Thr Tyr Glu         580         581         581         582         583         583         584         585	Asn A	Arg			Gln	Leu	Pro			Lys	Ala	Ser			Val	Ser	
	Gly T	'hr		Val	Leu	Lys	Gly		Gly	Phe	Thr	Gly		Gly	Ile	Leu	
345       550       555       560         ier Pro Leu Th Gin Gin Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr       565         31y Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp         560       585         31y Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp         560       585         31y Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Gly Val Ser Ile Gly Asp         560       585         361 Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu         600       615         600       616         610       610 Glu Glu Ile Leu Thr Val Asn Ala Glu Gly         645       635         641       637         642       645         643       655         741 Asn Pro Ala Arg Glu Ala Glu Glu Glu Asp Leu Glu Ala Ala Lys Lys         660       665         741       Asn Pro Ala Arg Glu Ala Glu Glu Glu Asp Leu Glu Ala Ala Lys Lys         660       665         761       INFORMATION: Artificial sequence         7220       FENTHER:         7221       FORTHER:         72220       FENTHER:         7221       FORTHER:         72220       FENTHER:         7230       70         75 <td></td> <td></td> <td>Thr</td> <td>Thr</td> <td>Asn</td> <td>Glv</td> <td></td> <td>Phe</td> <td>Glv</td> <td>Thr</td> <td>Len</td> <td></td> <td>Val</td> <td>Thr</td> <td>Val</td> <td>Δen</td> <td></td>			Thr	Thr	Asn	Glv		Phe	Glv	Thr	Len		Val	Thr	Val	Δen	
565     570     575       Sly Am Phe Ser ILe Arg Val Leu Arg Gly Gly Val Ser ILe Gly App 590     590       Aral Arg Leu Gly Ser Thr Met Aen Arg Gly Gln Glu Leu Thr Tyr Glu 595     600       Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Aen Pro Pro 610     615       Ser Thr Gln Ala Gln Glu ILe Leu Thr Val Aen Ala Glu Gly 640     641       Yal Ser Thr Gln Ala Glu Glu ILe Leu Thr Val Aen Ala Glu Gly 640     642       Yal Ser Thr Gln Ala Glu Glu IVY TYr ILe Asp Arg TLE Glu ILE Val 645     645       Yal Aen Pro Ala Arg Glu Ala Glu Clu Asp Leu Glu Ala Ala Lys Lys 660     667       660     665     670       Vala     861     667       Vala     861     667       Vala     862     670       Vala     861     667       Vala     862     867       Vala     867     670       See Quin No 3     867     670       See Quin No 3     867	545	-				550			-		555	-				560	
580     585     590       Yal Arg Leu Gly Ser Thr Met Aan Arg Gly Gln Glu Leu Thr Tyr Glu     605       Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro     610       610     615     fr Thr Thr Gly Pro Phe Asn Pro Pro       625     630       626     640       7al Arg Glu Glu Qlu Tyr Tyr Ile Asp Arg Tle Glu Ile Val Pro       645       7al Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys       665       670       7al Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys       666       670       7al Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys       6670       7al Lawr Thr E iona       6671       7al Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys       6670       7al Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys       6671       6672       7210     FANTOR:       72110     ERMORT:       72110     ERMORT:       72120     TPRIOR:       721210     TROMATION: Artificial sequence       721210     ERMORT:       721210     ERMORT:       7210     TPRIOR:       7210     TPRIOR:       7210     TPRIOR:       7210     TPRIOR:       721	Ser P	ro	Leu	Thr		Gln	Tyr	Arg	Leu		Val	Arg	Phe	Ala		Thr	
$595 \qquad 600 \qquad 605$ see Phe Phe The The Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro $610 \qquad 615 \qquad 600 \qquad 605$ see Phe Phe Thr Gln Ala Glu Glu Ile Leu Thr Val Asn Ala Glu Gly $635 \qquad 640$ $7al Ser Thr Gly Gly Glu Tyr Tyr Ile Aap Arg Ile Glu Ile Val Pro 640 \qquad 655 7al Asn Pro Als Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys 660 \qquad 660 \qquad 670 what the theorem of theorem of theorem of the theorem of the$	Gly A	Asn	Phe		Ile	Arg	Val	Leu		Gly	Gly	Val	Ser		Gly	Asp	
610 615 620 the Thr Phe Thr Gln Ala Gln Glu Ile Leu Thr Val Asn Ala Glu Gly 635 640 Aal Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Arg Ile Glu Ile Val Pro 645 650 650 657 Aal Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys 660 665 667 Ala 2210- SEQ ID NO 3 2211- LENGTH: 2019 2212- TYPE: DNA 2213- ORGANISM: Artificial sequence 2205 FFATURE: 2223- OFHER INFCOMATION: Artificial sequence 164 2205 FFATURE: 2223- OFHER INFCOMATION: Artificial sequence 164 2205 FFATURE: 2223- OFHER INFCOMATION: Artificial sequence 2224 CHARMER: 2225 LOCATION: (1)(2019) 4400- SEQUENCE: 3 Attg aat cga aat aat cga atg gad tat gaa att att gat gcc ccc cat 48 let Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His 1 5 10 15 10 30 cca aat gca gcd ta caa aat atg aac tat aaa gat tac tta caa atg 144 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu Gln Met 40 45 50 40 45 50 10 75 70 Ser Lap App Val Arg Tyr Pro Ser Lau Ser Ile 50 cra gad gad tac act gat to tta tat aa at ct at tta tat 192 50 cra gad gad tac act gat to tta tat aa at ct at gtt tta tt 192 50 cra gad gad tac act gat to tta tat aa at ct at gtt gtt ggd gga ata 240 50 cra gad gad gad tac act gat to tta tat att ct att 192 50 cra gad gad gad tac act gad t gcd gtt gtt gtt gtd gdg gad ata 240 50 cra gad gad gad tac act gat gcd tto tgga caa ata gtd gdt ttat tat 288 10 cra gad gad gad tac act gad tto t gad caa gad tact gdt gdt gdt gdt gdt gdt gdt gdt gdt gd	Val A	arg		Gly	Ser	Thr	Met		Arg	Gly	Gln	Glu		Thr	Tyr	Glu	
125 6 630 635 640 Aal Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Arg Ile Glu Ile Val Pro 650 655 670 Aal Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys 665 670 Ala 14 2210 SEQ ID NO 3 2212 ILENGTH: 2019 2212 TYPE: DNA 2213 ORGANISM: Artificial sequence 2200 FEATURE: 2223 OTHER INFORMATION: Artificial sequence description: Cry9Cal Leu- 164 2200 FEATURE: 2223 OTHER INFORMATION: Artificial sequence description: Cry9Cal Leu- 164 2200 SEQUENCE: 3 4000 sp Asp Asp Cal Arg Tyr Pro Leu Ala Bro His 1 5 10 400 SEQUENCE: 3 400 Asp Arg Asp Asp Asp Val Arg Tyr Pro Leu Ala Ser Asp 3 400 Asp Asp Asp Asp Yal Arg Tyr Pro Leu Ala Cal Atg 144 405 Asp Asp Cal Cal Asp Asp Asp Yal Arg Tyr Pro Leu Ala Cal Atg 144 405 Asp Asp Asp Asp Yal Arg Tyr Pro Leu Ala Cal Atg 144 405 Asp Asp Asp Asp Asp Yal Arg Tyr Pro Leu Ala Cal Atg 144 405 Asp Asp Asp Asp Yal Arg Tyr Pro Leu Ser Ile 5 400 Asp Asp Asp Asp Yr Ile Asp Pro Ser Leu Ser Ile 5 400 Asp Asp Asp Asp Yr Ile Asp Pro Ser Leu Ser Ile 5 400 Asp Arg Asp Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Tip 80 400 Asp Tyr Thr Asp Ser Tyr Ile Asp Pro Ser Leu Ser Ile 6 50 Arg Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Tip 80 400 Asp Tyr Asp Asp Asp Cort tot gga caa ata gtg agt ttt tat 288 400 Asp Tyr Thr Asp Ser Tyr Ile Asp Pro Ser Leu Ser Ile 6 50 Arg Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Tip 80 400 Asp Tyr Asp Asp Asp Cort tot gga Caa Ata gtg agt ttt tat 288 400 Asp Tyr Thr Asp Ser Tyr Ile Asp Pro Ser Leu Ser Ile 60 400 Asp Tyr Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Tip 80 400 Asp Tyr Thr Asp Ser Tyr Ile Asp Pro Ser Leu Ser Ile 60 400 Asp Tyr Thr Asp Ser Tyr Ile Asp Pro Ser Leu Ser Ile 60 400 Asp Tyr Ala Leu Gly Val Pro Phe Ser Gly Gin Ile Val Ser Phe Tyr 70 400 Asp Ala Leu Gly Val Pro Phe Ser Gly Gin Ile Val Ser Phe Tyr 70 400 Asp Ala Leu Gly Val Pro Phe Ser Gly Gin Ile Val			Phe	Thr	Arg	Glu		Thr	Thr	Thr	Gly		Phe	Asn	Pro	Pro	
<pre>// Al Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Arg Ile Glu Ile Val Pro 645 650 670 // Al Aen Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys 660 665 670 // Ala // LineGTH: 2019 // LineGTH: 2019 // LineGTH: 2019 // LineGTH: 2019 // LineGTH: 2019 // LineGTH: Construct // LineGTH</pre>		'hr	Phe	Thr	Gln		Gln	Glu	Ile	Leu		Val	Asn	Ala	Glu	_	
Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys 660 $670$		Ser	Thr	Gly	_		Tyr	Tyr	Ile	_		Ile	Glu	Ile			
Ala 210 > SEQ ID NO 3 211 > LENGTH: 2019 212 > TPE: DNA 213 > ORGANISM: Artificial sequence 220 > FEATURE: 221 > INTER INFORMATION: Artificial sequence description: Cry9Cal Leu- 164 222 > FATURE: 222 > LOCATION: (1)(2019) 222 > LOCATION: (1)(2019) 2400 > SEQUENCE: 3 Attg aat cga aat aat caa aat gaa tat gaa att att gat gcc ccc cat 48 1 5 10 15 21 gg gg tgt cca tca gat gac gat gtg agg tat cct ttg gca agt gac 96 20 25 20 25 30 20 25 20	Val A	Asn	Pro	Ala		Glu	Ala	Glu	Glu		Leu	Glu	Ala	Ala		Lys	
2210> SEQ ID NO 3 2211> LENGTH: 2019 2212> TYPE: DNA 2213> ORGANISM: Artificial sequence 2205 PEATURE: 2225> COTHER INFORMATION: Artificial sequence description: Cry9Cal Leu- 164 2205 FEATURE: 2215> NAME/KEY: CDS 2222> LOCATION: (1)(2019) 2205 SEQUENCE: 3 Artg aat cga aat aat caa aat gaa tat gaa att att gat gcc ccc cat 48 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His 1 5 10 15 2014 15 2015 Sequence: 3 Artg agt gc cca tca gat gac gat gtg agg tat cct ttg gca agt gac 96 2015 201 2015 201 2016 2017 2017 2017 2017 2017 2017 2017 2017	۸la			660					665					670			
$\frac{4400}{5} \text{ SEQUENCE: } 3$ $\frac{4400}{5} \text{ SEQUENCE: } 3$ $\frac{4400}{5} \text{ SEQUENCE: } 3$ $\frac{4400}{5} \text{ SEQUENCE: } 3$ $\frac{440}{5} \text{ Asn } \frac{4}{5} \text{ Asn } \frac$	<213> <220> <223> <220> <221>	OR FE OT 16 FE NA	GANI ATUF HER 4 ATUF ME/F	SM: E: INFO RE: CEY:	CDS	TION:	: Art	-		seque	ence	desc	cript	ion:	: Cry	79Cal	Leu-
MetAsnAsnGlnAsnGluTyrGluIleIleAspAlaProHis15101515101596cgtgggtgtccatcagatgatgggtatcctttggcaagtgat96cysGlyCysProSerAspAspAspValArgTyrProLeuAlaSerAsp96ccaaatgcagcgttacaaaatatggattattatAsp20ccaaatgcagcgttacaaaatatggattattat144caaaataattataaagattattat144caagaggattataattataaagattat144caagaggattatcaaaattataaacattat144caagaggattatcaaattataaaccttattat192caagatgadgattatcatataatcctagttat192caagatgadgadgatgatgatgadata240cbr5570757580240240cbr7075807580240cbr7075 <td< td=""><td></td><td></td><td></td><td></td><td></td><td>••(20</td><td>,,,,</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></td<>						••(20	,,,,										
Eys Gly Cys Pro Ser Asp Asp Asp Asp Xal Arg Tyr Pro Leu Ala Ser Asp 2014420253020253020253020253020253020253020253020253020253020253020253020253020253020253020253020253020253020354020Asp Tyr Leu Gln Met 453540204520452025204520252025202620272027202820252025202520262025202621212223232424024257025262075202621212223232402402402570257026702780282829212021 <tr< td=""><td>Met A</td><td></td><td></td><td></td><td>Asn</td><td></td><td></td><td></td><td></td><td>Glu</td><td></td><td></td><td></td><td></td><td>Pro</td><td></td><td>48</td></tr<>	Met A				Asn					Glu					Pro		48
Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu Gln Met       35       40       45         aca gat gag gac tac act gat tct tat ata aat cct agt tta tct att       192         Chr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu Ser Ile       50       192         acg ggt aga gat gca gtt cag act gcg ctt act gtt gtt ggg aga ata       240         acer Gly Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Ile       80         act ggg gct tta ggt gtt ccg ttt tct gga caa ata gtg agt ttt tat       288				Pro					Val					Ala			96
Thr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu Ser Ile       50       55       60         Agt ggt aga gat gca gtt cag act gcg ctt act gtt gtt ggg aga ata       240         Ser Gly Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Ile       26       60         65       70       75       80         2tc ggg gct tta ggt gtt ccg ttt tct gga caa ata gtg agt ttt tat       288         Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Val Ser Phe Tyr       288			Āla					Met					Tyr				144
Ser Gly Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Ile 65 70 75 80 etc ggg gct tta ggt gtt ccg ttt tct gga caa ata gtg agt ttt tat 288 Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Val Ser Phe Tyr	Thr A	Asp					Asp					Pro					192
Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Val Ser Phe Tyr	Ser G					Val					Thr					Ile	240
	ctc c	aa	gct														288

-continued

and the other that all according tog or conget that get according the target of target of													con	tin	ued		
Glu Åla Phe Hei Arg Glu Val Glu Glu Leu Val Aan Gln Gln Ile Thr         gaa tit goa aga aat car got cit goa aga tig caa gga tia gga gac       432         gaa tit goa aga aat car got cit goa aga tig caa gga tia gga gac       432         fli phe Ala Arg Aan Gln Ala Leu Ala Arg fee Gln Gly Leu Gly Aap       440         itt aat gfa tat caa crit co cit ciaa aat tig tig got gat cag       450         aat gat aca tia aat tia agi gti git grit ggi got caa tit ata got tia       528         aat gat aca tia aat tia agi gti git grit ggi got caa tit ata got tia       528         aat gat aca tia aat tia agi gti git grit ggi got gat ata ggi cag       576         aga of u gat tit gta at got att coa cit git tig cag gat ggi cag       576         aga ggi to coa tit att git agt at goa cag gat ggi gga tig gga tig ggi gat cag       624         con ggi gti coa tit act gita git att goa caa ggat ggg gga tig gga gat cag       672         con gi ggi gea att coa cat tit gaa caa gig tig gga tig gga gat caa       672         con gi ggi gea att coa cat tit gga act ggi tag gat tig gaa caa act ggi tig git gat ta gat gat       720         col gig gea att coa cat at ta ta gat git git git git git git git gan ta ac ggit fig       720         col gag gaa aca act act at at ta tig gat git gat aat aca ggit fig       720         col gag gaa aca act act at ta ta gat git ggit ggit gat gag tig gag       720         col gag ggaa aca aca act act at ta ta gat ggit ggi ggi ta aga cag cac ac				Leu					Pro					Ala			336
Glu phe Åla Arg Aen Glú Åla Leu Åla Arg Leë Glu Glu Ju Leu Öly Ásp         130         tot tit at at gta tat caa ogt too ott caa aat tgg tig got gat oga       480         sor Phe Aen Val Tyr Glu Arg Ser Leu Glu Am Trp Leu Ala Aep Arg       155         aat gat aca tta aat tta eqt git git ogt oca tit at ogat tta       528         Ann Aep Thr Leu Aan Leu Ser Val Val Arg Ala Glu Phe Tie Ala Leu       528         gat cit gat tit git ogt got coa tit at ogat tta       576         age git coa tta cit a dit git got gat caa aat gog cag       624         oli Val Pro Leu Leu Ser Val Val Arg Ala Glu Ala Val Aen Leu       528         log git coa tta cit og tot cot tit tit gga gaa gat tit ca cit tit       624         oli Val Pro Leu Leu Ser Val Yr Ala Glu Ala Val Aen Leu His Leu       525         tit ta tta aaa gat goa tot tit tit gga gaa gat tig gga tto aca       672         Leu Leu Lue Lys Aap Ala Ser Leu Phe Gly Glu Gly Trp Gly Phe Thr       720         210       225       720         ag ta ca cat aat tac tig tag act tig the tat aca git tig gat cit agat cyt       768         Lys Tyr Thr An Tyr Cyr Glu Thr Trp Tyr Asp Arg Glu Leu Glu Leu Thr Ala       816         uo Arg Gly gaa act act aat act gat agt tit ty tig grig cit at aca act cy git aga act act act act act gat agt tit tig the grig cit at tit ca to ca act gat gat cyr       864         Lys Tyr Thr An Thr Leu Val Val Leu Aap Val Val Ala Leu Phe Pro Tyr <t< td=""><td></td><td></td><td><math>\mathtt{Phe}</math></td><td></td><td></td><td></td><td></td><td>Glu</td><td></td><td></td><td></td><td></td><td>Gln</td><td></td><td></td><td></td><td>384</td></t<>			$\mathtt{Phe}$					Glu					Gln				384
Ser Phe Aen Val Tyr Gin Arg Ser Leu Ch Aen Trp Leu Ala Aep Arg 145 145 145 145 145 145 145 145		$\mathtt{Phe}$					Āla					Gln					432
Ann Àsp Thr Leu Ann Leu Ser Val Val Arg Ala Gln Phe I had Leu 175gac ott gat ttt gtt aat got att oca ttg ttt got gat aat gga cag Asp Leu Asp Phe Val Asn Ala II has 180576cag gtt oca tta otg toa gta tat goa caa got gtg aat tta oat ttg Oln Val Pro Leu Leu Ser Val Tyr Ala Gln Ala Val Asn Gly Gln 190624cag gtt oca tta otg toa gta tat goa caa got gtg aat tta cat ttg I val Pro Leu Leu Ser Val Tyr Ala Gln Ala Val Asn Gly gap tta cac 200672cag gtg gaa att too cac tat att gao cgt ogt gan to cat act of the region of the ser try Tyr Tyr Asp Arg Gln Leu Olu Leu Thr Ala 215672cag ggg gaa att coo act att att gao cgt ogt at a ta cag ggt tta gat cgt 220720cag gag aca att ad tgt gaa att tgg tta at aca ggt tta gat cgt 220768cag aga aca act aat tao tgt gaa att ggt ta gat gut cac cat cat ta 220816Leu Arg Oly Thr Asn Thr Cyr Tyr Asp Arg Cln Leu Phe Oly Clu Leu Thr Ala 220816Leu Arg Oly Thr Asn Thr Clu Ser Thr Tyr Tyr Asp Arg 225816Leu Arg Oly Thr Asn Thr Clu Ser Thr Deu Arg Tyr His Oln Phe Arg 260816Leu Arg Oly Thr Asn Thr Clu Ser Thr Deu Arg Tyr His Oln Phe Arg 200864car ga ga at a cct tat gtg gta tta gat gtg gt cat and cac act of t 210912aga gaa atg act tao cot ca cac agg gat cac act act act cac act of the Asp Pro Tyr 200864car gu ta ta sce gat cog att tat acc cac cac cag cat att gtg gag 200864car gu ta ta sce gat cog att tat act cac cac cac gad ta acc of the tac cac sot 200912aga gaa atg act ta pd gat ta at coc tat act act the tac acc dot di cac cac 200912cas gag gta tat acc ogt tag tat tag acc ca	Ser					Gln					Asn					Arg	480
App Leu Asp Phe Val Asn Àla II e Pro Leu Phe Àla Val Asn ĉiy Gln130130130130130130130130130130130135141415151515141515151515151516161718181818191					Asn					Arg					Ala		528
clnValProLeu Leu SerValTyrAlaGlnAlaValAsnLeuHisLeu195200205205205205672tatatatataaag gg ga tot ot otttt ty ga gaa gga tog gg tot oaca672Leu Leu Luy SapAlaSerLeu Phe Gly Glu Gly Trp Gly Phe Thr 215720cag gg ga att co acatat tat gac cgt cg tat dat acc got 235720720cag dat ga ca attot cacatat tat gac cgt caatag acc gat cgt 235768Lys TyrThr AsnThr Cly Glu Thr Trp Tyr AsnThr Gly Leu Asp Arg 25575tta aga gga acaat act gga tag tat cat caatca tco cgt 260816Leu Arg Gly 215Thr Asn Thr Glu SerTrp Leu Arg Tyr His Gln Phe Arg 265864Arg Glu Met Thr Leu Val Val Leu Asp Val Val Ala Leu Phe Pro Tyr 275864ag gg at attaca ga tot tat coa acg gga ta acc coa cgt ad ttt caa tat 300912Tyr Asp Val Arg Leu Tyr Pro Thr Oly Ser Asn Pro Gln Leu Thr Arg 300912gag gta tat aca gat cg at tgt at ta act coa cag gt aat ggt to gg cta 310315320cat to cg cgt tgg ggt act aat coc tat aat act ttt tc gag cgt 3201008gaa at goc tcat to cgo cca cca cat ct ttt gat agg cg aat agc 3451036gaa at goc tcat to cgo cca cca cat ct tt tg at agg cfg aat agc 3451056gaa at goc tcat tat cgg tac act cca tat tat at tat agg cfg aat agc 3451056gaa at goc tcat cat cg cca cca cat ct tt tg at agg cfg aat agc 345<	-		-	Phe	-		-		Pro	-		-	-	Asn		-	576
Let Leu Leu Lys Asp Ala Ser Leu Phe Gly Glu Gly Trp Gly Phe Thr 210 220 220 Cag gdg gaa att toc aca tat tat gac cgt caa ttg gaa cta acc got 720 Gln Gly Glu Tle Ser Thr Tyr Tyr Asp Arg Gln Leu Glu Leu Thr Ala 235 240 aag tac act aat tac tg ga gaa act gg tat aat aca ggt tta gat cgt 768 Lys Tyr Thr Asn Tyr Cys Glu Thr Trp Tyr Asn Thr Gly Leu Asp Arg 245 255 tta aga gga aca aat act gaa agt tgg tta aga tac cac act cc cgt 816 Leu Arg Gly Thr Asn Thr Glu Ser Trp Leu Asp Tyr His Gln Phe Arg 260 275 tat aga gga act tag tg gt at ta gat gtt gtg gg cta ttt cc at tt Arg Glu Me Thr Leu Val Val Leu Asp Val Val Ala Leu Phe Pro Tyr 275 280 275 280 285 28			$\operatorname{Pro}$					Tyr					Asn				624
Glin Giý Glu Ile Ser Thr Tyr Tyr Asp Arg Glin Leu Glu Leu Thr Ála225230aag tac act aat tac tgt gaa act tgg tat aat aca ggt tta gat cgt768Lys Tyr Thr Asn Tyr Cys Glu Thr Trp Tyr Asn Thr Gly Leu Asp Arg 245768Leu Arg Gly Thr Asn Thr Glu Ser Trp Leu Arg Tyr His Gln Phe Arg 260816aag tag act tta gtg gta tta gat gtt gtg gcg cta ttt cca tat Arg Glu Met Thr Leu Val Val Leu Asp Val Val Ala Leu Phe Pro Tyr 2908642902802859121y Asp Val Arg Leu Tyr Pro Thr Gly Ser Asn Pro Gln Leu Thr Arg 290912291295295292295285293291295294280285295285285290291295291295295292295285293295295294295285295285285296285285297295285290295285290295285290295295290295295290295295290295295290295295290295295290295295290295295290295295290295295290295295291295295292295295293295295		Leu					Ser					Gly					672
LysTyrThrAsnTyrQyrGluThrTipTyrAsnThrGlyLeuAspAspAsp255ttaagagcaactgcaactgcagcggcaactgcagcagcaactgca </td <td>Gln</td> <td></td> <td></td> <td></td> <td></td> <td>Thr</td> <td></td> <td></td> <td>-</td> <td></td> <td>Gln</td> <td></td> <td></td> <td></td> <td></td> <td>Ála</td> <td>720</td>	Gln					Thr			-		Gln					Ála	720
Leu Arg Cly 260Thr Asn Thr Clu Ser Trp Leu Arg Tyr His Cln Phe Arg 265Cln Phe Arg 270aga gaa atg act tta gtg gta tta gat gtt gtg gcg cta ttt cca tat Arg Clu Met Thr Leu Val Val Leu Asp Val Val Ala Leu Phe Pro Tyr 275864tat gat gta cga ctt tat cca acg gga tca aac cca cag ctt aca cgt 290912tat gat gta cga ctt tat cca acg gga tca aac cca cag ctt aca cgt 290912gag gta tat aca gat ccg att gta ttt aat cca cca gct aat gtt gga 300960Glu Val Tyr Thr Asp Pro The Oly Pro Pro Pro Pro Pro Pro Pro Pro Ala Asn Val Gly 310960ctt tgc cga cgt tgg ggt act aat ccc tat aat act ttt tct gag ctc 3201008ctt tgc cga cgt tgg ggt act aat ccc tat aat act ttt gat agg ctg aat agc 3251008gaa aat gcc ttc att cgc cca cca cat ctt tt gat agg ctg aat acg 3201056fu Asn Ala Phe Tle Arg Pro Pro Val Ser Ser Asn Phe Met Asp 3551104Leu Thr Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe Met Asp 3501104tat tgg tca gga cat acg tta cgc cgt agt tat ctg acc 3601152tat tgg tca gga cat acg tta cgc cgt agt tat ctg acc 3701152tat tgg tca gga cat acg tta cgc cgt agt tat ctg acc 3701152tat tgg tca gga cat acg tta ggc cta att aca cac aca gg gta caca act 3701200					Tyr					Tyr					Āsp		768
Arg Glu Met 275Thr Leu Val Val Leu Asp Val Val Ala Leu Phe Pro Tyr 280912tat gat gta cga ctt tat cca acg gga toa aac cca cag ctt aca cgt 290912Tyr Asp Val Arg Leu Tyr Pro Thr Gly Ser Asn Pro Gln Leu Thr Arg 290960gag gta tat aca gat ccg att gta ttt aat cca cca gct aat gtt gga 310960Glu Val Tyr Thr Asp Pro Ile Val Phe Asn Pro Pro Ala Asn Val Gly 315960ctt tgc cga cgt tgg ggt act aat ccc tat aat act ttt tct gag ctc 3251008ctt tgc cga cgt tgg ggt act aat ccc tat aat act ttt tct gag ctc 3251008gaa aat gcc ttc att cgc cca cca cat ctt ttt gat agg ctg aat agc 3451056gaa aat gcc ttc att cgc cca cca cat ctt ttt gat agg ctg aat agc 3451056tta aca atc agc agt aat cga ttt cca gtt tca tct aat ttt atg gat 3551104Leu Thr Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe Met Asp 3551104Leu Thr Ile Ser Ser Asn Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg				Thr					Trp					Gln			816
TyrAspValArgLeuTyrProThrGlySerAsnProGlnLeuThrArg290gaggtatatacagatccgattgtatttaatccaccagatgfagfaGluValTyrThrAspProIleValPheAsnProProAlaAsnValGlygfa305ValTyrThrAspProIleValPheAsnProProAlaAsnValGlygfa305ValTyrThrAspProIleValPheAsnProProAlaAsnValGlygfa305ValTyrThrAspProIleValPheAsnProProAlaAsnValGly320ctttgccgacgttggggaactaatcccttttttttttttdgadgactcLeuCysArgArgTyrGlyThrAsnProTyrTyrAsnAsnValGluLeuJagaaaatgccttcattcccccaccacctttttttfagaaaatgccttcattccaccaccaccatttfagcfagaaaatgccttca			Met					Leu					Leu				864
GluValTyrThrAspProIleValPheAsnProProAlaAsnValGly305310310315315320320ctttgccgacgttggggtactaatccctttttcdgggtcd1008LeuCysArgArgTrpGlyThrAsnProTyrAsnThrPheSerGluLeu3351008gaaaatgccttattcgcccaccacatcttttttdtdtdGluAsnAlaPheTheSerThrAsnPheSerSerSerSerSerSer355gaaaatcgccgtagttattcgcgtagttattattagtagtataggtadagcgaaaatagcaatcgccgaaatcgcserAsnAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerAsnSerSerAsnSerSerAsnSerSerAsnSerSerSerAsnSerSerSerSerAsnSerSerSerSerAsnSerSerSerAsnSerSerSerSerAsnSerSerSerAsn		Asp					$\operatorname{Pro}$					Pro					912
Leu Cys Arg Arg Trp Gly Thr Asn Pro Tyr Asn Thr Phe Ser Glu Leu 3251056gaa aat goc ttc att cgc cca cca cat ctt ttt gat agg ctg aat agc Glu Asn Ala Phe Ile Arg Pro Pro His Leu Phe Asp Arg Leu Asn Ser 3401056tta aca atc agc agt aat cga ttt cca gtt tca tct aat ttt atg gat 3551104Leu Thr Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe Met Asp 3601104tat tgg tca gga cat acg tta cgc cgt agt tat ctg aac gat tca gca 3701152tat tgg tca gga cat acg tta gcc cgt agt tat cag acc aca aga gca aca att Qat Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Arg Ala Thr Ile1200	Glu				Asp	Pro	Ile			Asn	Pro						960
Glu Asn Ala Phe Ile Arg Pro Pro His Leu Phe Asp Arg Leu Asn Ser 3401104tta aca atc agc agt aat cga ttt cca gtt tca tct aat ttt atg gat Leu Thr Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe Met Asp 3551104tat tgg tca gga cat acg tta cgc cgt agt tat ctg aac gat tca gca 3701152tat tgg tca gga cat acg tta ggc cta att aca acc aca aga gca aca att Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Arg Ala Thr Ile1200					Trp					Tyr					Glu		1008
Leu Thr Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe Met Asp 3553603651152tat tgg tca gga cat acg tta cgc cgt agt tat ctg aac gat tca gca 3701152Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Asn Asp Ser Ala 3701152gta caa gaa gat agt tat ggc cta att aca acc aca aga gca aca att Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala Thr Ile1200				Phe					His					Leu			1056
Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Asn Asp Ser Ala 370 375 380 gta caa gaa gat agt tat ggc cta att aca acc aca aga gca aca att 1200 Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala Thr Ile			Ile					Phe					Asn				1104
Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala Thr Ile		Trp					Leu					Leu					1152
	Val					Tyr					Thr					Ile	1200

-continued

-continued	
aat ccc gga gtt gat gga aca aac cgc ata gag tca acg gca gta gat Asn Pro Gly Val Asp Gly Thr Asn Arg Ile Glu Ser Thr Ala Val Asp 405 410 415	1248
ttt cgt tct gca ttg ata ggt ata tat ggc gtg aat aga gct tct ttt Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe 420 425 430	1296
gtc cca gga ggc ttg ttt aat ggt acg act tct cct gct aat gga gga Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly 435 440 445	1344
tgt aga gat ctc tat gat aca aat gat gaa tta cca cca gat gaa agt Cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp Glu Ser 450 455 460	1392
acc gga agt tca acc cat aga cta tct cat gtt acc ttt ttt agc tttThr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe465470475480	1440
caa act aat cag gct gga tct ata gct aat gca gga agt gta cct act Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr 485 490 495	1488
tat gtt tgg acc cgt cgt gat gtg gac ctt aat aat acg att acc cca Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro 500 505 510	1536
aat aga att aca caa tta cca ttg gta aag gca tct gca cct gtt tcg Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 515 520 525	1584
ggt act acg gtc tta aaa ggt cca gga ttt aca gga ggg ggt ata ctc Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu 530 535 540	1632
cga aga aca act aat ggc aca ttt gga acg tta aga gta acg gtt aat Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn 545 550 555 560	1680
tca cca tta aca caa caa tat cgc cta aga gtt cgt ttt gcc tca aca Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr 565 570 575	1728
gga aat ttc agt ata agg gta ctc cgt gga ggg gtt tct atc ggt gat Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp 580 585 590	1776
gtt aga tta ggg agc aca atg aac aga ggg cag gaa cta act tac gaa Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu 595 600 605	1824
tcc ttt ttc aca aga gag ttt act act act ggt ccg ttc aat ccg cct Ser Phe Phe Thr Arg Glu Phe Thr Thr Gly Pro Phe Asn Pro Pro 610 615 620	1872
ttt aca ttt aca caa gct caa gag att cta aca gtg aat gca gaa ggt Phe Thr Phe Thr Gln Ala Gln Glu Ile Leu Thr Val Asn Ala Glu Gly 625 630 635 640	1920
gtt agc acc ggt ggt gaa tat tat ata gat aga att gaa att gtc cct Val Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Arg Ile Glu Ile Val Pro 645 650 655	1968
gtg aat ccg gca cga gaa gcg gaa gag gat tta gaa gcg gcg aag aaa Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys 660 665 670	2016
gcg Ala	2019
210. CEO TE NO 4	

<210> SEQ ID NO 4 <211> LENGTH: 673 <212> TYPE: PRT <213> ORGANISM: Artificial sequence

<220> FEATURE: <223> OTHER IN 164		: Artificial	sequence	description:	Cry9Cal Leu-
<400> SEQUENCE	: 4				
Met Asn Arg As 1	n Asn Gln 5	Asn Glu Ty	r Glu Ile 10	Ile Asp Ala	Pro His 15
Cys Gly Cys P:	o Ser Asp 20		l Arg Tyr 5	Pro Leu Ala 30	Ser Asp
Pro Asn Ala A 35	.a Leu Gln	Asn Met As 40	n Tyr Lys	Asp Tyr Leu 45	Gln Met
Thr Asp Glu A: 50	sp T <b>y</b> r Thr	Asp Ser Ty 55	r Ile Asn	Pro Ser Leu 60	Ser Ile
Ser Gly Arg A 65	p Ala Val 70		a Leu Thr 75	Val Val Gly	Arg Ile 80
Leu Gly Ala Lo	eu Gly Val 85	Pro Phe Se	r Gly Gln 90	Ile Val Ser	Phe Tyr 95
Gln Phe Leu La 10		Leu Trp Pr 10		Asp Thr Ala 110	Ile Trp
Glu Ala Phe Ma 115	et Arg Gln	Val Glu Gl 120	u Leu Val	Asn Gln Gln 125	Ile Thr
Glu Phe Ala A 130	g Asn Gln	Ala Leu Al 135	-	Gln Gly Leu 140	Gly Asp
Ser Phe Asn Va 145	al T <b>y</b> r Gln 150		u Gln Asn 155	Trp Leu Ala	Asp Arg 160
Asn Asp Thr Le	eu Asn Leu 165	Ser Val Va	l Arg Ala 170	Gln Phe Ile	Ala Leu 175
Asp Leu Asp Pl 1		Ala Ile Pr 18		Ala Val Asn 190	Gly Gln
Gln Val Pro Lo 195	eu Leu Ser	Val Tyr Al 200	a Gln Ala	Val Asn Leu 205	His Leu
Leu Leu Leu L <u></u> 210	vs Asp Ala	Ser Leu Ph 215		Gly Trp Gly 220	Phe Thr
Gln Gly Glu I 225	e Ser Thr. 230		p Arg Gln 235	Leu Glu Leu	Thr Ala 240
Lys Tyr Thr A	n Tyr Cys 245	Glu Thr Tr	p Tyr Asn 250	Thr Gly Leu	Asp Arg 255
Leu Arg Gly T 20		Glu Ser Tr 26		Tyr His Gln 270	Phe Arg
Arg Glu Met Tl 275	ır Leu Val	Val Leu As 280	p Val Val	Ala Leu Phe 285	Pro Tyr
Tyr Asp Val A 290	ng Leu Tyr	Pro Thr Gl 295		Pro Gln Leu 300	Thr Arg
Glu Val Tyr Tl 305	nr Asp Pro 310		e Asn Pro 315	Pro Ala Asn	Val Gly 320
Leu Cys Arg A:	g Trp Gly 325	Thr Asn Pr	o Tyr Asn 330	Thr Phe Ser	Glu Leu 335
Glu Asn Ala Pl 34		Pro Pro Hi 34		Asp Arg Leu 350	Asn Ser
Leu Thr Ile So 355	er Ser Asn	Arg Phe Pr 360	o Val Ser	Ser Asn Phe 365	Met Asp
Tyr Trp Ser G	y His Thr	Leu Arg Ar	g Ser Tyr	Leu Asn Asp	Ser Ala

```
-continued
```

370 375 380
Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Arg Ala Thr Ile385390395400
Asn Pro Gly Val Asp Gly Thr Asn Arg Ile Glu Ser Thr Ala Val Asp 405 410 415
Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe 420 425 430
Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly 435 440 445
Cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp Glu Ser 450 455 460
Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe 465 470 475 480
Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr 485 490 495
Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro 500 505 510
Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 515 520 525
Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu 530 535 540
Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn 545 550 555 560 Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr
Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp
Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu
595 600 605 Ser Phe Phe Thr Arg Glu Phe Thr Thr Gly Pro Phe Asn Pro Pro
610 615 620 Phe Thr Ghn Ala Ghn Glu Ile Leu Thr Val Asn Ala Glu Gly
625 630 635 640 Val Ser Thr Gly Glu Tyr Tyr Ile Asp Arg Ile Glu Ile Val Pro
645 650 655 Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys
660 665 670 Ala
<pre>&lt;210&gt; SEQ ID NO 5 &lt;211&gt; LENGTH: 2019 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal Phe- 164 &lt;220&gt; FEATURE: &lt;221&gt; NAME/KEY: CDS &lt;222&gt; LOCATION: (1)(2019) &lt;&lt;400&gt; SEQUENCE: 5</pre>
atg aat cga aat aat caa aat gaa tat gaa att att gat gcc ccc cat 48 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His 1 5 10 15

tgt ggg Cys Gly cca aat Pro Asr aca gat Thr Asp 50 agt ggt	y C t g n A t g 0 t a	ys ca la 35 ag	Pro 20 gcg Ala	Ser tta	Asp	Asp								96
Pro Asn aca gat Ihr Asp 50	n Ā t g p G t a	la 35 ag	Ala			+					30		нор	
Thr Asp 50	p G 0 t a		qac		GIII									144
agt ggt														192
Ser Gly 65	уА													240
ctc ggg Leu Gly														288
caa tto Gln Phe		eu												336
gaa gct Glu Al <i>a</i>	a P													384
gaa ttt Glu Phe 130	еĀ				-	-		-	-				-	432
tct ttt Ser Phe 145														480
aat gat Asn Asp														528
gac ctt Asp Leu		sp												576
cag gtt Gln Val	1 P													624
tta tta Leu Leu 210	u L													672
cag ggg Gln Gly 225										-			-	720
aag tac Lys Tyr														768
tta aga Leu Arg		ly												816
aga gaa Arg Glu	u M													864
tat gat Iyr Asp 290	рÝ												2	912
gag gta Glu Val 305				-	-		-			-		-		960

-continued

<pre>stt top cgs cgs cgs cgs cgs can are one take as at effect the term of the ser cgu cgs cgs cgs cgs cgs can are grown of the cgs cgs cgs cgs cgs cgs cgs cgs cgs cgs</pre>	_												_	COII	υIII	ueu				
Glu Aen Ala Phe Tle Arg Pro Pro His Leu Phe Asp Arg Leu Aen Ser         340         1111         1111         1111          1111         1111         1111         1						Trp					Tyr					Glu		1008		
Lee Thr ILe Ser Ser Aan Arg Phe Pro Val Ser Ser Aan Phe Met Aap 365 tat tgg toa gga cat aog tta ogo ogt agt tat ctg aao gat toa goa Tyr Try Ser Giy His Thr Leu Arg Arg Ser Tyr Lee Aan Aap Ser Ala 370 ta caa gaa gat agt tat gu c cta att ara aco aco aga gca aca att 410 tat cor gga gtt ggt gga aca aac ogo ata gag toa acg gca gta gat Aan Pro Giy Val Aap Giy Thr Aan Arg Ind Gu Ser Thr Ala Val Aap Aan Pro Giy Val Aap Giy Thr Aan Arg go dtg aat aga gct tot ttt 940 tat cor gga gtt gat gga aca aac ogo atg gga gtt gat aga ggt gg 410 tat cor gga gtt gat gga aca aac gor atg agg toa atg gg cg tat gag 410 tat cor gag agg ctg tg tt aat ggt acg act ctc cct gt aat gga gga 410 tat cor giv Giy Leu IDe Giy ID Thr Far Arg Jia Aan Arg Ala Ser Phe 420 tat cor gga ggr tg tt aat ggt acg act ctc cct gt aat gga gga 410 tag agg agg ct g tt aat ggt acg act ctc cct gt aat gga gga 410 tag agg agg ct g tt aat ggt ang at a co ao on gg tga adt 430 tg kag gat cac acc cat aga cta tot cat gt acc tt tt tag ttt Hr Giy Ser Arn Aap Leu Thr Aan Arg Giu Leu Pro Pro Nap Giu Ser 450 tr Gag agg tt ca acc cat aga cta tot cat gt ta act with arg twi 450 ta far ga gat cac cat aga cta tot cat gt ta cot with arg twi 450 tag agg agt tca acc cat aga cta tot cat gt ta cot with arg twi 450 ta far Gig Agg tat caa cc at aga cta tat at ang ga agt gta cot act 450 tat gt tg ag cot cy cg qa gtg tg tg ta at agg gat ct act cor fill 450 tat gt tg ag cot cy cg qa fg tg gga tat at acg at acg agt agt acc acc act 450 tat gt tg acc cgt cgt gat gga gat ta acg ga gg gg gg gg gg gg ggi ggi ggi ggi g					Phe					His					Leu			1056		
Tyr Trj Ser Gly His Thr Leu Arg Arg Ser Tyr Leu An Arp Ser Ala 370 375 375 376 376 377 377 377 377 377 377				Ile					Phe					Asn				1104		
Val Chi Chi Xap SerTyr Ci J Leu II e Thr Thr Thr Arg Ala Thr IIe385396397386397398387398398aat coo gag gt gat gat gag aca aac ogo at gag to a co go at gag to a da gag to a da gag to a da gag to a da gag to a da gag to a da gag to a da gag to a da gag of to t tit1248tit cgt tot goa tig at ggt at a tat ggt acg act tot of got and gag gag acg tig tit at ggt acg act tot cot got at gag got to da gag do tig tit at ggt acg act tot cot got at gag gag acg to a da gag do to at a gag acg acg a do to a da gag acg to a da gag acg acg a do to a da gag acg acg to a da gag acg acg a do to a da ga acg acg acg a do to a da ga acg acg acg a do to a da ga acg acg acg a do to a da ga acg acg acg a do to a da ga acg acg acg acg acg acg a do to a da ga acg acg acg acg acg a do to a da ga acg acg acg acg acg acg acg acg acg			Trp					Leu					Leu					1152		
Aen Pro Čiy Val Aep Čiy Thr Aen Arg IIe člů ser Thř Åla Val Åep111112114115114115114115114115115116116117117117117118118118119119111 <t< td=""><td>v</td><td>al</td><td></td><td></td><td></td><td></td><td>Tyr</td><td></td><td></td><td></td><td></td><td>Thr</td><td></td><td></td><td></td><td></td><td>Ile</td><td>1200</td><td></td><td></td></t<>	v	al					Tyr					Thr					Ile	1200		
Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe 420420gtc cca gqa gqc tfg ttt aat gqt acg act tot cct gct aat gqa qga 13441344Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly 4351344tgt qda gat ctc tat gat aca aat gat gaa tta cca cca gat gaa agt 4501392cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp Glu Ser 4501400acc gga agt tca acc cat aga cta tct cat gtt acc ttt tt agc ttt 4701440Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe 4701400aca act aat cca gct gga tct ata gct aat gca gga gat gt acc act 4701488Gin Thr Asn Gli Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr 4901488Ass Arg Jtt gac cct gt gat gtg gac ctt ata act acg at acc cca 5001536Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro 5001536aat aga at acc aca ta cca gtg ca gga tt acc gga ggg gg at a ctc 5151632ggt act acg gto tta aaa ggt cca gga tta aca gga gg ggg gg at a ctc 5101632ggt act acg gto tta aaa ggt cca gga tta acg gta acg gtt aat 5151684Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 5101632ggt act acg gto tta aaa ggt cca gga tta acg gta acg gtt aat 5301680Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn 5551680star acg aca act aat gc cca gg agg gga gt acg gtt aat 5401632gga at tca aca caa caat at gc cca gga gga gt cg tta gc aca 5751680fig aar the ga aca act aat gc cca gg agg gg agg gga gga gg gg gg gg1776gga aat tca gat ac					-	Asp				-	Ile			-	-	Val	-	1248		
ValProCityCityThrThrSerProAla AsnCityCity4354404454454454451392CysArgAspLeuTyrAspThrAsnAspGluLeuProProAspGluSer1392accggaagttcaacccatagactatctctaggttcaaccttaacacatacccatagactatcacatacttatcatactagattatattcaggttatt					Ala					Tyr					Ala			1296		
Cýs Arg ásp Leu Tyr Asp Thr Asn Ásp Glu Leu Pro Pro Ásp Glu Sér 4504601440acc gga agt toa acc oat aga cta tot oat gtt acc ttt ttt agc ttt Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe 4751440465470475480aca act aat cag gt gga tot ata got aat goa gga agt gta cot act at 4851488Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro 5001488at ggt tgg acc cgt gg gat gtg gac ctt aat aat acg att acc cca 5001536at agt att aca caa tta cca ttg gta aag gca tot gca cct gtt cg 5101584Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 5101632ggt act acg gto tta aaa ggt cat tt gga agg gt ata ctc 5101632ggt act acg gto cat atg gca catt gga agt tta agg gta act act 5101632gga aga ca act aat ggc aca ttt gga acg tta agg gta acg gt aat 5451680reg agg aca act at acg cat atg cct aga gtt cgt ttt gct ta aca 5501776gga aat tto agt ata agg gta ctc cgt gga ggg ggt tct at cc ggt gat 5501776gga aat tto agt ata agg gta ctc cgt gga ggg gg gg gg gg gt 5551776gga aat tto agt ata agg gta ctc cgt gga ggg gg ttot atc gga gga 5851776gga aat tto agt ata agg gta ctc cgt gga ggg gg cag gaa cta act act gga 5851776gga aat tto agt ata agg gta ctc cgt gga ggg cag gaa cta act act aga 5851776gga aat tto agt ata agg gta ctc cgt gga ggg cag gaa cta act act aga 5851776gga aat tto agt ata agg gta cto cgt gga ggg cag gaa cta act act aga 6001824val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu<				Gly					Gly					Ala				1344		
Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe 475480caa act aat cag gct gga tot ata gct aat gca gga agt gta cot act Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr 4851488cat gtt tgg acc cgt ogt gat gtg gac ott aat aat acg att acc cca Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro 5001536aat aga att aca cca tta cca ttg gta aag gca tot gca cct gtt cg S101584Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 5151632gqt act acg gtc tta aaa ggt cca gga tt aca gga ggg ggt at a ctc S301632Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Gly Ile Leu 5301680cqa aga aca act aat ggc aca ttt gga acg tta aga gta acg gtt aat Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn 5551728cqa aga aca act aat ggc aca ttg cca gga ggg ggt tot at cg gt ta S451728cqa aga ttc aca ga gt acc cct gt gga ggg ggg gg gg gg gat at ctc 5301728cga aga aca tta agg ta ccc cgt gg gg gg gg gg gg gg gg gg gt tca cca tta aca caa tat cgc cta aga gt cct gg gg ggg gg gg gg gg gg gg gg gg gg g			Arg					Thr					Pro					1392		
Gin Thr Asn Gin Ala Giy Ser Ile Ala Asn Ala Giy Ser Val Pro Thr 485Here Thr 485Here The Asn Ala Giy Ser Val App 490Pro Thr 495tat gtt tgg acc cgt cgt gat gtg gac ctt aat aat acg att acc cca 50015361536Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro 5101584aat aga att aca caa tta cca ttg gta aag gca tct gca cct gtt tcg Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 5151632ggt act acg gtc tta aaa ggt cca gga ttt aca gga ggg ggt ata ctc S301632cly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu S301680Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn 5501680Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn 5601728gga aat tc agt ata agg gta ctc cgt gga ggg ggt tt ct atc ggt ggt ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr 5701776gga aat tc agt ata agg gta ctc cgt gga ggg ggt act act tac gga gga gt ses The Asn Phe Ser Tle Arg Val Leu Arg Gly Gly Val Ser Tle Gly Asp 5851824gt aga tta ggg agc aca attg aca aga aga gga gga cca act tac caga ses The Arg Val Leu Arg Gly Gln Glu Leu Thr Tyr Glu 6001824gt aga tta ggg agg agt tact act act act act act tac gaa Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu 6051872	т	hr					His					Val					Phe	1440		
Tyr Val Trp Thr Arg Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro 500Image: Step Step Step Step Step Step Step Step						Ala					Asn					Pro		1488		
Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 525ggt act acg gtc tta aaa ggt cca gga ttt aca gga ggg ggt ata ctc Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu 5301632cga aga aca act aat ggc aca ttt gga acg tta aga gta acg gt aat Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn 5501680tca cca tta aca caa caa tat cgc cta aga gtt cgt tt gcc tca aca Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr 5701728gga aat ttc agt ata agg gta ctc cgt gga ggg ggt tct atc atc ggt ggt S801776gta aga tta ggg agc aca atg aac aga ggg cag gaa cta act tac gaa Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu 6001824tcc ttt ttc aca aga ggt tt act act act ggt ccg ttc aat ccg cct Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro1872					Thr					Asp					Ile			1536		
Gly Thr Thr Val Leu Lys Cly Pro Gly Phe Thr Gly Gly Gly Gly Ile Leu 5301680cga aga aca act aat ggc aca ttt gga acg tta aga gta acg gtt aat Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn 5501680tca cca tta aca caa caa tat cgc cta aga gtt cgt tt gcc tca aca Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr 5651728gga aat ttc agt ata agg gta ctc cgt gga ggg gtt tct atc ggt gat Gly Asn Phe Ser Ile Arg Val Leu Arg Cly Gly Val Ser Ile Gly Asp 5851776gtt aga tta ggg agc aca atg aac aga ggg cag gaa cta act tac gaa Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu 6001824tcc ttt ttc aca aga gag ttt act act act ggt ccg ttc aat ccg cct Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro1872			-	Ile					Leu	-	-	-		Ala		-	-	1584		
Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn 545Sto1728tca cca tta aca caa caa tat cgc cta aga gtt cgt ttt gcc tca aca Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr 5651728gga aat ttc agt ata agg gta ctc cgt gga ggg gtt tct atc ggt gat Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp 5801776gtt aga tta ggg agc aca atg aac aga ggg cag gaa cta act tac gaa Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu 6001824tcc ttt ttc aca aga gag ttt act act act ggt ccg ttc aat ccg cct Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro1872			Thr					Gly					Gly					1632		
Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr 5651776gga aat ttc agt ata agg gta ctc cgt gga ggg ggt tct atc ggt ggt Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp 5801776gtt aga tta ggg agc aca atg aac aga ggg cag gaa cta act tac gaa Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu 6001824tcc ttt ttc aca aga gag ttt act act act ggt ccg ttc aat ccg cct Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro1872	А	rg					Gly					Leu					Asn	1680		
Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp         gtt aga tta ggg agc aca atg aac aga ggg cag gaa cta act tac gaa         Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu         595         tcc ttt ttc aca aga gag ttt act act act ggt ccg ttc aat ccg cct         Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro						Gln					Arg					Ser		1728		
Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu 595 600 605 tcc ttt ttc aca aga gag ttt act act act ggt ccg ttc aat ccg cct 1872 Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro	_	-			Ser					Arg			-		Ile		-	1776		
Ser Phe Phe Thr Arg Glu Phe Thr Thr Gly Pro Phe Asn Pro Pro	-		-	Leu		-		-	Asn	-		-	-	Leu			-	1824		
						-							-			-		1872		

-continued

												0011	C T II	ucu			
						caa Gln										1920	
						tat Tyr										1968	
				Arg		gcg Ala										2016	
gcg Ala																2019	
<21 <21 <21 <22	0> FI 3> O:	ENGTH YPE: RGANI EATUH	H: 6' PRT ESM: RE:	73 Art:		al se : Art	-		seque	ence	desc	cript	tion	: Crj	79Cal	Phe-	
	0> SI Asn				Gln	Asn	Glu	Tvr	Glu	Ile	Ile	Asp	Ala	Pro	His		
1		-		5				-	10			-		15			
Сув	GLÀ	Сув	Pro 20	Ser	Asp	Asp	Asp	Val 25	Arg	Tyr	Pro	Leu	Ala 30	Ser	Авр		
Pro	Asn	Ala 35	Ala	Leu	Gln	Asn	Met 40	Asn	Tyr	Lys	Asp	Tyr 45	Leu	Gln	Met		
Thr	Asp 50	Glu	Asp	Tyr	Thr	Asp 55	Ser	Tyr	Ile	Asn	Pro 60	Ser	Leu	Ser	Ile		
Ser 65	Gly	Arg	Asp	Ala	Val 70	Gln	Thr	Ala	Leu	Thr 75	Val	Val	Gly	Arg	Ile 80		
Leu	Gly	Ala	Leu	Gly 85	Val	Pro	Phe	Ser	Gly 90	Gln	Ile	Val	Ser	Phe 95	Tyr		
Gln	Phe	Leu	Leu 100	Asn	Thr	Leu	Trp	Pro 105	Val	Asn	Asp	Thr	Ala 110	Ile	Trp		
Glu	Ala	Phe 115	Met	Arg	Gln	Val	Glu 120	Glu	Leu	Val	Asn	Gln 125	Gln	Ile	Thr		
Glu	Phe 130	Ala	Arg	Asn	Gln	Ala 135	Leu	Ala	Arg	Leu	Gln 140	Gly	Leu	Gly	Asp		
Ser 145		Asn	Val	Tyr	Gln 150	Arg	Ser	Leu	Gln	Asn 155		Leu	Ala	Asp	Arg 160		
	Asp	Thr	Phe		Leu	Ser	Val	Val	-		Gln	Phe	Ile				
Asp	Leu	Asp				Ala	Ile	Pro	170 Leu	Phe	Ala	Val	Asn	175 Gly	Gln		
Gln	Val	Pro	180 Leu		Ser	Val	Tvr	185 Ala	Gln	Ala	Va]	Asn	190 Leu	His	Leu		
		195					200					205					
	210		-	-		Ser 215			-		220	-	-				
Gln 225	Gly	Glu	Ile	Ser	Thr 230	Tyr	Tyr	Asp	Arg	Gln 235	Leu	Glu	Leu	Thr	Ala 240		
Lys	Tyr	Thr	Asn	<b>Tyr</b> 245	Суз	Glu	Thr	Trp	<b>Ty</b> r 250	Asn	Thr	Gly	Leu	<b>A</b> sp 255	Arg		
Leu	Arg	Gly	Thr 260		Thr	Glu	Ser	Trp 265	Leu	Arg	Tyr	His	Gln 270	Phe	Arg		

Arg Glu Met Thr Leu Val Val Leu Asp Val Val Ala Leu Phe Pro Tyr 275 280 285 Tyr Asp Val Arg Leu Tyr Pro Thr Gly Ser Asn Pro Gln Leu Thr Arg 290 295 300 Glu Val Tyr Thr Asp Pro Ile Val Phe Asn Pro Pro Ala Asn Val Gly 310 305 315 Leu Cys Arg Arg Trp Gly Thr Asn Pro Tyr Asn Thr Phe Ser Glu Leu 325 330 335 Glu Asn Ala Phe Ile Arg Pro Pro His Leu Phe Asp Arg Leu Asn Ser 340 345 350 Leu Thr Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe Met Asp 355 360 365 Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Asn Asp Ser Ala 375 380 Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala Thr Ile 385 390 395 400 395 Asn Pro Gly Val Asp Gly Thr Asn Arg Ile Glu Ser Thr Ala Val Asp 405 410 415 Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe 420 425 430 Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly 435 440 445 Cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp Glu Ser 450 455 460 Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe 465 470 475 Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr 485 490 495 Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro 500 505 510 Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 515 520 525 520 515 Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu 535 540 Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn545550550555 Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr 565 570 575 570 Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp 580 585 590 Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu 595 600 605 Ser Phe Phe Thr Arg Glu Phe Thr Thr Gly Pro Phe Asn Pro Pro 610 615 620 

 Phe Thr Phe Thr Gln Ala Gln Glu Ile Leu Thr Val Asn Ala Glu Gly

 625
 630
 635
 640

 630 Val Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Arg Ile Glu Ile Val Pro 645 650 655 Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys 660 665 670

Ala

<pre>&lt;210&gt; SEQ ID NO &lt;211&gt; LENGTH: 20 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFO 164 &lt;220&gt; FEATURE: &lt;221&gt; NAME/KEY: &lt;222&gt; LOCATION: &lt;400&gt; SEQUENCE:</pre>	D19 Artificial se DRMATION: Art CDS (1)(2019)	-	description: Cr	y9Cal Glu-
atg aat cga aat Met Asn Arg Asn 1				) His
tgt ggg tgt cca Cys Gly Cys Pro 20				
cca aat gca gcg Pro Asn Ala Ala 35				
aca gat gag gac Thr Asp Glu Asp 50				
agt ggt aga gat Ser Gly Arg Asp 65			Val Val Gly Arc	
ctc ggg gct tta Leu Gly Ala Leu				e Tyr
caa ttc ctt tta Gln Phe Leu Leu 100				
gaa gct ttc atg Glu Ala Phe Met 115				
gaa ttt gca aga Glu Phe Ala Arg 130				
tct ttt aat gta Ser Phe Asn Val 145			Trp Leu Ala Asp	
aat gat aca gaa Asn Asp Thr Glu				a Leu
gac ctt gat ttt Asp Leu Asp Phe 180				
cag gtt cca tta Gln Val Pro Leu 195				-
tta tta tta aaa Leu Leu Leu Lys 210				
cag ggg gaa att Gln Gly Glu Ile 225			Leu Glu Leu Thi	

-continued

ag to c act act to ty ga act typ the act see out to gat to gat out gat out       763         Lys Tyr Thr Am Tyr Cyc Glu Thr Thr Tyr Yam Am Thr Gly Les App Arg 255       763         Law Arg Gly Am a act act gas act gat gat to gat act cot act act act to gat cot act act act act act act and y and y and y all y all the gat gyr file file the Arg 276       814         agg gaa act act the gyg gat tat act cot act act act act act act act act act ac
Lee Arg Gly The Aam The Glu See Try Lee Arg Tyr His Gln Phe Arg 270 age gam atg act the gtg gtm tha gat gtt gtg gog ctm the com that arg and the mean of the second secon
Arg Cut Het Thr Leu Val Val Leu App Val Val Ala Leu Dhe Pro Tyr 280280280tat gat gat oga ott tat oca aog gga toa aac ota cag ott aca ogt 290912ag gta tat aca gat cog att gta tat ac oca coa got att gga 310960Gil Val Tyr Thr Ap Pro The Oll Val Phe Aen Pro Tyr Aan Thr Phe Ser Glu Leu 1935960Gil Val Tyr Thr Ap Pro To He Val Phe Aen Pro Tyr Aan Thr Phe Ser Glu Leu 
Tyr App Val Arg Leu Tyr Pro Thr Ĝly Ser Am Pro Gln Leu Thr Arg 290300gag qta tat aca gat cog att gta ttt aat coa coa got aat gtt gga Glu Val Tyr Thr App Pro ILe Val Phe Am Pro Pro Ala Am Val Gly 315960ctt tgo cga cgt tgg ggt act aat coc tat aat act tit tot gag rtc 12201008cu Cyr Arg Arg Trp Gly Thr Am Pro Tyr Am Thr Phe Ser Glu Leu 3201008gaa att goc tct att cgc cca cca cat ctt tit gat agg ctg aat agc Glu Am Ala Phe TLA Arg Pro Pro His Leu Phe App Arg Leu Am Ser 3401056gaa att goc tc att cgc cca cca cat ctt tit gat agg ctg aat agc 3401014Leu Thr ILe Ser Sar Am Ang Phe Pro Val Ser Ser Am Phe Met App 3551104Leu Thr ILe Ser Sar Am Ang The Pro Val Ser Ser Am Phe Met App 3551104set agg cag agt agt gg ct a dt aca acc aca agc gca aca att 13751200Tyr Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Am Am Sep Ser Ala 3801152Tyr Try Ser Gly His Thr Leu Arg Arg Ser Tyr Gly Leu ILe Thr Thr Arg Ala Thr ILe 3851200aat coc gga gt gat gga aca acc ogc ata gag tca acg gca gta gat 4001248Am Pro Gly Val App Gly Thr Am Arg ILe Glu Ser Thr Ala Val Alap 4001344401405400405400405400405400405405406405407405408405409405409405400405405405405405405406407408408409<
Gli Val Tyr Thr Asp Pro Ile Val Phe Am Pro Pro Ala Asn Val Gly       310         310       310       310         Gtt tgo oga ogt tgg ggt act aat occ tat aat act tit tot gag otc       1009         322       330       330         gaa aat goc tto att ogo cac cac act ott tit gat agg otg aat agc       1056         Gli Aan Ala Phe Tile Arg Pro Pro Bis Leu Phe Asp Arg Leu Aan Ser       340         340       345       345         355       365       1104         Leu Thr Tile Ser Ser Aan Arg Phe Pro Val Ser Ser Aan Phe Ret Pag       1104         350       350       365         70       115 Ser Ger Aan Arg Phe Pro Val Ser Ser Aan Phe Ret Pag Org Aat Agg       1152         717       Try Gag Cat arg gt act ogg ct at at ca acc cac ag gat ca gat ca gat agt at 375       310         718       Try Try Gag att agg tat agg caca acc gat agt tat agg ogg adg adg at agg age act at 400       1200         325       320       320       1200         326       320       320       1200         326       320       320       1200         320       320       320       320       1200         320       320       320       320       1200         320       320       320       320
Leu Cys Arg Arg Trỳ Gly Thr Aen Pro Tyr Aen Thr Phe Ser Glu Leu 330 gaa aat goo tho att ogo con oon at oft titt gat agg oft gaat ago Glu Aen Ala Phe Ile Arg Pro Pro His Leu Phe Asp Arg Ieu Aen Ser 340 1056 tita aca ato ago agt aat ogo tit con gut too tot aat tit atg gat Leu Thr Ile Ser Ser Aan Arg Phe Pro Val Ser Ser Aen Phe Met Asp 355 tat tgg toa gga cat acg tit ocgo cgt agt tat oft oa goa Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Aen Aep Ser Ala 370 gta caa gaa gat agt tat ggo ota att aca aco aca aga goa aca att Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala Thr Ile Asp Pro Gly Val Asp Ggy oft at at aca aco cac agg of a gat Asp Pro Gly Val Asp Gly Thr Aen Arg Ile Glu Ser Thr Ala Val Asp 400 aat coo gga gft gat gga aca aac cgo ata ggg toa agg oft tot tit Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe 420 420 420 420 420 420 420 420
Glu Aan ÅLa Phe ILe Arg Pro Pro His Leu Phe Åsp Arg Leu Aan Ser 340JAO1tta aca atc agt agt aat cga tt cca gtt toa tot att tt atg gat 1551104Leu Thr ILe Ser Ser Aan Arg Phe Pro Val Ser Ser Aan Phe Met Asp 3551152Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Aan Asp Ser Ala 3701152Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Aan Asp Ser Ala 3701152gta cca gaa gat agt tat ggc cta att acca acc acca aga gca aca att val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala Thr Ile 3951200aat ccc gga gtt gat gga aca aac cgc ata gag toa acg gca gta gat 4001248Aan Pro Gly Val Asp Gly Thr Asn Arg Ile Glu Ser Thr Ala Val Asp 4151248ttt cgt tot gra atg agt ata tat ggc gtg aat aga got tot ttt 4251296gtc cca gga gge tgt that ggt aga aca at gat gag att cc ccd gat aga gg 4101344val Pro Gly Gly Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe 4251344val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly 4551344val Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe 4251392ctg aga gat cct at gat aca at gat gaa tta ccc aga gag aga 4501392ca aga gat cta acc at aga tat ct cat gt acc tot cat gaa agt 4551392ca ac gga agt toa acc ot aga cta tot cat gt acc ttt ttt age ttt 4701440460475475ca act acg cgt ggt gat tat ag ct at act aga gga agt gta ct act 4551440dt for Nan Gln Ala Gly Ser Ile Ala Ash Ala Gly Ser Val Pro Thr 5051536ttt gt gac cct ct gt gat gtg gac ctt aat act acg gta cca dt acc cca 5051536 </td
Leu Thr ILe Ser Ser Ann Arg Phe Pro Val Ser Ser Ann Phe Met Asp 355 14t tgg tca gga cat acg tta cgc ggt agt tat ctg acc gat tca gca Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Ann Asp Ser Ala 370 370 370 370 370 370 370 370
Tyr Trip Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Asn Asp Ser Ala 380gta caa gaa gat agt tat ggc cta att aca acc aca aga gca aca att Val Gln Glu Asp Ser Tyr Gly Leu IIe Thr Thr Thr Arg Ala Thr TIe 3951200aat coc gga gtt gat gga aca aac cgc ata gag tca acg gca gta gat Asn Pro Gly Val Asp Gly Thr Asn Arg IIe Glu Ser Thr Ala Val Asp 4051248tt cgt tct gca ttg ata ggt ata tat ggc gtg aat aga gct tct ttt Phe Arg Ser Ala Leu TIe Gly Thr Thr Ser Pro Ala Asn Arg Ala Ser Phe 4201248gtc cca gga ggc ttg tt aat ggt acg act tct cct gct aat gga gga Val Pro Cly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Cly Gly 4351344val Fro Cly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Cly Gly 4351392gta caa ga gat tca act cat ag ag ta ta tc ca gca gat gaa agt 4551392cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp Glu Ser 4551440acc gga agt tca acc cat aga ct at cct at gt acc ttt ttt agc ttt 4551440Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe 4651488dcf Thr Asn Gln Ala Gly Ser IIe Ala Asn Ala Gly Ser Val Pro Thr 4851488dcn Thr Asn Gln Ala Gly Ser IIe Ala Asn Ala Gly Ser Val Pro Thr 4851488dcn Thr Asn Gln Ala Gly Ser IIe Ala Asn Ala Gly Ser Val Pro Thr 4851536gta taca cat aca tat cca ttg gt aga ct tc gca cct gtt tcg Sto1536gra cat act act act at ga ga ga ct tc gca cct gtt tcg Sto1536gra act aca cat aca cat ta cca ttg gta agg cat tc gca cct gtt tcg Sto1536gra act aca cat aca cat ta cca ttg gta agg cat tc gca cct gtt tcg Sto1536gra act aca gat ct aca cat ga gga ggt
ValGluÅspSerTyrGlyLeuIleThrThrThrArgÅlaThrIle385390390395395400400400400Aatcccggagtgtgagaacaacccccacdgagtgaga1248Ann ProGlyValAspGlyThrAnnArgIleGly12481296405410415415415129613441451296tttcgserAlaEquIleGlyGlyFile1344PheArgSerAlaLeuFileAlaSerFileyalproGlygatacaacaacaccgagad1344ValProGlyGlyLeuProAlaSerProAlaSeryalacagatctacaacagaatatcccccagaga1344ValAspLeuProProProAspGluSer13921344yalacagatctacacacacagaagaa140yalacagatctac
Asn Pro Gly Val Asp Gly Thr Asn Arg IIe Glu Ser Thr Ala Val Asp 4101296ttt cgt tct goa ttg ata ggt ata tat ggc gtg aat aga got tct ttt 4201296Phe Arg Ser Ala Leu IIe Gly IIe Tyr Gly Val Asn Arg Ala Ser Phe 4201344gtc cca gga ggc ttg ttt aat ggt acg act tct cct gct aat gga gga Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly 4351344tgt aga gat ctc tat gat aca aat gat gaa tta cca cca gat gaa agt 4551392gtc cy gga agt tca acc cat aga cta tct cat gtt acc ttt ttt agc ttt 4551400acc gga agt tca acc cat aga cta tct cat gtt acc ttt ttt agc ttt 4701440Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe 4701440caa act aat cag gct gga tct ata gct aat gca gga agt gta cct act 4851488caa act aat cag gct gg tgt gat ggg gac ctt aat aat acg att acc cca 5051536tat gtt tgg acc cgt cgt gat gtg gac ctt aat aat acg att acc cca 5051536at aga att aca caa tta cca ttg gta aag gca tct gca cct gtt tcg S101584at aga att aca caa tta cca ttg gta agg gga gt gta act gca 5051632
Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe 4201344gto cca gga ggc ttg ttt aat ggt acg act tct cct gct aat gga gga 4351344Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly 4451342tgt aga gat ctc tat gat aca aat gat gaa tta cca cca gat gaa agt 4501392cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp Glu Ser 4501392acc gga agt tca acc cat aga cta tct cat gtt acc ttt ttt agc ttt 4701440Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe 4751440acc aga qgt tgg tct ata gct aat gca gga agt gta cct act 4701488Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr 4851488tat gtt tgg acc cgt cgt gat gtg gac ctt aat aat acg att acc cca 5051536tat gtt tgg acc act aca ta cca ttg gta aag gca tct gca cct gtt tcg 5051584aat aga att aca caa tta cca ttg gta aag gca ttt aca gga ggg ggt ata ctc 5151584Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 5251632ggt act acg gtc tta aag ggt cca gga ttt aca gga ggg ggt ata ctc 5251632
ValProGiyGiyLeuPheAsnGiyThrSerProAlaAsnGiyGiytgtagagatctctatgatacaaatgatgattatccagatgaaagt1392CysArgAspLeuTyrAspThrAsnAspGluLeuProProAspGluSer1440accggaagttcaacccatagactatctcatgttaccttt1440ThrGlySerSerThrHisArgLeuSerHisArg1440ThrGlySerSerThrHisArgLeuSerHis1440ThrGlySerSerThrHisArgLeuSerHis1440Afo470475475480148014801480caaactcadcadgatgtagtagtagtagtaGlnThrAsnGlySerValProThr1488GlnThrAsnGlySerValProThr485GuSerSofSofSof16321536TyrValTrpThrAspAspLeuAspAspAspYrValTrpThrGlyAspLeuAspAspSof
CysArgAspLeuTyrAspThrAsnAspGluLeuProProAspGluSeraccggaagttcaacccatagactatctcatggtttt1440ThrGlySerSerThrHisArgLeuSerHisValThrPhePheSerPhe465470SerHisValSerHisValThrPhePheSerPhe465470475ThPhePheSerPheAso1440caaactcaggctggaactggaagtgtacctactGlnThrAsnGlySerIleAlaAsnAlaGlySerValProThr485485YelgaccctactactactactactactactactGlnThrAsnGlySerIleAlaAsnAlaGlySerValAp51636tatgttgaccctactgcacctgcacctgcagcacctgca1536TyrValTrpThrArgArgAspValAspAsnAsnThrFhe1632aatagaattacacaatacatacagcacctgcacctgta1632<
Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe 4654701488caa act aat cag gct gga tct ata gct aat gca gga agt gta cct act Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr 4851488tat gtt tgg acc cgt cgt gat gtg gac ctt aat aat acg att acc cca 5001536Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro 5051584aat aga att aca caa tta cca ttg gta aag gca tct gca cct gtt tcg 5151584asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 5201632ggt act acg gtc tta aaa ggt cca gga ttt aca gga ggg ggt ata ctc Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu1632
Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr 485tat gtt tgg acc cgt cgt gat gtg gac ctt aat aat acg att acc ccaTyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro 500aat aga att aca caa tta cca ttg gta aag gca tct gca cct gtt tcg S10aat aga att aca caa tta cca ttg gta aag gca tct gca cct gtt tcg S10ggt act acg gtc tta aaa ggt cca gga ttt aca gga ggg ggt ata ctc Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu
Tyr Val Trp Thr Arg Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro         500       505       510         aat aga att aca caa tta cca ttg gta aag gca tct gca cct gtt tcg       1584         Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser       525         ggt act acg gtc tta aaa ggt cca gga ttt aca gga ggg ggt ata ctc       1632         Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu       1632
Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser 515 520 525 ggt act acg gtc tta aaa ggt cca gga ttt aca gga ggg ggt ata ctc 1632 Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu
Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu

cga aga aca act aat ggc aca ttt gga acg tta aga gta acg gtt aat Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn 545 550 555 560	1680
tca cca tta aca caa caa tat cgc cta aga gtt cgt ttt gcc tca aca Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr 565 570 575	1728
gga aat ttc agt ata agg gta ctc cgt gga ggg gtt tct atc ggt gat Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp 580 585 590	1776
gtt aga tta ggg agc aca atg aac aga ggg cag gaa cta act tac gaa Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu 595 600 605	1824
tcc ttt ttc aca aga gag ttt act act act ggt ccg ttc aat ccg cct Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro 610 615 620	1872
ttt aca ttt aca caa gct caa gag att cta aca gtg aat gca gaa ggt Phe Thr Phe Thr Gln Ala Gln Glu Ile Leu Thr Val Asn Ala Glu Gly 625 630 635 640	1920
gtt agc acc ggt ggt gaa tat tat ata gat aga att gaa att gtc cct Val Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Arg Ile Glu Ile Val Pro 645 650 655	1968
gtg aat ccg gca cga gaa gcg gaa gag gat tta gaa gcg gcg aag aaa Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys 660 665 670	2016
gcg Ala	2019
010 FR TR NO 0	
<210> SEQ ID NO 8 <211> LENGTH: 673 <212> TYPE: PRT <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: Cry9Cal 164	. Glu-
<211> LENGTH: 673 <212> TYPE: PRT <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: Cry9Ca1	. Glu-
<pre>&lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal</pre>	. Glu-
<pre>&lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal</pre>	. Glu-
<pre>&lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal</pre>	. Glu-
<pre>&lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal</pre>	. Glu-
<pre>&lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;220&gt; SEQUENCE: 8 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His 1 5 10 15 Cys Gly Cys Pro Ser Asp Asp Asp Val Arg Tyr Pro Leu Ala Ser Asp 20 25 30 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu Gln Met 35 40 40 Thr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu Ser Ile 50 55 60 Ser Gly Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Ile 65 70 75 80</pre>	. Glu-
<pre>&lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal 164 &lt;400&gt; SEQUENCE: 8 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His 1 5 10 15 Cys Gly Cys Pro Ser Asp Asp Asp Val Arg Tyr Pro Leu Ala Ser Asp 20 25 30 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu Gln Met 35 40 40 45 Thr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu Ser Ile 50 55 70 75 80 Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Val Ser Phe Tyr 85 90 95</pre>	. Glu-
<pre>&lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;220&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal 164 &lt;400&gt; SEQUENCE: 8 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His 1 5 10 15 Cys Gly Cys Pro Ser Asp Asp Asp Val Arg Tyr Pro Leu Ala Ser Asp 20 25 30 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu Gln Met 35 40 45 Thr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu Ser Ile 50 55 70 75 80 Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Val Ser Phe Tyr 85 90 Gln Phe Leu Leu Asn Thr Leu Trp Pro Val Asn Asp Thr Ala Ile Trp 100 105 110</pre>	. Glu-
<pre>&lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;220&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal 164 &lt;400&gt; SEQUENCE: 8 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His 1 5 10 15 Cys Gly Cys Pro Ser Asp Asp Asp Val Arg Tyr Pro Leu Ala Ser Asp 20 25 30 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu Gln Met 35 40 met 40 45 Thr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu Ser Ile 50 55 70 75 80 Leu Gly Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Ile 65 70 90 95 Gln Phe Leu Leu Asn Thr Leu Trp Pro Val Asn Asp Thr Ala Ile Trp 100 Glu Ala Phe Met Arg Gln Val Glu Glu Leu Val Asn Gln Gln Ile Thr 120 125</pre>	. Glu-
<pre>&lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal 164 &lt;400&gt; SEQUENCE: 8 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His 1 5 10 15 Cys Gly Cys Pro Ser Asp Asp Asp Val Arg Tyr Pro Leu Ala Ser Asp 20 25 30 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu Gln Met 35 40 40 45 Thr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu Ser Ile 50 Ser Gly Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Ile 65 70 75 80 Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Val Ser Phe Tyr 85 90 90 95 Gln Phe Leu Leu Asn Thr Leu Trp Pro Val Asn Asp Thr Ala Ile Trp 100 Glu Ala Phe Met Arg Gln Val Glu Glu Leu Val Asn Gln Gln Ile Thr</pre>	. Glu-

													CIII	ueu	
Asn	Asp	Thr	Glu	Asn 165	Leu	Ser	Val	Val	<b>A</b> rg 170	Ala	Gln	Phe	Ile	<b>Ala</b> 175	Leu
Asp	Leu	Asp	Phe 180	Val	Asn	Ala	Ile	Pro 185	Leu	Phe	Ala	Val	Asn 190	Gly	Gln
Gln	. Val	Pro 195	Leu	Leu	Ser	Val	<b>Ty</b> r 200	Ala	Gln	Ala	Val	Asn 205	Leu	His	Leu
Leu	Leu 210	Leu	Lys	Asp	Ala	Ser 215	Leu	Phe	Gly	Glu	Gly 220	Trp	Gly	Phe	Thr
Gln 225	Gly	Glu	Ile	Ser	Thr 230	Tyr	Tyr	Asp	Arg	Gln 235	Leu	Glu	Leu	Thr	Ala 240
Lys	Tyr	Thr	Asn	<b>Ty</b> r 245	Cys	Glu	Thr	Trp	T <b>y</b> r 250	Asn	Thr	Gly	Leu	<b>As</b> p 255	Arg
Leu	Arg	Gly	Thr 260		Thr	Glu	Ser	Trp 265		Arg	Tyr	His	Gln 270		Arg
Arg	Glu	Met 275		Leu	Val	Val	Leu 280		Val	Val	Ala	Leu 285		Pro	Tyr
Tyr	Asp		Arg	Leu	Tyr			Gly	Ser	Asn			Leu	Thr	Arg
	290 . Val	Tyr	Thr	Asp		295 Ile	Val	Phe	Asn		300 Pro	Ala	Asn	Val	_
305 Leu	Cys	Arg	Arg	Trp	310 Gly	Thr	Asn	Pro	Tyr	315 Asn	Thr	Phe	Ser	Glu	320 Leu
Glu	Asn	Ala	Phe	325 Ile	Arq	Pro	Pro	His	330 Leu	Phe	Asp	Arq	Leu	335 Asn	Ser
	Thr		340		-			345			-	-	350		
		355				-	360					365			-
Tyr	Trp 370	Ser	Gly	His	Thr	Leu 375	Arg	Arg	Ser	Tyr	Leu 380	Asn	Asp	Ser	Ala
Val 385	Gln	Glu	Asp	Ser	Tyr 390	Gly	Leu	Ile	Thr	Thr 395	Thr	Arg	Ala	Thr	Ile 400
Asn	Pro	Gly	Val	Asp 405	Gly	Thr	Asn	Arg	Ile 410	Glu	Ser	Thr	Ala	Val 415	Asp
Phe	Arg	Ser	Ala 420	Leu	Ile	Gly	Ile	T <b>y</b> r 425	Gly	Val	Asn	Arg	Ala 430	Ser	Phe
Val	Pro	Gly 435	Gly	Leu	Phe	Asn	Gly 440	Thr	Thr	Ser	Pro	Ala 445	Asn	Gly	Gly
Суз	Arg 450	Asp	Leu	Tyr	Asp	Thr 455	Asn	Asp	Glu	Leu	Pro 460	Pro	Asp	Glu	Ser
Thr 465	Gly	Ser	Ser	Thr	His 470	Arg	Leu	Ser	His	Val 475	Thr	Phe	Phe	Ser	Phe 480
	Thr	Asn	Gln	Ala 485		Ser	Ile	Ala	Asn 490		Gly	Ser	Val	Pro 495	
Tyr	Val	Trp	Thr 500		Arg	Asp	Val	Asp 505		Asn	Asn	Thr	Ile 510		Pro
Asn	. Arg			Gln	Leu	Pro			Lys	Ala	Ser			Val	Ser
Gly	. Thr		Val	Leu	Lys		520 Pro	Gly	Phe	Thr		525 Gly	Gly	Ile	Leu
Arg	530 Arg		Thr	Asn	Gly	535 Thr	Phe	Gly	Thr	Leu	540 Arg	Val	Thr	Val	Asn
545					550					555					560
Set	FIO	шец	1111	9111	9111	тут	лгу	ыец	лıу	vai	лгу	- ne	лıd	Der	1111

											-	con	tin	ued		
				565					570					575		
Gly	Asn	Phe	Ser 580	Ile	Arg	Val	Leu	Arg 585	Gly	Gly	Val	Ser	Ile 590	Gly	Asp	
Val	Arg	Leu 595	Gly	Ser	Thr	Met	Asn 600	Arg	Gly	Gln	Glu	Leu 605	Thr	Tyr	Glu	
Ser	Phe 610	Phe	Thr	Arg	Glu	Phe 615	Thr	Thr	Thr	Gly	Pro 620	Phe	Asn	Pro	Pro	
Phe 625	Thr	Phe	Thr	Gln	Ala 630	Gln	Glu	Ile	Leu	Thr 635	Val	Asn	Ala	Glu	Gl <b>y</b> 640	
Val	Ser	Thr	Gly	Gly 645	Glu	Tyr	Tyr	Ile	Asp 650	Arg	Ile	Glu	Ile	Val 655	Pro	
Val	Asn	Pro	Ala 660	Arg	Glu	Ala	Glu	Glu 665	Asp	Leu	Glu	Ala	Ala 670	Lys	Lys	
Ala																
<211 <212 <213 <220 <223 <220 <221	> LE > TY > OF > FE > OT > FE > NA	EATUR THER EATUR ME/P	H: 20 DNA ISM: RE: INFO RE: RE: KEY:	D19 Art: DRMAT	ificia FION: ••(20	: Art			seque	ence	desc	cript	tion:	: Crj	79Cal	-100%
<400	> SE	IQUEI	ICE :	9												
					caa Gln											48
					gaa Glu											96
					caa Gln											144
					act Thr											192
					tta Leu 70											240
					tta Leu											288
					aca Thr											336
					cag Gln											384
					cag Gln			-							-	432
					caa Gln 150											480
aat	gaa	aca	cga	aat	tta	agt	tta	tta	cgt	gct	caa	ttt	ata	gct	tta	528

-continued

												con		ueu		
Asn	Glu	Thr	Arg	Asn 165	Leu	Ser	Leu	Leu	<b>A</b> rg 170	Ala	Gln	Phe	Ile	Ala 175	Leu	
-		-				-			ttg Leu		-				-	576
-				-				-	caa Gln	-					-	624
				-	-				gga Gly	-						672
-		-						-	cgt Arg		-	-			-	720
									tat Tyr 250							768
	_						_		tta Leu							816
									tta Leu							864
									tca Ser							912
									aat Asn							960
	-	-	-						tat Tyr 330							1008
-		-			-				ctt Leu		-		-		-	1056
									tta Leu							1104
									agt Ser							1152
									aca Thr							1200
									ata Ile 410							1248
									ggc Gly							1296
									act Thr							1344
									gaa Glu							1392
acc	gga	agt	tca	acc	cat	aga	cta	tct	cat	tta	acc	ttt	ttt	agc	ttt	1440

-continued

The lay ber ser The file Ary Leu der His Leo The Phe Phe Ser Phe 405 405 405 405 405 405 405 405
Sin The Aen Gin Ála Giy See Ile Ála Aen Ála Giy See Leu Pro Thr 455 tat tha tht acc ogt ogt ga atta gaa ott aat aat acg att acc oca 1536 Sys Leu Me Thr Ary Arg Gil Leu Gil Leu Aen Aen Thr 116 Thr Pro 555 520 520 520 520 520 520 520
Pyr Leu Phe Thr Arýg Arg Gu Leu Giu Leu Aen Aen Thr ILE Thr Pro         Sto         Sto         Stan Arg Tile Thr Gin Leu Pro Leu Luy Palla Ser Alla Pro Leu Ser         Sto
ham Arg Ile Thr Gln Leu Pro Leu Leu Lye Ala Ser Ala Pro Leu Ser 510 520 10 mr thr Leu Leu Ly Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu 530 530 arg aca act at agg cac tt gga acg tt agg tt agg tt agt to 1632 1600 1600 1600 1610
Sign Thr Thr Leu Leu Luy Gip pro Gip Phe Thr Gip Gip Gip Gip Ile Leu         Sign Thr Thr Am Gip Thr Phe Gip Thr Leu Arg Leu Thr Leu Am         Sign Thr Thr Am Gip Thr Phe Gip Thr Leu Arg Leu Thr Leu Am         Sign Thr Thr Am Gip Thr Phe Gip Thr Leu Arg Leu Thr Leu Am         Sign Thr Thr Am Gip Thr Phe Gip Thr Leu Arg Leu Thr Leu Am         Sign Thr Thr Am Gip Thr Phe Gip Thr Leu Arg Leu Thr Leu Am         Sign Thr Thr Am Gip Thr Phe Gip Thr Leu Arg Leu Thr Leu Am         Sign Thr Thr Am Gip Thr Phe Gip Thr Leu Arg Leu Arg Gip Gip Sign Thr Sign Thr Sign Thr Sign Thr Sign Thr Sign Thr Sign Thr Thr Gip Gip Gip Gip Gip Gip Gip Gip Gip Gip
Arg Arg Thr Thr Am Gly Thr Phe Gly Thr Lew Arg Lew Thr Lew Am       550         Set a coa tat act cog of a agat ta coft tit got to a aca 1728         Ser Pro Lew Thr Gln Gln Tyr Arg Lew Arg Phe Ala Ser Thr         565         ga aat tic agt ata agg tta cic coft gag agg gg tat to tat coft gig aga 1776         Sly Am Phe Ser ILe Arg Lew Lew Arg Gly Gly Lew Ser ILe Gly Glu         580         tata aga tta gg agc aca atg aca aga ggg cag gaa cta act tac gaa 1824         Eeu Arg Lew Gly Ser Thr Met Aan Arg Gly Gln Glu Lew Thr Tyr Glu         600         590         tta aga tta gg agc aca atg aca gag ggt coft o aat cog oct 1872         600         600         600         601         602         603         604         605         605         606         606         607         610         628         629         630         631         632         633         635         645         666         670         667         670         668         670         670
Ser Pro Leu Tr Gin Gin Tyr Arg Leu Arg Leu Arg Phe Åla ser Thr 565 570 575 gga aat tto agt ata agg tta oto ogt gga ggg tta tot ato ggt gaa 1776 Sly Aen Phe Ser Ile Arg Leu Leu Arg Gily Gily Leu Ser Ile Gily Glu 580 590 tta agg tta ggg ago aca atg aac aga ggg cag gga ota act act tao gaa 1824 Leu Arg Leu Gily Ser Thr Met Aen Arg Gily Gin Giu Leu Thr Tyr Glu 695 tto ttt tto aca aga gag tt act act act ggt cog to aat cog oct 1872 Strong Thr Met Aen Arg Gily Cin Contact at a cog gdt 1872 tto ttt ta aca aga gag tt cat act act ggt cog to aat cog oct 1872 1872 tto att aca caa got caa gga att cta aca tta ast gca gaa ggt 1920 Her Thr Phe Thr Gin Ala Gin Giu Ile Leu Thr Leu Asn Ala Giu Gily 620 630 640 640 640 640 655 640 655 640 655 655 656 657 658 659 659 659 659 650 659 650 650 650 650 650 650 655 657 650 659 659 650 650 650 650 650 650 650 650
Sily Asn Phe Ser Ile Arg Leu Leu Arg Gily Gily Leu Ser Ile Gily Gu       580         S80       585         Leu Arg Leu Gily Ser Thr Met Asn Arg Gly Gln Giu Leu Thr Tyr Glu       1824         Leu Arg Leu Gily Ser Thr Met Asn Arg Gly Gln Giu Leu Thr Tyr Glu       1872         S55       600       615         tco ttt ttc aca aga gag ttt act act act gqt ccg ttc aat ccg cct       1872         Ser Phe Phe Thr Arg Glu Phe Thr Thr Gly Pro Phe Asn Pro Pro       620         Fob Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro       640         610       615       620         Feb Thr Phe Thr Gln Alg Glu Glu Ile Leu Thr Leu Asn Ala Glu Glu Glu       1920         625       630       635       640         625       630       655       640         626       655       640       655         tta agc acc gqt ggt gat att at ata gaa aga tt gaa att tta cct       1968       962         Leu Asn Pro Ala Arg Glu Ala Glu Glu Glu Leu Glu Ala Ala Lys Lys       660       665       670         grg Ala       2019       635       670       2019         Ala       221> OPKENTH F073       221>       221>       Cypectrificial sequence       222>       700       15         2210       GRANISM: Artificial sequence       10       15
Leu Àrg Leu Glý Ser Thr Mei Aan Arg Glý Gln Glu Leu Thr Tyr Glu 595 600 605 1872 tco ttt tto aca aga gag ttt act act act ggt ccg tto aat oog cot 610 610 610 615 620 620 620 630 630 630 630 630 630 630 63
Ser Phe Phe Thr Arg Glu Phe Thr Thr Gly Pro Phe Asn Pro Pro 610 611 620 620 620 620 620 620 620 620 620 620
Phe Thr Phe Thr Gln Åla Gln Glu Ile Leu Thr Leu Asn Åla Glu Gjy 635 640 tta agc acc ggt ggt gaa tat tat ata gaa aga att gaa att tta cct 1968 Leu Ser Thr Gly Glu Tyr Tyr Ile Glu Arg Ile Glu Ile Leu Pro 645 650 650 655 2016 tta aat ccg gca cga gaa gcg gaa gag gaa ta gaa gcg gcg aag aaa 2016 Leu Asn Pro Ala Arg Glu Ala Glu Glu Glu Leu Glu Ala Ala Lys Lys 660 665 670 2019 Ala 2019 2
Leu Ser Thr Gly Gly Glu Tyr Tyr Ile Glu Arg Ile Glu Ile Leu Pro 645 650 655 655 655 655 655 655 655 655 65
Leu Asn Pro Ala Arg Glu Ala Glu Glu Glu Leu Glu Ala Ala Lys Lys 660 665 670 2019 Ala 2019 Ala 2019 Ala 2019 4211> LENGTH: 673 4212> TYPE: PRT 4213> ORGANISM: Artificial sequence 4220> FEATURE: 4223> OTHER INFORMATION: Artificial sequence description: Cry9Cal-100% 400> SEQUENCE: 10 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Glu Ala Pro His 1 5 10 15 Cys Gly Cys Pro Ser Glu Glu Glu Leu Arg Tyr Pro Leu Ala Ser Glu 20 25 30 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Glu Tyr Leu Gln Met 35 40 45 Thr Glu Glu Glu Tyr Thr Glu Ser Tyr Ile Asn Pro Ser Leu Ser Ile
Ala <210> SEQ ID NO 10 <211> LENGTH: 673 <212> TYPE: PRT <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: Cry9Cal-100% <<400> SEQUENCE: 10 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Glu Ala Pro His 1 5 10 15 Cys Gly Cys Pro Ser Glu Glu Glu Leu Arg Tyr Pro Leu Ala Ser Glu 20 25 30 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Glu Tyr Leu Gln Met 35 40 45 Thr Glu Glu Glu Tyr Thr Glu Ser Tyr Ile Asn Pro Ser Leu Ser Ile
<pre>&lt;211&gt; LENGTH: 673 &lt;212&gt; TYPE: PRT &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal-100% &lt;400&gt; SEQUENCE: 10 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Glu Ala Pro His 1 5 10 15 Cys Gly Cys Pro Ser Glu Glu Glu Leu Arg Tyr Pro Leu Ala Ser Glu 20 25 30 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Glu Tyr Leu Gln Met 35 40 45 Thr Glu Glu Glu Tyr Thr Glu Ser Tyr Ile Asn Pro Ser Leu Ser Ile</pre>
Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Glu Ala Pro His 1 $5$ $10$ $10$ $15$ $15$ Cys Gly Cys Pro Ser Glu Glu Glu Leu Arg Tyr Pro Leu Ala Ser Glu 20 $20$ $20$ $20$ $20$ $20$ $20$ $20$ $2$
1       5       10       15         Cys Gly Cys Pro Ser Glu Glu Glu Leu Arg Tyr Pro Leu Ala Ser Glu 20       20       20       20         Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Glu Tyr Leu Gln Met 35       40       45         Thr Glu Glu Glu Tyr Thr Glu Ser Tyr Ile Asn Pro Ser Leu Ser Ile
20 25 30 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Glu Tyr Leu Gln Met 35 40 45 Thr Glu Glu Glu Tyr Thr Glu Ser Tyr Ile Asn Pro Ser Leu Ser Ile
35 40 45 Thr Glu Glu Tyr Thr Glu Ser Tyr Ile Asn Pro Ser Leu Ser Ile

_													con		ueu	
	er 55	Gly	Arg	Glu	Ala	Leu 70	Gln	Thr	Ala	Leu	Thr 75	Val	Ile	Arg	Arg	Ile 80
Le	eu	Gly	Ala	Leu	Gly 85	Leu	Pro	Phe	Ser	Gly 90	Gln	Ile	Leu	Ser	Phe 95	Tyr
Gl	.n	Phe	Leu	Leu 100	Asn	Thr	Leu	Phe	Pro 105	Leu	Asn	Glu	Thr	Ala 110	Ile	Phe
Gl	.u	Ala	Phe 115	Met	Arg	Gln	Leu	Glu 120	Glu	Leu	Leu	Asn	Gln 125	Gln	Ile	Thr
Gl		Phe 130	Ala	Arg	Asn	Gln	Ala 135	Leu	Ala	Arg	Leu	Gln 140	Gly	Leu	Gly	Glu
Se 14		Phe	Asn	Leu	Tyr	Gln 150	Arg	Ser	Leu	Gln	Asn 155	Phe	Leu	Ala	Glu	Arg 160
As	n	Glu	Thr	Arg	Asn 165	Leu	Ser	Leu	Leu	<b>A</b> rg 170	Ala	Gln	Phe	Ile	Ala 175	Leu
Gl	.u	Leu	Glu	Phe 180	Leu	Asn	Ala	Ile	Pro 185	Leu	Phe	Ala	Leu	Asn 190	Gly	Gln
Gl	.n	Leu	Pro 195	Leu	Leu	Ser	Leu	<b>Ty</b> r 200	Ala	Gln	Ala	Leu	Asn 205	Leu	His	Leu
Le		Leu 210	Leu	Lys	Glu	Ala	Ser 215	Leu	Phe	Gly	Glu	Gly 220	Phe	Gly	Phe	Thr
G1 22		Gly	Glu	Ile	Ser	Thr 230	Tyr	Tyr	Glu	Arg	Gln 235	Leu	Glu	Leu	Thr	Ala 240
Ly	75	Tyr	Thr	Asn	<b>Ty</b> r 245	Cys	Glu	Thr	Phe	<b>Ty</b> r 250	Asn	Thr	Gly	Leu	Glu 255	Arg
Le	eu	Arg	Gly	Thr 260	Asn	Thr	Glu	Ser	Phe 265	Leu	Arg	Tyr	His	Gln 270	Phe	Arg
Ar	g	Glu	Met 275	Thr	Leu	Leu	Leu	Leu 280	Glu	Leu	Leu	Ala	Leu 285	Phe	Pro	Tyr
Ту		Glu 290		Arg	Leu	Tyr	Pro 295		Gly	Ser	Asn	Pro 300		Leu	Thr	Arg
G1 30	.u		Tyr	Thr	Glu	Pro 310		Leu	Phe	Asn	Pro 315		Ala	Asn	Leu	Gly 320
		Cys	Arg	Arg	Phe 325		Thr	Asn	Pro	Tyr 330		Thr	Phe	Ser	Glu 335	
Gl	.u	Asn	Ala		Ile	Arg	Pro	Pro			Phe	Glu	Arg	Leu 350		Ser
Le	eu	Thr		340 Ser	Ser	Asn	Arg		345 Pro	Leu	Ser	Ser			Met	Glu
Ту			355 Ser	Gly	His	Thr		360 Arg	Arg	Ser	Tyr		365 Asn	Glu	Ser	Ala
	eu	370 Gln	Glu	Glu	Ser			Leu	Ile	Thr		380 Thr	Arg	Ala	Thr	
38 As		Pro	Gly	Leu	Glu			Asn	Arg		395 Glu	Ser	Thr	Ala		400 Glu
Ph	ıe	Arg	Ser		405 Leu		Gly	Ile		410 Gly	Leu	Asn	Arg		415 Ser	Phe
Le	eu	Pro	Gly	420 Gly	Leu	Phe	Asn	Gly	425 Thr	Thr	Ser	Pro	Ala	430 Asn	Gly	Gly
			435	-	Tyr			440					445		-	_
		450			Thr		455					460				
	-	1		- 01		0	9	u		0	_0u					

-continued

			-001	ItIllied	
465	470		475		480
Gln Thr Asn Gln A	la Gly Ser 85	Ile Ala Asn 490	Ala Gly Ser	r Leu Pro 495	Thr
Tyr Leu Phe Thr A 500	rg Arg Glu	Leu Glu Leu 505	Asn Asn Thi	r Ile Thr 510	Pro
Asn Arg Ile Thr 0 515		Leu Leu L <b>y</b> s 520	Ala Ser Ala 525		Ser
Gl <b>y</b> Thr Thr Leu I 530	eu Lys Gly 535	Pro Gly Phe	Thr Gly Gly 540	y Gly Ile	Leu
Arg Arg Thr Thr A 545	sn Gly Thr 550	Phe Gly Thr	Leu Arg Leu 555	ı Thr Leu	Asn 560
Ser Pro Leu Thr C	ln Gln T <b>y</b> r 65	Arg Leu Arg 570	Leu Arg Phe	e Ala Ser 575	Thr
Gly Asn Phe Ser 1 580	le Arg Leu	Leu Arg Gly 585	Gly Leu Sei	f Ile Gly 590	Glu
Leu Arg Leu Gly S 595		Asn Arg Gly 600	Gln Glu Leu 605		Glu
Ser Phe Phe Thr A 610	rg Glu Phe 615	Thr Thr Thr	Gly Pro Phe 620	e Asn Pro	Pro
Phe Thr Phe Thr 0 625	ln Ala Gln 630	Glu Ile Leu	Thr Leu Ası 635	n Ala Glu	Gl <b>y</b> 640
Leu Ser Thr Gly G	ly Glu Tyr 45	Tyr Ile Glu 650	Arg Ile Glu	ı Ile Leu 655	Pro
Leu Asn Pro Ala A 660	arg Glu Ala	Glu Glu Glu 665	Leu Glu Ala	a Ala Lys 670	Lys
Ala					
<pre>&lt;210&gt; SEQ ID NO 1 &lt;211&gt; LENGTH: 201 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: A &lt;220&gt; FEATURE: &lt;220&gt; FEATURE: &lt;220&gt; FEATURE: &lt;220&gt; FEATURE: &lt;221&gt; NAME/KEY: C &lt;222&gt; LOCATION: (</pre>	9 rtificial se MATION: Art DS	-	nce descrip	otion: Cry	9Ca1-25%
<400> SEQUENCE: 1	1				
atg aat cga aat a Met Asn Arg Asn A 1					
tgt ggg tgt cca t Cys Gly Cys Pro S 20					
cca aat gca gcg t Pro Asn Ala Ala I 35		-	-	r Leu Gln	-
aca gat gag gac t Thr Asp Glu Asp T 50					
agt ggt aga gaa g Ser Gly Arg Glu A 65					
ctc ggg gct tta c Leu Gly Ala Leu C					

						ctg Leu										336
						gtg Val										384
						gca Ala 135										432
						cgt Arg										480
	-		-			agt Ser			-	-				-		528
						gct Ala										576
	-					gta Val		-		-						624
						tct Ser 215										672
						tat Tyr										720
-					-	gaa Glu								-	-	768
	-					gaa Glu	-			-					-	816
						gta Val										864
						cca Pro 295										912
						att Ile										960
	-	-	-			act Thr										1008
					-	cca Pro					-		-		-	1056
						cga Arg										1104
						tta Leu 375										1152
-		-	-	-		ggc Gl <b>y</b>						-	-			1200

		gga Gly													1248
		tct Ser													1296
		gga Gly 435													1344
		gat Asp													1392
		agt Ser			_										1440
		aat Asn													1488
		tgg Trp													1536
		att Ile 515													1584
		acg Thr													1632
		aca Thr													1680
		tta Leu				-		-		-		-			1728
		ttc Phe													1776
-	-	tta Leu 595	 -		-		-		-	-				-	1824
		ttc Phe													1872
		ttt Phe		-								-	-		1920
-	-	acc Thr	 					-			-		-		1968
		ccg Pro													2016
gcg															2019

gcg Ala 2019

<210> SEQ ID NO 12 <211> LENGTH: 673 <212> TYPE: PRT

<213> ORGANISM: Artificial sequence														
<pre>&lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: Cry9Cal-25%</pre>														
<400> SEQUENCE:		inciai sequence	description. crystar-23%											
Met Asn Arg Asn	Asn Gln Asn	Glu Tyr Glu Ile	Ile Asp Ala Pro His											
1	5	10	15											
Cys Gly Cys Prc		Asp Val Arg Tyr	Pro Leu Ala Ser Asp											
20		25	30											
Pro Asn Ala Ala	Leu Gln Asn I	Met Asn Tyr Lys	Asp Tyr Leu Gln Met											
35		40	45											
Thr Asp Glu Asp	o Tyr Thr Asp	Ser Tyr Ile Asn	Pro Ser Leu Ser Ile											
50	55		60											
Ser Gly Arg Glu Ala Leu Gln Thr Ala Leu Thr Leu Leu Gly Arg Ile         65       70       75       80         Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Leu Ser Phe Tyr														
85 90 95														
Gln Phe Leu Leu Asn Thr Leu Trp Pro Val Asn Asp Thr Ala Ile Trp 100 105 110														
Glu Ala Phe Met	-	Glu Glu Leu Val	Asn Gln Gln Ile Thr											
115		120	125											
Glu Phe Ala Arg	Asn Gln Ala 1	Leu Ala Arg Leu	Gln Gly Leu Gly Glu											
130	135		140											
Ser Phe Asn Val	. Tyr Gln Arg	Ser Leu Gln Asn	Trp Leu Ala Asp Arg											
145	150	155	160											
Asn Asp Thr Arg	Asn Leu Ser 1	Leu Leu Arg Ala	Gln Phe Ile Ala Leu											
	165	170	175											
Asp Leu Asp Phe		Ile Pro Leu Phe	Ala Val Asn Gly Gln											
180		185	190											
Gln Val Pro Leu		Tyr Ala Gln Ala	Leu Asn Leu His Leu											
195		200	205											
Leu Leu Leu Lys	Glu Ala Ser 1	Leu Phe Gly Glu	Gly Trp Gly Phe Thr											
210	215		220											
Gln Gly Glu Ile	e Ser Thr Tyr	Tyr Glu Arg Gln	Leu Glu Leu Thr Ala											
225	230	235	240											
Lys Tyr Thr Asn	1 Tyr Cys Glu	Thr Trp Tyr Asn	Thr Gl <b>y</b> Leu Glu Arg											
	245	250	255											
Leu Arg Gly Thr		Ser Phe Leu Arg	Tyr His Gln Phe Arg											
260		265	270											
Arg Glu Met Thr		Leu Asp Val Val	Ala Leu Phe Pro Tyr											
275		280	285											
Tyr Asp Val Arg	Leu Tyr Pro	Thr Gly Ser Asn	Pro Gln Leu Thr Arg											
290	295		300											
Glu Val Tyr Thr	Asp Pro Ile	Val Phe Asn Pro	Pro Ala Asn Leu Gly											
305	310	315	320											
Leu Cys Arg Arg	Trp Gly Thr .	Asn Pro Tyr Asn	Thr Phe Ser Glu Leu											
	325	330	335											
Glu Asn Ala Phe		Pro His Leu Phe	Glu Arg Leu Asn Ser											
340		345	350											
Leu Thr Ile Ser	-	Phe Pro Val Ser	Ser Asn Phe Met Glu											
355		360	365											
Tyr Phe Ser Gly	'His Thr Leu .	Arg Arg Ser Tyr	Leu Asn Asp Ser Ala											

$\sim$	$\sim$	n	+	п.	n	11	Δ	d

											-	con	tin	ued					
	370					375					380								
Val 385	Gln	Glu	Asp	Ser	T <b>y</b> r 390	Gly	Leu	Ile	Thr	Thr 395	Thr	Arg	Ala	Thr	Ile 400				
Asn	Pro	Gly	Val	Asp 405	Gly	Thr	Asn	Arg	Ile 410	Glu	Ser	Thr	Ala	Val 415	Asp				
Phe	Arg	Ser	Ala 420	Leu	Ile	Gly	Ile	T <b>y</b> r 425	Gly	Val	Asn	Arg	Ala 430	Ser	Phe				
Val	Pro	Gly 435	Gly	Leu	Phe	Asn	Gly 440	Thr	Thr	Ser	Pro	Ala 445	Asn	Gly	Gly				
Cys	Arg 450	Asp	Leu	Tyr	Asp	Thr 455	Asn	Asp	Glu	Leu	Pro 460	Pro	Asp	Glu	Ser				
Thr 465	Gly	Ser	Ser	Thr	His 470	Arg	Leu	Ser	His	Leu 475	Thr	Phe	Phe	Ser	Phe 480				
Gln	Thr	Asn	Gln	Ala 485	Gly	Ser	Ile	Ala	Asn 490	Ala	Gly	Ser	Val	Pro 495	Thr				
Tyr	Val	Trp	Thr 500	Arg	Arg	Asp	Val	<b>A</b> sp 505	Leu	Asn	Asn	Thr	Ile 510		Pro				
Asn	Arg	Ile 515		Gln	Leu	Pro	Leu 520		Lys	Ala	Ser	Ala 525		Val	Ser				
Gly	Thr 530		Val	Leu	Lys	Gly 535		Gly	Phe	Thr	Gly 540		Gly	Ile	Leu				
		Thr	Thr	Asn			Phe	Gly	Thr			Val	Thr	Val					
545 Ser	Pro	Leu	Thr	Gln	550 Gln	Tyr	Arg	Leu		555 Leu	Arg	Phe	Ala		560 Thr				
Gly	Asn	Phe	Ser	565 Ile	Arg	Val	Leu	Arg	570 Gly	Gly	Val	Ser	Ile	575 Gl <b>y</b>	Asp				
	_	_	580	_			_	585				_	590	_					
Val	Arg	Leu 595	GIY	Ser	Thr	Met	Asn 600	Arg	GIY	Gln	Glu	Leu 605	Thr	Tyr	Glu				
Ser	Phe 610	Phe	Thr	Arg	Glu	Phe 615	Thr	Thr	Thr	Gly	Pro 620	Phe	Asn	Pro	Pro				
Phe 625	Thr	Phe	Thr	Gln	Ala 630	Gln	Glu	Ile	Leu	Thr 635	Val	Asn	Ala	Glu	Gly 640				
Val	Ser	Thr	Gly	Gly 645	Glu	Tyr	Tyr	Ile	Asp 650	Arg	Ile	Glu	Ile	Val 655	Pro				
Val	Asn	Pro	Ala 660	Arg	Glu	Ala	Glu	Glu 665	Asp	Leu	Glu	Ala	Ala 670	Lys	Lys				
Ala																			
<211 <212 <213 <220	)> FE	NGTH PE: RGANI ATUP	H: 30 DNA ESM: RE:	Art			-		seque	ence	desc	ript	ion:	: mut	ant 1				
<400	)> SE	QUEI	NCE :	13															
gaat	taaa	atg a	aatt	ttta	aa t	ttaa	gtgt	t								30			
<211 <212 <213	)> SE L> LE 2> TY 3> OF )> FE	ENGTH PE: RGANI	H: 30 DNA ESM:		Lficia	al se	equer	nce											

54

	-continued	
<223> OTHER INFORMATION: Artificial	sequence description: mutant 2	
<400> SEQUENCE: 14		
gaattaaatg aattattaaa tttaagtgtt		30
<210> SEQ ID NO 15 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial	sequence description: mutant 3	
<400> SEQUENCE: 15		
gaattattag aatttttatt attaagtgtt		30
<210> SEQ ID NO 16 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial	sequence description: mutant 4	
<400> SEQUENCE: 16		
gaattattag aattattatt attaagtgtt		30
<210> SEQ ID NO 17 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial	sequence description: mutant 5	
<400> SEQUENCE: 17		
gaattattag aagaattatt attaagtgtt		30
<210> SEQ ID NO 18 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial <400> SEQUENCE: 18		
gaacgattag aatttttatt attaagtgtt		30
<210> SEQ ID NO 19 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial		
<400> SEQUENCE: 19		
gaacgattag aattattatt attaagtgtt		30
<210> SEQ ID NO 20 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial <400> SEQUENCE: 20	sequence description: mutant 8	

-continued	
gaattagaag aattattatt attaagtgtt	30
<210> SEQ ID NO 21 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 9	
<400> SEQUENCE: 21	
gaattattag aagaagaaga attaagtgtt	30
<210> SEQ ID NO 22 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 10	
<400> SEQUENCE: 22	
tttttattaa atttatttt tttaccatta ctg	33
<210> SEQ ID NO 23 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant11	
<400> SEQUENCE: 23	
tttttattaa atttagaaga attaccatta ctg	33
<210> SEQ ID NO 24 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 12 <400> SEQUENCE: 24	
tttgaagaaa atttagaaga attaccatta ctg	33
<pre>&lt;210&gt; SEQ ID NO 25 &lt;211&gt; LENGTH: 33 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: mutant 13 &lt;400&gt; SEQUENCE: 25</pre>	
tttgaagaaa attttttatt atttccatta ctg	33
<210> SEQ ID NO 26 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 14 <400> SEQUENCE: 26	
tttgaagaaa attttgaaga atttccatta ctg	33
<210> SEQ ID NO 27	

56

-continued	
<211> LENGTH: 33	
<212> TYPE: DNA <213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: mutant 15	
<400> SEQUENCE: 27	
tttttattaa attttgaaga atttccatta ctg	33
<210> SEQ ID NO 28	
<211> LENGTH: 33	
<212> TYPE: DNA <213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: mutant 16	
<400> SEQUENCE: 28	
tttttattaa atgaattttt tgaaccatta ctg	33
-210. CEO TE NO 20	
<210> SEQ ID NO 29 <211> LENGTH: 24	
<212> TYPE: DNA <213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: mutant 17	
<400> SEQUENCE: 29	
cttttttag aattatttt attc	24
<210> SEQ ID NO 30 <211> LENGTH: 24	
<212> TYPE: DNA	
<pre>&lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE:</pre>	
<223> OTHER INFORMATION: Artificial sequence description: mutant 18	
<400> SEQUENCE: 30	
cttttttat tattattttt attc	24
<210> SEQ ID NO 31 <211> LENGTH: 24	
<212> TYPE: DNA	
<pre>&lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE:</pre>	
<223> OTHER INFORMATION: Artificial sequence description: mutant 19	
<400> SEQUENCE: 31	
cttttttag aagaatttga atta	24
<210> SEQ ID NO 32	
<211> LENGTH: 24 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<pre>&lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: mutant 20</pre>	
<400> SEQUENCE: 32	
ctttttgaag aagaatttga atta	24
<210> SEQ ID NO 33	
<211> LENGTH: 24 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	

<223> OTHER INFORMATION: Artificial sequence description: mutant 21 <400> SEQUENCE: 33 24 ctttttgaag aattatttga agaa <210> SEQ ID NO 34 <211> LENGTH: 15 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 22 <400> SEQUENCE: 34 ttattagaat taaat 15 <210> SEO ID NO 35 <211> LENGTH: 15 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 23 <400> SEQUENCE: 35 ttattatttt taaat 15 <210> SEQ ID NO 36 <211> LENGTH: 15 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 24 <400> SEQUENCE: 36 ttagaattat taaat 15 <210> SEQ ID NO 37 <211> LENGTH: 15 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 25 <400> SEQUENCE: 37 ttattatttt ttaat 15 <210> SEQ ID NO 38 <211> LENGTH: 15 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 26 <400> SEQUENCE: 38 ttagaagaat taaat 15 <210> SEQ ID NO 39 <211> LENGTH: 15 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 27 <400> SEOUENCE: 39

ttagaatttt taaat 15 <210> SEQ ID NO 40 <211> LENGTH: 15 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 28 <400> SEQUENCE: 40 ttagaatttg aaaat 15 <210> SEQ ID NO 41 <211> LENGTH: 15 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: mutant 29 <400> SEQUENCE: 41 ttagaagaag aaaat 15 <210> SEQ ID NO 42 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 1 <400> SEQUENCE: 42 gatcgaaatg atacattaaa tttaagtgtt gtt 33 <210> SEQ ID NO 43 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 2 <400> SEQUENCE: 43 33 gatcgaaatg atacatttaa tttaagtgtt gtt <210> SEQ ID NO 44 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 3 <400> SEQUENCE: 44 gatcgaaatg atacagaaaa tttaagtgtt gtt 33 <210> SEQ ID NO 45 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 4 <400> SEQUENCE: 45

59

-continued	
cgaaatgata cacgattatt aagtgttgtt cgt	33
<210> SEQ ID NO 46	
<211> LENGTH: 33 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 5	
-	
<400> SEQUENCE: 46	
cgaaatgata cacgagaatt aagtgttgtt cgt	33
<210> SEQ ID NO 47 <211> LENGTH: 39	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<pre>&lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	
oligonucleotide 6	
<400> SEQUENCE: 47	
ttggctgatc gaaatgaatt tttaaattta agtgttgtt	39
<210> SEQ ID NO 48	
<211> LENGTH: 39	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence <220> FEATURE:	
<pre>&lt;220&gt; FEALORE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	
oligonucleotide 7	
<400> SEQUENCE: 48	
ttggctgatc gaaatgaatt tttattatta agtgttgtt	39
<210> SEQ ID NO 49 <211> LENGTH: 39	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 8	
<400> SEQUENCE: 49	
ttggctgatc gaaatgaatt attaaattta agtgttgtt	39
<210> SEQ ID NO 50	
<211> LENGTH: 39	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence <220> FEATURE:	
<pre>&lt;220&gt; FLAIORE: &lt;220&gt; OTHER INFORMATION: Artificial sequence description:</pre>	
<400> SEQUENCE: 50	
ttggctgatc gaaatgaatt attattatta agtgttgtt	39
<210> SEQ ID NO 51	
<211> LENGTH: 39	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence <220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description:	
oligonucleotide 10	

<400> SEQUENCE: 51	
ttggctgatc gaaatgaaga agaagaatta agtgttgtt 3	39
<210> SEQ ID NO 52 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 11	
<400> SEQUENCE: 52	
ttggctgatc gaaatgaaga attattatta agtgttgtt 3	39
<210> SEQ ID NO 53 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 12	
<400> SEQUENCE: 53	
caaaattggt tggctgaatt aaatgaatta ttaaat	36
<210> SEQ ID NO 54 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 13	
<400> SEQUENCE: 54	
caaaattggt tggctgaatt aaatgaattt ttaaat	36
<210> SEQ ID NO 55 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 14	
<400> SEQUENCE: 55	
caaaattggt tggctgaatt attagaattt ttattatta 3	39
<210> SEQ ID NO 56 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 15	
<400> SEQUENCE: 56	
caaaattggt tggctgaatt attagaatta ttattatta	39
<210> SEQ ID NO 57 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description:	

-continued	
oligonucleotide 16	
<400> SEQUENCE: 57	
caaaattggt tggctgaatt attagaagaa ttattatta	39
<210> SEQ ID NO 58 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 17	
<400> SEQUENCE: 58	
caaaattggt tggctgaacg attagaattt ttattatta	39
<210> SEQ ID NO 59 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 18	
<400> SEQUENCE: 59	
caaaattggt tggctgaacg attagaatta ttattatta	39
<210> SEQ ID NO 60 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 19	
<400> SEQUENCE: 60	
caaaattggt tggctgaatt agaagaatta ttattatta	39
<210> SEQ ID NO 61 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 20 <400> SEQUENCE: 61	
caaaattggt tggctgaatt attagaagaa gaagaatta	39
<pre>&lt;210&gt; SEQ ID NO 62 &lt;211&gt; LENGTH: 36 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description: oligonucleotide 21 &lt;400&gt; SEQUENCE: 62</pre>	
gctattccat tgtttttatt aaatggacag caggtt	36
<210> SEQ ID NO 63 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence	

62

-continued	
<220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 22	
<400> SEQUENCE: 63	
gctattccat tgtttgaaga aaatggacag caggtt	36
<210> SEQ ID NO 64 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 23	
<400> SEQUENCE: 64	
ttattaaatg gacagcagtt accattactg tcagta	36
<210> SEQ ID NO 65 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 24	
<400> SEQUENCE: 65	
ttattaaatg gacagcagtt tccattactg tcagta	36
<210> SEQ ID NO 66 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 25	
<400> SEQUENCE: 66	
ttattaaatg gacagcagga accattactg tcagta	36
<210> SEQ ID NO 67 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 26	
<400> SEQUENCE: 67	
gaagaaaatg gacagcagtt accattactg tcagta	36
<210> SEQ ID NO 68 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 27	
<400> SEQUENCE: 68	
gaagaaaatg gacagcagtt tccattactg tcagta	36
<210> SEQ ID NO 69 <211> LENGTH: 45	

-concinued	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<pre>&lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	
oligonucleotide 28	
<400> SEQUENCE: 69	
ccattgtttt tattaaattt attttttta ccattactgt cagta	45
<pre>&lt;210&gt; SEQ ID NO 70 &lt;211&gt; LENGTH: 45</pre>	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 29	
<400> SEQUENCE: 70	
ccattgtttt tattaaattt agaagaatta ccattactgt cagta	45
010 GTO TR NO 71	
<210> SEQ ID NO 71 <211> LENGTH: 45	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 30	
Silgonacieotiae SV	
<400> SEQUENCE: 71	
ccattgtttg aagaaaattt agaagaatta ccattactgt cagta	45
contrycting adyadaatti ayaayaatta coattactyt cayta	45
<210> SEQ ID NO 72	
<pre>&lt;211&gt; LENGTH: 45 &lt;212&gt; TYPE: DNA</pre>	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description:	
oligonucleotide 31	
<400> SEQUENCE: 72	
	45
ccattgtttg aagaaaattt tttattattt ccattactgt cagta	45
<210> SEQ ID NO 73	
<211> LENGTH: 45 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 32	
<400> SEQUENCE: 73	
ccattgtttg aagaaaattt tgaagaattt ccattactgt cagta	45
<210- STO ID NO 74	
<210> SEQ ID NO 74 <211> LENGTH: 45	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<pre>&lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	
oligonucleotide 33	
<400> SEQUENCE: 74	
ccattgtttt tattaaattt tgaagaattt ccattactgt cagta	45
-	

<210> SEQ ID NO 75 <211> LENGTH: 45 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 34 <400> SEQUENCE: 75 ccattgtttt tattaaatga attttttgaa ccattactgt cagta 45 <210> SEQ ID NO 76 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 35 <400> SEQUENCE: 76 gatgcatctc tttttttaga aggatgggga ttc 33 <210> SEQ ID NO 77 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 36 <400> SEQUENCE: 77 gatgcatctc ttttttatt aggatgggga ttcaca 36 <210> SEQ ID NO 78 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 37 <400> SEQUENCE: 78 gatgcatctc tttttgaaga aggatgggga ttc 33 <210> SEQ ID NO 79 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 38 <400> SEQUENCE: 79 33 ttagaaggat ggggattaac acagggggaa att <210> SEQ ID NO 80 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 39 <400> SEQUENCE: 80 33 gaagaaggat ggggagaaac acagggggaa att

<210> SEQ ID NO 81 <211> LENGTH: 45 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 40 <400> SEQUENCE: 81 gcatctcttt ttttagaatt atttttattc acacaggggg aaatt 45 <210> SEQ ID NO 82 <211> LENGTH: 45 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 41 <400> SEQUENCE: 82 gcatctcttt ttttattatt atttttattc acacaggggg aaatt 45 <210> SEQ ID NO 83 <211> LENGTH: 45 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 42 <400> SEQUENCE: 83 gcatctcttt ttttagaatt atttttattc acacaggggg aaatt 45 <210> SEQ ID NO 84 <211> LENGTH: 45 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 43 <400> SEQUENCE: 84 45 gcatctcttt ttgaagaatt atttttattc acacaggggg aaatt <210> SEQ ID NO 85 <211> LENGTH: 45 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 44 <400> SEQUENCE: 85 gcatctcttt ttgaagaatt atttttagaa acacaggggg aaatt 45 <210> SEQ ID NO 86 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 45 <400> SEQUENCE: 86

ggtttagatc gtttattaga attaaatact gaaagttgg	39
<210> SEQ ID NO 87 <211> LENGTH: 39 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description:	
oligonucleotide 46 <400> SEQUENCE: 87	
ggtttagatc gtttattatt tttaaatact gaaagttgg	39
<210> SEQ ID NO 88 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description:	
oligonucleotide 47 <400> SEQUENCE: 88	
ggtttagatc gtttagaatt attaaatact gaaagttgg	39
<210> SEQ ID NO 89 <211> LENGTH: 39 <212> TYPE: DNA	
<pre>&lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	
<400> SEQUENCE: 89	
ggtttagatc gtttattatt ttttaatact gaaagttgg	39
<pre>&lt;210&gt; SEQ ID NO 90 &lt;211&gt; LENGTH: 39 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	
<400> SEQUENCE: 90	
ggtttagatc gtttagaaga attaaatact gaaagttgg	39
<210> SEQ ID NO 91 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 50	
<400> SEQUENCE: 91	
ggtttagatc gtttagaatt tttaaatact gaaagttgg	39
<210> SEQ ID NO 92 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 51	
51130AU01000140 51	

<400> SEQUENCE: 92		
ggtttagatc gtttagaatt tgaaaatact gaaagttgg	39	
<210> SEQ ID NO 93 <211> LENGTH: 39 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 52		
<400> SEQUENCE: 93		
ggtttagatc gtttagaaga agaaaatact gaaagttgg	39	
<210> SEQ ID NO 94 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 53 <400> SEQUENCE: 94		
tgaatatgaa attattgaag ccccccattg	30	
<210> SEQ ID NO 95 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 54		
<400> SEQUENCE: 95		
<400> SEQUENCE: 95 tgggtgtcca tcagaagaag aattaaggta tootttggca	40	
<pre>tgggtgtcca tcagaagaag aattaaggta tcctttggca &lt;210&gt; SEQ ID NO 96 &lt;211&gt; LENGTH: 27 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	40	
<pre>tgggtgtcca tcagaagaag aattaaggta tcctttggca &lt;210&gt; SEQ ID NO 96 &lt;211&gt; LENGTH: 27 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>		
<pre>tgggtgtcca tcagaagaag aattaaggta tcctttggca &lt;210&gt; SEQ ID NO 96 &lt;211&gt; LENGTH: 27 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	40 27	
<pre>tgggtgtcca tcagaagaag aattaaggta tcctttggca &lt;210&gt; SEQ ID NO 96 &lt;211&gt; LENGTH: 27 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>		
<pre>tgggtgtcca tcagaagaag aattaaggta tcctttggca &lt;210&gt; SEQ ID NO 96 &lt;211&gt; LENGTH: 27 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>		
<pre>tgggtgtcca tcagaagaag aattaaggta tcctttggca &lt;210&gt; SEQ ID NO 96 &lt;211&gt; LENGTH: 27 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>		

-continued <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 57 <400> SEQUENCE: 98 26 caaatgacag aagaggaata cactga <210> SEQ ID NO 99 <211> LENGTH: 20 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 58 <400> SEQUENCE: 99 tacactgaat cttatataaa 20 <210> SEQ ID NO 100 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 59 <400> SEOUENCE: 100 tattagtggt agagaagcat tacagactgc gcttac 36 <210> SEQ ID NO 101 <211> LENGTH: 37 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 60 <400> SEQUENCE: 101 cagactgcgc ttactgttat tagggagaat actcggg 37 <210> SEQ ID NO 102 <211> LENGTH: 25 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 61 <400> SEQUENCE: 102 25 gggctttagg tttaccgttt tctgg <210> SEQ ID NO 103 <211> LENGTH: 28 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 62 <400> SEQUENCE: 103 ttctggacaa atattaagtt tttatcaa 28 <210> SEO ID NO 104

<211> LENGTH: 40 <212> TYPE: DNA

<213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 63 <400> SEQUENCE: 104 cttttaaata cactgtttcc attaaatgaa acagctatat 40 <210> SEQ ID NO 105 <211> LENGTH: 24 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 64 <400> SEQUENCE: 105 acagctatat ttgaagcttt catg 24 <210> SEQ ID NO 106 <211> LENGTH: 26 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 65 <400> SEQUENCE: 106 ctttcatgcg acagttagag gaactt 26 <210> SEQ ID NO 107 <211> LENGTH: 26 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 66 <400> SEQUENCE: 107 26 gaggaacttt taaatcaaca aataac <210> SEQ ID NO 108 <211> LENGTH: 21 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 67 <400> SEQUENCE: 108 21 ggattaggag aatcttttaa t <210> SEQ ID NO 109 <211> LENGTH: 23 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 68 <400> SEQUENCE: 109 tettttaatt tatatcaacg tte 23 <210> SEQ ID NO 110

<211> LENGTH: 21 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 69 <400> SEQUENCE: 110 ccttcaaaat tttttggctg a 21 <210> SEQ ID NO 111 <211> LENGTH: 17 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 70 <400> SEQUENCE: 111 17 ttggctgaac gaaatga <210> SEQ ID NO 112 <211> LENGTH: 23 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 71 <400> SEQUENCE: 112 23 cgaaatgaaa cacgaaattt aag <210> SEQ ID NO 113 <211> LENGTH: 37 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 72 <400> SEQUENCE: 113 acacgaaatt taagtttatt acgtgctcaa tttatag 37 <210> SEQ ID NO 114 <211> LENGTH: 48 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 73 <400> SEOUENCE: 114 gctcaattta tagctttaga acttgaattt ttaaatgcta ttccattg 48 <210> SEQ ID NO 115 <211> LENGTH: 27 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 74 <400> SEQUENCE: 115 27 ccattgtttg cattaaatgg acagcag

								- 1
-c	$\sim$	n	t.	п.	n	11	ρ	a
~	~		~	_		~	~	~

<210> SEQ ID NO 116 <211> LENGTH: 27 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 75 <400> SEQUENCE: 116 ccattgtttg cattaaatgg acagcag 27 <210> SEQ ID NO 117 <211> LENGTH: 27 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 76 <400> SEQUENCE: 117 ccattactgt cattatatgc acaagct 27 <210> SEQ ID NO 118 <211> LENGTH: 27 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 77 <400> SEQUENCE: 118 tatgcacaag ctttaaattt acatttg 27 <210> SEQ ID NO 119 <211> LENGTH: 23 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 78 <400> SEQUENCE: 119 23 ttattaaaag aagcatctct ttt <210> SEQ ID NO 120 <211> LENGTH: 25 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 79 <400> SEQUENCE: 120 tggagaagga tttggattca cacag 25 <210> SEQ ID NO 121 <211> LENGTH: 24 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 80 <400> SEQUENCE: 121

	-continue	a	
cacatattat gaacgtcaat tgga		24	
<pre>&lt;210&gt; SEQ ID NO 122 &lt;211&gt; LENGTH: 28 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence oligonucleotide 81</pre>	uence description:		
<400> SEQUENCE: 122			
tactgtgaaa ctttttataa tacaggtt		28	
<pre>&lt;210&gt; SEQ ID NO 123 &lt;211&gt; LENGTH: 25 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial seq oligonucleotide 82 </pre>	uence description:		
<400> SEQUENCE: 123		25	
tacaggttta gaacgtttaa gagga		25	
<pre>&lt;210&gt; SEQ ID NO 124 &lt;211&gt; LENGTH: 28 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence oligonucleotide 83</pre>	uence description:		
<400> SEQUENCE: 124			
aatactgaaa gttttttaag atatcatc		28	
<210> SEQ ID NO 125 <211> LENGTH: 51 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial seq oligonucleotide 84	uence description:		
<400> SEQUENCE: 125			
gtagagaaat gactttatta ttattagaat tatta	gcgct atttccatat t	51	
<210> SEQ ID NO 126 <211> LENGTH: 27 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial seq oligonucleotide 85	uence description:		
<400> SEQUENCE: 126			
atattatgaa ttacgacttt atccaac		27	
<210> SEQ ID NO 127 <211> LENGTH: 23 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial seq oligonucleotide 86	uence description:		

100 CEOUENCE 107	
<400> SEQUENCE: 127	
cttacacgtg agttatatac aga	23
<210> SEQ ID NO 128 <211> LENGTH: 29 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial s oligonucleotide 87	sequence description:
<400> SEQUENCE: 128	
tatacagaac cgattttatt taatccacc	29
<210> SEQ ID NO 129 <211> LENGTH: 28 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial s oligonucleotide 88	sequence description:
<400> SEQUENCE: 129	
ccaccagcta atttaggact ttgccgac	28
<210> SEQ ID NO 130 <211> LENGTH: 27 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial s oligonucleotide 89	sequence description:
<400> SEQUENCE: 130	
ctttgccgac gttttggtac taatccc	27
<pre>&lt;210&gt; SEQ ID NO 131 &lt;211&gt; LENGTH: 23 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial s oligonucleotide 90</pre>	sequence description:
<400> SEQUENCE: 131	
catctttttg aaaggctgaa tag	23
<pre>&lt;210&gt; SEQ ID NO 132 &lt;211&gt; LENGTH: 30 &lt;212&gt; TYPE: DNA &lt;213&gt; ORGANISM: Artificial sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial s oligonucleotide 91</pre>	sequence description:
<400> SEQUENCE: 132	
taatcgattt ccattatcat ctaattttat	30
<210> SEQ ID NO 133 <211> LENGTH: 36 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial s	sequence description:

-continued	
oligonucleotide 92	
<400> SEQUENCE: 133	
ctaattttat ggaatatttt tcaggacata cgttac	36
<210> SEQ ID NO 134 <211> LENGTH: 33	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence <220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 93	
<400> SEQUENCE: 134	
tagttatctg aacgaatcag cattacaaga aga	33
<210> SEQ ID NO 135	
<211> LENGTH: 20 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description:	
oligonucleotide 94	
<400> SEQUENCE: 135	
caagaagaaa gttatggcct	20
<210> SEQ ID NO 136	
<211> LENGTH: 35 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<pre>&lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	
oligonucleotide 95	
<400> SEQUENCE: 136	
caattaatcc cggattagaa ggaacaaacc gcata	35
<210> SEQ ID NO 137	
<211> LENGTH: 30 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<pre>&lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	
oligonucleotide 96	
<400> SEQUENCE: 137	
gagtcaacgg cattagaatt tcgttctgca	30
<210> SEQ ID NO 138	
<211> LENGTH: 26 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 97	
<400> SEQUENCE: 138	
ggtatatatg gcttaaatag agcttc	26
<210> SEQ ID NO 139	
<211> LENGTH: 30 <212> TYPE: DNA	
<112> IIFE. DNA <213> ORGANISM: Artificial sequence	

<220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 98 <400> SEOUENCE: 139 tagagettet tttttaceag gaggettgtt 30 <210> SEQ ID NO 140 <211> LENGTH: 31 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 99 <400> SEQUENCE: 140 31 ctgctaatgg aggatgtaga gaactctatg a <210> SEQ ID NO 141 <211> LENGTH: 17 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 100 <400> SEQUENCE: 141 ctctatgaaa caaatga 17 <210> SEQ ID NO 142 <211> LENGTH: 20 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 101 <400> SEOUENCE: 142 acaaatgaag aattaccacc 20 <210> SEQ ID NO 143 <211> LENGTH: 27 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 102 <400> SEQUENCE: 143 27 attaccacca gaagaaagta ccggaag <210> SEQ ID NO 144 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 103 <400> SEQUENCE: 144 30 agactatctc atttaacctt ttttagcttt <210> SEQ ID NO 145 <211> LENGTH: 27

<212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 104 <400> SEQUENCE: 145 gctaatgcag gaagtttacc tacttat 27 <210> SEQ ID NO 146 <211> LENGTH: 26 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 105 <400> SEOUENCE: 146 cctacttatt tatttacccg tcgtga 26 <210> SEQ ID NO 147 <211> LENGTH: 33 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 106 <400> SEOUENCE: 147 acccgtcgtg aattagaact taataatacg att 33 <210> SEQ ID NO 148 <211> LENGTH: 24 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 107 <400> SEOUENCE: 148 attaccattg ttaaaggcat ctgc 24 <210> SEQ ID NO 149 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 108 <400> SEOUENCE: 149 aaggcatctg cacctttatc gggtactacg 30 <210> SEQ ID NO 150 <211> LENGTH: 29 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 109 <400> SEQUENCE: 150 29 tcgggtacta cgttattaaa aggtccagg

<210> SEQ ID NO 151 <211> LENGTH: 40 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 110 <400> SEQUENCE: 151 acatttggaa cgttaagatt aacgttaaat tcaccattaa 40 <210> SEQ ID NO 152 <211> LENGTH: 37 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 111 <400> SEQUENCE: 152 cacaacaata tcgcctaaga ttacgttttg cctcaac 37 <210> SEQ ID NO 153 <211> LENGTH: 31 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 112 <400> SEQUENCE: 153 31 aaatttcagt ataaggttac tccgtggagg g <210> SEQ ID NO 154 <211> LENGTH: 35 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 113 <400> SEQUENCE: 154 ataagggtac tccgtggagg gttatctatc ggtga 35 <210> SEQ ID NO 155 <211> LENGTH: 29 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 114 <400> SEQUENCE: 155 29 tctatcggtg aattaagatt agggagcac <210> SEQ ID NO 156 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 115 <400> SEQUENCE: 156 30 caagagattc taacattaaa tgcagaaggt

<210> SEQ ID NO 157 <211> LENGTH: 32	
<212> TYPE: DNA <213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 116	
<400> SEQUENCE: 157	
aatgcagaag gtttaagcac cggtggtgaa ta	32
<210> SEQ ID NO 158	
<211> LENGTH: 32 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE: <223> OTHER INFORMATION: Artificial sequence description:	
oligonucleotide 117	
<400> SEQUENCE: 158	
gtggtgaata ttatatagaa agaattgaaa tt	32
<210> SEQ ID NO 159	
<211> LENGTH: 37 <212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<pre>&lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Artificial sequence description:</pre>	
oligonucleotide 118	
<400> SEQUENCE: 159	
agaattgaaa ttttaccttt aaatccggca cgagaag	37
<210> SEQ ID NO 160	
<211> LENGTH: 30	
<212> TYPE: DNA <213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Artificial sequence description: oligonucleotide 119	
<400> SEQUENCE: 160	
cgagaagcgg aagaggaatt agaagcggcg	30

**1**. A pepsin-sensitive modified Cry protein, characterized in that it has at least one additional pepsin cleavage site.

2. The modified Cry protein as claimed in claim 1, characterized in that the additional pepsin cleavage site is represented by an amino acid residue chosen from leucine, phenylalanine and glutamic acid residues.

**3**. The modified Cry protein as claimed in either of claims **1** and **2**, characterized in that it is selected from the Cry1, Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 proteins.

**4**. The modified Cry protein as claimed in claim 3, characterized in that it is a Cry9C protein.

**5**. The modified Cry protein as claimed in claim 4, characterized in that it is the Cry9Ca1 protein.

6. The modified Cry protein as claimed in one of claims 1 to 5, characterized in that it has at least one additional pepsin cleavage site in at least one of the inter- $\alpha$ -helix loops of domain I.

7. The modified Cry protein as claimed in one of claims 1 to 6, characterized in that it has at least one additional pepsin cleavage site in the inter- $\alpha$ -helix loop linking the  $\alpha$ 3 and  $\alpha$ 4 helices of domain I.

**8**. The modified Cry protein as claimed in one of claims 5 to 7, characterized in that it has an additional pepsin cleavage site at position 164.

**9**. The modified Cry protein as claimed in claim 8, characterized in that it is selected from the Cry proteins, the sequences of which are represented by the identifiers SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:8.

**10**. The modified Cry protein as claimed in one of claims 1 to 5, characterized in that the additional pepsin cleavage sites are introduced by substituting aspartic acid residues with glutamic acid residues, substituting tryptophan residues with phenylalanine residues, and substituting valine or isoleucine residues with leucine residues.

**11**. The modified Cry protein as claimed in claim 11, characterized in that the degree of substitutions which said Cry protein possesses is 25%.

**12**. A method for increasing the pepsin sensitivity of the Cry proteins, characterized in that at least one additional pepsin or cleavage site is introduced into said Cry proteins.

**13**. The method as claimed in claim 12, characterized in that the additional pepsin cleavage site introduced is represented by an amino acid chosen from leucine, phenylalanine and glutamic acid residues.

14. The method as claimed in either of claims 12 and 13, characterized in that it applies to the Cry proteins selected from the Cry1, Cry3, Cry4, Cry7, Cry8, Cry9, Cry10, Cry16, Cry17, Cry19 and Cry20 proteins.

**15**. The method as claimed in claim 14, characterized in that it applies to the Cry9C protein.

**16**. The method as claimed in claim 15, characterized in that it applies to the Cry9Ca1 protein.

17. The method as claimed in one of claims 12 to 16, characterized in that at least one additional pepsin cleavage site is introduced into at least one of the inter- $\alpha$ -helix loops of domain I of said Cry proteins.

18. The method as claimed in one of claims 12 to 17, characterized in that at least one additional pepsin cleavage site is introduced into the inter- $\alpha$ -helix loop linking the  $\alpha$  and  $\alpha$ 4 helices of domain I.

**19**. The method as claimed in one of claims 16 to 18, characterized in that an additional pepsin cleavage site is introduced at position 164.

**20**. The method as claimed in one of claims 12 to 16, characterized in that the additional pepsin cleavage sites are introduced by substituting aspartic acid residues with glutamic acid, substituting tryptophan residues with phenylalanine residues, and substituting value or isoleucine residues with leucine residues.

**21**. The method as claimed in claim 20, characterized in that the degree of substitution which said Cry protein possesses is less than or equal to 25%.

**22.** A polynucleotide encoding a modified Cry protein as claimed in one of claims 1 to 11.

23. A chimeric gene comprising, functionally linked to one another, at least:

(a) one promoter which is functional in a host organism

(b) a polynucleotide as claimed in claim 22

(c) a terminator element which is functional in a host organism.

**24**. The chimeric gene as claimed in claim 23, characterized in that the promoter and the terminator element are functional inplants.

25. An expression or transformation vector containing a chimeric gene as claimed in either of claims 23 and 24.

**26**. The vector as claimed in claim 27, characterized in that it is a plasmid, a phase or a virus.

27. A host organism transformed with one of the vectors as claimed in either of claims 25 and 26.

**28**. The host organism as claimed in claim 27, characterized in that it is a plant.

**29.** The plant as claimed in claim 28, characterized in that it contains, in addition to a chimeric gene as claimed in either of claims **23** and **24**, at least one other chimeric gene containing a polynucleotide encoding a protein of interest.

**30**. A part of a plant as claimed in claim 29.

**31**. A seed from a plant as claimed in claim 29.

**32**. A method for producing the modified Cry proteins as claimed in one of claims 1 to 11, characterized in that it comprises at least the steps of:

- (a) culturing a transformed host organism according to the invention in a culture medium suitable for the growth and for the multiplication of said organism,
- (b) (b) extracting the Cry proteins produced by the transformed organism cultured in step (a).

**33**. The method as claimed in claim 32, characterized in that it comprises a step (c) of purification of the Cry proteins extracted in step (b).

**34**. The method as claimed in either of claims **32** and **33**, characterized in that the host organism is a microorganism.

**35**. The method as claimed in claim 34, characterized in that the host organism is a *Bacillus thuringiensis* bacterium.

**36**. A monoclonal or polyclonal antibody, characterized in that it is directed against a modified Cry protein as claimed in one of claims 1 to 11.

\* \* \* \* \*