

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
1 June 2006 (01.06.2006)

PCT

(10) International Publication Number
WO 2006/057917 A2

(51) International Patent Classification:
C07K 14/705 (2006.01)

(US). **ALLENZA, Paul** [US/US]; 16 Pleasant View Way,
Flemington, NJ 08822 (US).

(21) International Application Number:
PCT/US2005/041893

(74) Agent: **SHEEHAN, John**; FMC Corporation, 1735 Mar-
ket Street, Philadelphia, PA 19103 (US).

(22) International Filing Date:
18 November 2005 (18.11.2005)

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/630,548 23 November 2004 (23.11.2004) US

(71) Applicant (*for all designated States except US*): **FMC
CORPORATION** [US/US]; 1735 Market Street, Philadel-
phia, PA 19103 (US).

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

(71) Applicant (*for US only*): **DIERKS, Connie, Perretz** (ex-
ecutrix for the deceased inventor) [US/US]; 262 Daleview
Drive, Yardley, PA 19067 (US).

(72) Inventor: **DIERKS, Peter** (deceased).

Published:

— *without international search report and to be republished
upon receipt of that report*

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **CHEN, Rui-
hua** [CN/US]; 16 Labrador Drive, Kendell Park, NJ
08824 (US). **HUANG, Cancan** [CN/US]; 16 Labrador
Drive, Kendall Park, NJ 08824 (US). **DITOLVO, Kevin**
[US/US]; 423 Oliver Street, Bordentown, NJ 08505

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

(54) Title: HEMIPTERAN 5HT1 RECEPTOR

(57) Abstract: Novel nucleic acid sequences encoding 5HT1 receptor, and recombinant expressions and host cells comprising the same are disclosed. Isolated 5HT1 receptor, host cells expressing 5HT1 receptor, methods of producing the 5HT1 receptor and antibodies specific for 5HT1 receptor are also disclosed. Methods of identifying modulators and/or inhibitors of 5HT1 receptor are disclosed.



WO 2006/057917 A2

HEMIPTERAN 5HT1 RECEPTOR

This application claims the benefit of U.S. Provisional Application No. 60/630548, filed November 23, 2004, which is incorporated herein by reference.

5

Field of the Invention

The present invention relates to compositions that are useful in agrochemical, veterinary or pharmaceutical fields. In particular, the invention relates to nucleotide sequences that encode polypeptides that are useful in the identification or
10 development of compounds that affect 5HT1 receptor activity, indicating that such compounds may be useful as pesticides or as pharmaceuticals.

Background of the Invention

Serotonin is a neuromodulator that plays an important role in learning,
15 feeding, locomotion, circadian rhythms, and defense behaviors in invertebrate species (Angers et al., J. Neurosci., 18, pp. 5586-5593, 1998; Brownlee et al., Parasitology, 111 (pt 3), pp. 379-384, 1995; Mendel et al., Science, 267, pp. 1652-1655, 1995). Serotonin activates multiple receptor subtypes that are coupled to different intracellular signaling pathways, including, but not limited to, 5HT1,
20 5HT2, 5HT3, and 5HT7 receptors (Tierney, Compar. Biochem. Physiol., 128 (pt A), pp. 791-804, 2001;). 5HT1 receptor subtype is a Gi coupled receptor and is linked to inhibition of adenylate cyclase (Tierney, Compar. Biochem. Physiol., 128 (pt A), pp. 791-804, 2001). Targeting serotonin signaling pathway, specifically 5HT1 receptor may provide a mean for identifying biologically active compounds,
25 including insecticides.

Several insect 5HT1 receptors have been cloned and functionally expressed. Drosophila 5HT1A and 5HT1B receptors were cloned by Saudou et al. (EMBO J., 11, pp. 7-17, 1992). When expressed in NIH3T3 cells, both receptors inhibited adenylate cyclase in response to serotonin. 5HT₁ receptor was then cloned by
30 Sugamori et al. (Proc. Natl. Acad. Sci. USA, 90, pp. 11-15, 1993). This protein has

structure features that resembles Gi coupled receptors suggesting 5Htlym may inhibit adenylate cyclase but actual data on signaling are not available. The C. elegans 5HT1 receptor was cloned by Olde and McCombie (J. Mol. Neurosci., 8, pp. 53-62, 1997). When expressed in Ltk cells and stimulated with 5HT, the receptor
5 caused a reduction in the level of forskolin-induced cAMP accumulation. The Aplysia 5HT1 (termed 5HTap1) was cloned by Angers et al. (J. Neurosci., 18, pp. 5586-5593) and functionally expressed in HEK293 cells. It was shown to inhibit adenylate cyclase.

10 Summary of the Invention

One embodiment of the invention relates to nucleotide sequences that encode polypeptides with 5HT1 receptor activity. In preferred embodiments, nucleotide sequences that encode polypeptides with aphis gossypii 5HT1 receptor activity. Such nucleotide sequences may be used to express amino acid sequences that are
15 useful in the identification or development of compounds that affect 5HT1 receptor activity. Such compounds may be useful as pesticides or as pharmaceuticals. These nucleotide sequences that encode polypeptides with 5HT1 receptor activity, preferably aphis gossypii 5HT1 receptor activity, including mutants and fragments thereof, which will be further described below, will also be referred to herein as
20 “nucleotide sequences of the invention”. The polypeptides with 5HT1 receptor activity, preferably aphis gossypii 5HT1 receptor activity, including mutants and fragments thereof, which will be further characterized below, will also be described as “proteins of the invention,” “amino acid sequences of the invention,” or “polypeptide of the invention”.

25 The present invention further relates to the use of the nucleotide sequences of the invention, preferably in the form of a suitable genetic construct as described below, in the transformation of host cells or host organisms, for example for the expression of the amino acid sequences of the invention and to such genetic constructs and host cells.

Another aspect of the invention relates to methods for the identification or development of compounds that can modulate and/or inhibit the biological activity of the amino acid sequences of the invention, in which one or more of the above mentioned nucleotide sequences, amino acid sequences, genetic constructs, host
5 cells or host organisms are used. Such methods, which will usually be in the form of an assay or screen, will also be further described below.

A further aspect of the invention relates to compounds that can modulate the biological activity of, or that can otherwise interact with, an amino acid sequence of the invention, either *in vitro* or preferably (also) *in vivo*. The invention also relates
10 to compositions that contain such compounds, and to the use of such compounds in the preparation of these compositions and the control of pests.

Brief Description of the Figures

Figure 1 shows serotonin inhibited forskolin-stimulated cAMP production in
15 a dose dependent manner in HEK-293 cells expressing aphid 5HT1 receptors.

Definitions

Collectively, the nucleic acids of the present invention will be referred to herein as "*nucleic acids of the invention*". Also, where appropriate in the context of
20 the further description of the invention below, the terms "*nucleotide sequence of the invention*" and "*nucleic acid of the invention*" may be considered essentially equivalent and essentially interchangeable.

Also, for the purposes of the present invention, a nucleic acid or amino acid sequence is considered to be "*(in) essentially isolated (form)*" – for example, from
25 its native biological source - when it has been separated from at least one other nucleic acid molecule and/or sequence with which it is usually associated. Similarly, a protein or polypeptide of the invention is considered to be "*(in) essentially isolated (form)*" – for example, from its native biological source - when it has been effectively separated from other polypeptide molecules with which it is normally
30 associated. In particular, a nucleic acid or polypeptide of the invention is considered

“essentially isolated” when it has been purified at least 2-fold, in particular at least 10-fold, more in particular at least 100-fold, and up to 1000-fold or more.

Detailed Description of the Invention

5 The present invention was established from the finding that the amino acid sequences of the invention can be used as “target(s)” in assays to identify chemical compounds and other factors (with the term “*target*” having its usual meaning in the art, provide for example the definition given in WO 98/06737) which interact with them *in vitro* or *in vivo*. Consequently, compounds or factors that have been
10 identified as interacting with the amino acid sequences of the invention (e.g. by the methods as described herein below) may be useful as active agents in the agrochemical, veterinary or pharmaceutical fields.

 In one embodiment, the invention relates to a nucleic acid, preferably in essentially isolated form, which nucleic acid comprises a nucleotide sequence of the
15 invention, and in particular the nucleotide sequence of SEQ ID NO: 1 and SEQ ID NO:3 and mutants and fragments thereof. The nucleotide sequence of SEQ ID NO: 1 and SEQ ID NO:3 were derived or isolated from *Aphis gossypii*, in the manner as further described in the Examples below. The present invention also relates to a polypeptides of the invention, and in particular those comprising SEQ ID NO:2 and
20 SEQ ID NO: 4 and mutants and fragments thereof. SEQ ID NO: 1 encodes SEQ ID NO: 2, and SEQ ID NO: 3 encodes SEQ ID NO: 4.

 Generally, the nucleotide sequences of the invention, when in the form of a nucleic acid, may be DNA or RNA, and may be single stranded or double stranded. For example, the nucleotide sequences of the invention may be genomic DNA,
25 cDNA or synthetic DNA (such as DNA with a codon usage that has been specifically adapted for expression in the intended host cell or host organism, which may for instance be designed using suitable computer programs such as the BackTranslate analysis tool in Vector NTI (InforMax, Inc., Bethesda, MD). Thus, the nucleotide sequences of the invention may contain intron sequences, and also
30 generally comprises different splice variants.

Yet another embodiment relates to a double stranded RNA molecule directed against a nucleotide sequence of the invention (one strand of which will usually comprise at least part of a nucleotide sequence of the invention). Such double stranded RNA molecules have particular utility in RNA interference studies of gene function (Zamore et al, Cell 101:25-33 (2000)). The invention also relates to genetic constructs that can be used to provide such double stranded RNA molecules (e.g. by suitable expression in a host cell or host organism, or for example in a bacterial strain such as *E.coli*). For such constructs, reference is made to Maniatis et al., *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Press, 1989).

10 In a broader sense, the term “*nucleotide sequence of the invention*” also comprises:

- parts or fragments of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:3, preferably SEQ ID NO: 1
- (natural or synthetic) mutants, variants, alleles, analogs, orthologs (herein below collectively referred to as “*mutants*”) of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:3, preferably SEQ ID NO: 1 as further described below as well as any nucleotide sequence that encodes a polypeptide that comprises the amino acid sequence of SEQ ID NO: 2, or SEQ ID NO:4 and fragments thereof and DNA sequences that detectably hybridize to such DNA sequences or their complementary strands under conditions of moderate or high stringency;
- 20 - as further described below.
- parts or fragments of such (natural or synthetic) mutants;
- nucleotide fusions of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:3, preferably SEQ ID NO: 1;
- 25 - (or a part or fragment thereof) with at least one further nucleotide sequence;
- nucleotide fusions of (natural or synthetic) mutants (or a part or fragment thereof) with at least one further nucleotide sequence; in which such mutants, parts, fragments or fusions are preferably as further described below. The invention also comprises different splice variants of the above nucleotide sequences.
- 30

- In some embodiments, the nucleotide sequence of the invention is a fragment of a nucleic acid molecule that encodes a 5HT1 receptor. Preferably, a nucleotide sequence of the invention will have a length of at least 500 nucleotides, preferably at least 1,000 nucleotides, more preferably at least 1,200 nucleotides; and up to a length of at most 3,500 nucleotides, preferably at most 3,000 nucleotides, more preferably at most, 2,240 nucleotides. Examples of parts or fragments of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:3, preferably SEQ ID NO: 1;
or a part or fragment of a (natural or synthetic) mutant thereof include, but are not limited to, 5' or 3' truncated nucleotide sequences, or sequences with an introduced in frame startcodon or stopcodon. Also, two or more such parts or fragments of one or more nucleotide sequences of the invention may be suitably combined (e.g. ligated in frame) to provide a further nucleotide sequence of the invention.
- In some embodiments, such parts or fragments comprise at least one continuous stretch of at least 100 nucleotides, preferably at least 250 nucleotides, more preferably at least 500 nucleotides, even more preferably more than 1,000 nucleotides, of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:3, preferably SEQ ID NO: 1;
- Also, it is expected that - based upon the disclosure herein - the skilled person will be able to identify, derive or isolate natural "mutants" (as mentioned above) of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:3, preferably SEQ ID NO: 1;
from (other individuals of) the same species (for example from an individual of a different strain or line, including but not limited to mutant strains or lines). It is also expected that - based upon the disclosure herein - the skilled person will be able to provide or derive synthetic mutants (as defined hereinabove) of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:3, preferably SEQ ID NO: 1; It is also expected that - based upon the disclosure herein - the skilled person will be able to provide or derive polypeptide having polypeptide sequences of SEQ ID NO: 2 or SEQ ID NO:4 by means of protein expression.

In one specific embodiment, the mutant is such that it encodes the nucleotide sequence of SEQ ID NO: 1, or a part or fragment thereof. In another specific embodiment, the mutant is such that it encodes the nucleotide sequence of SEQ ID NO: 3 or a part or fragment thereof;

5 Preferably, any mutants as described herein will have one or more, and preferably all, of the structural characteristics or conserved features referred to below for the nucleotide sequences of SEQ ID NO: 1 or SEQ ID NO:3;

In particular, any mutants, parts or fragments as described herein may be such that they at least encode the active or catalytic site of the corresponding amino
10 acid sequence of the invention and a binding domain of the corresponding amino acid sequence of the invention.

Also, any mutants, parts or fragments as described herein will preferably have a degree of "sequence identity", at the nucleotide level, with the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:3, preferably SEQ ID NO: 1, of at least
15 75%, preferably at least 80%, more preferably at least 85%, and in particular more than 90%, and up to 95% or more.

Also, preferably, any mutants, parts or fragments of the nucleotide sequence of the invention will be such that they encode an amino acid sequence which has a degree of "sequence identity", at the amino acid level, with the amino acid sequence
20 of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2; of at least 55%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90% and up to 95% or more, in which the percentage of "sequence identity" is calculated as described below.

For this purpose, the percentage of "sequence identity" between a given
25 nucleotide sequence and the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3 may be calculated by dividing the number of nucleotides in the given nucleotide sequence that are identical to the nucleotide at the corresponding position in the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3 by the total number of nucleotides in the given nucleotide sequence and multiplying by 100%, in which
30 each deletion, insertion, substitution or addition of a nucleotide - compared to the

sequence of SEQ ID NO: 1 or SEQ ID NO: 3 - is considered as a difference at a single nucleotide position.

The preferred computer program for performing global sequence alignments and determining sequence identity is ClustalW (Higgins et al., Nucleic Acids Research 22:4673-4680 (1994)), which is publicly available for a variety of computer platforms. Preferably the parameters used with the ClustalW program for protein sequence alignments are ktuple = 1, diagonals = 5, windows = 5, gap = 3, score = PERCENTAGE, matrix = BLOSUM, open penalty = 10.0 and extension penalty = 0.05.

Also, in a preferred aspect, any mutants, parts or fragments as described herein will encode proteins or polypeptides having biological activity that is essentially similar to the biological activity described above for the sequences of SEQ ID NO: 1 or SEQ ID NO:3, preferably SEQ ID NO: 1, i.e. to a degree of at least 55%, preferably at least 75%, and up to 90%, as measured by standard assay techniques as described below.

Any mutants, parts or fragments as described herein are preferably such that they are capable of hybridizing with the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:3, preferably SEQ ID NO: 1, i.e. under conditions of "moderate stringency", and preferably under conditions of "high stringency". Such conditions will be clear to the skilled person, for example from the standard handbooks, such as Sambrook et al. and Ausubel et al., mentioned above, as well as in EP 0 967 284, EP 1 085 089 or WO 00/55318.

It is also within the scope of the invention to use a fusion of a nucleotide sequence of the invention (as described above) with one or more further nucleotide sequence(s), including but not limited to one or more coding sequences, non-coding sequences or regulatory sequences. Preferably, in such fusions, the one or more further nucleotide sequences are operably connected (as described below) to the nucleotide sequence of the invention (for example so that, when the further nucleotide sequence is a coding sequence, the nucleotide fusion encodes a protein fusion as described below).

In another embodiment, the invention relates to an antisense molecule against a nucleotide sequence of the invention.

A nucleic acid, preferably in essentially isolated form, can be used to express an amino acid sequence of the invention, for example, the amino acid sequence of
5 SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2;

On the basis of the above, and although the invention is not specifically limited to any specific explanation or mechanism, the nucleotide sequences of some embodiments of the invention encode proteins that have biological activity or ligand binding properties of 5HT1 receptor from insects of the order *Hemiptera*, which are
10 aphids, leafhoppers, whiteflies, scales and true bugs that have mouthparts adapted to piercing and sucking.

. The nucleic acids of the invention may also be in the form of a genetic construct, again as further described below. Genetic constructs of the invention will generally comprise at least one nucleotide sequence of the invention, optionally
15 linked to one or more elements of genetic constructs known per se, as described below. Such genetic constructs may be DNA or RNA, and are preferably double-stranded DNA. The constructs may also be in a form suitable for transformation of the intended host cell or host organism, in a form suitable for integration into the genomic DNA of the intended host cell or in a form suitable independent replication,
20 maintenance and inheritance in the intended host organism. For instance, the genetic construct may be in the form of a vector, such as for example a plasmid, cosmid, a yeast artificial chromosome ("YAC"), a viral vector or transposon. In particular, the vector may be an expression vector, i.e. a vector that can provide for expression *in vitro* or *in vivo* (e.g. in a suitable host cell or host organism as
25 described below). An expression vector comprising a nucleotide sequence of the invention is also referred to herein as a recombinant expression vector. These constructs will also be referred to herein as "*genetic constructs of the invention*".

In a preferred embodiment, such a construct a recombinant expression vector which will comprise:

30 a) the nucleotide sequence of the invention; operably connected to:

b) one or more regulatory elements, such as a promoter and optionally a suitable terminator;

and optionally also:

5 c) one or more further elements of genetic constructs known per se; in which the terms "*regulatory element*", "*promoter*", "*terminator*", "*further elements*" and "*operably connected*" have the meanings indicated herein below.

As the one or more "further elements" referred to above, the genetic construct(s) of the invention may generally contain one or more suitable regulatory elements (such as a suitable promoter(s), enhancer(s), or terminator(s)), 3'- or 5'-
10 untranslated region(s) ("UTR") sequences, leader sequences, selection markers, expression markers or reporter genes, or elements that may facilitate or increase (the efficiency of) transformation or integration. These and other suitable elements for such genetic constructs will be clear to the skilled person, and may for instance depend upon the type of construct used, the intended host cell or host organism; the
15 manner in which the nucleotide sequences of the invention of interest are to be expressed (e.g. via constitutive, transient or inducible expression); and the transformation technique to be used.

Preferably, in the genetic constructs of the invention, the one or more further elements are "*operably linked*" to the nucleotide sequence(s) of the invention or to
20 each other, by which is generally meant that they are in a functional relationship with each other. For instance, a promoter is considered "*operably linked*" to a coding sequence if said promoter is able to initiate or otherwise control or regulate the transcription or the expression of a coding sequence (in which said coding sequence should be understood as being "*under the control of*" said promoter)

25 Generally, when two nucleotide sequences are operably linked, they will be in the same orientation and usually also in the same reading frame. They will usually also be essentially contiguous, although this may also not be required.

Preferably, the optional further elements of the genetic construct(s) used in the invention are such that they are capable of providing their intended biological
30 function in the intended host cell or host organism.

For instance, a promoter, enhancer or terminator should be “operable” in the intended host cell or host organism, by which is meant that (for example) said promoter should be capable of initiating or otherwise controlling or regulating the transcription or the expression of a nucleotide sequence - e.g. a coding sequence - to which it is operably linked (as defined above).

Such a promoter may be a constitutive promoter or an inducible promoter, and may also be such that it (only) provides for expression in a specific stage of development of the host cell or host organism, or such that it (only) provides for expression in a specific cell, tissue, organ or part of a multicellular host organism.

Some particularly preferred promoters include, but are not limited to, constitutive promoters, such as cytomegalovirus (“CMV”), Rous sarcoma virus (“RSV”), simian virus-40 (“SV40”), for example, pSVL SV40 Late Promoter Expression Vector (Pharmacia Biotech Inc., Piscataway, NJ), or herpes simplex virus (“HSV”) for expression in mammalian cells or insect constitutive promoters such as the immediate early baculovirus promoter described by Jarvis et al. (Methods in Molecular Biology Vol. 39 Baculovirus Expression Protocols, ed. C. Richardson., Humana Press Inc., Totowa, NJ (1995)) available in pIE vectors from Novagen (Novagen, Inc. Madison, WI) or insect inducible promoters such as the *Drosophila metallothionein* promoter described by Bunch et al. (Nucleic Acids Research, Vol. 6, No. 3 1043-106, (1988)) available in vectors from Invitrogen (Invitrogen Corporation, Carlsbad, CA).

A selection marker should be such that it allows - i.e. under appropriate selection conditions - host cells or host organisms that have been (successfully) transformed with the nucleotide sequence of the invention to be distinguished from host cells or organisms that have not been (successfully) transformed. Some preferred, but non-limiting examples of such markers are genes that provide resistance against antibiotics (such as geneticin or G-418 (GIBCO- BRL, Grand Island, NY), kanamycin or ampicillin), genes that provide for temperature resistance, or genes that allow the host cell or host organism to be maintained in the

absence of certain factors, compounds or (food) components in the medium that are essential for survival of the non-transformed cells or organisms.

A leader sequence should be such that - in the intended host cell or host organism - it allows for the desired post-translational modifications or such that it directs the transcribed mRNA to a desired part or organelle of a cell such as a signal peptide. A leader sequence may also allow for secretion of the expression product from said cell. As such, the leader sequence may be any pro-, pre-, or prepro-sequence operable in the host cell or host organism, including, but not limited to, picornavirus leaders, potyvirus leaders, a human immunoglobulin heavy-chain binding protein ("BiP"), a tobacco mosaic virus leader ("TMV"), and a maize chlorotic mottle virus leader ("MCMV").

An expression marker or reporter gene should be such that - in the host cell or host organism - it allows for detection of the expression of (a gene or nucleotide sequence present on) the genetic construct. An expression marker may optionally also allow for the localization of the expressed product, e.g. in a specific part or organelle of a cell or in (a) specific cell(s), tissue(s), organ(s) or part(s) of a multicellular organism. Such reporter genes may also be expressed as a protein fusion with the amino acid sequence of the invention. Some preferred, but non-limiting examples include fluorescent proteins, such as GFP, antibody recognition proteins, for example, V5 epitope or poly Histidine available in vectors and antibodies supplied by Invitrogen, or purification affinity handles such as polyhistidine which allows for purification on nickel columns or dihydrofolate reductase which allows for purification on methotrexate column, or markers which allow for selection of cells expressing the gene such as the *E. coli* beta-galactosidase gene.

For some non-limiting examples of the promoters, selection markers, leader sequences, expression markers and further elements that may be present or used in the genetic constructs of the invention - such as terminators, transcriptional or translational enhancers or integration factors - reference is made to the general handbooks such as Sambrook et al. and Ausubel et al. mentioned above, to W.B.

Wood et al., "*The nematode Caenorhabditis elegans*", Cold Spring Harbor Laboratory Press (1988) and D.L. Riddle et al., "*C. ELEGANS II*", Cold Spring Harbor Laboratory Press (1997), as well as to the examples that are given in WO 95/07463, WO 96/23810, WO 95/07463, WO 95/21191, WO 97/11094, WO 97/42320, WO 98/06737, WO 98/21355, U.S. Patent 6,207,410, U.S. Patent 5,693,492 and EP 1 085 089. Other examples will be clear to the skilled person.

Another embodiment of the invention relates to a host cell or host organism that has been transformed or contains with a nucleotide sequence, with a nucleic acid or with a genetic construct of the invention. The invention also relates to a host cell or host organism that expresses, or (at least) is capable of expressing (e.g. under suitable conditions), an amino acid sequence of the invention. Collectively, such host cells or host organisms will also be referred to herein as "*host cells or host organisms of the invention*".

The host cell may be any suitable (fungal, prokaryotic or eukaryotic) cell or cell line, for example:

- a bacterial strain, including but not limited to strains of *E. coli*, *Bacillus*, *Streptomyces* or *Pseudomonas*;
- a fungal cell, including but not limited to cells from species of *Aspergillus* or *Trichoderma*;
- a yeast cell, including but not limited to cells from species of *Kluyveromyces* or *Saccharomyces*;
- an amphibian cell or cell line, such as *Xenopus* oocytes.

In one specific embodiment, which may particularly useful when the nucleotide sequences of the invention are (to be) used in the discovery and development of insecticidal compounds, the host cell may be an insect-derived cell or cell line, such as:

- cells or cell lines derived from *Hemipteran*, including, but not limited to, *Spodoptera* SF9 and Sf21 cells and cells or cell lines derived from *Aphis*;
- cells or cell lines derived from *Drosophila*, such as Schneider and Kc cells; and

- cells or cell lines derived from a pest species of interest (as mentioned below), such as from *Aphis gossypii*.

The host cell may also be a mammalian cell or cell line, including but not limited to CHO- and BHK-cells and human cells or cell lines such as HEK, HeLa
5 and COS.

The host organism may be any suitable multicellular (vertebrate or invertebrate) organism, including but not limited to:

- a nematode, including but not limited to nematodes from the genus *Caenorhabditis*, such as *C. elegans*,
- 10 - an insect, including but not limited to species of *Aphis*, *Drosophila*, *Heliothis*, or a specific pest species of interest (such as those mentioned above);
- other well known model organisms, such as zebrafish;
- a mammal such as a rat or mouse;

Other suitable host cells or host organisms will be clear to the skilled person,
15 for example from the handbooks and patent applications mentioned above.

It should be noted that when a nucleotide sequence of the invention is expressed in a multicellular organism, it may be expressed throughout the entire organism, or only in one or more specific cells, tissues, organs or parts thereof, for example by expression under the control of a promoter that is specific for said
20 cell(s), tissue(s), organ(s) or part(s).

The nucleotide sequence may also be expressed during only a specific stage of development or life cycle of the host cell or host organism, again for example by expression under the control of a promoter that is specific for said stage of development or life cycle. Also, as already mentioned above, said expression may
25 be constitutive, transient or inducible.

Preferably, these host cells or host organisms are such that they express, or are (at least) capable of expressing (e.g. under suitable conditions), an amino acid sequence of the invention (and in case of a host organism: in at least one cell, part, tissue or organ thereof). The invention also includes further generations, progeny

and offspring of the host cell or host organism of the invention, which may for instance be obtained by cell division or by sexual or asexual reproduction.

In yet another aspect, the invention relates to a nucleic acid, preferably in (essentially) isolated form, which nucleic acid encodes or can be used to express an amino acid sequence of the invention (as defined herein), and in particular the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4. The amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, may be isolated from the species mentioned above, using any technique(s) for protein isolation and purification known to one skilled in the art. Alternatively, the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2, may be obtained by suitable expression of a suitable nucleotide sequence - such as the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:3 or a suitable mutant thereof - in an appropriate host cell or host organism, as further described below.

In another aspect, the invention relates to a protein or polypeptide, preferably in (essentially) isolated form, said protein or polypeptide comprising an amino acid sequence of the invention (as defined above), in particular the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2.

In a broader sense, the term "*amino acid sequence of the invention*" also comprises:

- 20 - parts or fragments of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2;
- (natural or synthetic) mutants, variants, alleles, analogs, orthologs (herein below collectively referred to as "*analogs*") of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2;
- 25 - parts or fragments of such analogs;
- fusions of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2; (or a part or fragment thereof) with at least one further amino acid residue or sequence;
- fusions of the amino acid sequence of an analog (or a part or fragment thereof)
- 30 with at least one further amino acid residue or sequence;

in which such mutants, parts, fragments or fusions are preferably as further described below.

The term “*amino acid sequence of the invention*” also comprises “immature” forms of the above-mentioned amino acid sequences, such as a pre-, pro- or prepro-
5 forms or fusions with suitable leader sequences. Also, the amino acid sequences of the invention may have been subjected to post-translational processing or be suitably glycosylated, depending upon the host cell or host organism used to express or produce said amino acid sequence; or may be otherwise modified (e.g. by chemical techniques known per se in the art).

10 Examples of parts or fragments of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2, or a part or fragment of a (natural or synthetic) analog thereof mutant thereof include, but are not limited to, N- and C-truncated amino acid sequence. Also, two or more parts or fragments of one or more amino acid sequences of the invention may be suitably combined to provide an
15 amino acid sequence of the invention.

Preferably, an amino acid sequence of the invention has a length of at least 100 amino acids, preferably at least 250 amino acids, more preferably at least 300 amino acids; and up to a length of at most 1,000 amino acids, preferably at most 750 amino acids, more preferably at most 600 amino acids.

20 Preferably, any such parts or fragments will be such that they comprise at least one continuous stretch of at least 5 amino acids, preferably at least 10 amino acids, more preferably at least 20 amino acids, even more preferably more than 30 amino acids, of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2.

25 In particular, any parts or fragments as described herein are such that they (at least) comprise the active or catalytic site of the corresponding amino acid sequence of the invention or a binding domain of the corresponding amino acid sequence of the invention. As will be clear to the skilled person, such parts or fragments may find particular use in assay- and screening techniques (as generally described below)

and (when said part or fragment is provided in crystalline form) in X-ray crystallography.

Also, it is expected that - based upon the disclosure herein - the skilled person will be able to identify, derive or isolate natural "analogs" (as mentioned
5 above) of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2. Such mutants could be derived from (other individuals of) the same species (for example from an individual of a different strain or line, including but not limited to mutant strains or lines); or from (individuals of) other species. For example, such analogs could be derived from the insect species mentioned above.

10 It is also expected that - based upon the disclosure herein - the skilled person will be able to provide or derive synthetic "analogs" (as mentioned above) of the amino sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2.

Preferably, any mutants as described herein will have one or more, and preferably all, of the structural characteristics or conserved features referred to
15 below for the sequences of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2.

Preferably, any analogs, parts or fragments as described herein will be such that they have a degree of "sequence identity", at the amino acid level, with the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2,
20 of at least 55%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90% and up to 95 % or more.

For this purpose, the percentage of "sequence identity" between a given amino acid sequence and the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2 may be calculated by dividing the number of amino
25 acid residues in the given amino acid sequence that are identical to the amino acid residue at the corresponding position in the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2 by the total number of amino acid residues in the given amino acid sequence and multiplying by 100%, in which each deletion, insertion, substitution or addition of an amino acid residue - compared to
30 the sequence of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2- is

considered as a difference at a single amino acid (position). As mentioned above, the preferred method of performing pairwise global sequence alignments for such calculations is with the program ClustalW.

Also, preferably, any analogs, parts or fragments as described herein will
5 have a biological activity that is essentially similar to the biological activity described above for the sequences of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2, i.e. to a degree of at least 10%, preferably at least 50% more preferably at least 75%, and up to 90%, as measured by standard assay techniques as described below.

10 It is also within the scope of the invention to use a fusion of an amino acid sequence of the invention (as described above) with one or more further amino acid sequences, for example to provide a protein fusion. Generally, such fusions may be obtained by suitable expression of a suitable nucleotide sequence of the invention - such as a suitable fusion of a nucleotide sequence of the invention with one or more
15 further coding sequences - in an appropriate host cell or host organism, as further described below.

One particular embodiment, such fusions may comprise an amino acid sequence of the invention fused with a reporter protein such as glutathione S-transferase ("GST"), green fluorescent protein ("GFP"), luciferase or another
20 fluorescent protein moiety. As will be clear to the skilled person, such fusions may find particular use in expression analysis and similar methodologies.

In another embodiment, the fusion partner may be an amino acid sequence or residue that may be used in purification of the expressed amino acid sequence, for example using affinity techniques directed against said sequence or residue.
25 Thereafter, said sequence or residue may be removed (e.g. by chemical or enzymatical cleavage) to provide the nucleotide sequence of the invention (for this purpose, the sequence or residue may optionally be linked to the amino acid sequence of the invention via a cleavable linker sequence). Some preferred, but non-limiting examples of such residues are multiple histidine residues and glutathione
30 residues.

In one preferred, but non-limiting aspect, any such fusion will have a biological activity that is essentially similar to the biological activity described above for the sequences of SEQ ID NO: 2 or SEQ ID NO:4, preferably SEQ ID NO: 2, i.e. to a degree of at least 10%, preferably at least 50 % more preferably at least 75%, and up to 90%, as measured by standard assay techniques as described below.

The nucleotide sequences and amino acid sequences of the invention may generally be characterized by the presence of one or more of the following structural characteristics or conserved features:

For the gene *aphis gossypii* 5HT1 receptor: SEQ ID NO:1 or SEQ ID NO:3 are cDNA sequences encompassing the open reading frame; SEQ ID NO: 2 or SEQ ID NO:4 are the protein encoded by SEQ ID NO: 1 or SEQ ID NO:3. The *aphis gossypii* 5HT1 receptor protein sequence is related to other 5HT1 receptor.

By analogy to other 5HT1 receptor, it is likely that the functional protein is monomeric. (See, e.g., Hannan and Hall, In Comparative Molecular Neurobiology, Y. Pichon, 1993, Birkhuaser Verlag Basel Switzerland).

On the basis of the above, and although the invention is not specifically limited to any specific explanation or mechanism, the nucleotide sequences and amino acid sequences have (biological) activity as G protein (Gi) coupled receptor that inhibits cAMP production. In particular, the present invention has shown activity as a 5HT1 receptor from insects of the order Hemiptera, which are aphids, leafhoppers, whiteflies, scales and true bugs that have mouthparts adapted to piercing and sucking.

As is known in the art, cAMP assay can be used to determine functional activity of 5HT1 receptor. It is preferred that when performing test assays using test compounds, the test compounds be prepared as serial dilutions such that multiple test assays are performed with the same compound at different concentrations. In some embodiments, assays optionally employ positive controls such as compounds known to have a particular biological activity. In some embodiments, the positive control may be an antibody specific for the polypeptide of the invention. In some embodiments, assays optionally employ negative controls such as compounds

known not to have a particular biological activity. These positive control assays and/or negative control assays may be run side by side with test assays to help establish and confirm that the compounds identified as having biological activity in the test assays are in fact biologically active. In addition to positive and negative
5 control assays, additional control assays may be performed using other known 5HT1 receptor from other species. By performing these types of assays, it may be established that the compound found to be active in the test assay is not active in assays using 5HT1 receptor from other species. This result would indicate that the compound has some degree of specificity. Kits can be provided to allow for such
10 assays to be performed. Kits can include containers containing some or all of the reagents needed to perform the assay, samples of a polypeptide of the invention, a nucleic acid of the invention, a genetic construct of the invention or a host cell of the invention.

Another embodiment of the invention relates to a nucleic acid probe that is
15 capable of hybridizing with a nucleotide sequence of the invention under conditions of moderate stringency, preferably under conditions of high stringency, and in particular under stringent conditions (all as described above). Such nucleotide probes may for instance be used for detecting or isolating a nucleotide sequence of the invention or as a primer for amplifying a nucleotide sequence of the invention;
20 all using techniques known per se, for which reference is again made to the general handbooks such as Sambrook et al. and Ausubel et al., mentioned above.

Preferably, when to be used for detecting or isolating another nucleotide sequence of the invention, such a nucleotide probe will usually have a length of between 15 and 100 nucleotides, and preferably between 20 and 80 nucleotides.
25 When used as a primer for amplification, such a nucleotide probe will have a length of between 25 and 75 nucleotides, and preferably between 20 and 40 nucleotides.

Generally, such probes can be designed by the skilled person starting from a nucleotide sequence or amino acid sequence of the invention - and in particular the sequence of SEQ ID NO: 1 or SEQ ID NO:3 - optionally using a suitable computer
30 algorithm. Also, as will be clear to the skilled person, such probes may be

degenerate probes. Probes and primers preferably have sequences that include unique sequences. A unique sequence is a sequence that is not found on other DNA molecules. The presence of unique sequences ensures that the probe or primer will not cross hybridize to identical sequences found on other genes. One skilled in the art can readily determine if a probe or primer contains unique sequences by first
5 designing the probe or primer and then comparing the sequences thereof with sequences in databases of known nucleic acid sequences. Such comparisons are routinely performed by those skilled in the art.

In a further aspect, the invention relates to methods for preparing mutants
10 and genetic constructs of the nucleotide sequences of the present invention.

Natural mutants of the nucleotide sequences of the present invention may be obtained in a manner essentially analogous to the method described in the Examples, or alternatively by:

- construction of a DNA library from the species of interest in an appropriate
15 expression vector system, followed by direct expression of the mutant sequence;
 - construction of a DNA library from the species of interest in an appropriate expression vector system, followed by screening of said library with a probe of the invention (as described below) or with a nucleotide sequence of the invention;
 - 20 - isolation of mRNA that encodes the mutant sequence from the species of interest, followed by cDNA synthesis using reverse transcriptase;
- or by any other suitable method(s) or technique(s) known per se, for which reference is for instance made to the standard handbooks, such as Sambrook et al., "Molecular Cloning: A Laboratory Manual" (2nd.ed.), Vols. 1-3, Cold Spring Harbor
25 Laboratory Press (1989) and F. Ausubel et al., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987).

Techniques for generating such synthetic sequences of the nucleotide sequences of the present invention will be clear to the skilled person and may for instance include, but are not limited to, automated DNA synthesis; site-directed
30 mutagenesis; combining two or more parts of one or more naturally occurring

sequences, introduction of mutations that lead to the expression of a truncated expression product; introduction of one or more restriction sites (e.g. to create cassettes or regions that may easily be digested or ligated using suitable restriction enzymes), and the introduction of mutations by means of a PCR reaction using one or more “mismatched” primers, using for example a sequence of other techniques will be clear to the skilled person, and reference is again made to the standard handbooks, such naturally occurring 5HT₁ receptor as a template. These and as Sambrook et al. and Ausubel et al., mentioned above.

The genetic constructs of the invention may generally be provided by suitably linking the nucleotide sequence(s) of the invention to the one or more further elements described above, for example using the techniques described in the general handbooks such as Sambrook et al. and Ausubel et al., mentioned above.

Often, the genetic constructs of the invention will be obtained by inserting a nucleotide sequence of the invention in a suitable (expression) vector known per se.

Some preferred, but non-limiting examples of suitable expression vectors include:

- vectors for expression in mammalian cells: pSVL SV40 (Pharmacia), pMAMneo (Clontech), pcDNA3 (Invitrogen), pMC1neo (Stratagene), pSG5 (Stratagene), EBO-pSV2-neo (ATCC 37593), pBPV-1 (8-2) (ATCC 37110), pDBPV-MMTneo (342-12) (ATCC 37224), pRSVgpt (ATCC37199), pRSVneo (ATCC37198), pSV2-dhfr (ATCC 37146), pUCtag (ATCC 37460) and 1ZD35 (ATCC 37565);
- vectors for expression in bacterial cells: pET vectors (Novagen), pGEX vectors (Pharmacia) and pQE vectors (Qiagen);
- vectors for expression in yeast or other fungal cells: pYES2 (Invitrogen) and Pichia expression vectors (Invitrogen);
- vectors for expression in insect cells: pBlueBacII (Invitrogen), pEI1 (Novagen), pMT/V5His (Invitrogen).

In a further aspect, the invention relates to methods for transforming a host cell or a host organism with a nucleotide sequence, with a nucleic acid or with a genetic construct of the invention. The invention also relates to the use of a

nucleotide sequence, of a nucleic acid or of a genetic construct of the invention transforming a host cell or a host organism.

According to one specific embodiment, the expression of a nucleotide sequence of the invention in a host cell or host organism may be reduced, compared
5 to the original (e.g. native) host cell or host organism. This may for instance be achieved in a transient manner using antisense or RNA-interference techniques well known in the art, or in a constitutive manner using random, site specific or chemical mutagenesis of the nucleotide sequence of the invention.

Suitable transformation techniques will be clear to the skilled person and
10 may depend on the intended host cell or host organism and the genetic construct to be used. Some preferred, but non-limiting examples of suitable techniques include ballistic transformation, (micro-)injection, transfection (e.g. using suitable transposons), electroporation and lipofection. For these and other suitable techniques, reference is again made to the handbooks and patent applications
15 mentioned above.

After transformation, a step for detecting and selecting those host cells or host organisms that have been successfully transformed with the nucleotide sequence or genetic construct of the invention may be performed. This may for instance be a selection step based on a selectable marker present in the genetic
20 construct of the invention or a step involving the detection of the amino acid sequence of the invention, e.g. using specific antibodies.

The transformed host cell (which may be in the form of a stable cell line) or host organisms (which may be in the form of a stable mutant line or strain) form further aspects of the present invention.

25 In yet another aspect, the invention relates to methods for producing an amino acid sequence of the invention.

To produce or obtain expression of the amino acid sequences of the invention, a transformed host cell or transformed host organism may generally be kept, maintained or cultured under conditions such that the (desired) amino acid
30 sequence of the invention is expressed or produced. Suitable conditions will be

clear to the skilled person and will usually depend upon the host cell or host organism used, as well as on the regulatory elements that control the expression of the (relevant) nucleotide sequence of the invention. Again, reference is made to the handbooks and patent applications mentioned above in the paragraphs on the genetic
5 constructs of the invention.

Generally, suitable conditions may include the use of a suitable medium, the presence of a suitable source of food or suitable nutrients, the use of a suitable temperature, and optionally the presence of a suitable inducing factor or compound (e.g. when the nucleotide sequences of the invention are under the control of an
10 inducible promoter); all of which may be selected by the skilled person. Again, under such conditions, the amino acid sequences of the invention may be expressed in a constitutive manner, in a transient manner, or only when suitably induced.

It will also be clear to the skilled person that the amino acid sequence of the invention may (first) be generated in an immature form (as mentioned above), which
15 may then be subjected to post-translational modification, depending on the host cell or host organism used. Also, the amino acid sequence of the invention may be glycosylated, again depending on the host cell or host organism used.

The amino acid sequences of the invention may then be isolated from the host cell or host organism or from the medium in which said host cell or host
20 organism was cultivated, using protein isolation and purification techniques known per se, such as (preparative) chromatography and electrophoresis techniques, differential precipitation techniques, affinity techniques (e.g. using a specific, cleavable amino acid sequence fused with the amino acid sequence of the invention) and preparative immunological techniques (i.e. using antibodies against the amino
25 acid sequence to be isolated).

In one embodiment, the amino acid sequence thus obtained may also be used to generate antibodies specifically against said sequence or an antigenic part or epitope thereof.

In one embodiment, the present invention relates to antibodies, for example
30 monoclonal and polyclonal antibodies, that are generated specifically against amino

acid sequences of the present invention, preferably SEQ ID NO. 2, or an analog, variant, allele, ortholog, part, fragment or epitope thereof, or SEQ ID NO: 4, or an analog, variant, allele, ortholog, part, fragment or epitope thereof. In preferred embodiments, the antibody does not cross reactive with other 5HT1 receptor polypeptides.

Such antibodies, which form a further aspect of the invention, may be generated in a manner known per se, for example as described in GB-A-2 357 768, USA 5,693,492, WO 95/32734, WO 96/23882, WO 98/02456, WO 98/41633 and WO 98/49306. Often, but not exclusively, such methods will involve as immunizing a immunocompetent host with the pertinent amino acid sequence of the invention or an immunogenic part thereof (such as a specific epitope), in amount(s) and according to a regimen such that antibodies against said amino acid sequence are raised, and then harvesting the antibodies thus generated, e.g. from blood or serum derived from said host.

For instance, polyclonal antibodies can be obtained by immunizing a suitable host such as a goat, rabbit, sheep, rat, pig or mouse with (an epitope of) an amino acid sequence of the invention, optionally with the use of an immunogenic carrier (such as bovine serum albumin or keyhole limpet hemocyanin) or an adjuvant such as Freund's, saponin, aluminium hydroxide or a similar mineral gel, or keyhole limpet hemocyanin or a similar surface active substance. After a suitable immune response has been raised (usually within 1-7 days), the antibodies can be isolated from blood or serum taken from the immunized animal in a manner known per se, which optionally may involve a step of screening for an antibody with desired properties (i.e. specificity) using known immunoassay techniques, for which reference is again made to for instance WO 96/23882.

Monoclonal antibodies may for example be produced using continuous cell lines in culture, including hybridoma-based and similar techniques, again essentially as described in the above cited references. Accordingly, cells and cell lines that produce monoclonal antibodies against an amino acid sequence of the invention form a further aspect of the invention, as do methods for producing antibodies against amino acid

sequences of the invention, which methods may generally involve cultivating such a cell and isolating the antibodies from the culture or medium, again using techniques known per se.

Also, Fab-fragments against the amino acid sequences of the invention (such as F(ab)₂, Fab' and Fab fragments) may be obtained by digestion of an antibody with pepsin or another protease, reducing disulfide-linkages and treatment with papain and a reducing agent, respectively. Fab-expression libraries may for instance be obtained by the method of Huse et al., 1989, Science 245:1275-1281.

In another embodiment, the amino acid sequence of the invention, or a host cell or host organism that expresses such an amino acid sequence, may also be used to identify or develop compounds or other factors that can modulate the (biological) activity of, or that can otherwise interact with, the amino acid sequences of the invention, and such uses form further aspects of the invention. As will be clear to the skilled person, in this context, the amino acid sequence of the invention will serve as a target for interaction with such a compound or factor.

In this context, the terms “*modulate*”, “*modulation*”, “*modulator*” and “*target*” will have their usual meaning in the art, for which reference is *inter alia* made to the definitions given in WO 98/06737. Generally, a modulator is a compound or factor that can enhance, inhibit or reduce or otherwise alter, influence or affect (collectively referred to as “*modulation*”) a functional property of a biological activity or process (for example, the biological activity of an amino acid sequence of the invention). As noted above, it is known in the art that 5HT1 receptor activity can be measured using standard assay techniques such as cAMP assays. Such technology, as well as other well-known technology, can be adapted to methods of finding biologically active compounds that specifically effect the compositions of the present invention, particularly, the polypeptide of the invention. In preferred embodiments, the biological activity is the inhibition 5HT1 receptor activity. It is preferred that when performing test assays using test compounds, the test compounds be prepared as serial dilutions such that multiple test assays are performed with the same compound at different concentrations. In some

embodiments, assays optionally employ positive controls such as compounds known to have a particular biological activity. In some embodiments, the positive control may be an antibody specific for the polypeptide of the invention. In some embodiments, assays optionally employ negative controls such as compounds
5 known not to have a particular biological activity. These positive control assays and/or negative control assays may be run side by side with test assays to help establish and confirm that the compounds identified as having biological activity in the test assays are in fact biologically active. In addition to positive and negative control assays, additional control assays may be performed using other known 5HT1
10 receptor from other species. By performing these types of assays, it may be established that the compound found to be active in the test assay is not active in assays using 5HT1 receptor from other species. This result would indicate that the compound has some degree of specificity. Kits can be provided to allow for such assays to be performed. Kits can include containers containing some or all of the
15 reagents needed to perform the assay, samples of a polypeptide of the invention, a nucleic acid of the invention, a genetic construct of the invention or a host cell of the invention.

In this context, the amino acid sequence of the invention may serve as a target for modulation *in vitro* (e.g. as part of an assay or screen) or for modulation *in*
20 *vivo* (e.g. for modulation by a compound or factor that is known to modulate the target, which compound or factor may for example be used as an active compound for agrochemical, veterinary or pharmaceutical use).

For example, the amino acid sequences, host cells or host organisms of the invention may be used as part of an assay or screen that may be used to identify or
25 develop modulators of the amino acid sequence of the invention, such as a primary screen (e.g. a screen used to identify modulators of the target from a set or library of test chemicals with unknown activity with respect to the target) or a secondary assay (e.g. an assay used for validating hits from a primary screen or used in optimizing hit molecules, e.g. as part of hits-to-leads chemistry).

For instance, such an assay or screen may be configured as an *in vitro* assay or screen. Suitable techniques for such *in vitro* screening will be clear to the skilled person, and are for example described in Eldefrawi et al., (1987). FASEB J., Vol.1, pages 262-271 and Rauh et al., (1990), Trends in Pharmacol. Sci., vol.11, pages 325-
5 329.

Such an assay or screen may also be configured as a cell-based assay or screen, in which a host cell of the invention is contacted with or exposed to a test chemical, upon which at least one biological response by the host cell is measured.

Also, such an assay or screen may also be configured as an whole animal
10 screen, in which a host organism of the invention is contacted with or exposed to a test chemical, upon which at least one biological response (such as a phenotypical, behavioral or physiological change, including but not limited to paralysis or death) by the host organism is measured.

Thus, generally, the assays and screens described above will comprise at
15 least one step in which the test chemical is contacted with the target (or with a host cell or host organism that expresses the target), and in particular in such a way that a signal is generated that is representative for the modulation of the target by the test chemical. In a further step, said signal may then be detected.

Accordingly, in one aspect, the invention relates to a method for generating a
20 signal that is representative for the interaction of an amino acid sequence of the invention with a test chemical, said method at least comprising the steps of:

- a) contacting the amino acid sequence of the invention, or a host cell or host organism containing or expressing an amino acid sequence, with said test chemical, in such a way that a signal may be generated that is representative for
25 the interaction between said test chemical and said amino acid sequence; and optionally
- b) detecting the signal that may thus be generated.

In another aspect, the invention relates to a method for identifying modulators and/or inhibitors of an amino acid sequence of the invention (e.g. from a
30 set or library of test chemicals), said method at least comprising the steps of:

- a) contacting the amino acid sequence of the invention, or a host cell or host organism containing or expressing an amino acid sequence, with a test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said the target; and optionally
- 5 b) detecting the signal that may thus be generated, said signal identifying the modulator and/or inhibitor of said amino acid sequence.

Accordingly, the present invention provides methods of identifying a modulator and/or inhibitor of a 5HT1 receptor activity. In preferred embodiments, the 5HT1 receptor protein used in the methods has an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, a mutant thereof, a fragment thereof,

10 SEQ ID NO: 4, a mutant thereof, and a fragment thereof, preferably SEQ ID NO: 2. In some embodiments, the nucleic acid sequence that encodes the 5HT1 receptor is SEQ ID NO: 1 a fragment thereof, SEQ ID NO: 3, or a fragment thereof, preferably SEQ ID NO: 1.

15 A test chemical may be part of a set or library of compounds, which may be a diverse set or library or a focussed set or library, as will be clear to the skilled person. The libraries that may be used for such screening can be prepared using combinatorial chemical processes known in the art or conventional means for chemical synthesis.

20 The assays and screens of the invention may be carried out at medium throughput to high throughput, for example in an automated fashion using suitable robotics. In particular, in this embodiment, the method of the invention may be carried out by contacting the target with the test compound in a well of a multi-well plate, such as a standard 24, 96, 384, 1536 or 3456 well plate.

25 Usually, in a screen or assay of the invention, for each measurement, the target or host cell or host organism will be contacted with only a single test compound. However, it is also within the scope of the invention to contact the target with two or more test compounds - either simultaneously or sequentially - for example to determine whether said combination provides a synergistic effect.

Once a test chemical has been identified as a modulator and/or inhibitor for an amino acid sequence of the invention (e.g. by means of a screen or assay as described hereinabove), it may be used per se as a modulator and/or inhibitor of the relevant amino acid sequence of the invention, preferably, an amino acid sequence of SEQ ID NO:2, a mutant thereof, a fragment thereof, SEQ ID NO: 4, a mutant thereof, and a fragment thereof, more preferably SEQ ID NO: 2 (e.g. as an active substance for agrochemical, veterinary or pharmaceutical use), or it may optionally be further optimized for final use, e.g. to improve properties such as solubility, adsorption, bio-availability, toxicity, stability, persistence, environmental impact, etc.. It will be clear to the skilled person that the nucleotide sequences, preferably SEQ ID NO: 1 or SEQ ID NO: 3, more preferably SEQ ID NO: 1, amino acid sequences, host cells or host organisms and methods of the invention may find further use in such optimization methodology, for example as (part of) secondary assays.

The invention is not particularly limited to any specific manner or mechanism in or via which the modulator and/or inhibitor (e.g. the test chemical, compound or factor) modulates, inhibits, or interacts with, the target (*in vivo* or *in vitro*). For example, the modulator and/or inhibitor may be an agonist, an antagonist, an inverse agonist, a partial agonist, a competitive inhibitor, a non-competitive inhibitor, a cofactor, an allosteric inhibitor or other allosteric factor for the target, or may be a compound or factor that enhances or reduces binding of target to another biological component associated with its (biological) activity, such as another protein or polypeptide, a receptor, or a part of organelle of a cell. As such, the modulator and/or inhibitor may bind with the target (at the active site, at an allosteric site, at a binding domain or at another site on the target, e.g. covalently or via hydrogen bonding), block and/or inhibit the active site of the target (in a reversible, irreversible or competitive manner), block and/or inhibit a binding domain of the target (in a reversible, irreversible or competitive manner), or influence or change the conformation of the target.

As such, the test chemical, modulator and/or inhibitor may for instance be:

- an analog of a known substrate of the target;
- an oligopeptide, e.g. comprising between 2 and 20, preferably between 3 and 15 amino acid residues;
- an antisense or double stranded RNA molecule;
- 5 - a protein, polypeptide;
- a cofactor or an analog of a cofactor.

The test chemical, modulator and/or inhibitor may also be a reference compound or factor, which may be a compound that is known to modulate, inhibit or otherwise interact with the target (e.g. a known substrate or inhibitor for the target)
10 or a compound or factor that is generally known to modulate, inhibit or otherwise interact with other members from the general class to which the target belongs (e.g. a known substrate or inhibitor of said class).

Preferably, however, the test chemical, modulator and/or inhibitor is a small molecule, by which is meant a molecular entity with a molecular weight of less than
15 1500, preferably less than 1000. This may for example be an organic, inorganic or organometallic molecule, which may also be in the form of a suitable salt, such as a water-soluble salt. The term "small molecule" also covers complexes, chelates and similar molecular entities, as long as their (total) molecular weight is in the range indicated above.

20 As already mentioned above, the compounds or factors that have been identified or developed as modulators and/or inhibitors of the amino acid sequences of the invention, preferably, an amino acid sequence of SEQ ID NO: 2, a mutant thereof, a fragment thereof, SEQ ID NO: 4, a mutant thereof, and a fragment thereof, more preferably SEQ ID NO: 2, (and precursors for such compounds) may be useful
25 as active substances in the agrochemical, veterinary or pharmaceutical fields, for example in the preparation of agrochemical, veterinary or pharmaceutical compositions, and both such modulators as well as compositions containing them further aspects of the invention.

For example, in the agrochemical field, the modulators and/or inhibitors of
30 the invention may be used as an insecticide, nematocide, molluscicide, helminthicide,

acaricide or other types of pesticides or biocides, e.g. to prevent or control (infestations with) harmful organisms, both as contact agents and as systemic agents. As such, the modulators and/or inhibitors may for example be used as a crop protection agent, as a pesticide for household use, or as an agent to prevent or treat
5 damage caused by harmful organisms (e.g. for the protection of seed, wood or stored crops or fruits). Preferably, the modulators and/or inhibitors of the invention are used as insecticides.

For any such application, one or more modulators and/or inhibitors of the invention may be suitably combined with one or more agronomically acceptable
10 carriers, adjuvants or diluents - and optionally also with one or more further compounds known per se with activity as (for example) a plant protection agent (to broaden the spectrum of action and optionally to provide a synergistic effect), herbicide, fertilizer or plant growth regulator - to provide a formulation suitable for the intended final use. Such a formulation may for example be in the form of a
15 solution, emulsion, dispersion, concentrate, aerosol, spray, powder, flowable, dust, granule, pellet, fumigation candle, bait or other suitable solid, semi-solid or liquid formulation, and may optionally also contain suitable solvents, emulsifiers, stabilizers, surfactants, antifoam agents, wetting agents, spreading agents, sticking agents, attractants or (for a bait) food components. Reference is made to the
20 standard manuals, such as "Pesticidal Formulation Research", ACS-publications (1969) and "Pesticide Formulations", Wade van Valkenburg Ed, Marcel Dekker publications (1973).

Such compositions may generally contain one or more modulators and/or inhibitors of the invention in a suitable amount, which generally may be between 0.1
25 and 99 %, and in particular between 10 and 50 %, by weight of the total composition.

The modulators and/or inhibitors and compositions of the invention may be particularly useful as insecticides, for example to combat or control undesired or harmful insects (both adult and immature forms, such as larvae) from following
30 orders:

- *Coleoptera*, such as *Pissodes strobi*, *Diabrotica undecimpunctata howardi*, and *Leptinotarsa decemlineata*;
- *Diptera*, such as *Rhagoletis pomonella*, *Mayetiola destructor*, and *Liriomyza huidobrensis*;
- 5 - *Hymenoptera*, such as *Neodiprion taedae tsugae*, *Camponotus pennsylvanicus*, and *Solenopsis wagneri*;
- *Hemiptera*, such as *Pseudatomoscelis seriatus*, *Lygus lineolaris* (Palisot de Beauvois), *Acrosternum hilare*, and *Aphis gossypii*
- *Homoptera*; and
- 10 - *Lepidoptera* such as *Heliothis virescens*.

When used to control harmful or undesired organisms, these organisms may be directly contacted with the modulators, inhibitors, or compositions of the invention in an amount suitable to control (e.g. kill or paralyze) the organism. This amount may be readily determined by the skilled person (e.g. by testing the
15 compound on the species to be controlled) and will usually be in the region of between particular between 10 and 500 g/ha, in particular between 100 and 250 g/ha.

The modulators, inhibitors, or compositions of the invention may also be applied systemically (e.g. to the habitat of the organism to be controlled or to the soil), and may also be applied to the plant, seed, fruit etc. to be protected, again in
20 suitable amounts, which can be determined by the skilled person. The modulators and/or inhibitors of the invention may also be incorporated - e.g. as additives - in other compositions known per se, for example to replace other pesticidal compounds normally used in such compositions.

In one specific embodiment, the modulators and/or inhibitors and
25 compositions of the invention may be used in the fields of agrochemical, veterinary or human health to prevent or treat infection or damage or discomfort caused by parasitic organisms, and in particular by parasitic arthropods, nematodes and helminths such as:

- ectoparasitic arthropods such as ticks, mites, fleas, lice, stable flies, horn flies,
30 blowflies and other biting or sucking ectoparasites;

- endoparasites organisms such as helminths;

and also to prevent or treat diseases that are caused or transferred by such parasites.

For such purposes, the modulators and/or inhibitors of the invention may for example be formulated as a tablet, an oral solution or emulsion, an injectable

5 solution or emulsion, a lotion, an aerosol, a spray, a powder, a dip or a concentrate.

In the fields of animal and human health, the modulators, inhibitors, and compositions of the invention may also be used for the prevention or treatment of diseases or disorders in which the amino acid sequence of the invention may be involved as a target. For this purpose, the modulators and/or inhibitors of the
10 invention may be formulated with one or more additives, carriers or diluents acceptable for pharmaceutical or veterinary use, which will be clear to the skilled person.

Thus, in a further aspect, the invention relates to the use of a modulator and/or inhibitor of the invention in the preparation of a composition for
15 agrochemical, veterinary or pharmaceutical use, as described hereinabove. The invention relates to the use of the modulators, inhibitors and compositions of the invention in controlling harmful organisms and in preventing infestation or damage caused by harmful organisms, again as described above.

The invention will now be further illustrated by means of the following non-
20 limiting Examples.

Example 1. Cloning of *Aphis gossypii* 5HT1 receptor

1. Isolation of poly(A⁺) RNA.

25 Cotton aphids were collected from cotton plants and placed in ice-chilled glass centrifuge tubes which had been cleaned and baked for 6 hours at 180°C prior to use. Aliquots of approximately 0.4 gram of cotton aphids was used for isolation of poly(A⁺) RNA.

30 Diethyl pyrocarbonate (DEPC)-treated water was made by incubating DEPC (Aldrich Chemical Co., Inc. Milwaukee, WI) in water at concentration of 0.1% (v/v)

for 16 hours at room temperature, followed by autoclaving. The microprobe of a Braun homogenizer (B. Brawn Biotech International, Allentown, PA) was soaked in 100% ethanol and dried prior to use.

5 RNA isolation was done using QuickPrep mRNA Purification kit (Amersham Pharmacia biotech, Piscataway, NJ) according to the manufacturer's instruction. All the buffers and solutions mentioned here are included in the kit. An aliquot of 0.4 gram of cotton aphid was homogenized at full speed in 1.5 ml chilled extraction buffer until it is in a uniform suspension. After adding 3 ml of elution
10 buffer, the sample was homogenized again briefly and the resulting mixture was centrifuged at approximately 12000 x g for 10 minutes at room temperature. The supernatant was used for poly(A⁺) RNA isolation. After application of supernatant to the resin of oligo(dT)-cellulose spun column, washing with high salt and low salt buffers, the bound poly(A⁺) RNA was eluted with three washes of 0.25 ml elution
15 buffer pre-warmed to 65°C. To precipitate the mRNA, 50 µl of K Acetate solution, 10 µl of Glycogen solution, and 1 ml of 95% Ethanol were added to 0.5 ml of elute. The mixture was placed at -20°C for one hour and then centrifuged at maximal speed at room temperature in an eppendorf microcentrifuge. Precipitated poly(A⁺) RNA was then dissolved in 50 µl DEPC-treated water and stored at -80°C until use.

20

2. Reverse Transcription and PCR amplification (RT-PCR), RACE reactions.

Reverse transcription (RT) was first carried out using oligo dT as the primer and the ThermoScript RT-PCR system (Gibco). The resulting RT reaction mixture was used in the following PCR reaction.

25 Initial degenerate primers were designed using CODEHOP program. CODEHOP primers (table 1) designed from TM6 and TM7 were first used for nest PCR reaction. First round of amplification used primer 1 and primer 2 in a touchdown cycling program, with annealing temperature of first 20 cycles decreasing 0.5°C per cycle from 60°C to 50°C, followed by additional 20 cycles of
30 amplification at an annealing temperature of 55°C. The resulting PCR reaction

mixture was then subjected to a second round of amplification (same touchdown cycling conditions) using nested primer pair (primers 3 and 4).

Table 1

Primer ID	Primer sequence	Primer orientation
1	SEQ ID NO:5 AAACACTGGAAGCCAAAAGAgarmgraargc	forward
2	SEQ ID NO:6 AAATAACTGGGTTTCAGCGTAGAGttraartadcc	reverse
3	SEQ ID NO:7 CGGTGCCTTCGTGATCtgytggytrcc	forward
4	SEQ ID NO:8 TGGGTTTCAGCGTAGAGTTGaartadccyarc	reverse

From above-described RT-PCR reaction we obtained a very small quantity of PCR product which was not sufficient for sequencing. We then use pCRII-TOPO vector (Invitrogen, Carlsbad, CA) for sub-cloning. We directly used PCR product from Taq amplification in TOPO cloning; or before TOPO cloning, we incubated PCR product with Taq polymerase at 72°C to add A's to the PCR product. TOPO cloning was done according to manufacturer's instruction. The resulting plasmid was sequenced using T7 or SP6 as sequencing primers. BLAST analysis of sequencing results suggested that we had obtained a fragment of aphid 5HT1 (0.15 kb) partially spanning TM6 and TM7.

We then design a gene specific primer (primer 6, table 2) from this 0.15 kb fragment to pair with a CODEHOP primer from TM2 (Primer 5, table 2) and conduct PCR amplification using primers 5 and primer 6 in a touchdown cycling program, with annealing temperature of first 20 cycles decreasing 0.5°C per cycle from 61°C to 51°C, followed by additional 20 cycles of amplification at an annealing temperature of 56°C. This amplification yielded a DNA fragment of approximately 1kb corresponding to aphid 5HT1 between TM 2 and TM7, by blast analysis.

Table 2

Primer ID	Primer sequence	Primer orientation
5	SEQ ID NO:9 ACCTGATGGTGGCCTGCytrgtnatgcc	forward

6	SEQ ID NO:10TTGGGTTTCAGCGTAGAGTTG	reverse
---	--------------------------------------	---------

By designing more gene specific primers to pair with CODEHOP primers in TM1 region we were able to further expand toward the 5' end of the 5HT1 gene up to the TM1. We then used RACE technique to complete the 5' end of the gene, but
5 did not succeed in that effort, however.

We used RACE technique (Gibco Life technologies) to obtain DNA sequence downstream to TM7. Per manufacturer's instruction, After annealing oligo(dT)-containing Adapter Primer to mRNA and extending AP using Superscript RT, we degraded RNA template with RNaseH set up PCR amplification with nested
10 primers. First round of PCR used primer 7 (table 3) and primer 9 (UAP, supplied by Gibco Life technologies) using a touchdown cycling conditions with annealing temperature of first 20 cycles decreasing 0.5°C per cycle from 65°C to 55°C, followed by additional 20 cycles of amplification at an annealing temperature of 60°C. The resulting reaction mixture was subjected to a second round of PCR using
15 primer 8 (table 3) and primer 9 with the same cycling conditions as first round. This 3' RACE yielded a fragment of approximately 0.6 Kb which comprises primarily the 3'UTR.

Table 3

Primer ID	Primer sequence	Primer orientation
7	SEQ ID NO:11 GTCAGTCAAGACATTGGCTA	forward
8	SEQ ID NO:12 GGCGTTCACGATCACTA	forward
9	SEQ ID NO:13 (CUA) ₄ GGCCACGCGTCGACTAGTAC	reverse

20

To summarize the above RT-PCR and RACE attempt to clone the aphid 5HT1 gene, we were able to obtain >85% of the DNA sequence except the 5' end. Indeed, when we used the gene sequence information we obtained we are able to obtain a large fragment of the gene which is shown in SEQ ID #2. To obtain the full
25 length gene, we proceeded with library selection(below).

4. Construction of cDNA libraries and sequence databases

A cDNA libraries enriched for full-length clones was constructed by Invitrogen Corporation (Carlsbad, CA) from poly(A)-containing RNA purified from mixed life stage *Aphis gossypii* (CA cDNA library). This library was constructed in plasmid vector pCMV-SPORT6.1 and transformed into *E. coli* strain DH10B (TonA). The library was stored without further amplification as a glycerol culture at -80°C. A portion of the cDNA library was amplified, normalized and arrayed in 384 well plates by Invitrogen Corporation. The arrayed clones were subjected to automated sequence analysis using vector primers that flanked the 5' and 3' termini of the cloned cDNA (Genome Therapeutics, Waltham, MA). The sequence reads were trimmed for quality and vector contamination, assembled into contigs and installed into a BLAST-compatible database. Access to the Basic Local Alignment Sequence Tool (BLAST) suite of search tools and instructions for the use of these tools and the construction of BLAST-compatible databases can be obtained from the National Center for Biotechnology Information at the National Institutes for Health (Bethesda, MD).

5. RecA cloning of the full length 5HT1 gene

A full-length cDNAs corresponding to *Aphis Gossypii* 5HT1 receptor partial cDNA sequence, was isolated from the cotton aphid cDNA library by RecA-mediated gene enrichment using the RecActive™ Gene Enrichment Kit and protocol from Active Motif (Carlsbad, CA). Briefly, a biotinylated probe was synthesized from the cDNA fragment by a PCR reaction containing the appropriate primers B and C (see Table 4) and biotin-dNTP mix from RecActive™. The biotinylated probe was then gel purified and used to clone full length Cotton Aphid cDNAs from the non-normalized Cotton Aphid cDNA library by recA-mediated gene selection, as implemented in the RecActive™ Gene Enrichment Kit. PCR using the appropriate primers A and D were used to screen isolated clones. Resulting clones of the correct

size for a full-length insert for 5HT1 receptor were sequenced to confirm completeness.

Table 4.

Primer	Sequence
A	SEQ ID NO:14 CGGTCGCCGATTTAATGGT
B	SEQ ID NO:15 TGGTGTGGTTTGTGCGACTA
C	SEQ ID NO:16 ATGACCGTCGATCAGACCTA
D	SEQ ID NO:17 ATCGTGAACGCCGTGGTCT

5

Example 2. Stable expression of *Aphis gossypii* 5HT1 receptor in HEK293 cells

PCR was used to introduce the Kozak consensus sequence to 5HT1 receptor (SEQ ID NO 1) for optimal mammalian expression. The resulting Kozak-containing full length gene was cloned into pcDNA3.1(+) vector by restriction digestion, in which process the 3' UTR was deleted. The recombinant vector was then transfected into HEK-293 cells with lipofectamine. G418 (0.8 mg/ml) was used to enrich the population of cells that express 5HT1. Single clones of cells were picked and assayed for 5HT1 receptor activity. The positive cell clones were maintained in culture for at least six months, and 5HT1 receptor assays were done to confirm stable activity.

Example 3. Functional assays of *Aphis gossypii* 5HT1 receptor

The 5HT1 receptor activity was determined by cAMP production in HEK-293 cells.

Concentration of cAMP in HEK cells was measured using HTRF (Homogeneous Time Resolved Fluorescence) method per manufacturer's instruction

25

(CIS bio international). It is known that upon serotonin stimulation, 5HT1 receptor mediates the inhibition of forskolin-stimulated cAMP production. As shown in Figure 1, serotonin stimulation caused HEK-293 cells expressing aphid 5HT1 receptor gene to decrease the production of cAMP, in a concentration-dependent manner, indicating it is a Gi coupled receptor, as expected. Please note: higher fluorescent reading (Delta F%) indicates lower cAMP concentration.

5

CLAIMS

1. A substantially pure protein having the amino acid sequence selected from the group consisting of: SEQ ID NO: 2, a mutant thereof, a fragment thereof, SEQ ID NO: 4, a mutant thereof, and a fragment thereof, wherein said protein has 5HT1 receptor activity OR wherein said protein BINDS to serotonin.
2. The protein of claim 1 wherein said protein has the amino acid sequence selected from the group consisting of: SEQ ID NO: 2, a mutant thereof, and a fragment thereof.
3. The protein of claim 1 wherein said protein has the amino acid sequence selected from the group consisting of: SEQ ID NO: 4, a mutant thereof, and a fragment thereof.
4. An isolated nucleic acid molecule comprising a nucleic acid sequence encoding a protein of claim 1, wherein the nucleic acid sequence is selected from the group consisting of:
 - (a) a nucleic acid sequence encoding a polypeptide comprising SEQ ID NO:
5 2;
 - (b) DNA sequences that detectably hybridize to the nucleic acid sequences of
(a) or their complementary strands under conditions of moderate stringency;
 - (c) fragments thereof.
 - (d) a nucleic acid sequence encoding a polypeptide comprising SEQ ID NO:
10 4;
 - (e) DNA sequences that detectably hybridize to the nucleic acid sequences of
(d) or their complementary strands under conditions of moderate stringency; and
 - (f) fragments thereof.
5. An isolated nucleic acid molecule that comprises a nucleic acid sequence that encodes the protein of claim 1.

6. A recombinant expression vector comprising the nucleic acid molecule of claim 5.
7. A host cell comprising the recombinant expression vector of claim 6.
8. An isolated nucleic acid molecule having a nucleic acid sequence selected from the group consisting of: SEQ ID NO: 1, a fragment thereof having at least 10 nucleotides, SEQ ID NO: 3, and a fragment thereof having at least 10 nucleotides.
9. A recombinant expression vector comprising the nucleic acid molecule of claim 8.
10. A host cell comprising the recombinant expression vector of claim 9.
11. A method of identifying a modulator of 5HT₁ receptor protein activity by performing a test assay comprising the steps of:
 - a) contacting the protein of claim 1, or a host cell or host organism containing or expressing said protein, with a substrate in the presence of said test compound;
 - b) detecting whether said substrate is processed at a rate less than processing rate of said substrate in the presence of said protein and absence of said test compound, wherein reduction in substrate processing in the presence of said test compound indicates said test compound is an inhibitor of said protein's activity AND wherein an increase in substrate processing in the presence of said test compound indicates said test compound is an enhancer of said protein's activity.

OR

- a) contacting the protein of claim 1, or a host cell or host organism containing or expressing said protein, with a ligand of said protein in the presence of a test compound; and
- b) detecting the said ligand binds to said protein at a level less than the binding level observed when said protein is contacted with said ligand in the absence of said test compounds, wherein reduction in level of said ligand binding to said protein in the presence of said test compound indicates said test compound is an inhibitor of ligand/protein binding activity and wherein an increase in level of said ligand binding to said protein in the presence of said test compound indicates said test compound is an enhancer of ligand/protein binding activity

OR

- a) contacting the protein of claim 1, or a host cell or host organism containing or expressing said protein, with a test compound, in such a way that a signal may be generated that is representative for the interaction between said test compound and said target; and
- b) detecting the signal that may thus be generated, said signal identifying the modulator of said protein.

12. The method of claim 11 wherein said 5HT1 receptor protein has an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, a mutant thereof, a fragment thereof, SEQ ID NO: 4, a mutant thereof, and a fragment thereof.

13. The method of claim 11 further comprising the performance of a positive control assay and/or a negative control assay and/or a control assay using a non-5HT1 receptor

14. A method of preparing an isolated protein having the amino acid sequence selected from the group consisting of: SEQ ID NO: 2, a mutant thereof, a fragment thereof, SEQ ID NO: 4, a mutant thereof, and a fragment thereof comprising the step of isolating said protein from a host cell of claim 7.

15. A method of controlling an insect, comprising contacting the insect with the modulator of the protein of claim 1.
16. A method of controlling a hemipteran insect, comprising contacting the hemipteran insect with the modulator of claim 1.
17. A method of controlling a hemipteran insect, comprising contacting the hemipteran insect with the inhibitor of claim 1.
18. An antibody that specifically binds to the protein of claim 1.
19. The antibody of claim 18 wherein said antibody is a monoclonal antibody.
20. The antibody of claim 18 wherein said antibody specifically binds to a protein that has the amino acid sequence selected from the group consisting of: SEQ ID NO: 2, a mutant thereof, a fragment thereof, SEQ ID NO: 4, a mutant thereof, and a fragment thereof
21. The antibody of claim 18 which binds to an epitope on SEQ ID NO:2 or an epitope on SEQ ID NO: 4.

1/1

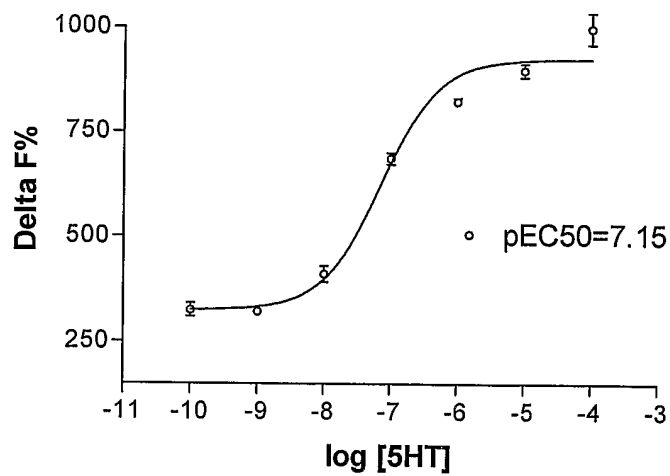


Figure 1. Serotonin inhibited forskolin-stimulated cAMP production in a dose dependent manner in HEK-293 cells expressing aphid 5HT1 receptors.

FMC-ADV-010.ST25.txt
SEQUENCE LISTING

<110> FMC CORPORATION
Chen, Ruihua
Dierks, Peter
Huang, Cancan
DitoIvo, Kevin
Allenza, Paul

<120> Hemipteran 5HT1 receptor

<130> FMC-ADV-010

<150> 60/630,548

<151> 2004-11-23

<160> 17

<170> PatentIn version 3.3

<210> 1

<211> 2440

<212> DNA

<213> APHIS GOSSYPPII

<400> 1

gctgaggtag	gaacgcgtat	ccgccgcttg	atgtggacca	ccgtttttcc	gctttccatc	60
cggtgtgatca	taactgctct	cgtatgcacg	cggtgaaaat	aactgtcaat	aattaatcat	120
tcaaattata	aacaatagtc	gatttaactg	tataatatac	gtttttaata	aataaaccac	180
cgccgtttcc	gtagacgtac	gtgtgtgaga	acaattttccg	ccgttagatc	tggtatgcac	240
ggaaagttac	aataatgttc	ggtttttttt	ttaataatat	tggtattttt	ttctcaccga	300
tcaatgaaag	aaacatgaac	catttggtgg	acccaaaaaa	ttgtatgggt	aaggggtgaat	360
tcataccgtg	cactgcagtc	gaaaacgact	taataatgac	gggctacgac	gaattttgat	420
ccggcgggccg	gtcaacaaga	ccgcgacgag	tatcatatgg	ccctgctcgg	agcacttgcg	480
gccaccggca	gtagacgcca	gggctgcgtg	ctgcggggcg	cggccgacgc	gataacgttc	540
tactgcggcg	gcgacgtggt	cagtggcggc	ggcagcagtg	gcggcggtgg	tggcggtggc	600
agtggcgggcg	acggtgcagc	cgccgaatac	gacggggccg	cggccgcgtc	cggccggtcg	660
tccgcccggcg	gcgatcgcg	actcggcgac	tacttccggc	ggcccgtccg	gtcgggtggcc	720
atcggcgacg	gcggttccgg	ggccggcggc	tactactacg	gcaccgaaaa	cggttcgtcg	780
gcggccgccc	tgctggccaa	cgcgacggcc	ggcgtggaac	cgttcctgaa	cggttcgtac	840
aatcacacgc	tcggcgccga	ggtagaggac	ctcaccaatt	acatattcgt	catcaccatc	900
ggcataatat	tgggcgcgat	gatactgacg	accatttgtg	gtaacgcgtt	ggtgatagca	960
gcaattttgt	tggagaggaa	tttacaaaac	gtggccaact	acctcattgt	ctcgctagcg	1020
gtcgccgatt	taatggtagc	atgttttagtc	atgcccttag	gagctgttta	tgagataagc	1080
agaggatgga	tattaggtcc	agaactatgt	gacatgtgga	cttcaagtga	cgtactatgt	1140

FMC-ADV-010.ST25.txt

```

tgtacagctt ccatacttca cctagtcgca atagctgttg atagatattg ggctgttact 1200
aatgttgatt atatacacac gagaaacggg agaagaatag taggcatgat tattgtggtg 1260
tggtttgctg cactagtcgt gtcattagct ccacagtttg gctggaaaga cccggattac 1320
ctagatcgta tcaatattca acaaagatgt atggtcagtc aagacattgg ctatcagata 1380
tttgccacat gttcaacgtt ttacttgcct ctactgggtca tcctgtttct ttattggaag 1440
atattcaaga tagccaggag aaggattcgc agaagaaggg ctacagagaaa cgctatgata 1500
gaacattcaa ggtcgaaaaa gccagcggac gaagcggccg catcggttga gcccaaaaaa 1560
tgcggtttct tccggaaggc cgttttcctg cggatcaaga aaaagaagga agtggaagag 1620
acggccgtgt gctcgagtat aggtctgac gacggtcatt cgagctgctc gatgcccagag 1680
tcacgcaag acggcgggtg aggggggtggc ggcggtcagt cgaaggcggg caacggcagc 1740
ggaagcgggtg gcggaagcgg tgggtggaagc ggtggcgga gttgtgggtg aagcgggtggc 1800
gaagacaaga ccacggcgtt cacgatcact accaatcacg acctcgggtg tacgtcgacg 1860
ctccagtgcc atcagcagca gcaacaccac caccaccacc accaacacca ccaacaccaa 1920
aaccagaatc agctccagcc tccggtgatc gaggacgacg gaccacaat cacggatccg 1980
tcggcgtccc ggcacacggt catcgagatg aacgtgaacc ggccttgag cgtgccggcg 2040
cccgttcggc cgaagacgaa gcgcgagaag aaagagtcgc tggaggcgaa acgcgaacgg 2100
aaagcggcca agacgttggc catcatcacc ggggcctttg tcgtgtgctg gctgccgttc 2160
ttcatcatgg ctctggtcat gccctgtgc cagggttgcg tgatcaacga gtacgtatcc 2220
agcttcttcc tgtggttggg ttacttcaat tccaccctga acccgatcat ctacaccgtt 2280
ttcagtcccg agttccggtt ggcattcacc aaaatgattt gcggcactac gtataggaga 2340
taatgaaatg cgcgacatt attattatta ttattattat tattattaca taaacaatac 2400
atataatatt attatattaa aaaaaaaaaa aaaaaaaaaa 2440

```

```

<210> 2
<211> 628
<212> PRT
<213> APHIS GOSSYPHI

```

```

<400> 2

```

```

Met Ala Leu Leu Gly Ala Leu Ala Ala Thr Gly Ser Arg Arg Gln Gly
1           5           10          15

```

```

Cys Val Leu Arg Ala Ala Ala Asp Ala Ile Thr Phe Tyr Cys Gly Gly
          20          25          30

```

```

Asp Val Val Ser Gly Gly Gly Ser Ser Gly Gly Gly Gly Gly Gly Gly
          35          40          45

```

FMC-ADV-010.ST25.txt

Ser Gly Gly Asp Gly Ala Ala Ala Glu Tyr Asp Gly Ala Ala Ala Ala
 50 55 60

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

Arg Arg Pro Val Arg Ser Val Ala Ile Gly Asp Gly Gly Ser Gly Ala
 85 90 95

Gly Gly Tyr Tyr Tyr Gly Thr Glu Asn Gly Ser Ser Ala Ala Ala Val
 100 105 110

Leu Ala Asn Ala Thr Ala Gly Val Glu Pro Phe Leu Asn Gly Ser Tyr
 115 120 125

Asn His Thr Leu Gly Ala Glu Val Glu Asp Leu Thr Asn Tyr Ile Phe
 130 135 140

Val Ile Thr Ile Gly Ile Ile Leu Gly Ala Met Ile Leu Thr Thr Ile
 145 150 155 160

Cys Gly Asn Ala Leu Val Ile Ala Ala Ile Leu Leu Glu Arg Asn Leu
 165 170 175

Gln Asn Val Ala Asn Tyr Leu Ile Val Ser Leu Ala Val Ala Asp Leu
 180 185 190

Met Val Ala Cys Leu Val Met Pro Leu Gly Ala Val Tyr Glu Ile Ser
 195 200 205

Arg Gly Trp Ile Leu Gly Pro Glu Leu Cys Asp Met Trp Thr Ser Ser
 210 215 220

Asp Val Leu Cys Cys Thr Ala Ser Ile Leu His Leu Val Ala Ile Ala
 225 230 235 240

Val Asp Arg Tyr Trp Ala Val Thr Asn Val Asp Tyr Ile His Thr Arg
 245 250 255

Asn Gly Arg Arg Ile Val Gly Met Ile Ile Val Val Trp Phe Val Ala
 260 265 270

Leu Val Val Ser Leu Ala Pro Gln Phe Gly Trp Lys Asp Pro Asp Tyr
 275 280 285

Leu Asp Arg Ile Asn Ile Gln Gln Arg Cys Met Val Ser Gln Asp Ile
 Page 3

FMC-ADV-010.ST25.txt
300

290

295

Gly Tyr Gln Ile Phe Ala Thr Cys Ser Thr Phe Tyr Leu Pro Leu Leu
305 310 315 320

Val Ile Leu Phe Leu Tyr Trp Lys Ile Phe Lys Ile Ala Arg Arg Arg
325 330 335

Ile Arg Arg Arg Arg Ala Gln Arg Asn Ala Met Ile Glu His Ser Arg
340 345 350

Ser Lys Lys Pro Ala Asp Glu Ala Ala Ala Ser Ala Glu Pro Lys Lys
355 360 365

Cys Gly Phe Phe Arg Lys Ala Val Phe Leu Arg Ile Lys Lys Lys Lys
370 375 380

Glu Val Glu Glu Thr Ala Val Cys Ser Ser Ile Gly Leu Ile Asp Gly
385 390 395 400

His Ser Ser Cys Ser Met Pro Glu Ser Ser Gln Asp Gly Gly Gly Gly
405 410 415

Gly Gly Gly Gly Gln Ser Lys Ala Gly Asn Gly Ser Gly Ser Gly Gly
420 425 430

Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Val Gly Gly Ser Gly Gly
435 440 445

Glu Asp Lys Thr Thr Ala Phe Thr Ile Thr Thr Asn His Asp Leu Gly
450 455 460

Ala Thr Ser Thr Leu Gln Cys His Gln Gln Gln Gln His His His His
465 470 475 480

His His Gln His His Gln His Gln Asn Gln Asn Gln Leu Gln Pro Pro
485 490 495

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

His Thr Val Ile Glu Met Asn Val Asn Arg Pro Cys Ser Val Pro Ala
515 520 525

Pro Val Arg Pro Lys Thr Lys Arg Glu Lys Lys Glu Ser Leu Glu Ala
530 535 540

FMC-ADV-010.ST25.txt

Lys Arg Glu Arg Lys Ala Ala Lys Thr Leu Ala Ile Ile Thr Gly Ala
 545 550 555 560

Phe Val Val Cys Trp Leu Pro Phe Phe Ile Met Ala Leu Val Met Pro
 565 570 575

Leu Cys Gln Gly Cys Val Ile Asn Glu Tyr Val Ser Ser Phe Phe Leu
 580 585 590

Trp Leu Gly Tyr Phe Asn Ser Thr Leu Asn Pro Ile Ile Tyr Thr Val
 595 600 605

Phe Ser Pro Glu Phe Arg Leu Ala Phe Thr Lys Met Ile Cys Gly Thr
 610 615 620

Thr Tyr Arg Arg
 625

<210> 3
 <211> 1452
 <212> DNA
 <213> APHIS GOSSYPPII

<400> 3
 tgacgaccat ttgtggtaac gcgttggtga tagcagcaat tttggtggag aggaatttac 60
 aaaacgtggc caactacctc attgtctcgc tagcggtcgc cgatttaatg gtagcatggt 120
 tagtcatgcc cttaggagct gtttatgaga taagcagagg atggatatta ggtccagaac 180
 tatgtgacat gtggacttca agtgacgtac tatgttgtag agcttccata cttcacttag 240
 tcgcaatagc tggtgataga tattgggctg ttactaatgt tgattatata cacacgagaa 300
 acggtagaag aatagtaggc atgattattg tgggtgtggt tgtcgacta gtcgtgtcat 360
 tagctccaca gtttggtg aaagaccg attacctaga tcgtatcaat attcaacaaa 420
 gatgtatggt cagtcaagac attggctatc agatatttgc cacatgttca acgttttact 480
 tgcctctact ggtcatcctg tttctttatt ggaagatatt caagatagcc aggagaagga 540
 ttcgcagaag aagggtcag agaaacgcta tgatagaaca ttcaagggtc aaaaagccag 600
 cggacgaagc ggccgcatcg gctgagccca aaaaatgcgg tttcttccgg aaggccgttt 660
 tcctgcggat caagaaaaag aaggaagtgg aagagacggc cgtgtgctcg agtataggtc 720
 tgatcgacgg tcattcgagc tgctcgatgc ccgagtcac gcaagacggc ggtggagggg 780
 gtggcgcgcg ccagtcgaag gcgggcaacg gcagcggaag cggtggcgga agcgggtggtg 840
 gaagcgggtg cggaagtgtg ggtggaagcg gtggcgaaga caagaccacg gcgttcacga 900
 tcactacca tcacgacctc ggtgctacgt cgacgtcca gtgccatcag cagcagcaac 960
 accaccacca ccaccaacac caccaacacc aaaaccagaa tcagctccag cctccggtga 1020
 Page 5

FMC-ADV-010.ST25.txt

```

tcgaggacga cggacccaca atcacggatc cgcccgttcg gccgaagacg aagcgcgaga 1080
agaaagagtc gctggaggcg aaacgcgaac ggaaagcggc caagacgttg gccatcatca 1140
ccggggcctt tgtcgtgtgc tggctgccgt tcttcatcat ggctctggtc atgcccctgt 1200
gccaggggttg cgtgatcaac gagtacgtat ccagcttctt cctgtgggttg ggttacttca 1260
attccaccct gaacccgatc atctacaccg ttttcagtcc cgagttccgg ttggcattca 1320
ccaaaatgat ttgcggcact acgtatagga gataatgaaa tgcgcgcaca ttattattat 1380
tattattatt attattatta cataaacaat acatataata ttattatatt aaaaaaaaaa 1440
aaaaaaaaaa aa 1452

```

```

<210> 4
<211> 628
<212> PRT
<213> APHIS GOSSYPPII

```

```

<400> 4

```

```

Met Ala Leu Leu Gly Ala Leu Ala Ala Thr Gly Ser Arg Arg Gln Gly
1          5          10          15

```

```

Cys Val Leu Arg Ala Ala Ala Asp Ala Ile Thr Phe Tyr Cys Gly Gly
20          25          30

```

```

Asp Val Val Ser Gly Gly Gly Ser Ser Gly Gly Gly Gly Gly Gly Gly
35          40          45

```

```

Ser Gly Gly Asp Gly Ala Ala Ala Glu Tyr Asp Gly Ala Ala Ala Ala
50          55          60

```

```

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

```

```

Arg Arg Pro Val Arg Ser Val Ala Ile Gly Asp Gly Gly Ser Gly Ala
85          90          95

```

```

Gly Gly Tyr Tyr Tyr Gly Thr Glu Asn Gly Ser Ser Ala Ala Ala Val
100          105          110

```

```

Leu Ala Asn Ala Thr Ala Gly Val Glu Pro Phe Leu Asn Gly Ser Tyr
115          120          125

```

```

Asn His Thr Leu Gly Ala Glu Val Glu Asp Leu Thr Asn Tyr Ile Phe
130          135          140

```

```

Val Ile Thr Ile Gly Ile Ile Leu Gly Ala Met Ile Leu Thr Thr Ile
145          150          155          160

```

FMC-ADV-010.ST25.txt

Cys Gly Asn Ala Leu Val Ile Ala Ala Ile Leu Leu Glu Arg Asn Leu
 165 170 175
 Gln Asn Val Ala Asn Tyr Leu Ile Val Ser Leu Ala Val Ala Asp Leu
 180 185 190
 Met Val Ala Cys Leu Val Met Pro Leu Gly Ala Val Tyr Glu Ile Ser
 195 200 205
 Arg Gly Trp Ile Leu Gly Pro Glu Leu Cys Asp Met Trp Thr Ser Ser
 210 215 220
 Asp Val Leu Cys Cys Thr Ala Ser Ile Leu His Leu Val Ala Ile Ala
 225 230 235 240
 Val Asp Arg Tyr Trp Ala Val Thr Asn Val Asp Tyr Ile His Thr Arg
 245 250 255
 Asn Gly Arg Arg Ile Val Gly Met Ile Ile Val Val Trp Phe Val Ala
 260 265 270
 Leu Val Val Ser Leu Ala Pro Gln Phe Gly Trp Lys Asp Pro Asp Tyr
 275 280 285
 Leu Asp Arg Ile Asn Ile Gln Gln Arg Cys Met Val Ser Gln Asp Ile
 290 295 300
 Gly Tyr Gln Ile Phe Ala Thr Cys Ser Thr Phe Tyr Leu Pro Leu Leu
 305 310 315 320
 Val Ile Leu Phe Leu Tyr Trp Lys Ile Phe Lys Ile Ala Arg Arg Arg
 325 330 335
 Ile Arg Arg Arg Arg Ala Gln Arg Asn Ala Met Ile Glu His Ser Arg
 340 345 350
 Ser Lys Lys Pro Ala Asp Glu Ala Ala Ala Ser Ala Glu Pro Lys Lys
 355 360 365
 Cys Gly Phe Phe Arg Lys Ala Val Phe Leu Arg Ile Lys Lys Lys Lys
 370 375 380
 Glu Val Glu Glu Thr Ala Val Cys Ser Ser Ile Gly Leu Ile Asp Gly
 385 390 395 400
 His Ser Ser Cys Ser Met Pro Glu Ser Ser Gln Asp Gly Gly Gly Gly
 Page 7

FMC-ADV-010.ST25.txt
410

405

415

Gly Gly Gly Gly Gln Ser Lys Ala Gly Asn Gly Ser Gly Ser Gly Gly
420 425 430Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Val Gly Gly Ser Gly Gly
435 440 445Glu Asp Lys Thr Thr Ala Phe Thr Ile Thr Thr Asn His Asp Leu Gly
450 455 460Ala Thr Ser Thr Leu Gln Cys His Gln Gln Gln Gln His His His His
465 470 475 480His His Gln His His Gln His Gln Asn Gln Asn Gln Leu Gln Pro Pro
485 490 495Val Ile Glu Asp Asp Gly Pro Thr Ile Thr Asp Pro Ser Ala Ser Arg
500 505 510His Thr Val Ile Glu Met Asn Val Asn Arg Pro Cys Ser Val Pro Ala
515 520 525Pro Val Arg Pro Lys Thr Lys Arg Glu Lys Lys Glu Ser Leu Glu Ala
530 535 540Lys Arg Glu Arg Lys Ala Ala Lys Thr Leu Ala Ile Ile Thr Gly Ala
545 550 555 560Phe Val Val Cys Trp Leu Pro Phe Phe Ile Met Ala Leu Val Met Pro
565 570 575Leu Cys Gln Gly Cys Val Ile Asn Glu Tyr Val Ser Ser Phe Phe Leu
580 585 590Trp Leu Gly Tyr Phe Asn Ser Thr Leu Asn Pro Ile Ile Tyr Thr Val
595 600 605Phe Ser Pro Glu Phe Arg Leu Ala Phe Thr Lys Met Ile Cys Gly Thr
610 615 620Thr Tyr Arg Arg
625

<210> 5
 <211> 31
 <212> DNA
 <213> APHIS GOSSYPPII

FMC-ADV-010.ST25.txt

<400> 5
aaacactgga agccaaaaga garmgraarg c 31

<210> 6
<211> 34
<212> DNA
<213> APHIS GOSSYPPII

<400> 6
aaataactgg gttcagcgta gagttraart adcc 34

<210> 7
<211> 27
<212> DNA
<213> APHIS GOSSYPPII

<400> 7
cgggtgccttc gtgatctgyt ggytrcc 27

<210> 8
<211> 31
<212> DNA
<213> APHIS GOSSYPPII

<400> 8
tggggttcagc gtagagttga artadccyar c 31

<210> 9
<211> 28
<212> DNA
<213> APHIS GOSSYPPII

<220>
<221> misc_feature
<222> (23)..(23)
<223> n is a, c, g, or t

<400> 9
acctgatggt ggcctgcytr gtnatgcc 28

<210> 10
<211> 20
<212> DNA
<213> APHIS GOSSYPPII

<400> 10
ttgggttcag cgtagagttg 20

<210> 11
<211> 20
<212> DNA
<213> APHIS GOSSYPPII

<400> 11
gtcagtcgaag acattggcta 20

FMC-ADV-010.ST25.txt

<210> 12
<211> 17
<212> DNA
<213> APHIS GOSSYPPII

<400> 12
ggcgttcacg atcacta 17

<210> 13
<211> 22
<212> DNA
<213> APHIS GOSSYPPII

<400> 13
caggccacgc gtcgactagt ac 22

<210> 14
<211> 19
<212> DNA
<213> APHIS GOSSYPPII

<400> 14
cggtcgccga tttaatggt 19

<210> 15
<211> 20
<212> DNA
<213> APHIS GOSSYPPII

<400> 15
tggtgtggtt tgtcgacta 20

<210> 16
<211> 20
<212> DNA
<213> APHIS GOSSYPPII

<400> 16
atgaccgtcg atcagaccta 20

<210> 17
<211> 19
<212> DNA
<213> APHIS GOSSYPPII

<400> 17
atcgtgaacg ccgtggtct 19