

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
24 February 2005 (24.02.2005)

PCT

(10) International Publication Number
WO 2005/017103 A2

(51) International Patent Classification⁷: **C12N**
(21) International Application Number:
PCT/US2004/017406
(22) International Filing Date: 2 June 2004 (02.06.2004)
(25) Filing Language: English
(26) Publication Language: English
(30) Priority Data:
60/475,111 2 June 2003 (02.06.2003) US

(74) Agents: **GOLDMAN, Michael, L.** et al.; Nixon Peabody, LLP, Clinton Square, P.O. Box 31051, Rochester, NY 14603-1051 (US).

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

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, 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) Applicants (*for all designated States except US*): **CORNELL RESEARCH FOUNDATION, INC.** [US/US]; Cornell Business & Technology Park, 20 Thornwood Drive, Suite 105, Ithaca, NY 14850 (US). **THE UNITED STATES OF AMERICA AS REPRESENTED BY THE SECRETARY OF AGRICULTURE** [US/US]; 1815 N. University Street, Peoria, IL 61604 (US).

(72) Inventors: **LORIA, Rosemary**; 111 Northway Road, Ithaca, NY 14850 (US). **CRANE, Brian, R.**; 105 McIntyre Place, Ithaca, NY 14850 (US). **KERS, Johan**; 2250 North Triphammer Road, C1, Ithaca, NY 14850 (US). **WATCH, Michael, J.**; 316 Turkey Hill Road, Ithaca, NY 14850 (US). **GIBSON, Donna, M.**; 2021 Ellis Hollow Road, Ithaca, NY 14850 (US).

Published:

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

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

(54) Title: BACTERIAL NITRIC OXIDE SYNTHASES AND USES THEREOF

(57) Abstract: The present invention relates to isolated nucleic acid molecules encoding nitric oxide synthases. The isolated nucleic acid molecules and their encoded protein or polypeptides are useful in methods for attaching a nitrogen group to a target moiety of a compound and for synthesizing a nitrogen-modified compound in a transgenic host cell. The present invention also relates to expression systems and host cells containing the nucleic acids of the present invention, as well as a method of recombinantly producing the nitric oxide synthases of the present invention.



WO 2005/017103 A2

- 1 -

BACTERIAL NITRIC OXIDE SYNTHASES AND USES THEREOF

[0001] This application claims the benefit of U.S. Provisional Patent Application Serial No. 60/475,111, filed June 2, 2003.

[0002] The subject matter of this application was made with support from the United States Government under USDA Grant No. 99-35303-8084. The U.S. Government may have certain rights.

FIELD OF THE INVENTION

[0003] The present invention relates to isolated nucleic acid molecules encoding nitric oxide synthases, the encoded isolated nitric oxide synthases, and the uses of these isolated nucleic acids and nitric oxide synthases for catalyzing nitration and nitrosylation reactions.

BACKGROUND OF THE INVENTION

[0004] Nitric oxide synthases ("NOSs") are highly regulated enzymes that synthesize the potent cytotoxin and signal molecule nitric oxide (NO) from L-arginine (L-arg) (Alderton et al., *Biochem J.* 357:593-615 (2001)). In mammals, NOSs are responsible for many functions that range from protection against pathogens and tumor cells to blood pressure regulation and nerve transmission. The production of NO in mammals is catalyzed solely by three highly regulated isozymes of NOS (Alderton et al., *Biochem J.* 357:593-615 (2001); Pfeiffer et al., *Angew. Chem. Int. Ed.* 38:1714-1731 (1999); and Stuehr, *Biochim. Biophys. Acta* 1411:217-230 (1999)). In particular, NOSs produce NO from oxidation of L-arginine to L-citrulline via the intermediate N-hydroxy-L-Arg (NHA) (Alderton et al., *Biochem J.* 357:593-615 (2001); Pfeiffer et al., *Angew. Chem. Int. Ed.* 38:1714-1731 (1999); and Stuehr, *Biochim. Biophys. Acta* 1411:217-230 (1999)). Mammalian NOSs are homodimers that contain an N-terminal heme oxygenase domain (NOS_{oxy}) and a C-terminal flavoprotein reductase domain (NOS_{red}). The oxygenase domain binds L-Arg, heme and the redox-active cofactor 6R-tetrahydrobiopterin (H₄B), whereas the reductase domain binds FAD, FMN and

- 2 -

NADPH. A calmodulin (CaM) binding sequence links the oxygenase and the reductase domains and regulates reduction of NOS_{oxy} by NOS_{red} in those isoforms that respond to Ca²⁺.

[0005] Genome sequencing has revealed truncated NOS proteins in some Gram-positive bacteria, including *Deinococcus radiodurans*, *Staphylococcus aureus*, *Bacillus subtilis*, *B. halodurans* and *B. anthracis*. In general, bacterial NOSs are homologous to the mammalian NOS_{oxy} but lack an associated NOS_{red} and N-terminal regions that bind Zn²⁺, the dihydroxypropyl side chain of H₄B, and the adjacent subunit of the dimer (Pant et al., *Biochemistry* 41:11071-11079 (2002) and Adak et al., *J. Biol. Chem.* 277:16167-16171 (2002)). Nevertheless, the *D. radiodurans* NOS (deiNOS) and *B. subtilis* NOS (bsNOS) are dimeric, have a heme liganded by cysteine thiol, bind substrate L-Arg, and produce nitrogen oxide species in a manner dependent on pterin (either with H₄B or with the related cofactor tetrahydrofolate (THF) (Adak et al., *Proc. Natl. Acad. Sci.* 99:107-12 (2002); Pant et al., *Biochemistry* 41:11071-11079 (2002); and Adak et al., *J. Biol. Chem.* 277:16167-16171 (2002)). The crystal structure of *B. subtilis* NOS, complexed with L-Arg, confirmed that bacterial NOS proteins are similar to mammalian NOSs and NO production has been demonstrated when a mammalian reductase domain is supplied (Adak et al., *J. Biol. Chem.* 277:16167-16171 (2002)). The redox mechanism by which the pterins H₄B and analog THF support NO synthesis in bacterial NOSs mirrors that of the mammalian enzymes (Adak et al., *J. Biol. Chem.* 277:16167-16171 (2002)). A reductase protein that supplies electrons for bacterial NOS catalysis has not yet been identified, although comparisons of bacterial and mammalian NOS structures suggest a common mode of interaction for such a redox partner (Adak et al., *J. Biol. Chem.* 277:16167-16171 (2002)).

[0006] Nitrated and nitrosylated natural products, although relatively rare, represent important herbicides, antibiotics, nematicides, fungicides, insecticides, and anti-cancer agents. Nitrated compounds are also key components of explosives, propellants. Despite the importance of nitration and nitrosylation processes for the chemical industry, conventional nitration reactions have drawbacks that include low specificity, low yields, difficult temperature control, difficult product workups, and the generation of waste acids (Pagoria et al.,

1 *Thermochim. Acta* 384:187-204 (2002); and Agrawal, *Prog. Energy Combust. Sci.*
24:1-30(1998)). For pyrotechnics, elevated reaction temperatures limit production
of unstable products. Regarding pharmaceuticals, greater control of reaction
selectivity is desired (Eaton et al., *J. Med. Chem.* 16:290-291 (1973); and Hazra et
5 al., *Org. Prep. Proced. Int.* 31:315-319 (1999)). Although nitrated and
nitrosylated compounds have vast importance in the pharmaceutical and high-
energy materials industries, the chemical processes responsible for the nitration
and nitrosylation processes have liabilities in efficacy, cost, and environmental
impact. Thus, there is a need for alternative enzymes and methods for conducting
10 nitration/nitrosylation reactions.

[0007] The present invention is directed to overcoming these and other
deficiencies in the art.

SUMMARY OF THE INVENTION

15 [0008] The present invention relates to an isolated nucleic acid molecule
encoding a nitric oxide synthase. The isolated nucleic acid molecule can (i)
include a nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, and/or SEQ ID
NO:5; (ii) include a nucleotide sequence that hybridizes to SEQ ID NO:1, SEQ ID
NO:3, or SEQ ID NO:5 under stringent conditions characterized by a
20 hybridization medium comprising about 5X SSC at a temperature of about 55°C;
(iii) include a nucleotide sequence having greater than 95 percent homology to a
nucleic acid according to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5; (iv)
include a nucleotide sequence that encodes a nitric oxide synthase protein having
an amino acid sequence that is at least 85 percent similar to either SEQ ID NO:2,
25 SEQ ID NO:4, or SEQ ID NO:6 by basic BLAST using default parameters
analysis; (v) encode a nitric oxide synthase comprising a protein or polypeptide
having an amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, and/or SEQ ID
NO:6; (vi) encode a nitric oxide synthase comprising a protein or polypeptide
having an amino acid motif of SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ
30 ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ
ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ
ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ

- 4 -

- ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, and/or combinations thereof; or (vii) encode a nitric oxide synthase protein or polypeptide having an amino acid sequence of SEQ ID NO:47. Expression vectors (e.g., expression systems) and host cells which include the nucleic acid molecules of the present invention are also disclosed. Isolated NOS proteins or polypeptides encoded by the isolated nucleic acid molecules of the present invention are also disclosed. The present invention also relates to a method of recombinantly producing in a host cell the nitric oxide synthases encoded by the nucleic acid molecules of the present invention.
- 10 **[0009]** The present invention also relates to a method of attaching a nitrogen group to a target moiety of a compound. This method involves providing a nitric oxide synthase and a compound having a target moiety. The nitric oxide synthase and the compound are combined in a reaction mixture under conditions effective to allow the nitric oxide synthase to catalyze a reaction whereby a
- 15 nitrogen group contained in the reaction mixture is attached to the target moiety. The method thereby yields a nitrogen-modified compound. Suitable nitric oxide synthases for use in this method can include: (i) a protein or polypeptide having an amino acid sequence that is at least 85 percent similar to either SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 by basic BLAST using default parameters
- 20 analysis; (ii) a protein or polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6; (iii) a protein or polypeptide having an amino acid motif such as the motif of SEQ ID NO:7, the motif of SEQ ID NO:8, the motif of SEQ ID NO:9, the motif of SEQ ID NO:10, the motif of SEQ ID NO:11, the motif of SEQ ID NO:12, the
- 25 motif of SEQ ID NO:13, the motif of SEQ ID NO:14, the motif of SEQ ID NO:15, the motif of SEQ ID NO:16, the motif of SEQ ID NO:17, the motif of SEQ ID NO:18, the motif of SEQ ID NO:19, the motif of SEQ ID NO:20, the motif of SEQ ID NO:21, the motif of SEQ ID NO:22, the motif of SEQ ID
- 30 SEQ ID NO:26, the motif of SEQ ID NO:27, the motif of SEQ ID NO:28, the motif of SEQ ID NO:29, the motif of SEQ ID NO:30, the motif of SEQ ID NO:31, the motif of SEQ ID NO:32, the motif of SEQ ID NO:33, the motif of SEQ ID NO:34, the motif of SEQ ID NO:35, the motif of SEQ ID NO:36, the

motif of SEQ ID NO:37, the motif of SEQ ID NO:38, the motif of SEQ ID NO:39, the motif of SEQ ID NO:40, the motif of SEQ ID NO:41, the motif of SEQ ID NO:42, the motif of SEQ ID NO:43, the motif of SEQ ID NO:44, the motif of SEQ ID NO:45, the motif of SEQ ID NO:46, and/or combinations thereof; or (iv) a protein or polypeptide having an amino acid sequence of SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, or SEQ ID NO:50.

[0010] The present invention further relates to a method of synthesizing a nitrogen-modified compound in a transgenic host cell. This method involves providing a transgenic host cell transformed with a DNA molecule encoding a nitric oxide synthase. This method also involves providing a non-modified compound containing a target moiety. The transgenic host cell and the non-modified compound are cultured in a culture medium under conditions effective to allow the nitric oxide synthase to be expressed and to catalyze a reaction whereby a nitrogen group contained in the host cell or in a culture medium attaches to the target moiety. The method thereby yields a nitrogen-modified compound.

Suitable DNA molecules for use in this method can include those that: (i) have a nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5; (ii) have a nucleotide sequence that hybridizes to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5 under stringent conditions characterized by a hybridization medium comprising about 5X SSC at a temperature of about 55°C; (iii) have a nucleotide sequence having greater than 95 percent homology to a nucleic acid according to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5; (iv) have a nucleotide sequence that encodes a nitric oxide synthase protein having an amino acid sequence that is at least 85 percent similar to either SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 by basic BLAST using default parameters analysis; (v) encode a nitric oxide synthase having a protein or polypeptide having an amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6; (vi) encode a nitric oxide synthase having a protein or polypeptide having amino acid motifs such as the motif of SEQ ID NO:7, the motif of SEQ ID NO:8, the motif of SEQ ID NO:9, the motif of SEQ ID NO:10, the motif of SEQ ID NO:11, the motif of SEQ ID NO:12, the motif of SEQ ID NO:13, the motif of SEQ ID NO:14, the motif of SEQ ID NO:15, the motif of SEQ ID NO:16, the motif of SEQ ID NO:17, the motif of SEQ ID NO:18, the motif of SEQ ID NO:19, the motif of SEQ ID NO:20, the

- 6 -

motif of SEQ ID NO:21, the motif of SEQ ID NO:22, the motif of SEQ ID NO:23, the motif of SEQ ID NO:24, the motif of SEQ ID NO:25, the motif of SEQ ID NO:26, the motif of SEQ ID NO:27, the motif of SEQ ID NO:28, the motif of SEQ ID NO:29, the motif of SEQ ID NO:30, the motif of SEQ ID NO:31, the motif of SEQ ID NO:32, the motif of SEQ ID NO:33, the motif of SEQ ID NO:34, the motif of SEQ ID NO:35, the motif of SEQ ID NO:36, the motif of SEQ ID NO:37, the motif of SEQ ID NO:38, the motif of SEQ ID NO:39, the motif of SEQ ID NO:40, the motif of SEQ ID NO:41, the motif of SEQ ID NO:42, the motif of SEQ ID NO:43, the motif of SEQ ID NO:44, the motif of SEQ ID NO:45, the motif of SEQ ID NO:46, and/or combinations thereof; or (vii) encode a nitric oxide synthase having a protein or polypeptide having an amino acid sequence of SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, or SEQ ID NO:50.

[0011] The nitric oxide synthases of the present invention are useful for catalyzing specific nitration and/or nitrosylation reactions that can be used in developing more efficient and safe methods of producing nitrated and/or nitrosylated compounds. The present invention could also be useful for bioengineering new bioactive products by incorporation of nitration and nitrosylation events into existing biosynthetic pathways. The nitric oxide synthases of the present invention can also be used to broaden substrate specificity and fine-tune selectivity for the introduction of nitroso and nitro functionalities.

[0012] An additional advantage of enzymatically controlled nitration (as provided by the present invention) is the potential for discovering entirely new compounds through bioengineering. For example, genetic incorporation of bacterial NOS into polyketide or non-ribosomal peptide biosynthetic pathways (Kleinkauf et al., *European Journal of Biochemistry* 236:335-351 (1996); and Watanabe et al., *J. Biol. Chem.* 278:42020-42026 (2003), which are hereby incorporated by reference in their entirety) will allow generation of novel natural products. Screened with suitable bioassays, nitrated compounds could produce novel leads for drug design.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] Figures 1A-1B show the chemical structure of thaxtomin A (Figure 1A) and the genetic organization of the *nos* region in the thaxtomin-producing species *Streptomyces turgidiscabies* (Figure 1B). Thaxtomin A is the
5 predominant thaxtomin congener produced by plant pathogenic *Streptomyces* spp. As shown in Figure 1B, the *txtAB* genes of *Streptomyces turgidiscabies* encode two similar peptide synthetases required for synthesis of the dipeptide (Healy et al., *Mol. Micro.* 38:794-804 (2000), which is hereby incorporated by reference in its entirety). The *txtC* gene encodes a P450 monooxygenase that is required for
10 post-cyclization hydroxylation of the dipeptide. The *nos* gene is upstream of characterized thaxtomin biosynthetic genes.

[0014] Figure 2 shows the amino acid sequence alignment of NOSs from *Streptomyces turgidiscabies* (stNOS), *Bacillus subtilis* (bsNOS), and murine iNOS (miNOS). Figure 2 demonstrates that stNOS is homologous to NOS proteins from
15 *Bacillus subtilis* (bsNOS) and the oxygenase domain of murine iNOS (miNOS). Residues involved in binding substrate L-Arg (bolded letters), pterin cofactor (underlined letters), heme iron and zinc cations (grey letters) are nearly completely conserved among these three classes of NOS_{oxy} proteins. Unlike other bacterial NOSs, stNOS contains an N-terminal extension involved in zinc ion and
20 pterin side-chain binding in miNOS (secondary structure elements and structural motifs for bsNOS shown above in grey, see Pant et al., *Biochemistry* 41:11071-11079 (2002), which is hereby incorporated by reference in its entirety, for nomenclature). The most C-terminal pterin binding residues (stNOS 385 and 386) vary among the proteins because interactions of these residues with pterin are
25 main-chain mediated. Boxed residues indicate positions of different residue character between mammalian and bacterial NOSs. These changes primarily map to the pterin binding site and a site likely involved in reductase protein interactions. The change from Val to Ile at stNOS residue 259 is conserved in all bacterial proteins and increases protection of the immediate distal heme pocket
30 compared to the mammalian proteins.

[0015] Figures 3A-3C show that deletion of *nos* reduces thaxtomin production by *S. turgidiscabies*. Figure 3A shows a model for *nos* replacement in

Streptomyces turgidiscabies car8 by the p2XNOS targeting vector using marker-exchange mutagenesis. Figure 3B shows Southern hybridization of an internal 673 bp α -³²P labeled *nos* probe (left) or the α -³²P labeled 4.7 kb HindIII/EcoRI p2XNOS insert (right) to EcoRI digests of *S. turgidiscabies* wild-type (lane 1), *Δnos* (lane 2) *Δnos* pIJ8600 (empty vector) (lane 3), *Δnos* pIJ8600NOS (complemented vector)(lane 4) genomic DNA. The thiostrepton resistance gene (*tsr*) in the p2XNOS insert probe also hybridizes to *tsr* on pIJ8600 (lanes 3, 4). Figure 3C shows confirmation of *nos* expression using RT-PCR. Dnase treated RNA, odd-numbered lanes, cDNA, even-numbered lanes. Wild-type *S. turgidiscabies*, lanes 1, 2; *Δnos*, lanes 3, 4; *Δnos* pIJ8600, lanes 5, 6; *Δnos* pIJ8600NOS, lanes 7, 8.

[0016] Figures 4A-4C shows that nitrite formation from N^ω-hydroxy-L-arginine by recombinant stNOS and sensitivity to inhibitors of mammalian NOSs. Figure 4A shows *Streptomyces turgidiscabies* NOS Δ 41 (residues 41-to carboxy-terminus) cloned into pet28 (Novagene) and expressed in *E. coli* BL-21 cells (lane 3). Full-length stNOS was rapidly degraded in *E. coli*, but the Δ 41 construct, which contains all NOS homology regions, produced high levels of protein that were comparable to levels of recombinant *Deinococcus radiodurans* NOS¹⁰ (lane 2) and significantly above levels of background proteins expressed from cells containing the pet28 vector with no NOS insert (lane 1). Sonicated cell lysates from 3 mL cultures were run on 12% SDS-PAGE gels after 2.5 hr induction with 0.1 mM IPTG. Arrow shows the positions of recombinant deiNOS, and stNOS in lanes 2 and 3. Figure 4B shows nitrite production by cell lysates depicted in Figure 4A measured by the Griess assay. The NOS intermediate N^ω-hydroxy-L-arginine (2 mM) and hydrogen peroxide (4 mM) were reacted with 20 μ L of cell lysates shown in Figure 4A and assayed for nitrite with the Griess reagents after 15 min. Lysates from cells overexpressing stNOS generate nitrite at levels comparable to that of overexpressed deiNOS and significantly above that of the empty vector control (pet28). The selective mammalian NOS inhibitor, nitro-L-arginine (active form of nitro-L-arginine methyl ester), inhibits nitrite formation by lysates expressing stNOS when added to a concentration of 1mM (same reaction conditions as above). Figure 4C shows the L-arginine based NOS inhibitors nitro-L-arginine methyl ester (NAME) (32 μ M), aminoguanidine (AG)

(270 μ M), NG-monomethyl-L-arginine (NMMA) (510 μ M), and 7-nitroindazole (7-IN) (1.2 μ M) each inhibit toxin product without affecting bacterial growth (Standard error reported, n=6).

[0017] Figures 5A-5B shows that the nitrate nitrogen derives from the
5 guanidine nitrogen of L-Arg. Nuclear magnetic resonance analysis of thaxtomin
A extracted from cultures fed either (Figure 5A) L-Arginine-guanido- $^{15}\text{N}_2\cdot\text{HCl}$
(Cambridge Isotope Laboratories), or (Figure 5B) $^{15}\text{NH}_4\text{NO}_3$ (Aldrich). Based on
literature values for ^{15}N atoms in electronic environments similar to those of the
four nitrogen atoms in thaxtomin A (Patel et al., *Biochemica. Biophys. Acta.*
10 1411:385-400 (1999), which is hereby incorporated by reference in its entirety),
the two most upfield peaks at 125 and 114.6 ppm can be assigned to the N-
methyls, the broad peak at 132.3 ppm can be assigned to the indole nitrogen, and
the downfield peak at 374.8 ppm is diagnostic for the NO_2 -nitrogen. Accordingly,
the single peak (see Figure 5A) in the spectrum of the L-Arg-guanidino- $^{15}\text{N}_2\cdot\text{HCl}$
15 supplied sample, at 375.0 ppm represents the signal of the NO_2 -nitrogen alone.

[0018] Figures 6A-6B show alignments of the amino acid sequences of the
nitric oxide synthases from the following organisms (with the ATCC Accession
Numbers representing the GenBank for the *nos* sequences): *Bacillus anthracis*
(ATCC Accession No. AAP29327); *Bacillus cereus* (ATCC Accession No.
20 AAP12306); *Bacillus subtilis* (ATCC Accession No. CAB12592); *Bacillus*
halodurans (ATCC Accession No. BAB04542); *Staphylococcus aureus* (ATCC
Accession No. BAB95720); *Staphylococcus epidermidis* (ATCC Accession No.
AE016749); *Deinococcus radiodurans* (ATCC Accession No. AE002088); *Mus*
musculus (ATCC Accession No. AAC52356); *Streptomyces acidiscabies* (ATCC
25 Accession No. AY204508); *Streptomyces scabies* (ATCC Accession No.
AY204507); *Streptomyces turgidiscabies* (ATCC Accession No. AY204509); and
Streptomyces avermitilis (ATCC Accession No. BAC69241).

DETAILED DESCRIPTION OF THE INVENTION

[0019] The present invention relates to isolated nucleic acid molecules encoding nitric oxide synthases ("NOSs"). The nitric oxide synthases of the present invention share a common function, in that they function as catalysts for
 5 nitration and/or nitrosylation reactions. Also disclosed are expression systems and host cells containing such nucleic acid molecules, as well as isolated proteins or polypeptides encoded by the nucleic acid molecules. Uses of the nucleic acid molecules and the proteins or polypeptides encoded by the nucleic acid molecules are disclosed.

10 [0020] The isolated nucleic acid molecules and their encoded proteins or polypeptides can be, without limitation, from various *Streptomyces* species. Suitable *Streptomyces* species include, for example, *Streptomyces acidiscabies*, *Streptomyces scabies*, *Streptomyces turgidiscabies*, *Streptomyces avermitilis*, and *Streptomyces ipomoea*.

15 [0021] A first isolated nucleic acid molecule of the present invention encodes a nitric oxide synthase ("NOS") from *Streptomyces acidiscabies* (see GenBank Accession No. AY204508) and has a nucleotide sequence according to SEQ ID NO:1 as follows:

20 gtgacctccgaagtcgctctgggcccttctcttgcccgcgccggtccccgacagcgtgcccggcactg
 gggcccgattcgtcccttggcccggtcccgctcggcggaaccggcgacgcgcagtcctgcccgcgtc
 gccgatccaaatgaggctgaggagttcctgcccagttccacgcggagcagtcctgatcagcccgtc
 ccgctcggccggcgccctggagcaggtccgcgcgcccatcgacgccacgggcacctaccggcacacc
 accgccgagctcgtgtacggtgcccgcgtcgcgtggcgcaactccagtcgctgcatcgccgcctg
 25 tactggaacagcctgcgcgtcctggaccgcgggacgccacagccccgatgagatccaccggcac
 ttgtgcacgcacctgcgccaggcgaccaacggcgggcgcatcaggccggtgatttcggtcttcgcc
 ccgactcccccgccggcccgcccgccgcaggtgtggaacgagcagctcatccggtacgcgggtac
 cgccgcgacgacggcacccgtgctcggtgacccgcgcaccgcgcacctaccgaggccatcctccgc
 ctcggtggcagggctgcccccaagggcggttcgacgtcctgcccctggtcatcgacacccccgac
 30 gacaaaccccggttcttcgagctgcgcggggagctggtcttgagggtccctatcacccaccccgac
 gtcccacgcctggccgaactgggcctgcgctggcacgcgcgtaccgctcatctccaacatgcgccta
 cgcacgcggcgatggactaccgcctcgcgccgttcaacggctggtacatgggcacggagatcggc
 gcccgcaacctcgtcgacgaggaccgctacaacatgctcccgcgcgtcgccgcctgctccagctg
 gacaccaccagcgagtcacacctgtggcgcgaccgcgccttggtcgagctcaacgtcgccgtcctg
 35 cactccttcgaggccgcaggtgtccggatcagcgaccaccacgaggagtcgccggcgttcctcgcc
 cacctggccaaggaggaacgccagggccgcaccgtatccgcagactggagctggtatcgctcccccg
 ctctccggcgcatcaccccggtgttcaccggttactacgacaacgtcgaccagcgcccaacttc
 taccaccaccagtga

The NOS protein or polypeptide encoded by this nucleic acid molecule has an amino acid sequence according to SEQ ID NO:2 as follows:

5 MTSEVALGPSLPAPSPTACPALGPDSSSLGPVPSAEPATPQSCGVADPNEAEFFLRQFHAEQSDQPV
PLARRLEQVRAAIDATGTYRHTTAELVYGARVAWRNSSRCIGRLYWNSLRVLDRRDATAPDEIHRH
LCTHLRQATNGGRI RPVISVFAPDSPGRPGPQVWNEQLIRYAGYRRDDGTVLGDPRADLTEAILR
LGWQGCPOGPFVDVLPVIDTPDDKPRFFELPRELVLEVPITHPDVPRLAELGLRWHAVPVISNMRL
RIGGMDYPLAPFNGWYMGTEIGARNLVDEDRYNMLPAVAACLQLDTTSESTLWRDRALVELNVAVL
HSFEAAGVRI SDHHEESRRFLAHLAKEERQGR TVSADWSWIVPPLSGGITPVFHRYYDNVDQRPNF
YPHQ

10

[0022] A second isolated nucleic acid molecule of the present invention encodes a NOS from *Streptomyces scabies* (see GenBank Accession No.

AY204507) and has a nucleotide sequence according to SEQ ID NO:3 as follows:

15 gtgacctccgaagtcgctctgggccccttccttgcccgcggcggtcccgacagcgtgcccggcgactg
gggcccgatctgctcccttgcccgggtcccgctcgggcggaacggcgagcgccgcagtcctgcggcgctc
gccgatccaaatgaggtgaggagttcctgcccaggtccacgcggagcagtcctgacgcccgtc
ccgctcgcccggcgccctggagcaggtccgcgcggccatcgacgccacgggcacctaccggcacacc
accgcccagctcgtgtacggtgcccgcgctcgcggtggcgcaactccagtcgctgcatcgccgcccgtg
tactggaaacagcctgcgcgtcctggacgcgggacgccacagcccccgatgagatccaccggcac
20 ttgtgcacgcacctgcgccaggcgaccaacggcgggcgcatcaggccggtgatttcggtcttcgcc
ccggactcccccgccggccccggccccgcaggtgtggaacgagcagctcatccggtacgccggctac
cgccgcgacgcagccaccgtgctcggtgacccgcgcacccgcgacctaccgaggccatcctccgc
ctcggtggtgagggctgcccccaagggccgttcgacgtcctgcccctggtcatcgaccccccgac
gacaaacccccggttcttcgagctgcccgcgggagctggtcttgagggtccctatccccacccccgac
25 gtcccacgcctggccgaactgggcctgcgctggcagccgtaccgctcatctccaacatgccccta
cgcatcgccgggatggactaccgctcgccccgttcaacggctggtacatgggcacggagatcggc
gcccgcacacctcgtcgacgaggacggctacaacatgctccccgcgctcgccgcctgctccagytg
gacaccaccagcgagtcacacctgtggcgcgaccgcgccttggtcgagctcaacgtcgccgtcctg
cactccttcgagggccgcaggtgtccggatcagcgaccaccacgaggagtcggcgccgttcctcgcc
30 cacctggccaaggaggaaacgccaggggccgcaccgatatccgcagactggagctggatcgccccccg
ctctccggcgccatcccccggtgttcaccggttactacgacaacgtcgaccagcgcccccaacttc
tccccccaccagtga

The NOS protein or polypeptide encoded by this nucleic acid molecule has an amino acid sequence according to SEQ ID NO:4 as follows:

35 MTSEVALGPSLPAPSPTACPALGPDSSSLGPVPSAEPATPQSCGVADPNEAEFFLRQFHAEQSDQPV
PLARRLEQVRAAIDATGTYRHTTAELVYGARVAWRNSSRCIGRLYWNSLRVLDRRDATAPDEIHRH
LCTHLRQATNGGRI RPVISVFAPDSPGRPGPQVWNEQLIRYAGYRRDDGTVLGDPRADLTEAILR
LGWQGCPOGPFVDVLPVIDTPDDKPRFFELPRELVLEVPITHPDVPRLAELGLRWHAVPVISNMRL
40 RIGGMDYPLAPFNGWYMGTEIGARNLVDEDRYNMLPAVAACLQLDTTSESTLWRDRALVELNVAVL
HSFEAAGVRI SDHHEESRRFLAHLAKEERQGR TVSADWSWIVPPLSGGITPVFHRYYDNVDQRPNF
YPHQ

[0023] A third isolated nucleic acid molecule of the present invention encodes a from *Streptomyces turgidiscabies* (see GenBank Accession No.

AY204509) and has a nucleotide sequence according to SEQ ID NO:5 as follows:

gtgactttcgaagtcgcccctgggccccttccttgcccgcggcggtcccgacagcgtgcccggcgctg
gcgcacgattcgcccctcagtcggcggtcccgctcgggcggaacggcgagcgtcgaggactgcccgcgtc
gccgatccagacgaggccgaggagttcctgcccaggtttcacgcggagcagtcggaccaggccgtc

- 12 -

ccgctcactcggcgccctggaccaggttcgcgctgccatcgacgccacgggcacctaccgtcatacc
 accgccgagctcgtgttcggtgcccggtgcgcgtggcgcaactccagtcgctgtatcggccgcctg
 tactggaacagcctgcgcgtcctggaccgcccgggacaccacagccccgaggtaatccacgggcac
 ctttgacgcacctgcgccaggcgaccaacgggtggcggtatcaggccgggtgatttcggtcttcgcc
 5 ccggacgcacccagccgacccggcccggggtgtggaacgagcaactcgtccggtagccggccac
 cgtcgcgacgacggcaccgtactcggcgaccccgcatctgccgacctcaccgaggccatccgcggc
 ctcggtatggcagggaggccgccaagggccggttcgacgtcctgcccctgggtcatcgacgccacgac
 gacaaaccgcggttcttcgagctgcccggggaggttgctctggaggtccctatcacccaccccgac
 gtcccacgactggccgaactctgcctgcgctggcagccggtaccggttatctccaacatgcgcctg
 10 cgtatcgccggggtggactacccccctcgccccgttcaacggctggtacatgggcacggagatcggc
 gtccgtaacctcgtcgacgaggcccgctacaacctgctccccgcggtggccgctgcctccagttg
 gacaccaccagcgagtcacccctgtggcggtgaccgcgctctggtcgaactcaacggttgcgctctg
 cactcttttcgcccgcgagggcggtccggatcagtgaccaccacgaggagtcggcgcttcctcgc
 cacctgaccaaggaggaacgccagggccgacccgtaccgcggactggagctggatcgtccctcgg
 15 ctttcacgagcgtacccccgtcttcacccgctactacgacaacgccgaccagcgccccaacttt
 taccctcatcagtga

The NOS protein or polypeptide encoded by this nucleic acid molecule has an amino acid sequence according to SEQ ID NO:6 as follows:

20 MTFEVALGPSLPAPSPTACPALAHDSPVPSAEPATSQDCGVADPDEAEFLRQFHAEQSDQAV
 PLTRRLDQVRAAIDATGTYRHHTAELVFGARVAWRNSSRCIGRLYWNLSRLVDRDRTTAPEVIHRH
 LCTHLRQATNGGRIRPVISVFAPDAPSRPGPRVWNEQLVRYAGHRRDDGTVLGDPRSADLTEAIRG
 LGWQGGRRQGFDFVLPVIDAHDDKPRFFELPREVVLEVPIITHPDVPRLAELCLRWHAVPVTINMRL
 RIGGVDYPLAPFNGWYMGTEIGVRNLVDEARYNLLPAVAACLQLDTTSESTLWRDRALVELNVAVL
 25 HSFAAAGVRI SDHHEESRRFLAHLTKERQGRTPADWSWIVPPLSSGITPVFHRYYDNADQRPNF
 YPHQ

[0024] Fragments of the above-identified proteins or polypeptides as well
 as fragments of full length proteins can also be used according to the present
 30 invention.

[0025] Suitable fragments can be produced by several means. Subclones
 of the gene encoding a known protein can be produced using conventional
 molecular genetic manipulation for subcloning gene fragments, such as described
 by Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Springs
 35 Laboratory, Cold Springs Harbor, New York (1989), and Ausubel et al. (ed.),
Current Protocols in Molecular Biology, John Wiley & Sons (New York, NY)
 (1999 and preceding editions), each of which is hereby incorporated by reference
 in its entirety. The subclones then are expressed *in vitro* or *in vivo* in bacterial
 cells to yield a smaller protein or polypeptide that can be tested for activity.
 40 Various other cloning protocols known in the art are suitable for use in the present
 invention, including, for example, those described for *Streptomyces* in Hopwood
 et al., *Genetic manipulation of Streptomyces: A Laboratory Manual*,
 Norwich:John Innes Foundation (1985) and Kieser et al., *Practical Streptomyces*

Genetics Norwich:John Innes Foundation (2000), which are hereby incorporated by reference in their entirety.

[0026] In another approach, based on knowledge of the primary structure of the protein, fragments of the protein-coding gene may be synthesized using the
5 PCR technique together with specific sets of primers chosen to represent particular portions of the protein. Erlich, H.A., et al., "Recent Advances in the Polymerase Chain Reaction," *Science* 252:1643-51 (1991), which is hereby incorporated by reference. These can then be cloned into an appropriate vector for expression of a truncated protein or polypeptide from bacterial cells as described
10 above.

[0027] As an alternative, fragments of a protein can be produced by digestion of a full-length protein with proteolytic enzymes like chymotrypsin or *Staphylococcus* proteinase A, or trypsin. Different proteolytic enzymes are likely to cleave different proteins at different sites based on the amino acid sequence of
15 the particular protein. Some of the fragments that result from proteolysis may be active virulence proteins or polypeptides.

[0028] Chemical synthesis can also be used to make suitable fragments. Such a synthesis is carried out using known amino acid sequences for the polypeptide being produced. Alternatively, subjecting a full length protein to high
20 temperatures and pressures will produce fragments. These fragments can then be separated by conventional procedures (e.g., chromatography, SDS-PAGE).

[0029] Another example of suitable fragments of the nucleic acids of the present invention are fragments of the genes which have been identified as conserved ("con") regions of the proteins, or alternatively, those portions of
25 nucleotide sequences that have been identified as variable ("var") regions. Conserved regions in accordance with the present invention are further described *infra*. Sequences identified using DNASTar Mega alignment program as either variable or conserved in a gene can be amplified using standard PCR methods using forward and reverse primers designed to amplify the region of choice and
30 which include a restriction enzyme sequence to allow ligation of the PCR product into a vector of choice. Combinations of amplified conserved and variable region sequences can be ligated into a single vector to create a "cassette" which contains a plurality of DNA molecules in one vector.

[0030] Also suitable as an isolated nucleic acid molecule according to the present invention is a nucleic molecule having a nucleotide sequence that encodes a NOS protein or polypeptide having an amino acid motif that is a conserved region of the NOS protein or polypeptide from *Streptomyces acidiscabies*,
 5 *Streptomyces scabies*, and *Streptomyces turgidiscabies* (as described in Table 1, below).

TABLE 1 — Conserved Regions of the NOS Proteins or Polypeptides From *Streptomyces acidiscabies*, *Streptomyces scabies*, and *Streptomyces turgidiscabies*

Location ¹	Motif
4-22	EVALGPSLPAPSPTACPAL (SEQ ID NO:7)
30-38	PVPSAEPAT (SEQ ID NO:8)
42-47	CGVADP (SEQ ID NO:9)
49-64	EAEFLRQFHAEQSDQ (SEQ ID NO:10)
74-93	QVRAAIDATGTYRHITTAELV (SEQ ID NO:11)
95-122	GARVAWRNSSRCIGRLYWNSLRVLDRRD (SEQ ID NO:12)
129-156	IHRHLCTHLRQATNGGRIRPVISVFAPD (SEQ ID NO:13)
165-170	VWNEQL (SEQ ID NO:14)
177-188	RRDDGTVLGDPR (SEQ ID NO:15)
190-196	ADLTEAI (SEQ ID NO:16)
199-203	LGWQG (SEQ ID NO:17)
206-217	QGPFDVLPLVID (SEQ ID NO:18)
220-231	DDKPRFFELPRE (SEQ ID NO:19)
233-249	VLEVPITHPDVPRLAEL (SEQ ID NO:20)
251-268	LRWHAVPVISNMRLRIGG (SEQ ID NO:21)
270-286	DYPLAPFNGWYMGTEIG (SEQ ID NO:22)
288-293	RNLVDE (SEQ ID NO:23)
299-333	LPAAVACLQLDTTSESTLWRDRALVELNVAVLHSF (SEQ ID NO:24)
335-354	AAGVRISDHHEESRRFLAHL (SEQ ID NO:25)
356-364	KEERQGRTV (SEQ ID NO:26)
366-389	ADWSWIVPPLSGGITPVFHRYYDN (SEQ ID NO:27)
391-400	DQRPNFYPHQ (SEQ ID NO:28)

¹The "Location" refers to the range of amino acid residues as corresponding to the amino acid sequence of the NOS from *Streptomyces acidiscabies* (i.e., SEQ ID NO:2).

- 15 -

[0031] Thus, one aspect of the present invention includes an isolated nucleic acid molecule that encode a NOS protein or polypeptide having an amino acid motif having an amino acid sequence of one of the following: SEQ ID NO:7; SEQ ID NO:8; SEQ ID NO:9; SEQ ID NO:10; SEQ ID NO:11; SEQ ID NO:12; SEQ ID NO:13; SEQ ID NO:14; SEQ ID NO:15; SEQ ID NO:16; SEQ ID NO:17; SEQ ID NO:18; SEQ ID NO:19; SEQ ID NO:20; SEQ ID NO:21; SEQ ID NO:22; SEQ ID NO:23; SEQ ID NO:24; SEQ ID NO:25; SEQ ID NO:26; SEQ ID NO:27; SEQ ID NO:28; and/or combinations thereof.

10 [0032] The information presented in Table 1 can be combined to define a NOS protein or polypeptide of the present invention as having an amino acid sequence of SEQ ID NO:47 (with X being any amino acid) as follows:

15 (3X-49X) EVALGPSLPAPSPATACPAL (7X) FVPSAEPAT (3X) CGVADP (1X) EAEEFLRQFHAE QSDQ (9X) QVRAAIDATGTYRHTTAELV (1X) GARVAWRNSSRCIGRLYWNLSLRVLDLRRD (6X) IHRHLCTHLRQATNGGRIRPVISVFAPD (8X) VWNEQL (6X) RRDDGTVLGDPR (1X) ADLTEAI (2X) LGWQG (2X) QGPFDVLPLVID (2X) DDKPRFFELPRE (1X) VLEVPIITHPDVPRLAEL (1X) LRWHAVPVISNMRLRIGG (1X) DYPLAPFNGWYMGTEIG (1X) RNLVDE (5X) LPAAVACLQLDTS SESTLWRDRALVELNVAVLHSF (1X) AAGVRI SDHHEESRRFLAHL (1X) KEERQGRV (1X) ADW 20 SWIVPPLSGGITPVFHRYYDN (1X) DQRPNFYPHQ

[0033] Mutations or variants of the above polypeptides or proteins are encompassed by the present invention. Variants may be made by, for example, the deletion or addition of amino acids that have minimal influence on the properties, secondary structure, and hydrophobic nature of a polypeptide or protein. For example, a polypeptide may be conjugated to a signal (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification, or identification of the polypeptide.

[0034] Also suitable as an isolated nucleic acid molecule according to the present invention is a nucleic acid molecule having a nucleotide sequence that is at least 55 percent similar, particularly at least 85 percent similar, and more particularly at least 90 percent similar to the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5 by basic BLAST using default parameters analysis. Another suitable isolated nucleic acid of the present invention is one

having a nucleotide sequence having at least 60 percent homology, particularly at least 70 percent homology, more particularly at least 80 percent homology, more particularly at least 90 percent homology, and still more particularly at least 95 percent homology to a nucleic acid according to SEQ ID NO:1, SEQ ID NO:3, or
5 SEQ ID NO:5.

[0035] Suitable nucleic acid molecules also include those that hybridize to a nucleic acid molecule comprising a nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5 under stringent conditions. For the purposes of defining the level of stringency, reference can conveniently be made to Sambrook
10 et al., *Molecular Cloning: a Laboratory Manual*, 2nd Edition, Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press, at 11.45 (1989), which is hereby incorporated by reference in its entirety). An example of low stringency conditions is 4-6X sodium citrate ("SSC")/0.1-0.5% w/v SDS at 37°-45°C for 2-3 hours. Depending on the source and concentration of the nucleic acid involved in
15 the hybridization, alternative conditions of stringency may be employed such as medium stringent conditions. Examples of medium stringent conditions include 1-4X SSC/0.25% w/v SDS at $\geq 45^{\circ}\text{C}$ for 2-3 hours. An example of high stringency conditions includes 0.1-1X SSC/0.1% w/v SDS at 60°C for 1-3 hours. The skilled artisan is aware of various parameters which may be altered during
20 hybridization and washing and which will either maintain or change the stringency conditions. Other examples of high stringency conditions include: 4-5X SSC/0.1% w/v SDS at 54° C for 1-3 hours and 4X SSC at 65°C, followed by a washing in 0.1X SSC at 65°C for about one hour. Alternatively, an exemplary stringent hybridization condition is in 50% formamide, 4X SSC, at 42°C. Still
25 another example of stringent conditions include hybridization at 62°C in 6X SSC, .05X BLOTTO, and washing at 2X SSC, 0.1% SDS at 62°C. In one particular embodiment, the stringent conditions are characterized by a hybridization medium including about 5X SSC at a temperature of about 55°C.

[0036] The precise conditions for any particular hybridization are left to
30 those skilled in the art because there are variables involved in nucleic acid hybridizations beyond those of the specific nucleic acid molecules to be hybridized that affect the choice of hybridization conditions. These variables include: the substrate used for nucleic acid hybridization (e.g., charged vs. non-

charged membrane); the detection method used (e.g., radioactive vs. chemiluminescent); and the source and concentration of the nucleic acid involved in the hybridization. All of these variables are routinely taken into account by those skilled in the art prior to undertaking a nucleic acid hybridization procedure.

5 **[0037]** A nitric oxide synthase protein or polypeptide of the present invention is preferably produced in purified form (e.g., at least about 80 percent, more preferably 90 percent pure) by conventional techniques. For example, a nitric oxide synthase protein or polypeptide of the present invention may be secreted into the growth medium of recombinant host cells. To isolate the nitric
10 oxide synthase protein or polypeptide, a protocol involving a host cell such as *Escherichia coli* may be used, in which protocol the *E. coli* host cell carrying a recombinant plasmid is propagated, homogenized, and the homogenate is centrifuged to remove bacterial debris. The supernatant is then subjected to sequential ammonium sulfate precipitation. The fraction containing the nitric
15 oxide synthase protein or polypeptide of the present invention is subjected to gel filtration in an appropriately sized dextran or polyacrylamide column to separate the proteins or polypeptides. If necessary, the protein fraction may be further purified by high performance liquid chromatography ("HPLC").

20 **[0038]** The present invention relates to a DNA construct that contains a DNA molecule encoding for a nitric oxide synthase protein or polypeptide of the present invention. This involves incorporating one or more of the nucleic acid molecules of the present invention, or a suitable portion thereof, into host cells using conventional recombinant DNA technology. Generally, this involves inserting the nucleic acid molecule into an expression system to which the nucleic
25 acid molecule is heterologous (i.e., not normally present). The expression system contains the necessary elements for the transcription and translation of the inserted protein-coding sequences.

30 **[0039]** The present invention also relates to an expression system (e.g., an expression vector) containing a nucleic acid molecule encoding a nitric oxide synthase protein or polypeptide of the present invention. The nucleic acid molecules of the present invention may be inserted into any of the many available expression vectors and cell systems using reagents that are well known in the art. In preparing a DNA vector for expression, the various DNA sequences may

normally be inserted or substituted into a bacterial plasmid. Any convenient plasmid may be employed, which will be characterized by having a bacterial replication system, a marker which allows for selection in a bacterium, and generally one or more unique, conveniently located restriction sites. Numerous
5 plasmids, referred to as transformation vectors, are available for transformation. The selection of a vector will depend on the preferred transformation technique and target cells for transfection.

[0040] Suitable vectors include, but are not limited to, the following viral vectors such as lambda vector system gt11, gt WES.tB, Charon 4, and plasmid
10 vectors such as pBR322, pBR325, pACYC177, pACYC1084, pUC8, pUC9, pUC18, pUC19, pLG339, pR290, pKC37, pKC101, SV 40, pBluescript II SK +/- or KS +/- (see "Stratagene Cloning Systems" Catalog (1993) from Stratagene, La Jolla, CA, which is hereby incorporated by reference in its entirety), pQE, pIH821, pGEX, pET series (see F.W. Studier et. al., "Use of T7 RNA Polymerase
15 to Direct Expression of Cloned Genes," *Gene Expression Technology*, Vol. 185 (1990), which is hereby incorporated by reference in its entirety), pCB201, and any derivatives thereof. Any appropriate vectors now known or later described for genetic transformation are suitable for use with the present invention. Recombinant molecules can be introduced into cells via transformation,
20 particularly transduction, conjugation, mobilization, or electroporation. The DNA sequences are cloned into the vector using standard cloning procedures in the art, as described by Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Press, NY (1989), and Ausubel, F. M. et al. (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York,
25 N.Y., which are hereby incorporated by reference in their entirety.

[0041] U.S. Patent No. 4,237,224 issued to Cohen and Boyer, which is hereby incorporated by reference in its entirety, describes the production of expression systems in the form of recombinant plasmids using restriction enzyme cleavage and ligation with DNA ligase. These recombinant plasmids are then
30 introduced by means of transformation and replicated in unicellular cultures including prokaryotic organisms and eukaryotic cells grown in tissue culture.

[0042] A variety of host-vector systems may be utilized to express the protein-encoding sequence(s). Primarily, the vector system must be compatible

with the host cell used. Host-vector systems include but are not limited to the following: bacteria transformed with bacteriophage DNA, plasmid DNA, or cosmid DNA; microorganisms such as yeast containing yeast vectors; mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell
5 systems infected with virus (e.g., baculovirus); and plant cells infected by bacteria. The expression elements of these vectors vary in their strength and specificities. Depending upon the host-vector system utilized, any one of a number of suitable transcription and translation elements can be used.

[0043] Thus, certain "control elements" or "regulatory sequences" are also
10 incorporated into the plasmid-vector constructs of the present invention. These include non-transcribed regions of the vector and 5' and 3' untranslated regions, which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and/or translation
15 elements, including constitutive, inducible, and repressible promoters, as well as minimal 5' promoter elements may be used. A constitutive promoter is a promoter that directs expression of a gene throughout the development and life of an organism. An inducible promoter is a promoter that is capable of directly or indirectly activating transcription of one or more DNA sequences or genes in
20 response to an inducer. In the absence of an inducer, the DNA sequences or genes will not be transcribed or will only be minimally transcribed.

[0044] The DNA sequences of eukaryotic promoters differ from those of prokaryotic promoters. Furthermore, eukaryotic promoters and accompanying genetic signals may not be recognized in or may not function in a prokaryotic
25 system, and, further, prokaryotic promoters are not recognized and do not function in eukaryotic cells.

[0045] Promoters vary in their "strength" (i.e. their ability to promote transcription). For the purposes of expressing a cloned gene, it is desirable to use strong promoters in order to obtain a high level of transcription and, hence,
30 expression of the gene. Depending upon the host cell system utilized, any one of a number of suitable promoters may be used. For instance, when cloning in *E. coli*, its bacteriophages, or plasmids, promoters such as the T7 phage promoter, *lac* promoter, *trp* promoter, *recA* promoter, ribosomal RNA promoter, the P_R and

P_L promoters of coliphage lambda and others, including but not limited, to *lacUV5*, *ompF*, *bla*, *lpp*, and the like, may be used to direct high levels of transcription of adjacent DNA segments. Additionally, a hybrid *trp-lacUV5 (tac)* promoter or other *E. coli* promoters produced by recombinant DNA or other synthetic DNA techniques may be used to provide for transcription of the inserted gene.

[0046] Other examples of some constitutive promoters that are widely used for inducing expression of transgenes include the nopoline synthase gene promoter from *Agrobacterium tumefaciens* (U.S. Patent No. 5,034,322 issued to Rogers et al., which is hereby incorporated by reference in its entirety), the cauliflower mosaic virus (CaMV) 35S and 19S promoters (U.S. Patent No. 5,352,605 issued to Fraley et al., which is hereby incorporated by reference in its entirety), the enhanced CaMV35S promoter ("enh CaMV35S"), the figwort mosaic virus full-length transcript promoter ("FMV35S"), those derived from any of the several actin genes, which are known to be expressed in most cells types (U.S. Patent No. 6,002,068 issued to Privalle et al., which is hereby incorporated by reference in its entirety), and the ubiquitin promoter, which is a gene product known to accumulate in many cell types. Examples of constitutive promoters for use in mammalian cells include the RSV promoter derived from Rous sarcoma virus, the CMV promoter derived from cytomegalovirus, β -actin and other actin promoters, and the EF1 α promoter derived from the cellular elongation factor 1 α gene.

[0047] Bacterial host cell strains and expression vectors may be chosen which inhibit the action of the promoter unless specifically induced. In certain operations, the addition of specific inducers is necessary for efficient transcription of the inserted nucleic acid. For example, the *lac* operon is induced by the addition of lactose or IPTG (isopropylthio-beta-D-galactoside). A variety of other operons, such as *trp*, *pro*, etc., are under different controls.

[0048] Other examples of some inducible promoters, induced, for examples by a chemical agent, such as a metabolite, growth regulator, herbicide or phenolic compound, or a physiological stress/physical means, such as cold, heat, salt, toxins, or through the action of a pathogen or disease agent such as a virus or fungus, include a glucocorticoid-inducible promoter (Skena et al., *Proc. Natl.*

- Acad. Sci.* 88:10421-5 (1991), which is hereby incorporated by reference in its entirety), the heat shock promoter ("Hsp"), IPTG or tetracycline ("Tet on" system), the metallothionine promoter, which is activated by heavy metal ions, and hormone-responsive promoters, which are activated by treatment of certain hormones. A host cell containing an inducible promoter may be exposed to an inducer by externally applying the inducer to the cell. In addition, "tissue-specific" promoters can be used, which are promoters that function in a tissue specific manner to regulate the gene of interest within selected tissues of the host. Examples of such tissue specific promoters include seed, flower, or root specific promoters as are well known in the field (e.g., U.S. Patent No. 5,750,385 to Shewmaker et al., which is hereby incorporated by reference in its entirety). Promoters of the nucleic acid construct of the present invention may be either homologous (derived from the same species as the host cell) or heterologous (derived from a different species than the host cell).
- 15 [0049] Specific initiation signals are also required for efficient gene transcription and translation in prokaryotic cells. These transcription and translation initiation signals may vary in "strength" as measured by the quantity of gene specific messenger RNA and protein synthesized, respectively. The DNA expression vector, which contains a promoter, may also contain any combination of various "strong" transcription and/or translation initiation signals. For instance, efficient translation in *E. coli* requires an SD sequence about 7-9 bases 5' to the initiation codon ("ATG") to provide a ribosome binding site. Thus, any SD-ATG combination that can be utilized by host cell ribosomes may be employed. Such combinations include but are not limited to the SD-ATG combination from the *cro* gene or the *N* gene of coliphage lambda, or from the *E. coli* tryptophan E, D, C, B or A genes. Additionally, any SD-ATG combination produced by recombinant DNA or other techniques involving incorporation of synthetic nucleotides may be used.
- 25 [0050] The constructs of the present invention also include an operable 3' regulatory region, selected from among those which are capable of providing correct transcription termination and polyadenylation of mRNA for expression in the host cell of choice, operably linked to a DNA molecule which encodes for a protein of choice. A number of 3' regulatory regions are known in the art.
- 30

Virtually any 3' regulatory region known to be operable in the host cell of choice would suffice for proper expression of the coding sequence of the nucleic acid of the present invention.

5 **[0051]** In one aspect of the present invention, the nucleic acid molecule of the present invention is incorporated into an appropriate vector in the sense direction, such that the open reading frame is properly oriented for the expression of the encoded protein under control of a promoter of choice. This involves the inclusion of the appropriate regulatory elements into the DNA-vector construct. These include non-translated regions of the vector, useful promoters, and 5' and 3' 10 untranslated regions which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used.

15 **[0052]** A nucleic acid molecule of the present invention, promoter of choice, an appropriate 3' regulatory region, and, if desired, a reporter gene, can be incorporated into a vector-expression system to contain a nucleic acid of the present invention, or a suitable fragment thereof, using standard cloning techniques as described in Sambrook et al., *Molecular Cloning: A Laboratory* 20 *Manual*, Second Edition, Cold Spring Harbor Press, NY (1989), and Ausubel et al. (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y., which are hereby incorporated by reference in their entirety. The transcriptional and translational elements are operably linked to the nucleic acid molecule of the present invention or a fragment thereof, meaning that the resulting 25 vector expresses the nitric oxide synthase protein or polypeptide when placed in a suitable host cell.

30 **[0053]** Once an isolated DNA molecule encoding a nitric oxide synthase protein or polypeptide has been cloned into an expression vector, it is ready to be incorporated into a host cell. Such incorporation can be carried out by the various forms of transformation noted above, depending upon the vector/host cell system. Recombinant molecules can be introduced into cells *via* transformation, particularly transduction, conjugation, mobilization, or electroporation. The nucleic acid sequences are cloned into the host cell using standard cloning

procedures known in the art, as described by Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Springs Laboratory, Cold Springs Harbor, New York (1989), which is hereby incorporated by reference in its entirety. Suitable host cells include, but are not limited to, bacteria, virus, yeast, mammalian cells, insect, plant, and the like. Examples of a suitable bacterial host cells include, without limitation, *Streptomyces*, *Bacillus*, *Escherichia*, *Brevibacterium*, *Microbacterium*, *Nocardia*, and *Rhodococcus* cells. Particular *Streptomyces* host cells can include *Streptomyces acidiscabies*, *Streptomyces scabies*, *Streptomyces turgidiscabies*, *Streptomyces avermitilis*, *Streptomyces lividans*, *Streptomyces coelicolor*, and *Streptomyces ipomoea*. Particular *Escherichia* host cells can include *Escherichia coli*. Particular *Bacillus* host cells can include *Bacillus subtilis*, *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus halodurans*. Examples of suitable fungal host cells include, without limitation, *Aspergillus*, *Cephalosporium*, and *Penicillium* cells. Suitable yeast host cells include, with limitation, *Saccharomyces* cells.

[0054] Thus, the present invention also relates to a host cell incorporating one or more of the isolated nucleic acid molecules of the present invention. In one embodiment, the isolated nucleic acid molecule is heterologous to the host cell. Such incorporation can be carried out by the various forms of transformation noted above, depending upon the vector/host system, and using the various host cells described above.

[0055] Methods of transformation may result in transient or stable expression of the DNA under control of the promoter. Preferably, the nucleic acid of the present invention is stably inserted into the genome of the host cell as a result of the transformation, although transient expression can serve an important purpose.

[0056] The present invention also relates to a method of attaching a nitrogen group to a target moiety of a compound. This method involves providing a nitric oxide synthase. The method also involves providing a compound having a target moiety to which can be attached a nitrogen group. The nitric oxide synthase and the compound are combined in a reaction mixture under conditions effective to allow the nitric oxide synthase to catalyze a reaction whereby a

nitrogen group contained in the reaction mixture is attached to the target moiety. The method thereby yields a nitrogen-modified compound.

[0057] The present invention further relates to a method of synthesizing a nitrogen-modified compound in a transgenic host cell. This method involves
5 providing a transgenic host cell transformed with a DNA molecule encoding a nitric oxide synthase. A non-modified compound containing a target moiety is also provided. The transgenic host cell and the non-modified compound are cultured in a culture medium under conditions effective to allow the nitric oxide synthase to be expressed and to catalyze a reaction whereby a nitrogen group
10 contained in the host cell or culture medium attaches to the target moiety. The method thereby yields a nitrogen-modified compound. The non-modified compound containing a target moiety can be provided exogenously. Alternatively, the non-modified compound containing a target moiety can be produced by the transgenic host cell. Suitable DNA molecules include those of
15 the present invention, as well as those encoding the various nitric oxide synthases described herein (and as specifically identified *infra*).

[0058] As used herein, the term "Nitrogen-Modifying Methods" refers to the above-referenced "method of attaching a nitrogen group to a target moiety of a compound" and "method of synthesizing a nitrogen-modified compound in a
20 transgenic host cell." Suitable host cells, nitrogen groups, target moieties, and nitrogen-modified compounds for use in the Nitrogen-Modifying Methods of the present invention are further described below.

[0059] In one example for conducting the Nitrogen-Modifying Methods of the present invention, the NOS-mediated nitration and/or nitrosylation can
25 proceed by incubating the compound (e.g., molecule) having the target moiety targeted for modification with the following: (i) a nitric oxide synthase of the present invention (which functions as a catalyst for nitration and/or nitrosylation); (ii) a nitrogen-containing substrate for NOS (also referred to herein as the "NOS substrate"); and (iii) a source of oxygen (e.g., reduced oxygen). A suitable NOS
30 substrate can include any nitrogen-containing molecule that can bind at the NOS heme center and be converted to an oxidized form of nitrogen capable of nitration and/or nitrosylation reactions. For example, the NOS substrate can contain either a guanidinium group, an hydroxy-guanidinium group, and/or an oxime group.

- 25 -

Oxygen can be supplied in the form of molecular oxygen (O_2), reduced oxygen (e.g., peroxide (O_2^{2-}) or superoxide (O_2^-)), or as an oxo-donor compound such as m-chloroperoxybenzoic acid. Depending on the oxygen and nitrogen-containing substrate used, additional reducing equivalents may be provided in the form of

5 small molecule electron donors (e.g., dithionite, reduced methyl viologen) or as a reductase protein (e.g., iron-sulfur cluster, heme, or flavin-containing enzymes).

[0060] Suitable nitric oxide synthases for use in the Nitrogen-Modifying Methods of the present invention can include any nitric oxide synthase (from any source) that can function as a catalyst in a nitration and/or nitrosylation reaction.

10 Examples of such suitable nitric oxide synthases are those described herein and/or encoded by the isolated nucleic acid molecules of the present invention (described *supra*). Other suitable examples of nitric oxide synthases or nucleic acid molecules for use in the Nitrogen-Modifying Methods include those that have (or encode) the following three conserved amino acid regions (i.e., motifs): (i) the

15 RCIGR (SEQ ID NO:44) motif (corresponding to amino acid residues 105-109 of the nitric oxide synthase of *Streptomyces acidiscabies* (SEQ ID NO:2)), which functions to coordinate the iron of the heme prosthetic group; (ii) the GWYXXXE (SEQ ID NO:45, where X can be any amino acid residue) motif (corresponding to amino acid residues 278-284 of the nitric oxide synthase of *Streptomyces*

20 *acidiscabies* (SEQ ID NO:2)), which functions to bind the guanidinium or hydroxy-guanidinium-containing NOS substrate; and (iii) the WSWXXXP (SEQ ID NO:46, where X can be any amino acid residue) motif (corresponding to amino acid residues 368-374 of the nitric oxide synthase of *Streptomyces acidiscabies* (SEQ ID NO:2)), which functions to interact with cofactors or substrates that can

25 be capable of reduction/oxidation reactions with the NOS heme center (e.g. reduced pterins such as tetrahydrofolate and tetrahydrobiopterin). The three conserved motifs described above can be combined to define a nitric oxide synthase protein or polypeptide for use in the Nitrogen-Modifying Methods of the present invention as having an amino acid sequence of SEQ ID NO:50 as follows:

30 (37X-174X) RCIGR (48X-67X) GWYXXXE (83X) WSWXXXP (20X-38X)

[0061] Still other suitable nitric oxide synthases for use in Nitrogen-Modifying Methods can include those having amino acid sequences that

- 26 -

contain amino acid motifs that are conserved in nitric oxide synthases from a variety of sources. Suitable sources of nitric oxide synthases can include, for example, *Bacillus anthracis*, *Bacillus cereus*, *Bacillus subtilis*, *Bacillus halodurans*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Deinococcus radiodurans*, *Mus musculus*, *Streptomyces acidiscabies*, *Streptomyces scabies*, *Streptomyces turgidiscabies*, and *Streptomyces avermitilis*. An alignment of the nitric oxide synthases from these organisms is shown in Figure 6.

[0062] The various conserved regions of the nitric oxide synthases from *Streptomyces acidiscabies*, *Streptomyces scabies*, *Streptomyces turgidiscabies*, and *Streptomyces avermitilis* are described below in Table 2 (below).

TABLE 2 — Conserved Regions of the Nitric Oxide Synthase Proteins or Polypeptides From *Streptomyces acidiscabies*, *Streptomyces scabies*, *Streptomyces turgidiscabies*, and *Streptomyces avermitilis*

Location ¹	Motif
82-89	TGTYRHTT (SEQ ID NO:29)
95-102	GARVAWRN (SEQ ID NO:30)
105-112	RCIGRLYW (SEQ ID NO:31)
152-156	VFAPD (SEQ ID NO:32)
166-170	WNEQL (SEQ ID NO:33)
208-214	PFDVLPL (SEQ ID NO:34)
251-257	LRWHAVP (SEQ ID NO:35)
274-286	APFNGWYMGTEIG (SEQ ID NO:36)
319-326	DRALVELN (SEQ ID NO:37)
329-333	VLHSF (SEQ ID NO:38)
366-374	ADWSWIVPP (SEQ ID NO:39)
380-385	TPVFHR (SEQ ID NO:40)

¹ The "Location" refers to the range of amino acid residues as corresponding to the amino acid sequence of the NOS from *Streptomyces acidiscabies* (i.e., SEQ ID NO:2).

The information presented in Table 2 can be combined to define a nitric oxide synthase protein or polypeptide for use in the Nitrogen-Modifying Methods of the present invention as having an amino acid sequence of SEQ ID NO:48 (with X being any amino acid) as follows:

(81X-151X) TGTYRHTT (5X) GARVAWRN (2X) RCIGRLYW (39X) VFAPD (9X) WNEQL (37X) PFDVLPL (36X) LRWHAVP (16X) APFNGWYMGTEIG (32X) DRALVELN (2X) VLHSF (32X) ADWSWIVPP (5X) TPVFHR (15X-30X)

[0063] The various conserved regions of the nitric oxide synthases from *Bacillus anthracis*, *Bacillus cereus*, *Bacillus subtilis*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Deinococcus radiodurans*, *Mus musculus*,
 5 *Streptomyces acidiscabies*, *Streptomyces scabies*, *Streptomyces turgidiscabies*, and *Streptomyces avermitilis* are described below in Table 3 (below).

TABLE 3 — Conserved Regions of the Nitric Oxide Synthase Proteins or Polypeptides From Eleven Different Sources³

10

Location ¹	Motif ²
97-103	<u>2</u> 3AWRN <u>4</u> (SEQ ID NO:41)
105-110	RC <u>5</u> GR <u>6</u> (SEQ ID NO:42)
169-174	QL <u>5</u> RYA (SEQ ID NO:43)

¹ The "Location" refers to the range of amino acid residues as corresponding to the amino acid sequence of the NOS from *Streptomyces acidiscabies* (i.e., SEQ ID NO:2).

15

² The numerals used in the Motifs represent alternative amino acid residues as follows: 2 = R or K; 3 = M or V; 4 = S or A; 5 = I or V; and 6 = L or I

³ The twelve sources include *Bacillus anthracis*, *Bacillus cereus*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Bacillus halodurans*, *Bacillus subtilis* subsp. *subtilis*, *Deinococcus radiodurans*, *Streptomyces acidiscabies*, *Streptomyces scabies*, *Streptomyces turgidiscabies*, *Streptomyces avermitilis*, and *Mus musculus*.

The information presented in Table 3 can be combined to define a nitric oxide synthase protein or polypeptide for use in the Nitrogen-Modifying Methods of the present invention as having an amino acid sequence of SEQ ID NO:49 (with X
 25 being any amino acid and 2, 3, 4, 5, and 6 having the same definitions as for Table 3) as follows:

(29X-166X) 23AWRN4 (1X) RC5GR6 (52X-58X) QL5RYA (220X-243X)

30

[0064] In view of the information contained herein regarding conserved regions of the various nitric oxide synthases, a suitable nitric oxide synthase for use in the Nitrogen-Modifying Methods of the present invention can include proteins or polypeptides having an amino acid motif having an amino acid
 35 sequence of one of the following: SEQ ID NO:7; SEQ ID NO:8; SEQ ID NO:9; SEQ ID NO:10; SEQ ID NO:11; SEQ ID NO:12; SEQ ID NO:13; SEQ ID

NO:14; SEQ ID NO:15; SEQ ID NO:16; SEQ ID NO:17; SEQ ID NO:18; SEQ ID NO:19; SEQ ID NO:20; SEQ ID NO:21; SEQ ID NO:22; SEQ ID NO:23; SEQ ID NO:24; SEQ ID NO:25; SEQ ID NO:26; SEQ ID NO:27; SEQ ID NO:28; SEQ ID NO:29; SEQ ID NO:30; SEQ ID NO:31; SEQ ID NO:32; SEQ ID NO:33; SEQ ID NO:34; SEQ ID NO:35; SEQ ID NO:36; SEQ ID NO:37; SEQ ID NO:38; SEQ ID NO:39; SEQ ID NO:40; SEQ ID NO:41; SEQ ID NO:42; SEQ ID NO:43; SEQ ID NO:44; SEQ ID NO:45; SEQ ID NO:46; and/or combinations thereof. More particularly, the nitric oxide synthase used in the Nitrogen-Modifying Methods of the present invention can have an amino acid sequence of SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, or SEQ ID NO:50.

[0065] As referenced in the Nitrogen-Modifying Methods of the present invention, a suitable nitrogen group can include a nitro group (i.e., NO₂) or a nitroso group (i.e., NO). The reaction catalyzed by the nitric oxide synthase, whereby the nitrogen group is added to the target moiety, can be a nitration reaction or a nitrosylation reaction.

[0066] The Nitrogen-Modifying Methods of the present invention can be used to yield various types of nitrogen-modified compounds. An example of a class of nitrogen-modified compounds of the present invention are secondary metabolites. As used herein, the term "secondary metabolite" refers to those compounds of an organism that are not essential for normal growth, development, or reproduction of that organism. Suitable examples of such secondary metabolites include alkaloids, terpenoids, aliphatic organic acids, aromatic organic acids, heteroaromatic organic acids, phenols, irridoids, steroids, volatile oils, resins, balsams, β -lactams, aminoglycosides, macrolides, tetracyclines, and saponins. Other suitable nitrogen-modified compounds can include polyketides, peptides (including, for example, non-ribosomal peptides), and phytotoxins (including, for example, thaxtomin). Still other suitable nitrogen-modified compounds can include herbicides, nematicides, fungicides, insecticides, antibiotics (e.g., chloramphenicol), and anticancer agents (e.g., L-alanosine). Additional suitable nitrogen-modified compounds can include high-energy materials. Examples of such high-energy materials include nitroglycerin, trinitrotoluene, pentaerythritoltetranitrate, cyclotrimethylenetrinitramine, and ammonium nitrate fertilizer.

- 29 -

[0067] Various moieties of compounds can be targeted for addition of a nitrogen group. An example of a target moiety (as defined by the present invention) can include an electron-rich aromatic group. Particular target moieties can include a phenyl moiety or an indole moiety, and more particularly a tryptophanyl moiety.

EXAMPLES

Example 1 – Thaxtomin Production

[0068] *Streptomyces turgidiscabies* cultures were grown in oat bran broth or oat meal broth (Healy et al., *Mol. Micro.* 38:794-804 (2000); and Goyer et al., *Phytopathology* 88:442-445 (1998), which are hereby incorporated by reference in their entirety) inoculated with spores. After shaking at 150 rpm and 25°C for 5-9 days the cultures were filtered and dry weight of the mycelia was measured. Thaxtomin was extracted from the filtrate with ethyl acetate, dried, redissolved in methanol, and quantified using HPLC (column: 5 μ m C18; 250 x 4.6 mm; mobile phase: MeCN: H₂O:TFA (40:60:0.1)). In experiments investigating the suppression of thaxtomin production in the presence of L-NAME, this NOS inhibitor was dissolved in water, filter-sterilized, and added at the time of inoculation.

Example 2 – ¹⁵N Feeding studies

[0069] NMR analysis was conducted on thaxtomin A extracted from cultures fed either L-Arginine-guanidino-¹⁵N₂•HCl (Cambridge Isotope Laboratories), or ¹⁵NH₄NO₃ (Aldrich). ¹⁵NH₄NO₃ was added to oat bran broth at the time of inoculation with spores of *S. turgidiscabies*, whereas L-Arg-guanidino-¹⁵N₂•HCl was added just prior to the onset of thaxtomin biosynthesis (4-5 days after inoculation). Purified thaxtomin A dissolved in CD₃OD was analyzed on a Varian VXR-400S spectrometer equipped with a Nalorac broad-band probe affording observation of ¹⁵N at 40.5 MHz. Spectra were referenced externally by observing the ¹⁵N signal of formamide (90% solution in DMSO) and then acquiring sample spectra with identical parameter sets. A conventional chemical

- 30 -

shift scale (liquid NH₃ = 0 ppm) was established by referencing the formamide signal to its reported value of 112 ppm (Martin et al., *J. Nat. Prod.* 63:543-585 (2000), which is hereby incorporated by reference in its entirety).

5 **Example 3 – Molecular biology**

[0070] DNA and RNA manipulation were performed using standard techniques. Transformation of *S. turgidiscabies* was performed using polyethylene glycol (PEG) mediated transformation of *S. turgidiscabies* protoplasts using plasmid vectors propagated in *E. coli* ET12567 (MacNeil et al., 10 *Gene* 111:61-68 (1992), which is hereby incorporated by reference in its entirety). *Streptomyces turgidiscabies* p2XNOS single cross-over transformants (apramycin resistant, thiostrepton resistant) were grown for three generations of growth and sporulation on minimal medium containing thiostrepton, after which colonies containing double cross-over (apramycin sensitive, thiostrepton resistant) 15 recombination events were screened. Expression of *nos* was induced in the complemented Δnos strain by addition of 10 $\mu\text{g ml}^{-1}$ thiostrepton in a 5 mL oat bran culture 3 days following inoculation.

20 **Example 4 – A Bacterial Nitric Oxide Synthase Functions to Nitrate a Peptide Phytotoxin**

[0071] Plant pathogenic *Streptomyces* species are the causal agents of potato scab disease, a globally important disease of potato. Pathogenicity depends on production of a class of dipeptide phytotoxins called thaxtomins (Healy et al., 25 *Mol. Micro.* 38:794-804 (2000), which is hereby incorporated by reference in its entirety). While investigating the molecular genetics of plant pathogenicity in *Streptomyces* species, it was discovered that a gene with high sequence similarity to mammalian and bacterial NOSs. The location of this gene on a pathogenicity island that mobilizes among species to confer thaxtomin biosynthetic ability and its proximity to two nonribosomal peptide synthases of the thaxtomin biosynthetic 30 pathway suggests that the NOS participates in nitration of thaxtomins (Figures 1A-1B). The DNA sequence of the *nos* genes in *S. turgidiscabies*, *S. acidiscabies* and *S. scabies*, all of which produce thaxtomin A, is highly conserved (Genbank

accessions AY204507-AY204509). Conservation of nearly all key residues known in mammalian and *B. subtilis* NOSs that participate in cofactor-binding, substrate-binding and catalysis by *S. turgidiscabies* NOS (stNOS) suggests that stNOS is capable of producing NO from L-Arg (Figure 2). As in other bacterial NOSs, the reductase domain and CaM binding site typical of mammalian NOSs are absent in stNOS. However, stNOS has an extended N-terminal region that is lacking in other bacterial proteins (Figure 2). A mammalian NOS zinc-binding motif absent in the N-termini of other bacterial NOSs may be conserved in some of the pathogenic *Streptomyces* spp. NOSs.

- 10 [0072] To determine if stNOS is required for production of thaxtomin A, *nos* was deleted from *S. turgidiscabies* (Figures 3A-3B). The Δnos strain produced only trace amounts of thaxtomin A ($0.20 \pm 0.02 \mu\text{g ml}^{-1}$; standard deviation reported, $n=3$ for each strain tested) compared to the wild-type strain ($18.31 \pm 0.92 \mu\text{g ml}^{-1}$). De-nitrothaxtomin was not detectable in the medium.
- 15 Deletion of *nos* did not affect bacterial growth, but did eliminate disease on potato tubers. Complementation of *nos* was achieved by expressing *nos* using the thiostrepton-inducible promoter on plasmid pIJ8600, which integrates into the chromosomal $\phi C31$ phage attachment site (Figure 3B) (Sun et al., *Microbiology* 145:2221-2227 (1999), which is hereby incorporated by reference in its entirety).
- 20 The complemented strain Δnos pIJ8600NOS increased the amount of thaxtomin A produced over 25 fold ($3.68 \pm 0.06 \mu\text{g ml}^{-1}$ thaxtomin A) compared to the empty vector control strain Δnos pIJ8600 ($0.13 \pm 0.01 \mu\text{g ml}^{-1}$). Expression of *nos* in the *S. turgidiscabies* wild-type strain was confirmed using reverse transcriptase (RT)-PCR (Figure 3C). *nos* was not expressed in the Δnos or Δnos pIJ8600 strains, but expression was restored in the complemented strain Δnos pIJ8600NOS.

- 25 [0073] For stNOS to be responsible for thaxtomin nitration, stNOS should have NO synthase activity. Further, this activity and thaxtomin production should be arginine dependent. As predicted, overexpressed stNOS in *E. coli* produced nitrite from N-hydroxy-L-arginine in a standard NOS assay (Figures 4A-4B). N-hydroxy-L-arginine is an intermediate unique to the NOS catalytic reaction and nitrite represents the end product of NO reacting in oxygenated solution. NO synthase activity of the recombinant stNOS was inhibited by the selective NOS
- 30

inhibitor nitro-L-arginine (Figure 4B); nitro-L-arginine is the active form of Nitro-L-arginine methyl ester (NAME). NOS inhibitors were also evaluated for their effect on thaxtomin A production by *S. turgidiscabies*. NAME greatly suppressed thaxtomin production by *S. turgidiscabies* ($IC_{50} = 15 \mu M$) but did not affect bacterial growth (Figure 4C). Inhibition of thaxtomin production occurred at concentrations commensurate with the binding affinities of NAME to mammalian NOSs (Alderton et al., *Biochem J.* 357:593-615 (2001); and Southan et al., *Biochem. Pharm.* 51,383-394 (1996), which are hereby incorporated by reference in their entirety). The other three NOS inhibitors tested also suppressed thaxtomin A production but to a lesser extent than NAME. Thus, inhibitors specific for the conserved NOS active center greatly curtail both NO synthase activity from recombinant stNOS *in vitro* and thaxtomin production *in vivo*.

[0074] Mammalian NOSs convert a terminal guanidino nitrogen of the L-Arg substrate to an oxidized nitrogen species. One model for the role of stNOS in nitration of thaxtomin predicts that the nitrate nitrogen derives from a terminal guanidino nitrogen of L-Arg. A ^{15}N feeding study was conducted; either L-Arg-guanidino- $^{15}N_2 \cdot HCl$ or $^{15}NH_4NO_3$ was added to cultures of *S. turgidiscabies* under conditions that induce thaxtomin production. ^{15}N -NMR analysis of thaxtomin A extracted from cultures fed the L-Arg-guanidino- $^{15}N_2 \cdot HCl$ detected label only at the 4-nitro position, while in thaxtomin A extracted from control cultures fed $^{15}NH_4NO_3$, all of the four nitrogens in thaxtomin were equally labeled (Figures 5A-5B). Electron impact ionization mass spectra of thaxtomin A also indicated specific incorporation of ^{15}N into the nitro group when the cultures were fed L-Arg-guanidino- $^{15}N_2 \cdot HCl$. There is no known enzymatic process other than NOS activity that converts L-Arg terminal guanidino nitrogen to an oxidized nitrogen species capable of nitration. These results provide definitive evidence for the role of stNOS in nitration of thaxtomin.

[0075] The specific nitration of a tryptophanyl moiety for biosynthesis is an unprecedented metabolic role for a NOS protein, although mammalian NOS production of nitroxyl (NO^-) and superoxide (O_2^-) is well characterized and such products could readily lead to nitrating species (Alderton et al., *Biochem J.* 357:593-615 (2001), which is hereby incorporated by reference in its entirety). Biosynthetic nitration reactions are rare and usually involve the oxidation of an

amine (Carter et al, *J. Chem. Soc. Chem. Commun.* 17:1271-1273 (1989), which is hereby incorporated by reference in its entirety). The chemical mechanism of a NOS-mediated nitration may be complex because NO is unlikely to react directly with indole (Patel et al., *Biochemica. Biophys. Acta.* 1411:385-400 (1999);
5 Koppenol, *Free Rad. Biol. Med.* 25:385-391 (1998); Hughes, *Biochim. Biophys. Acta* 1411:263-272 (1999); and Ridd, *Acta Chemica. Scand.* 52:11-22 (1998), which are hereby incorporated by reference in their entirety). Nevertheless, readily oxidized forms of NO, such as nitrosonium (NO^+), peroxynitrite (ONOO^-), nitronium (NO_2^+), or nitrogen dioxide (NO_2) actively nitrate aromatic amino acids
10 (Patel et al., *Biochemica. Biophys. Acta.* 1411:385-400 (1999); Koppenol, *Free Rad. Biol. Med.* 25:385-391 (1998); Hughes, *Biochim. Biophys. Acta* 1411:263-272 (1999); and Ridd, *Acta Chemica. Scand.* 52:11-22 (1998), which are hereby incorporated by reference in their entirety), and such reactions have been implicated in mammalian signaling (Packer, *Methods Enzymol.* 268 (1996); and
15 Ischiropoulos, *Free Rad. Biol. Med.* 33 (2002), which are hereby incorporated by reference in their entirety). Given the high reactivities of these nitrogen oxide species towards tryptophan, the substrate nitrated likely contains indole; whether this substrate is tryptophan, an immediate precursor of thaxtomin, or some other intermediate, is currently under investigation. Interestingly, tryptophan nitration
20 by peroxynitrite results primarily in modification at the indole 6-position (Alvarez et al., *Chem. Res. Toxicol.* 9:390-396 (1996), which is hereby incorporated by reference in its entirety). Thus, production of a 4-nitrotryptophanyl moiety (Figure 1A) suggests some enzymatic control of the nitration reaction. Nevertheless, minimal thaxtomin production in the Δnos strain indicates that
25 diffusible nitration agents produced by other cellular processes can react productively with a thaxtomin precursor.

[0076] Conservation of residues in the heme pocket, pterin site, and substrate access channel among the bacterial NOSs suggest a common function for prokaryotic enzymes and distinguish them from their mammalian counterparts
0 (Figure 2). A distinguishing catalytic feature between mammalian NOS and bsNOS is that the bacterial protein retains product NO coordinated to its heme iron 10-20X longer than the mammalian enzymes (Adak et al., *J. Biol. Chem.* 277:16167-16171 (2002), which is hereby incorporated by reference in its

entirety)). This correlates to a more sequestered immediate heme pocket that appears conserved in stNOS (Pant et al., *Biochemistry* 41:11071-11079 (2002), which is hereby incorporated by reference in its entirety). Perhaps a slower NO release allows bacterial proteins to direct NO reactivity by either sequestering NO close to the site of substrate binding, or by facilitating reaction with O₂. It is suggested that other bacterial NOSs may also participate in the biosynthesis of secondary metabolites through nitration of an amino acid or peptide substrate. Truncated NOSs have only been discovered in a subset of Gram-positive genomes and are absent from Gram-negative genomes, supporting a role for NOS proteins in secondary metabolism, rather than a signaling function essential for vitality. Nitrated compounds are produced by other bacteria and fungi (Carter et al, *J. Chem. Soc. Chem. Commun.* 17:1271-1273 (1989); Jalal et al., *Acta Crystallography* C42:733-738 (1986); and Ohba et al., *J. Antibiot.* 40:709-713 (1987), which are hereby incorporated by reference in their entirety) and at least one is likely to derive from an aromatic nitration reaction (Carter et al, *J. Chem. Soc. Chem. Commun.* 17:1271-1273 (1989), which is hereby incorporated by reference in its entirety). NOS-initiated nitrations would produce a class of compounds with unique chemical reactivity and diverse biological activities. From an evolutionary perspective, it is interesting that homologous enzymes are responsible for biosynthetic nitration reactions in bacteria and NO signaling in mammals.

[0077] Although preferred embodiments have been depicted and described in detail herein, it will be apparent to those skilled in the relevant art that various modifications, additions, substitutions, and the like can be made without departing from the spirit of the invention and these are therefore considered to be within the scope of the invention as defined in the claims which follow.

WHAT IS CLAIMED:

1. An isolated nucleic acid molecule encoding a nitric oxide synthase, wherein the nucleic acid molecule:
 - (i) comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5;
 - (ii) comprises a nucleotide sequence that hybridizes to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5 under stringent conditions characterized by a hybridization medium comprising about 5X SSC at a temperature of about 55°C;
 - (iii) comprises a nucleotide sequence having greater than 95 percent homology to a nucleic acid according to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5;
 - (iv) comprises a nucleotide sequence that encodes a nitric oxide synthase protein having an amino acid sequence that is at least 85 percent similar to either SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 by basic BLAST using default parameters analysis;
 - (v) encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6;
 - (vi) encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid motif selected from the group consisting of the motif of SEQ ID NO:7, the motif of SEQ ID NO:8, the motif of SEQ ID NO:9, the motif of SEQ ID NO:10, the motif of SEQ ID NO:11, the motif of SEQ ID NO:12, the motif of SEQ ID NO:13, the motif of SEQ ID NO:14, the motif of SEQ ID NO:15, the motif of SEQ ID NO:16, the motif of SEQ ID NO:17, the motif of SEQ ID NO:18, the motif of SEQ ID NO:19, the motif of SEQ ID NO:20, the motif of SEQ ID NO:21, the motif of SEQ ID NO:22, the motif of SEQ ID NO:23, the motif of SEQ ID NO:24, the motif of SEQ ID NO:25, the motif of SEQ ID NO:26, the motif of SEQ ID NO:27, the motif of SEQ ID NO:28, and combinations thereof; or
 - (vii) encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid sequence of SEQ ID NO:47.

- 36 -

2. The isolated nucleic acid molecule according to claim 1, wherein the nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5.

3. The isolated nucleic acid molecule according to claim 1, wherein the nucleic acid molecule comprises a nucleotide sequence that hybridizes to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5 under stringent conditions characterized by a hybridization medium comprising about 5X SSC at a temperature of about 55°C.

4. The isolated nucleic acid molecule according to claim 1, wherein the nucleic acid molecule comprises a nucleotide sequence having greater than 95 percent homology to a nucleic acid according to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5.

5. The isolated nucleic acid molecule according to claim 1, wherein the nucleic acid molecule comprises a nucleotide sequence that encodes a nitric oxide synthase protein having an amino acid sequence that is at least 85 percent similar to either SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 by basic BLAST using default parameters analysis

6. The isolated nucleic acid molecule according to claim 1, wherein the nucleic acid molecule encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6.

7. The isolated nucleic acid molecule according to claim 1, wherein the nucleic acid molecule encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid motif selected from the group consisting of the motif of SEQ ID NO:7, the motif of SEQ ID NO:8, the motif of SEQ ID NO:9, the motif of SEQ ID NO:10, the motif of SEQ ID NO:11, the motif of SEQ ID NO:12, the motif of SEQ ID NO:13, the motif of SEQ ID NO:14, the motif of SEQ ID NO:15, the motif of SEQ ID NO:16, the motif of SEQ ID NO:17, the motif of SEQ ID NO:18, the motif of SEQ ID NO:19, the motif of SEQ ID NO:20, the motif of SEQ ID NO:21, the motif of SEQ ID NO:22, the motif of SEQ ID NO:23, the motif of SEQ ID NO:24, the motif of SEQ ID

NO:25, the motif of SEQ ID NO:26, the motif of SEQ ID NO:27, the motif of SEQ ID NO:28, and combinations thereof.

8. The isolated nucleic acid molecule according to claim 1, wherein the nucleic acid molecule encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid sequence of SEQ ID NO:47.

9. The isolated nucleic acid molecule according to claim 1, wherein the nitric oxide synthase is from a *Streptomyces* species selected from the group consisting of *Streptomyces acidiscabies*, *Streptomyces scabies*, *Streptomyces turgidiscabies*, *Streptomyces avermitilis*, and *Streptomyces ipomoea*.

10. The isolated nucleic acid molecule according to claim 1, wherein the nucleic acid is DNA.

11. An expression system comprising a vector into which is inserted the nucleic acid molecule according to claim 1.

12. The expression system according to claim 11, wherein the nucleic acid molecule is inserted in sense orientation.

13. A host cell comprising the nucleic acid molecule according to claim 1.

14. The host cell according to claim 13, wherein the host cell is selected from the group consisting of a bacterial cell, a yeast cell, a fungal cell, an insect cell, a plant cell, and a mammalian cell.

15. The host cell according to claim 14, wherein the bacterial cell is *Streptomyces*, *Bacillus*, *Escherichia*, *Brevibacterium*, *Microbacterium*, *Nocardia*, or *Rhodococcus*.

16. An isolated protein or polypeptide encoded by the nucleic acid molecule according to claim 1.

17. The isolated protein or polypeptide according to claim 16, wherein the protein or polypeptide comprises an amino acid sequence that is at least 85 percent similar to either SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 by basic BLAST using default parameters analysis.

18. The isolated protein or polypeptide according to claim 16, wherein the protein or polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6.

19. The isolated protein or polypeptide according to claim 16, wherein the protein or polypeptide comprises an amino acid motif selected from the group consisting of the motif of SEQ ID NO:7, the motif of SEQ ID NO:8, the motif of SEQ ID NO:9, the motif of SEQ ID NO:10, the motif of SEQ ID NO:11, the motif of SEQ ID NO:12, the motif of SEQ ID NO:13, the motif of SEQ ID NO:14, the motif of SEQ ID NO:15, the motif of SEQ ID NO:16, the motif of SEQ ID NO:17, the motif of SEQ ID NO:18, the motif of SEQ ID NO:19, the motif of SEQ ID NO:20, the motif of SEQ ID NO:21, the motif of SEQ ID NO:22, the motif of SEQ ID NO:23, the motif of SEQ ID NO:24, the motif of SEQ ID NO:25, the motif of SEQ ID NO:26, the motif of SEQ ID NO:27, the motif of SEQ ID NO:28, and combinations thereof.

20. The isolated protein or polypeptide according to claim 16, wherein the protein or polypeptide comprises an amino acid sequence of SEQ ID NO:47.

21. A method of attaching a nitrogen group to a target moiety of a compound, said method comprising:

- providing a nitric oxide synthase;
- providing a compound having a target moiety; and
- combining the nitric oxide synthase and the compound in a reaction mixture under conditions effective to allow the nitric oxide synthase to catalyze a reaction whereby a nitrogen group contained in the reaction mixture is attached to the target

moiety, thereby yielding a nitrogen-modified compound, wherein the nitric oxide synthase comprises:

- (i) a protein or polypeptide having an amino acid sequence that is at least 85 percent similar to either SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 by basic BLAST using default parameters analysis;
- (ii) a protein or polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6;
- (iii) a protein or polypeptide having an amino acid motif selected from the group consisting of the motif of SEQ ID NO:7, the motif of SEQ ID NO:8, the motif of SEQ ID NO:9, the motif of SEQ ID NO:10, the motif of SEQ ID NO:11, the motif of SEQ ID NO:12, the motif of SEQ ID NO:13, the motif of SEQ ID NO:14, the motif of SEQ ID NO:15, the motif of SEQ ID NO:16, the motif of SEQ ID NO:17, the motif of SEQ ID NO:18, the motif of SEQ ID NO:19, the motif of SEQ ID NO:20, the motif of SEQ ID NO:21, the motif of SEQ ID NO:22, the motif of SEQ ID NO:23, the motif of SEQ ID NO:24, the motif of SEQ ID NO:25, the motif of SEQ ID NO:26, the motif of SEQ ID NO:27, the motif of SEQ ID NO:28, the motif of SEQ ID NO:29, the motif of SEQ ID NO:30, the motif of SEQ ID NO:31, the motif of SEQ ID NO:32, the motif of SEQ ID NO:33, the motif of SEQ ID NO:34, the motif of SEQ ID NO:35, the motif of SEQ ID NO:36, the motif of SEQ ID NO:37, the motif of SEQ ID NO:38, the motif of SEQ ID NO:39, the motif of SEQ ID NO:40, the motif of SEQ ID NO:41, the motif of SEQ ID NO:42, the motif of SEQ ID NO:43, the motif of SEQ ID NO:44, the motif of SEQ ID NO:45, the motif of SEQ ID NO:46, and combinations thereof combinations thereof; or
- (iv) a protein or polypeptide comprises an amino acid sequence of SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, or SEQ ID NO:50.

22. The method according to claim 21, wherein the nitric oxide synthase comprises a protein or polypeptide having an amino acid sequence that is at least 85 percent similar to either SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 by basic BLAST using default parameters analysis.

23. The method according to claim 21, wherein the nitric oxide synthase comprises a protein or polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6.

24. The method according to claim 21, wherein the nitric oxide synthase comprises a protein or polypeptide having an amino acid motif selected from the group consisting of the motif of SEQ ID NO:7, the motif of SEQ ID NO:8, the motif of SEQ ID NO:9, the motif of SEQ ID NO:10, the motif of SEQ ID NO:11, the motif of SEQ ID NO:12, the motif of SEQ ID NO:13, the motif of SEQ ID NO:14, the motif of SEQ ID NO:15, the motif of SEQ ID NO:16, the motif of SEQ ID NO:17, the motif of SEQ ID NO:18, the motif of SEQ ID NO:19, the motif of SEQ ID NO:20, the motif of SEQ ID NO:21, the motif of SEQ ID NO:22, the motif of SEQ ID NO:23, the motif of SEQ ID NO:24, the motif of SEQ ID NO:25, the motif of SEQ ID NO:26, the motif of SEQ ID NO:27, the motif of SEQ ID NO:28, the motif of SEQ ID NO:29, the motif of SEQ ID NO:30, the motif of SEQ ID NO:31, the motif of SEQ ID NO:32, the motif of SEQ ID NO:33, the motif of SEQ ID NO:34, the motif of SEQ ID NO:35, the motif of SEQ ID NO:36, the motif of SEQ ID NO:37, the motif of SEQ ID NO:38, the motif of SEQ ID NO:39, the motif of SEQ ID NO:40, the motif of SEQ ID NO:41, the motif of SEQ ID NO:42, the motif of SEQ ID NO:43, the motif of SEQ ID NO:44, the motif of SEQ ID NO:45, the motif of SEQ ID NO:46, and combinations thereof combinations thereof.

25. The method according to claim 21, wherein the nitric oxide synthase comprises a protein or polypeptide comprises an amino acid sequence of SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, or SEQ ID NO:50.

26. The method according to claim 21, wherein the nitrogen group is a nitro group or a nitroso group.

27. The method according to claim 21, wherein the reaction is a nitration reaction or a nitrosylation reaction.

28. The method according to claim 21, wherein the nitrogen-modified compound is a secondary metabolite.

29. The method according to claim 28, wherein the secondary metabolite is selected from the group consisting of an alkaloid, a terpenoid, an aliphatic organic acid, an aromatic organic acid, a heteroaromatic organic acid, a phenol, an irridoid, a steroid, a volatile oil, a resin, a balsam, and a saponin.

30. The method according to claim 21, wherein the nitrogen-modified compound is a polyketide.

31. The method according to claim 21, wherein the nitrogen-modified compound is a peptide.

32. The method according to claim 31, wherein the peptide is a non-ribosomal peptide.

33. The method according to claim 21, wherein the nitrogen-modified compound is a phytotoxin.

34. The method according to claim 33, wherein the phytotoxin is thaxtomin.

35. The method according to claim 21, wherein the nitrogen-modified compound is an herbicide, an antibiotic, a nematocide, a fungicide, an insecticide, or an anticancer agent.

36. The method according to claim 35, wherein the antibiotic is chloramphenicol.

37. The method according to claim 35, wherein the anticancer agent is L-alanosine.

- 42 -

38. The method according to claim 21, wherein the nitrogen-modified compound is a high-energy material selected from the group consisting of nitroglycerin, trinitrotoluene, pentaerythritoltetranitrate, cyclotrimethylenetrinitramine, and ammonium nitrate fertilizer.

39. The method according to claim 21, wherein the target moiety comprises an electron-rich aromatic group.

40. The method according to claim 21, wherein the target moiety comprises a phenyl moiety.

41. The method according to claim 21, wherein the target moiety comprises an indole moiety.

42. The method according to claim 21, wherein the target moiety comprises a tryptophanyl moiety.

43. A method of synthesizing a nitrogen-modified compound in a transgenic host cell, said method comprising:

providing a transgenic host cell transformed with a DNA molecule encoding a nitric oxide synthase;

providing a non-modified compound containing a target moiety; and

culturing the transgenic host cell and the non-modified compound in a culture medium under conditions effective to allow the nitric oxide synthase to be expressed and to catalyze a reaction whereby a nitrogen group contained in the host cell or culture medium attaches to the target moiety, thereby yielding a nitrogen-modified compound, wherein the DNA molecule:

(i) comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5;

(ii) comprises a nucleotide sequence that hybridizes to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5 under stringent conditions characterized by a hybridization medium comprising about 5X SSC at a temperature of about 55°C;

- 43 -

- (iii) comprises a nucleotide sequence having greater than 95 percent homology to a nucleic acid according to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5;
- (iv) comprises a nucleotide sequence that encodes a nitric oxide synthase protein having an amino acid sequence that is at least 85 percent similar to either SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 by basic BLAST using default parameters analysis;
- (v) encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6;
- (vi) encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid motif selected from the group consisting of the motif of SEQ ID NO:7, the motif of SEQ ID NO:8, the motif of SEQ ID NO:9, the motif of SEQ ID NO:10, the motif of SEQ ID NO:11, the motif of SEQ ID NO:12, the motif of SEQ ID NO:13, the motif of SEQ ID NO:14, the motif of SEQ ID NO:15, the motif of SEQ ID NO:16, the motif of SEQ ID NO:17, the motif of SEQ ID NO:18, the motif of SEQ ID NO:19, the motif of SEQ ID NO:20, the motif of SEQ ID NO:21, the motif of SEQ ID NO:22, the motif of SEQ ID NO:23, the motif of SEQ ID NO:24, the motif of SEQ ID NO:25, the motif of SEQ ID NO:26, the motif of SEQ ID NO:27, the motif of SEQ ID NO:28, the motif of SEQ ID NO:29, the motif of SEQ ID NO:30, the motif of SEQ ID NO:31, the motif of SEQ ID NO:32, the motif of SEQ ID NO:33, the motif of SEQ ID NO:34, the motif of SEQ ID NO:35, the motif of SEQ ID NO:36, the motif of SEQ ID NO:37, the motif of SEQ ID NO:38, the motif of SEQ ID NO:39, the motif of SEQ ID NO:40, the motif of SEQ ID NO:41, the motif of SEQ ID NO:42, the motif of SEQ ID NO:43, the motif of SEQ ID NO:44, the motif of SEQ ID NO:45, the motif of SEQ ID NO:46, and combinations thereof combinations thereof; or
- (vii) encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid sequence of SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, or SEQ ID NO:50.

44. The method according to claim 43, wherein the DNA molecule comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5.

45. The method according to claim 43, wherein the DNA molecule comprises a nucleotide sequence that hybridizes to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5 under stringent conditions characterized by a hybridization medium comprising about 5X SSC at a temperature of about 55°C.

46. The method according to claim 43, wherein the DNA molecule comprises a nucleotide sequence having greater than 95 percent homology to a nucleic acid according to SEQ ID NO:1, SEQ ID NO:3, or SEQ ID NO:5.

47. The method according to claim 43, wherein the DNA molecule comprises a nucleotide sequence that encodes a nitric oxide synthase protein having an amino acid sequence that is at least 85 percent similar to either SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 by basic BLAST using default parameters analysis

48. The method according to claim 43, wherein the DNA molecule encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:6.

49. The method according to claim 43, wherein the DNA molecule encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid motif selected from the group consisting of the motif of SEQ ID NO:7, the motif of SEQ ID NO:8, the motif of SEQ ID NO:9, the motif of SEQ ID NO:10, the motif of SEQ ID NO:11, the motif of SEQ ID NO:12, the motif of SEQ ID NO:13, the motif of SEQ ID NO:14, the motif of SEQ ID NO:15, the motif of SEQ ID NO:16, the motif of SEQ ID NO:17, the motif of SEQ ID NO:18, the motif of SEQ ID NO:19, the motif of SEQ ID NO:20, the motif of SEQ ID NO:21, the motif of SEQ ID NO:22, the motif of SEQ ID NO:23, the motif of SEQ ID NO:24, the motif of SEQ ID NO:25, the motif of SEQ ID NO:26, the motif of SEQ ID NO:27, the motif of SEQ ID NO:28, the motif of SEQ ID NO:29, the motif of SEQ ID NO:30, the motif of SEQ ID NO:31, the motif of SEQ ID NO:32, the motif of SEQ ID NO:33, the motif of SEQ ID NO:34, the motif of SEQ ID NO:35, the motif of SEQ ID NO:36, the motif of SEQ ID NO:37, the motif of SEQ ID NO:38, the motif of SEQ ID NO:39, the motif of SEQ ID NO:40, the motif of SEQ ID

- 45 -

NO:41, the motif of SEQ ID NO:42, the motif of SEQ ID NO:43, the motif of SEQ ID NO:44, the motif of SEQ ID NO:45, the motif of SEQ ID NO:46, and combinations thereof.

50. The method according to claim 43, wherein the nucleic acid molecule encodes a nitric oxide synthase comprising a protein or polypeptide having an amino acid sequence of SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, or SEQ ID NO:50.

51. The method according to claim 43, wherein the nitrogen group is a nitro group or a nitroso group.

52. The method according to claim 43, wherein the reaction is a nitration reaction or a nitrosylation reaction.

53. The method according to claim 43, wherein the nitrogen-modified compound is a secondary metabolite.

54. The method according to claim 53, wherein the secondary metabolite is selected from the group consisting of an alkaloid, a terpenoid, an aliphatic organic acid, an aromatic organic acid, a heteroaromatic organic acid, a phenol, an iridoid, a steroid, a volatile oil, a resin, a balsam, a β -lactam, an aminoglycoside, a macrolide, a tetracycline, and a saponin.

55. The method according to claim 43, wherein the nitrogen-modified compound is a polyketide.

56. The method according to claim 43, wherein the nitrogen-modified compound is a peptide.

57. The method according to claim 56, wherein the peptide is a non-ribosomal peptide.

- 46 -

58. The method according to claim 43, wherein the nitrogen-modified compound is a phytotoxin.

59. The method according to claim 58, wherein the phytotoxin is thaxtomin.

60. The method according to claim 43, wherein the nitrogen-modified compound is an herbicide, an antibiotic, or an anticancer agent.

61. The method according to claim 60, wherein the antibiotic is chloramphenicol.

62. The method according to claim 60, wherein the anticancer agent is L-alanosine.

63. The method according to claim 43, wherein the nitrogen-modified compound is a high-energy material selected from the group consisting of nitroglycerin, trinitrotoluene, pentaerythritoltetranitrate, cyclotrimethylenetrinitramine, and ammonium nitrate fertilizer.

64. The method according to claim 43, wherein the target moiety comprises an electron-rich aromatic group.

65. The method according to claim 43, wherein the target moiety comprises a phenyl moiety.

66. The method according to claim 43, wherein the target moiety comprises an indole moiety.

67. The method according to claim 43, wherein the target moiety comprises a tryptophanyl moiety.

68. The method according to claim 43, wherein the non-modified compound containing a target moiety is provided exogenously.

69. The method according to claim 43, wherein the non-modified compound containing a target moiety is produced by the transgenic host cell.

70. A method of recombinantly producing a nitric oxide synthase in a host cell, said method comprising:

transforming a host cell with at least one nucleic acid molecule according to claim 1, said transforming being performed under conditions suitable for expression of the nitric oxide synthase; and

isolating the nitric oxide synthase.

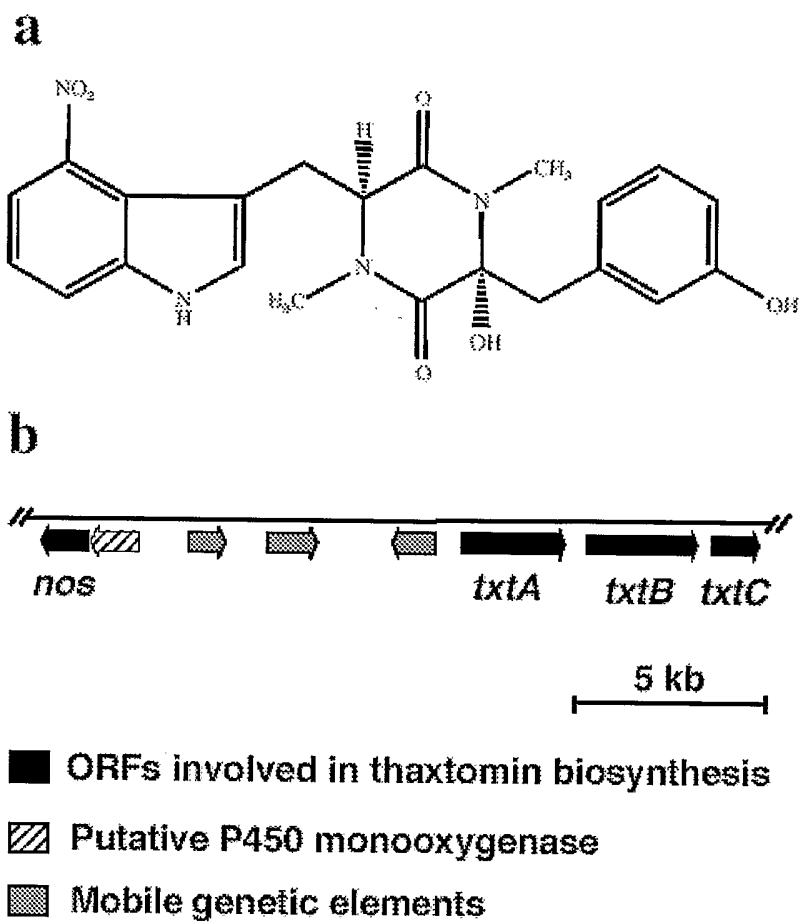
71. The method according to claim 70, wherein the host cell is selected from the group consisting of a bacterial cell, a yeast cell, a fungal cell, an insect cell, a plant cell, and a mammalian cell.

72. The method according to claim 71, wherein the bacterial cell is *Streptomyces*, *Bacillus*, *Escherichia*, *Brevibacterium*, *Microbacterium*, *Nocardia*, or *Rhodococcus*.

73. The method according to claim 71, wherein the yeast cell is *Saccharomyces*.

74. The method according to claim 71, wherein the fungal cell is *Aspergillus*, *Cephalosporium*, or *Penicillium*.

1/7

**FIG. 1**

2/7

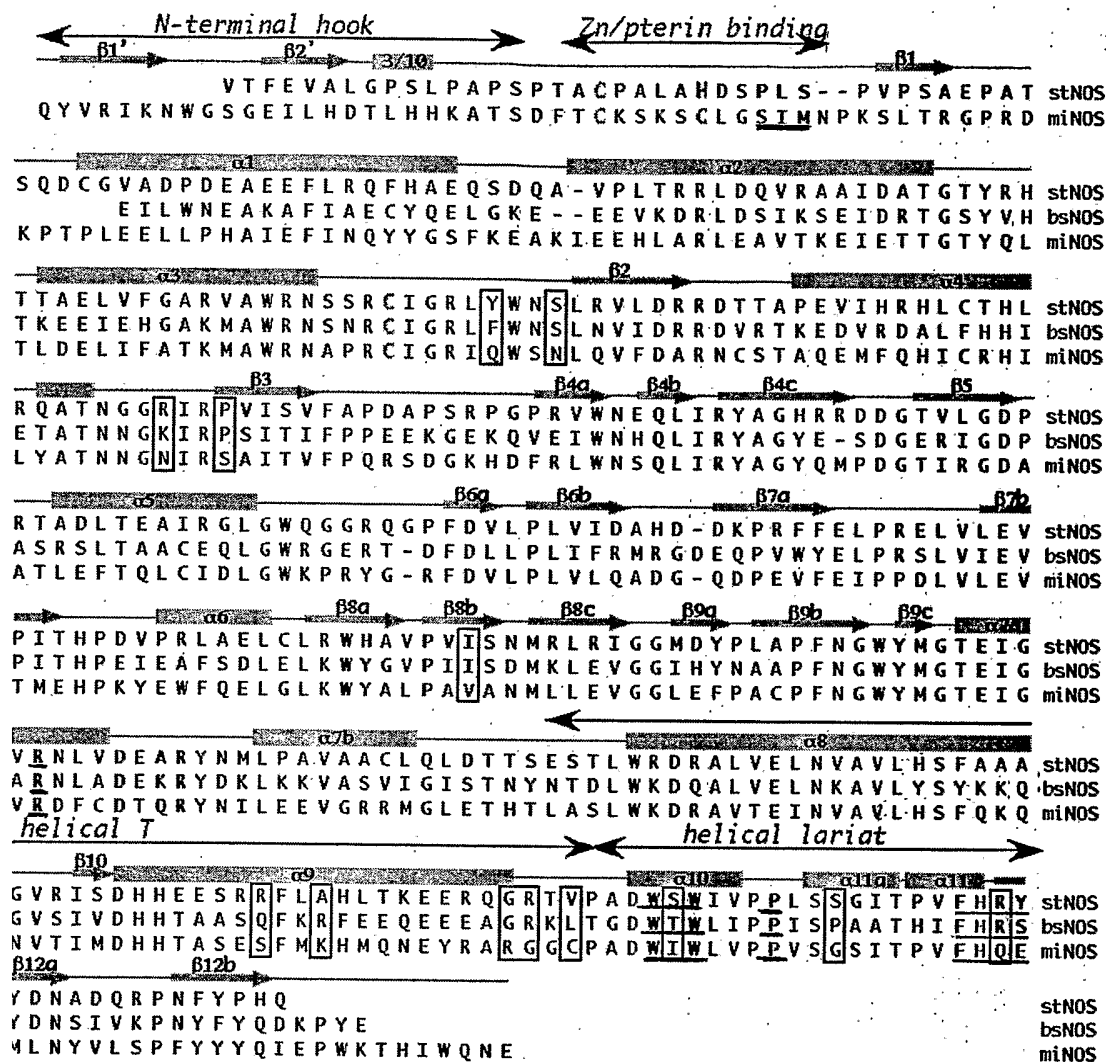
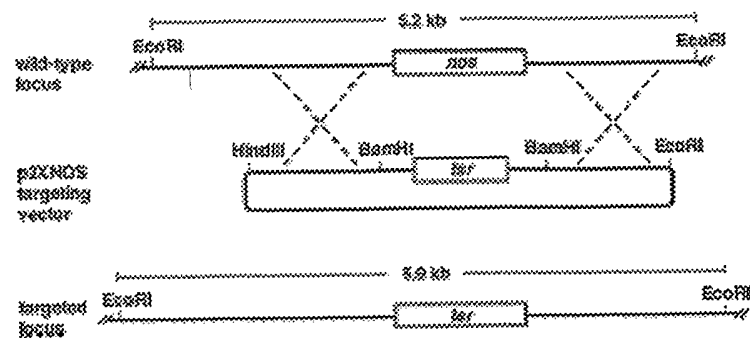


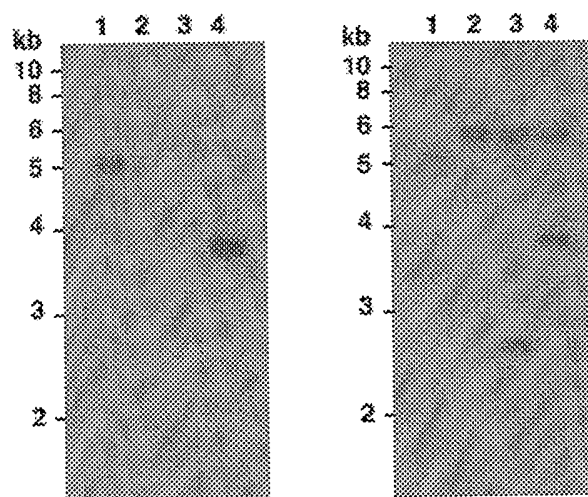
FIG. 2

3/7

a



b



c



FIG. 3

4/7

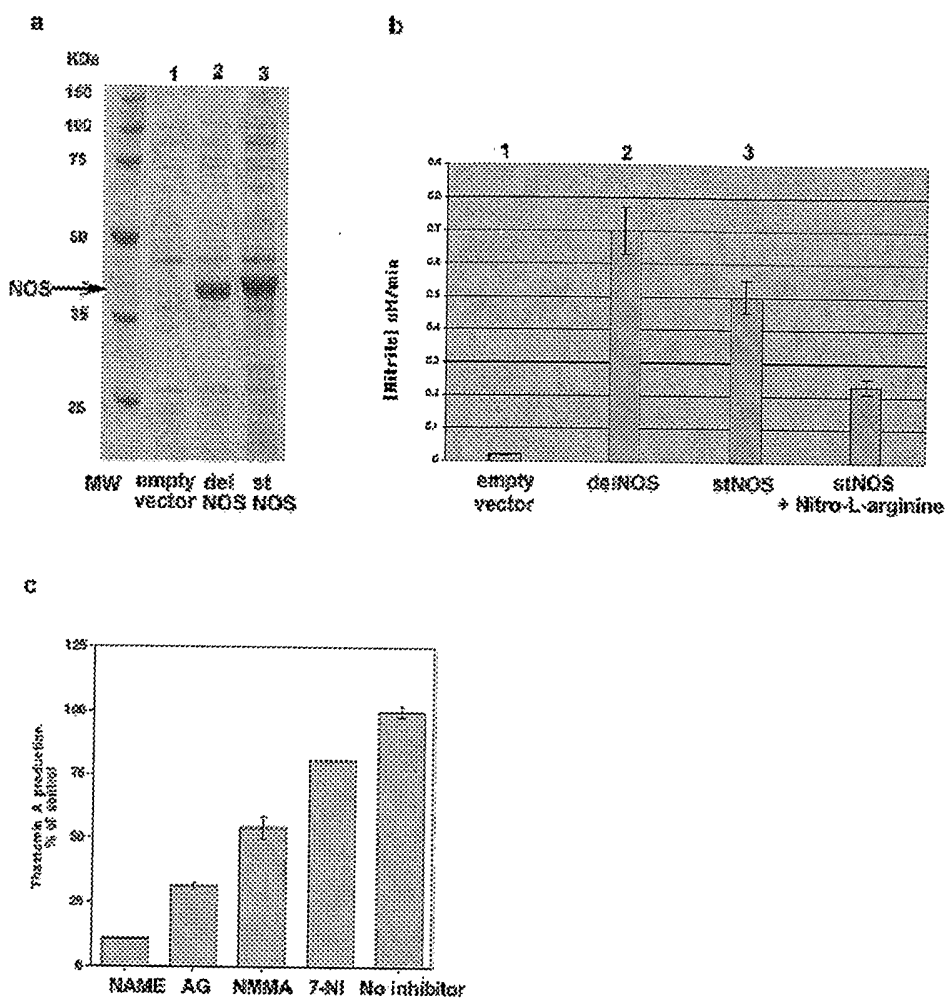
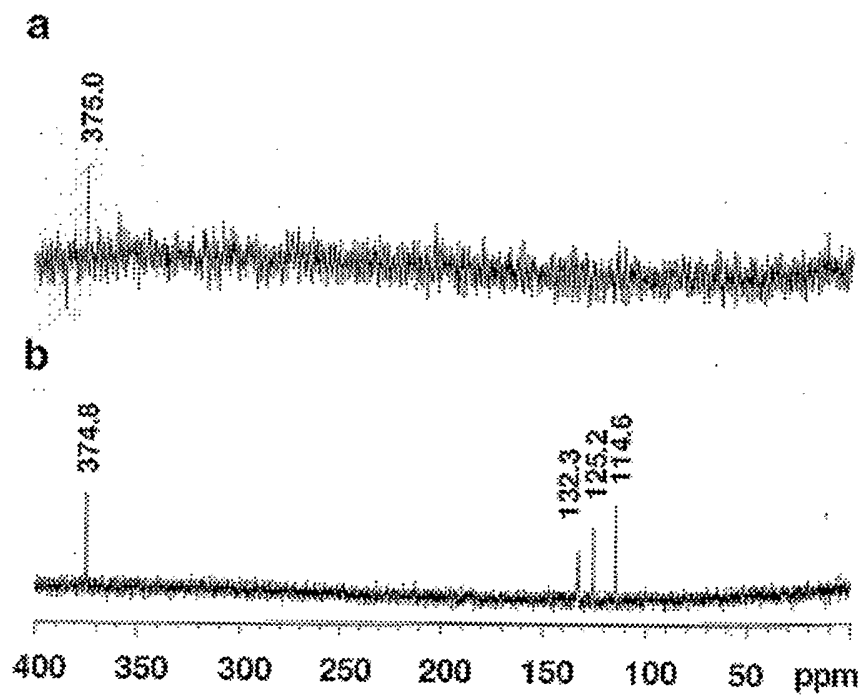


FIG. 4

8/7

**FIG. 5**

		1	
Bacillus anthracis (SEQ ID NO:51)	(1)	-----	50
Bacillus cereus (SEQ ID NO:52)	(1)	-----	
Staphylococcus aureus (SEQ ID NO:53)	(1)	-----	
Staphylococcus epidermidis (SEQ ID NO:54)	(1)	-----	
Bacillus halodurans (SEQ ID NO:55)	(1)	-----	
Bacillus subtilis (SEQ ID NO:56)	(1)	-----	
Streptomyces radiodurans (SEQ ID NO:57)	(1)	-----	
Streptomyces acidiscabies (SEQ ID NO:2)	(1)	-----	
Streptomyces scabies (SEQ ID NO:4)	(1)	-----	MTSE
Streptomyces turgidiscabies (SEQ ID NO:6)	(1)	-----	MTSE
Streptomyces avermitilis (SEQ ID NO:58)	(1)	-----	MTFE
Mus musculus (SEQ ID NO:59)	(1)	LADAPPESAASRPGGDARDRHGAPSRARGWSTETWCAPHGRKDTAGESEH	
		-----	QYVRILKNWGSGEI
		51	
Bacillus anthracis (SEQ ID NO:51)	(1)	-----	100
Bacillus cereus (SEQ ID NO:52)	(1)	-----	MSKTKQLIEEASH
Staphylococcus aureus (SEQ ID NO:53)	(1)	-----	MSKTKQLIEEASN
Staphylococcus epidermidis (SEQ ID NO:54)	(1)	-----	MLFKEAQA
Bacillus halodurans (SEQ ID NO:55)	(1)	-----	MLIDKARS
Bacillus subtilis (SEQ ID NO:56)	(1)	-----	MEEKERLQLEAES
Streptomyces radiodurans (SEQ ID NO:57)	(1)	-----	MSCPAAAVLTTPDMRA
Streptomyces acidiscabies (SEQ ID NO:2)	(5)	VALGPSLPAPSPPTACPALGPDSSLGP--VPSAEPATPQSCGVADPNEAEE	
Streptomyces scabies (SEQ ID NO:4)	(5)	VALGPSLPAPSPPTACPALGPDSSLGP--VPSAEPATPQSCGVADPNEAEE	
Streptomyces turgidiscabies (SEQ ID NO:6)	(5)	VALGPSLPAPSPPTACPALGPDSSLGP--VPSAEPATPQSCGVADPNEAEE	
Streptomyces avermitilis (SEQ ID NO:58)	(51)	RPPGPTLPARGWGAARTRGDRPDGHPDHPPTGSPPTDGPAPRPDRREE	
Mus musculus (SEQ ID NO:59)	(14)	LHDTLHHKATSDFTCKSKCLGSIINPKSLTRGPRDKPTPLEELLPHAI	
		101	
Bacillus anthracis (SEQ ID NO:51)	(14)	FITICYKELS-----	150
Bacillus cereus (SEQ ID NO:52)	(14)	FITICYKELH-----	KEHFIEERMKEIQAEIE
Staphylococcus aureus (SEQ ID NO:53)	(9)	FIENMYKECHY-----	KEQLIEERIKIEIQIEIE
Staphylococcus epidermidis (SEQ ID NO:54)	(9)	FIQTMYSSELKY-----	ETQIINKRLHDIIELEIK
Bacillus halodurans (SEQ ID NO:55)	(14)	FLTKCYEELG-----	NTNEIENRMKEIEQIEIN
Bacillus subtilis (SEQ ID NO:56)	(1)	-----	STGELSKRLEEVRKEID
Streptomyces radiodurans (SEQ ID NO:57)	(16)	FLRRFHEEMGE-----	MKDRADIKSEID
Streptomyces acidiscabies (SEQ ID NO:2)	(53)	FLRQFHAEQSD-----	PG-LPARLRAVE
Streptomyces scabies (SEQ ID NO:4)	(53)	FLRQFHAEQSD-----	QPVLARRLEQVRAAID
Streptomyces turgidiscabies (SEQ ID NO:6)	(53)	FLRQFHAEQSD-----	QPVLARRLEQVRAAID
Streptomyces avermitilis (SEQ ID NO:58)	(101)	GDGHDGHEGQDGHDLCAATAFLTLHHTEERLGDPAARRIAAAHAEIA	
Mus musculus (SEQ ID NO:59)	(64)	FINQYYSFKEA-----	KIEEHLARLEAVTKEIE
		151	
Bacillus anthracis (SEQ ID NO:51)	(41)	KTGYEHTFEELVHGSMAWRNSNRCIGRLFWSKMHILDAREVNDEEGVY	200
Bacillus cereus (SEQ ID NO:52)	(41)	KTGYEHTFEELVHGSMAWRNSNRCIGRLFWSKMHILDAREVNDEEGVY	
Staphylococcus aureus (SEQ ID NO:53)	(37)	ETGTYHTTEELIYGAKMAWRNSNRCIGRLFWDLSNVIDARDVTDEASFL	
Staphylococcus epidermidis (SEQ ID NO:54)	(37)	LTGSYHTTYEELS YGAKMAWRNSNRCIGRLFWSNLSNVKDARDVDCDEKEFI	
Bacillus halodurans (SEQ ID NO:55)	(41)	KTGYVHTTKELAHGARMMAWRNSNRCIGRLFWSNLSNVKDARDVDCDEKEFI	
Bacillus subtilis (SEQ ID NO:56)	(14)	LTGSYVHTKEELEHGAKMAWRNSNRCIGRLFWSNLSNVKDARDVDCDEKEFI	
Streptomyces radiodurans (SEQ ID NO:57)	(38)	EAGLWPTSAELTWGAKVAWRNS TRCVGRLYWEALSVRDLRELNTAQAVY	
Streptomyces acidiscabies (SEQ ID NO:2)	(81)	ATGTYRHTTAELVYGARVAWRNS SRICIGRLYWNLSRVLDRRDATAPDEIH	
Streptomyces scabies (SEQ ID NO:4)	(81)	ATGTYRHTTAELVYGARVAWRNS SRICIGRLYWNLSRVLDRRDATAPDEIH	
Streptomyces turgidiscabies (SEQ ID NO:6)	(81)	ATGTYRHTTAELVYGARVAWRNS SRICIGRLYWNLSRVLDRRDATAPDEIH	
Streptomyces avermitilis (SEQ ID NO:58)	(151)	ETGTYRHTTEELVFGARVAWRNSNRCIGRLYWNLSRVLDRRDATAPDEIH	
Mus musculus (SEQ ID NO:59)	(93)	TTGTYQLTDELIFATKMAWRNAPRCIGRIQWSNLQVFDARNCSTAQEMF	
		201	
Bacillus anthracis (SEQ ID NO:51)	(91)	HALIHIIKYATNDGKVKPTITITFKQYQGEENNIRIYNHQLIRYAGYKTEM	250
Bacillus cereus (SEQ ID NO:52)	(91)	NALIHIIKYATNDGKVKPTITITFKQYQGEENNIRIYNHQLIRYAGYKTEM	
Staphylococcus aureus (SEQ ID NO:53)	(87)	SSITYHITQATNEGKLPYITITYAP----KDGPKEFNQLIRYAGYDN--	
Staphylococcus epidermidis (SEQ ID NO:54)	(87)	KFIHITHIKEATNGGKIPYITIFSP----EDTPKIYNQLIRYAGYEN--	
Bacillus halodurans (SEQ ID NO:55)	(91)	EALVDHIYATNDGKILPTISVFRPRHPNKGDVRIWNQLIRYAGYEEGD	
Bacillus subtilis (SEQ ID NO:56)	(64)	DALFHIIETATNNGKIRPTITIFPPEEKGEKQVEIWNQLIRYAGYESDG	
Streptomyces radiodurans (SEQ ID NO:57)	(88)	EALLQHLDDAFCGGHIRPVISVFG-----PGVRLHNPQLIRYA-----	
Streptomyces acidiscabies (SEQ ID NO:2)	(131)	RHLCTHLRQATNGGRIRPVISVFPAPDSPGRPGPVWNEQLIRYAGYRRDD	
Streptomyces scabies (SEQ ID NO:4)	(131)	RHLCTHLRQATNGGRIRPVISVFPAPDSPGRPGPVWNEQLIRYAGYRRDD	
Streptomyces turgidiscabies (SEQ ID NO:6)	(131)	RHLCTHLRQATNGGRIRPVISVFPAPDSPGRPGPVWNEQLIRYAGYRRDD	
Streptomyces avermitilis (SEQ ID NO:58)	(201)	EASADHLREATRDGRIRALITVFAPDAPGRPGPRIWNEQLIRYAGYARP	
Mus musculus (SEQ ID NO:59)	(143)	QHICRHIIYATNNGNIRSAITVFQSRSDGKHDFRLWNSQLIRYAGYQMPD	

FIG. 6A

		251	300
Bacillus anthracis (SEQ ID NO:51)	(141)	-GVTGDSHSTAFDFCQELGWQG-EGTNFDVLPVFSID-GKAPIYKEIP	
Bacillus cereus (SEQ ID NO:52)	(141)	-GVIGDSHSATFTDFCQGLGWQG-EGTNFDVLPVFSIN-GKAPTYKEIP	
Staphylococcus aureus (SEQ ID NO:53)	(131)	---CGDPAEKEVTRLANHLGWKG-KGTNFDVLPVLIYQLP-NESVKFYEYP	
Staphylococcus epidermidis (SEQ ID NO:54)	(131)	---VGDPSEKKVTRLAHLGWKG-KGSNFDILPLIYQLP-NDTIKIHELP	
Bacillus halodurans (SEQ ID NO:55)	(141)	-QVIGDPISKTFTQACERLWGS-ERTPFVLPVLIQDG-SKPPKWFAPV	
Bacillus subtilis (SEQ ID NO:56)	(114)	-ERIGDPASCSTLAACEELGWRG-ERTDFDLPLIFRMKGDEQPVWYELP	
Streptomyces radiodurans (SEQ ID NO:57)	(126)	----DDPINADFVDKLRFGWQP-RGERFEVLPVLLIEVNG--RAELFSLP	
Streptomyces acidiscabies (SEQ ID NO:2)	(181)	GTVLGDPRTADLTAILRLGWGCGPQGFVLPVLPIDTDD-KPRFFELP	
Streptomyces scabies (SEQ ID NO:4)	(181)	GTVLGDPRTADLTAILRLGWGCGPQGFVLPVLPIDTDD-KPRFFELP	
Streptomyces turgidiscabies (SEQ ID NO:6)	(181)	GTVLGDPRTADLTAILRLGWGCGPQGFVLPVLPIDTDD-KPRFFELP	
Streptomyces avermitilis (SEQ ID NO:58)	(251)	GAVTGDPNAGLTALARLWPGGPGSPFDVLPVLIQVSGD-RPRWFTLP	
Mus musculus (SEQ ID NO:59)	(193)	GTIRGDAATLEFTQLCIDLGWKP-RYGRFDVLPVLIQAD-GQDPEVFEIP	
		301	350
Bacillus anthracis (SEQ ID NO:51)	(188)	KEEVKEVPIEHPEYPISS-LGAKWYGVPMISDMRLEIGGISYTAAPFNGW	
Bacillus cereus (SEQ ID NO:52)	(188)	REEVKEVPIEHPEYPISS-LGAKWYGVPMISDMRLEIGGISYTAAPFNGW	
Staphylococcus aureus (SEQ ID NO:53)	(176)	TSLIKEVPIEHNYPKLRKLNKLYAVPIISNMMLKIGGIYPTAPFNGW	
Staphylococcus epidermidis (SEQ ID NO:54)	(176)	NDIVKEVSIHHEHYPKLKLGLKLYAVPIISNMMLKIGGIYPTAPFNGW	
Bacillus halodurans (SEQ ID NO:55)	(188)	NESVKEVPLRHEPEYEFAGFQLKLYAVPIISNMMLKIGGIYPTAPFNGW	
Bacillus subtilis (SEQ ID NO:56)	(162)	RSLVIEVPIIHPDIEAFSDLELKWYGVPIISDMKLEVGGIHYNAAPFNGW	
Streptomyces radiodurans (SEQ ID NO:57)	(169)	PQAVQVEVATHPVCLGIGELGLRWHALPVISDMMLDGLGLHLP-FCSGW	
Streptomyces acidiscabies (SEQ ID NO:2)	(230)	RELVLVPIIHPDVERLAEGLRWHAHPVISNMMLRIGGMDYPLAPFNGW	
Streptomyces scabies (SEQ ID NO:4)	(230)	RELVLVPIIHPDVERLAEGLRWHAHPVISNMMLRIGGMDYPLAPFNGW	
Streptomyces turgidiscabies (SEQ ID NO:6)	(230)	RELVLVPIIHPDVERLAEGLRWHAHPVISNMMLRIGGMDYPLAPFNGW	
Streptomyces avermitilis (SEQ ID NO:58)	(300)	EDAVLEVALTHPEYPPWRSLSGLRWHAHPALAGMCLES GGICYPAAPFNGW	
Mus musculus (SEQ ID NO:59)	(241)	PDLVLEVTMEHPKYEFQELGLKLYALPAVANMLLEVGGLEFPACPFNGW	
		351	400
Bacillus anthracis (SEQ ID NO:51)	(237)	YMGTEIGARNLADHDYRNLPAVAEMMDLTSRNGTLWKDKALIELNVAV	
Bacillus cereus (SEQ ID NO:52)	(237)	YMGTEIGARNLADHDYRNLPAVAEMMDLTSRNGTLWKDKALIELNVAV	
Staphylococcus aureus (SEQ ID NO:53)	(226)	YMVTEIGVRNFIDYRYNLEKVAFAFEFDLKNNSFNKDRALVELNVAV	
Staphylococcus epidermidis (SEQ ID NO:54)	(226)	YMVTEIAVRNFIDYRYNLEKVAFAFEFDLKNNSFNKDRALVELNVAV	
Bacillus halodurans (SEQ ID NO:55)	(238)	YMGTEIGARNLADHDYRNLPAVAEMMDLTSRNGTLWKDKALIELNVAV	
Bacillus subtilis (SEQ ID NO:56)	(212)	YMGTEIGARNLADHDYRNLPAVAEMMDLTSRNGTLWKDKALIELNVAV	
Streptomyces radiodurans (SEQ ID NO:57)	(218)	YVQTEIAARDLADVGRYDQLPAVARALGLDTSRERTLWRDRALVELNVAV	
Streptomyces acidiscabies (SEQ ID NO:2)	(280)	YMGTEIGARNLADHDYRNLPAVAEMMDLTSRNGTLWKDKALIELNVAV	
Streptomyces scabies (SEQ ID NO:4)	(280)	YMGTEIGARNLADHDYRNLPAVAEMMDLTSRNGTLWKDKALIELNVAV	
Streptomyces turgidiscabies (SEQ ID NO:6)	(280)	YMGTEIGARNLADHDYRNLPAVAEMMDLTSRNGTLWKDKALIELNVAV	
Streptomyces avermitilis (SEQ ID NO:58)	(350)	YMGTEIGARNLADHDYRNLPAVAEMMDLTSRNGTLWKDKALIELNVAV	
Mus musculus (SEQ ID NO:59)	(291)	YMGTEIGVRDFCDTQRYNILEEVGRRMGLTHTLASLWKDRAVTEINVAV	
		401	450
Bacillus anthracis (SEQ ID NO:51)	(287)	LHSFKKQGVSI DVHHTAAQFQFQFEKQEAACGRVVTGNVWVLIPLSPAT	
Bacillus cereus (SEQ ID NO:52)	(287)	LHSFKKQGVSI DVHHTAAQFQFQFEKQEAACGRVVTGNVWVLIPLSPAT	
Staphylococcus aureus (SEQ ID NO:53)	(276)	YHSFKKEGVSIVDHLTAQKFELFERNEAQGRQVTGKWSWLPPLSPATL	
Staphylococcus epidermidis (SEQ ID NO:54)	(276)	YHSFKADGVSI DVHHTAAQFQFQFEKQEAACGRVVTGNVWVLIPLSPAT	
Bacillus halodurans (SEQ ID NO:55)	(288)	LYSYKQEGVSI DVHHTAAQFQFQFEKQEAACGRVVTGNVWVLIPLSPAT	
Bacillus subtilis (SEQ ID NO:56)	(262)	LHSYKKQGVSI DVHHTAAQFQFQFEKQEAACGRVVTGNVWVLIPLSPAT	
Streptomyces radiodurans (SEQ ID NO:57)	(268)	LHSFDAAGVKLADHHTVTAHVRFEEAREAGREVGRKWSWLPPLSPAT	
Streptomyces acidiscabies (SEQ ID NO:2)	(330)	LHSFEAAGVRI SDHHEESRRFLAHLAKEERQGRVTSADWSWLPPLSGGI	
Streptomyces scabies (SEQ ID NO:4)	(330)	LHSFEAAGVRI SDHHEESRRFLAHLAKEERQGRVTSADWSWLPPLSGGI	
Streptomyces turgidiscabies (SEQ ID NO:6)	(330)	LHSFEAAGVRI SDHHEESRRFLAHLAKEERQGRVTSADWSWLPPLSGGI	
Streptomyces avermitilis (SEQ ID NO:58)	(400)	LHSFDRAGVTVDHHTESLRFLTHLDREERKGRVGVADWSWLPPLSGGI	
Mus musculus (SEQ ID NO:59)	(341)	LHSFQKQNVTIMDHTASESFMKHMQNEFYRARGGCPADWILVLPVPPVSGSI	
		451	489
Bacillus anthracis (SEQ ID NO:51)	(337)	THIYHKPYPNEILKPNFFHK-----	
Bacillus cereus (SEQ ID NO:52)	(337)	THIYHKPYPNEILKPNFFHK-----	
Staphylococcus aureus (SEQ ID NO:53)	(326)	TSNYHHGYDNTVKDPNFFYKKKESNANQCPFFH-----	
Staphylococcus epidermidis (SEQ ID NO:54)	(326)	TSNYHHGYDNTMHHTNFFYKKEP--MKCPFH-----	
Bacillus halodurans (SEQ ID NO:55)	(338)	THIFHHEYDETVLPNYFYQAPYEDTF-----	
Bacillus subtilis (SEQ ID NO:56)	(312)	THIFHRSYDINSIVKPNFYQDKPYEB-TT-M-FF-RM--	
Streptomyces radiodurans (SEQ ID NO:57)	(318)	TPVHFRRYRARESPREFVRARCPFHTPTVHASTGHAPTG	
Streptomyces acidiscabies (SEQ ID NO:2)	(380)	TPVFHRYDNDVQRPNFYPHQ-----	
Streptomyces scabies (SEQ ID NO:4)	(380)	TPVFHRYDNDVQRPNFYPHQ-----	
Streptomyces turgidiscabies (SEQ ID NO:6)	(380)	TPVFHRYDNDVQRPNFYPHQ-----	
Streptomyces avermitilis (SEQ ID NO:58)	(450)	TPVFHRYETVERHPAYVHHPEALARAGEIDEILV---	
Mus musculus (SEQ ID NO:59)	(391)	TPVFHQEMLNYVLSPFYFYQIEPWKTHIWQNEK-----	

FIG. 6B

SEQUENCE LISTING

<110> Cornell Research Foundation, Inc.
United States Department of Agriculture

<120> BACTERIAL NITRIC OXIDE SYNTHASES AND USES THEREOF

<130> 19603/4402

<140>

<141>

<150> 60/475,111

<151> 2003-06-02

<160> 59

<170> PatentIn Ver. 2.1

<210> 1

<211> 1203

<212> DNA

<213> Streptomyces acidiscabies

<400> 1

```
gtgacctccg aagtcgctct gggcccttcc ttgcccgcgc cgtccccgac agcgtgcccc 60
gcactggggc ccgattcgtc ccttgggccg gtcccgtcgg cggaaccggc gacgccgcag 120
tcctgcggcg tcgccgatcc aaatgaggct gaggagtacc tgcgccagtt ccacgcggag 180
cagtccgatc agcccgtccc gctcgcccgg cgccctggagc aggtccgcgc cgccatcgac 240
gccacgggca cctaccggca caccaccgcc gagctcgtgt acggtgcccc cgtcgcgtgg 300
cgcaactcca gtgcgtgcat cgcccgccctg tactggaaca gcctgcgcgt cctggaccgc 360
cgggacgcca cagccccga tgagatccac cggcacttgt gcacgcacct gcgccaggcg 420
accaacggcg ggcgcacag gccggtgatt tcggtcttcg ccccgactc ccccgcccg 480
cccgccccgc aggtgtggaa cgagcagctc atccggtacg ccggtaccg ccgcgacgac 540
ggcacccgtg tcggtgacct gcgcaccgcc gacctaccg aggccatcct ccgcctcggc 600
tggcagggct gcccccaagg gccgttcgac gtcctgcccc tggtcacga ccccccgac 660
gacaaacccc gggtcttcga gctgccgcgg gagctggtct tggaggtccc tatcaccac 720
cccgacgtcc cagcctggc cgaactggc ctgcgctggc acgccgtacc cgtcatctcc 780
aacatgcgcc tacgcatcgg cgggatggac taccgctcg ccccgttcaa cggctggtac 840
atgggcacgg agatcggcgc ccgcaacctc gtgcacgagg accgctacaa catgctcccc 900
gccgtcgccg cctgcctcca gctggacacc accagcagat caaccctgtg gcgcgaccgc 960
gccctggctg agctcaacgt cgccgtcctg cactccttcg aggccgcagg tgtccggatc 1020
agcgaccacc acgaggagtc ccggcgcttc ctgcgccacc tggccaagga ggaacgccag 1080
ggcgcacccg tatccgcaga ctggagctgg atcgtcccc cgctctccgg cggcatcacc 1140
cccggtttcc accgttacta cgacaacgtc gaccagcgcc ccaacttcta cccccaccg 1200
tga
```

1203

<210> 2

<211> 400

<212> PRT

<213> Streptomyces acidiscabies

<400> 2

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

Thr Ala Cys Pro Ala Leu Gly Pro Asp Ser Ser Leu Gly Pro Val Pro
 20 25 30

Ser Ala Glu Pro Ala Thr Pro Gln Ser Cys Gly Val Ala Asp Pro Asn
 35 40 45

Glu Ala Glu Glu Phe Leu Arg Gln Phe His Ala Glu Gln Ser Asp Gln
 50 55 60

Pro Val Pro Leu Ala Arg Arg Leu Glu Gln Val Arg Ala Ala Ile Asp
 65 70 75 80

Ala Thr Gly Thr Tyr Arg His Thr Thr Ala Glu Leu Val Tyr Gly Ala
 85 90 95

Arg Val Ala Trp Arg Asn Ser Ser Arg Cys Ile Gly Arg Leu Tyr Trp
 100 105 110

Asn Ser Leu Arg Val Leu Asp Arg Arg Asp Ala Thr Ala Pro Asp Glu
 115 120 125

Ile His Arg His Leu Cys Thr His Leu Arg Gln Ala Thr Asn Gly Gly
 130 135 140

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

Pro Gly Pro Gln Val Trp Asn Glu Gln Leu Ile Arg Tyr Ala Gly Tyr
 165 170 175

Arg Arg Asp Asp Gly Thr Val Leu Gly Asp Pro Arg Thr Ala Asp Leu
 180 185 190

Thr Glu Ala Ile Leu Arg Leu Gly Trp Gln Gly Cys Pro Gln Gly Pro
 195 200 205

Phe Asp Val Leu Pro Leu Val Ile Asp Thr Pro Asp Asp Lys Pro Arg
 210 215 220

3/45

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

<210> 3

<211> 1203

<212> DNA

<213> Streptomyces scabies

<400> 3

gtgacctcgc aagtcgctct gggcccttcc ttgcccgccc cgtccccgac agcgtgcccg 60
 gcactggggc ccgattcgtc ccttgccccg gtcccgtcgg cggaaccggc gacgccgcag 120
 tcctgcgggc tcgccgatcc aaatgaggct gaggagtcc tgcgccagtt ccacgcggag 180
 cagtccgatc agcccgtccc gctcgccccg cgcctggagc aggtccgcgc cgccatcgac 240
 gccacgggca cctaccggca caccaccgcc gagctcgtgt acggtgcccc cgtcgcgtgg 300

4/45

```

cgcaactcca gtcgctgcat cggccgcctg tactggaaca gcctgcgctg cctggaccgc 360
cgggacgcca cagcccccca tgagatccac cggcacttgt gcacgcacct gcgccaggcg 420
accaacggcg ggcgcatcag gccggtgatt tcggtcttcg ccccggaact ccccgggccg 480
cccgggccgc aggtgtggaa cgagcagctc atccggtacg ccggctaccg ccgcgacgac 540
ggcaccgtgc tcggtgaccc gcgcaccgcc gacctcaccg aggccatcct ccgcctcggc 600
tggcagggct gcccccaagg gccgttcgac gtcttgcccc tggtcacga cacccccac 660
gacaaacccc gggtcttcga gctgccgcgg gagctggtct tggaggtccc tatcaccac 720
cccgacgtcc cagcctggc cgaactgggc ctgcgctggc acgccgtacc cgtcatctcc 780
aacatgcgcc tacgcatcgg cgggatggac taccgctcg ccccgttcaa cggctggtac 840
atgggcacgg agatcggcgc ccgcaacctc gtcgacgagg accgctacaa catgctcccc 900
gccgtcgccg cctgcctcca gytggacacc accagcgagt caaccctgtg gcgcgaccgc 960
gccctggtcg agtcaacgt cggcgtcctg cactccttcg aggcgcgagg tgtccggatc 1020
agcgaccacc acgaggagtc ccggcgcttc ctgcccacc tggccaagga ggaacgccag 1080
ggcgcaccg tatccgcaga ctggagctgg atcgctcccc cgctctccgg cggcatcacc 1140
cccggtgttc accgttacta cgacaacgtc gaccagcgcc ccaacttcta ccccaccag 1200
tga

```

1203

<210> 4

<211> 400

<212> PRT

<213> Streptomyces scabies

<400> 4

```

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

```

```

Thr Ala Cys Pro Ala Leu Gly Pro Asp Ser Ser Leu Gly Pro Val Pro
      20              25              30

```

```

Ser Ala Glu Pro Ala Thr Pro Gln Ser Cys Gly Val Ala Asp Pro Asn
    35              40              45

```

```

Glu Ala Glu Glu Phe Leu Arg Gln Phe His Ala Glu Gln Ser Asp Gln
    50              55              60

```

```

Pro Val Pro Leu Ala Arg Arg Leu Glu Gln Val Arg Ala Ala Ile Asp
    65              70              75              80

```

```

Ala Thr Gly Thr Tyr Arg His Thr Thr Ala Glu Leu Val Tyr Gly Ala
      85              90              95

```

```

Arg Val Ala Trp Arg Asn Ser Ser Arg Cys Ile Gly Arg Leu Tyr Trp
    100              105              110

```

```

Asn Ser Leu Arg Val Leu Asp Arg Arg Asp Ala Thr Ala Pro Asp Glu
    115              120              125

```

5/45

Ile His Arg His Leu Cys Thr His Leu Arg Gln Ala Thr Asn Gly Gly
 130 135 140

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

Pro Gly Pro Gln Val Trp Asn Glu Gln Leu Ile Arg Tyr Ala Gly Tyr
 165 170 175

Arg Arg Asp Asp Gly Thr Val Leu Gly Asp Pro Arg Thr Ala Asp Leu
 180 185 190

Thr Glu Ala Ile Leu Arg Leu Gly Trp Gln Gly Cys Pro Gln Gly Pro
 195 200 205

Phe Asp Val Leu Pro Leu Val Ile Asp Thr Pro Asp Asp Lys Pro Arg
 210 215 220

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

Pro Asp Val Pro Arg Leu Ala Glu Leu Gly Leu Arg Trp His Ala Val
 245 250 255

Pro Val Ile Ser Asn Met Arg Leu Arg Ile Gly Gly Met Asp Tyr Pro
 260 265 270

Leu Ala Pro Phe Asn Gly Trp Tyr Met Gly Thr Glu Ile Gly Ala Arg
 275 280 285

Asn Leu Val Asp Glu Asp Arg Tyr Asn Met Leu Pro Ala Val Ala Ala
 290 295 300

Cys Leu Gln Leu Asp Thr Thr Ser Glu Ser Thr Leu Trp Arg Asp Arg
 305 310 315 320

Ala Leu Val Glu Leu Asn Val Ala Val Leu His Ser Phe Glu Ala Ala
 325 330 335

Gly Val Arg Ile Ser Asp His His Glu Glu Ser Arg Arg Phe Leu Ala
 340 345 350

His Leu Ala Lys Glu Glu Arg Gln Gly Arg Thr Val Ser Ala Asp Trp
 355 360 365

Ser Trp Ile Val Pro Pro Leu Ser Gly Gly Ile Thr Pro Val Phe His
 370 375 380

6/45

Arg Tyr Tyr Asp Asn Val Asp Gln Arg Pro Asn Phe Tyr Pro His Gln
 385 390 395 400

<210> 5

<211> 1203

<212> DNA

<213> *Streptomyces turgidscabies*

<400> 5

```

gtgacttttcg aagtcgccct gggcccttcc ttgcccgcgc cgtccccgac agcgtgccccg 60
gcgctggcgc acgattcgcc cctcagtcgc gtcccgctcg cggaaccggc gacgtcgcag 120
gactgcggcg tcgccgatcc agacgaggcc gaggagttcc tgcgccagtt tcacgcggag 180
cagtccgacc aggccgtccc gctcactcgg cgccctggacc aggttcgcgc tgccatcgac 240
gccacgggca cctaccgtca taccaccgcc gagctcgtgt tcggtgcccg tgcgcgtgg 300
cgcaactcca gtcgctgtat cggccgcctg tactggaaca gcctgcgcgt cctggaccgc 360
cgggacacca cagccccga ggtaatccac cggcaccttt gcacgcacct gcgccaggcg 420
accaacggtg ggcgtatcag gccggtgatt tcggtcttcg ccccgacgc acccagccga 480
cccgcccgcc ggggtgtgaa cgagcaactc gtccggtacg ccggccaccg tcgcgacgac 540
ggcacggtac tcggcgaccc gcgatctgcc gacctaccg aggccatccg cggcctcgga 600
tggcagggag gccgccaagg gccgttcgac gtcctgcccc tggatcga cgcccacgac 660
gacaaaccgc ggttcttcga gctgcccgcg gaggttgtcc tggaggtccc tatcaccac 720
cccgacgtcc cactactggc cgaactctgc ctgcgctggc acgccgtacc cgttatctcc 780
aacatgcgcc tcggtatcgg cggggtggac taccctctcg ccccggtcaa cggctggtac 840
atgggcacgg agatcggcgt ccgtaacctc gtcgacgagg cccgctacaa cctgctcccc 900
gccgtggccg cctgcctcca gttggacacc accagcgagt ccacctgtg gcgtgaccgc 960
gctctggtcg aactcaacgt tgccgtcttg cactctttcg cggccgcagg cgtccggatc 1020
agtgaccacc acgaggagtc ccggcgcttc ctgcccacc tgaccaagga ggaacgccag 1080
ggccgcaccg taccgcgga ctggagctgg atcgctccctc cgctttccag cggcatcacc 1140
ccgctcttcc accgctacta cgacaacgcc gaccagcgcc ccaactttta ccctcatcag 1200
tga 1203

```

<210> 6

<211> 400

<212> PRT

<213> *Streptomyces turgidscabies*

<400> 6

Met Thr Phe Glu Val Ala Leu Gly Pro Ser Leu Pro Ala Pro Ser Pro
 1 5 10 15

Thr Ala Cys Pro Ala Leu Ala His Asp Ser Pro Leu Ser Pro Val Pro
 20 25 30

7/45

Ser Ala Glu Pro Ala Thr Ser Gln Asp Cys Gly Val Ala Asp Pro Asp
 35 40 45

Glu Ala Glu Glu Phe Leu Arg Gln Phe His Ala Glu Gln Ser Asp Gln
 50 55 60

Ala Val Pro Leu Thr Arg Arg Leu Asp Gln Val Arg Ala Ala Ile Asp
 65 70 75 80

Ala Thr Gly Thr Tyr Arg His Thr Thr Ala Glu Leu Val Phe Gly Ala
 85 90 95

Arg Val Ala Trp Arg Asn Ser Ser Arg Cys Ile Gly Arg Leu Tyr Trp
 100 105 110

Asn Ser Leu Arg Val Leu Asp Arg Arg Asp Thr Thr Ala Pro Glu Val
 115 120 125

Ile His Arg His Leu Cys Thr His Leu Arg Gln Ala Thr Asn Gly Gly
 130 135 140

Arg Ile Arg Pro Val Ile Ser Val Phe Ala Pro Asp Ala Pro Ser Arg
 145 150 155 160

Pro Gly Pro Arg Val Trp Asn Glu Gln Leu Val Arg Tyr Ala Gly His
 165 170 175

Arg Arg Asp Asp Gly Thr Val Leu Gly Asp Pro Arg Ser Ala Asp Leu
 180 185 190

Thr Glu Ala Ile Arg Gly Leu Gly Trp Gln Gly Gly Arg Gln Gly Pro
 195 200 205

Phe Asp Val Leu Pro Leu Val Ile Asp Ala His Asp Asp Lys Pro Arg
 210 215 220

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

Pro Asp Val Pro Arg Leu Ala Glu Leu Cys Leu Arg Trp His Ala Val
 245 250 255

Pro Val Ile Ser Asn Met Arg Leu Arg Ile Gly Gly Val Asp Tyr Pro
 260 265 270

Leu Ala Pro Phe Asn Gly Trp Tyr Met Gly Thr Glu Ile Gly Val Arg
 275 280 285

8/45

Asn Leu Val Asp Glu Ala Arg Tyr Asn Leu Leu Pro Ala Val Ala Ala
 290 295 300

Cys Leu Gln Leu Asp Thr Thr Ser Glu Ser Thr Leu Trp Arg Asp Arg
 305 310 315 320

Ala Leu Val Glu Leu Asn Val Ala Val Leu His Ser Phe Ala Ala Ala
 325 330 335

Gly Val Arg Ile Ser Asp His His Glu Glu Ser Arg Arg Phe Leu Ala
 340 345 350

His Leu Thr Lys Glu Glu Arg Gln Gly Arg Thr Val Pro Ala Asp Trp
 355 360 365

Ser Trp Ile Val Pro Pro Leu Ser Ser Gly Ile Thr Pro Val Phe His
 370 375 380

Arg Tyr Tyr Asp Asn Ala Asp Gln Arg Pro Asn Phe Tyr Pro His Gln
 385 390 395 400

<210> 7

<211> 19

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 7

Glu Val Ala Leu Gly Pro Ser Leu Pro Ala Pro Ser Pro Thr Ala Cys
 1 5 10 15

Pro Ala Leu

<210> 8

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 8

Pro Val Pro Ser Ala Glu Pro Ala Thr

1

5

<210> 9

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 9

Cys Gly Val Ala Asp Pro

1

5

<210> 10

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 10

Glu Ala Glu Glu Phe Leu Arg Gln Phe His Ala Glu Gln Ser Asp Gln

1

5

10

15

<210> 11

<211> 20

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 11

Gln Val Arg Ala Ala Ile Asp Ala Thr Gly Thr Tyr Arg His Thr Thr

1

5

10

15

Ala Glu Leu Val

20

<210> 12

<211> 28

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 12

Gly Ala Arg Val Ala Trp Arg Asn Ser Ser Arg Cys Ile Gly Arg Leu
1 5 10 15

Tyr Trp Asn Ser Leu Arg Val Leu Asp Arg Arg Asp
20 25

<210> 13

<211> 28

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 13

Ile His Arg His Leu Cys Thr His Leu Arg Gln Ala Thr Asn Gly Gly
1 5 10 15

Arg Ile Arg Pro Val Ile Ser Val Phe Ala Pro Asp
20 25

<210> 14

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 14

Val Trp Asn Glu Gln Leu
1 5

<210> 15

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 15

Arg Arg Asp Asp Gly Thr Val Leu Gly Asp Pro Arg
1 5 10

<210> 16

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 16

Ala Asp Leu Thr Glu Ala Ile
1 5

<210> 17

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 17

Leu Gly Trp Gln Gly
1 5

<210> 18

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 18

Gln Gly Pro Phe Asp Val Leu Pro Leu Val Ile Asp
1 5 10

<210> 19
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Peptide

<400> 19
Asp Asp Lys Pro Arg Phe Phe Glu Leu Pro Arg Glu
1 5 10

<210> 20
<211> 17
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Peptide

<400> 20
Val Leu Glu Val Pro Ile Thr His Pro Asp Val Pro Arg Leu Ala Glu
1 5 10 15

Leu

<210> 21
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Peptide

<400> 21
Leu Arg Trp His Ala Val Pro Val Ile Ser Asn Met Arg Leu Arg Ile
1 5 10 15

Gly Gly

<210> 22
<211> 17

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 22

Asp Tyr Pro Leu Ala Pro Phe Asn Gly Trp Tyr Met Gly Thr Glu Ile
1 5 10 15

Gly

<210> 23

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 23

Arg Asn Leu Val Asp Glu
1 5

<210> 24

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 24

Leu Pro Ala Val Ala Ala Cys Leu Gln Leu Asp Thr Thr Ser Glu Ser
1 5 10 15

Thr Leu Trp Arg Asp Arg Ala Leu Val Glu Leu Asn Val Ala Val Leu
20 25 30

His Ser Phe

35

<210> 25

<211> 20

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 25

Ala Ala Gly Val Arg Ile Ser Asp His His Glu Glu Ser Arg Arg Phe
1 5 10 15

Leu Ala His Leu
20

<210> 26

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 26

Lys Glu Glu Arg Gln Gly Arg Thr Val
1 5

<210> 27

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 27

Lys Glu Glu Arg Gln Gly Arg Thr Val
1 5

<210> 28

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 28

Asp Gln Arg Pro Asn Phe Tyr Pro His Gln
1 5 10

<210> 29

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 29

Thr Gly Thr Tyr Arg His Thr Thr
1 5

<210> 30

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 30

Gly Ala Arg Val Ala Trp Arg Asn
1 5

<210> 31

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 31

Arg Cys Ile Gly Arg Leu Tyr Trp
1 5

<210> 32

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 32

Val Phe Ala Pro Asp

1 5

<210> 33

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 33

Trp Asn Glu Gln Leu

1 5

<210> 34

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 34

Pro Phe Asp Val Leu Pro Leu

1 5

<210> 35

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 35

Leu Arg Trp His Ala Val Pro

1 5

<210> 36

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 36

Ala Pro Phe Asn Gly Trp Tyr Met Gly Thr Glu Ile Gly
1 5 10

<210> 37

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 37

Asp Arg Ala Leu Val Glu Leu Asn
1 5

<210> 38

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 38

Val Leu His Ser Phe
1 5

<210> 39

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 39

Ala Asp Trp Ser Trp Ile Val Pro Pro
1 5

<210> 40

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 40

Thr Pro Val Phe His Arg
1 5

<210> 41

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<220>

<221> PEPTIDE

<222> (1)

<223> Xaa at position 1 is either Arginine or Lysine

<220>

<221> PEPTIDE

<222> (2)

<223> Xaa at position 2 is either Methionine or Valine

<220>

<221> PEPTIDE

<222> (7)

<223> Xaa at position 7 is either Serine or Alanine

<400> 41

Xaa Xaa Ala Trp Arg Asn Xaa
1 5

<210> 42

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<220>

<221> PEPTIDE

<222> (3)

<223> Xaa at position 3 is either Isoleucine or Valine

<220>

<221> PEPTIDE

<222> (6)

<223> Xaa at position 6 is either Leucine or Isoleucine

<400> 42

Arg Cys Xaa Gly Arg Xaa

1

5

<210> 43

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<220>

<221> PEPTIDE

<222> (3)

<223> Xaa at position 3 is either Isoleucine or Valine

<400> 43

Gln Leu Xaa Arg Tyr Ala

1

5

<210> 44

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<400> 44

Arg Cys Ile Gly Arg

1

5

<210> 45

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<220>

<221> PEPTIDE

<222> (4)..(6)

<223> Xaa at positions 4-6 can be any amino acid

<400> 45

Gly Trp Tyr Xaa Xaa Xaa Glu

1

5

<210> 46

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<220>

<221> PEPTIDE

<222> (4)..(6)

<223> Xaa at positions 4-6 can be any amino acid

<400> 46

Trp Ser Trp Xaa Xaa Xaa Pro

1

5

<210> 47

<211> 398

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide,

<220>

<221> NON_TER
<222> (1)
<223> Xaa at position 1 represents from 3 up to 49
undefined residues that can be any amino acid

<220>
<221> UNSURE
<222> (21)..(27)
<223> Xaa at positions 21-27 can be any amino acid

<220>
<221> UNSURE
<222> (37)..(39)
<223> Xaa at positions 37-39 can be any amino acid

<220>
<221> UNSURE
<222> (46)
<223> Xaa at position 46 can be any amino acid

<220>
<221> UNSURE
<222> (63)..(71)
<223> Xaa at positions 63-71 can be any amino acid

<220>
<221> UNSURE
<222> (92)
<223> Xaa at position 92 can be any amino acid

<220>
<221> UNSURE
<222> (121)..(126)
<223> Xaa at positions 121-126 can be any amino acid

<220>
<221> UNSURE
<222> (155)..(162)
<223> Xaa at positions 155-162 can be any amino acid

<220>
<221> UNSURE
<222> (169)..(174)
<223> Xaa at positions 169-174 can be any amino acid

<220>
<221> UNSURE
<222> (187)

<223> Xaa at position 187 can be any amino acid

<220>

<221> UNSURE

<222> (195)..(196)

<223> Xaa at positions 195-196 can be any amino acid

<220>

<221> UNSURE

<222> (202)..(203)

<223> Xaa at positions 202-203 can be any amino acid

<220>

<221> UNSURE

<222> (216)..(217)

<223> Xaa at positions 216-217 can be any amino acid

<220>

<221> UNSURE

<222> (230)

<223> Xaa at position 230 can be any amino acid

<220>

<221> UNSURE

<222> (248)

<223> Xaa at position 248 can be any amino acid

<220>

<221> UNSURE

<222> (267)

<223> Xaa at position 267 can be any amino acid

<220>

<221> UNSURE

<222> (285)

<223> Xaa at position 285 can be any amino acid

<220>

<221> UNSURE

<222> (292)..(296)

<223> Xaa at positions 292-296 can be any amino acid

<220>

<221> UNSURE

<222> (332)

<223> Xaa at position 332 can be any amino acid

<220>

<221> UNSURE

<222> (353)

<223> Xaa at position 353 can be any amino acid

<220>

<221> UNSURE

<222> (362)

<223> Xaa at position 362 can be any amino acid

<220>

<221> UNSURE

<222> (388)

<223> Xaa at position 388 can be any amino acid

<400> 47

Xaa Glu Val Ala Leu Gly Pro Ser Leu Pro Ala Pro Ser Pro Thr Ala
1 5 10 15

Cys Pro Ala Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Val Pro Ser Ala
20 25 30

Glu Pro Ala Thr Xaa Xaa Xaa Cys Gly Val Ala Asp Pro Xaa Glu Ala
35 40 45

Glu Glu Phe Leu Arg Gln Phe His Ala Glu Gln Ser Asp Gln Xaa Xaa
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gln Val Arg Ala Ala Ile Asp Ala Thr
65 70 75 80

Gly Thr Tyr Arg His Thr Thr Ala Glu Leu Val Xaa Gly Ala Arg Val
85 90 95

Ala Trp Arg Asn Ser Ser Arg Cys Ile Gly Arg Leu Tyr Trp Asn Ser
100 105 110

Leu Arg Val Leu Asp Arg Arg Asp Xaa Xaa Xaa Xaa Xaa Xaa Ile His
115 120 125

Arg His Leu Cys Thr His Leu Arg Gln Ala Thr Asn Gly Gly Arg Ile
130 135 140

Arg Pro Val Ile Ser Val Phe Ala Pro Asp Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Val Trp Asn Glu Gln Leu Xaa Xaa Xaa Xaa Xaa Xaa Arg Arg
165 170 175

24/45

```

Asp Asp Gly Thr Val Leu Gly Asp Pro Arg Xaa Ala Asp Leu Thr Glu
      180                      185                      190

Ala Ile Xaa Xaa Leu Gly Trp Gln Gly Xaa Xaa Gln Gly Pro Phe Asp
      195                      200                      205

Val Leu Pro Leu Val Ile Asp Xaa Xaa Asp Asp Lys Pro Arg Phe Phe
      210                      215                      220

Glu Leu Pro Arg Glu Xaa Val Leu Glu Val Pro Ile Thr His Pro Asp
      225                      230                      235                      240

Val Pro Arg Leu Ala Glu Leu Xaa Leu Arg Trp His Ala Val Pro Val
      245                      250                      255

Ile Ser Asn Met Arg Leu Arg Ile Gly Gly Xaa Asp Tyr Pro Leu Ala
      260                      265                      270

Pro Phe Asn Gly Trp Tyr Met Gly Thr Glu Ile Gly Xaa Arg Asn Leu
      275                      280                      285

Val Asp Glu Xaa Xaa Xaa Xaa Xaa Leu Pro Ala Val Ala Ala Cys Leu
      290                      295                      300

Gln Leu Asp Thr Thr Ser Glu Ser Thr Leu Trp Arg Asp Arg Ala Leu
      305                      310                      315                      320

Val Glu Leu Asn Val Ala Val Leu His Ser Phe Xaa Ala Ala Gly Val
      325                      330                      335

Arg Ile Ser Asp His His Glu Glu Ser Arg Arg Phe Leu Ala His Leu
      340                      345                      350

Xaa Lys Glu Glu Arg Gln Gly Arg Thr Val Xaa Ala Asp Trp Ser Trp
      355                      360                      365

Ile Val Pro Pro Leu Ser Gly Gly Ile Thr Pro Val Phe His Arg Tyr
      370                      375                      380

Tyr Asp Asn Xaa Asp Gln Arg Pro Asn Phe Tyr Pro His Gln
      385                      390                      395

<210> 48
<211> 306
<212> PRT
<213> Artificial Sequence

```

<220>

<223> Description of Artificial Sequence: Peptide

<220>

<221> NON_TER

<222> (1)

<223> Xaa at position 1 represents from 81 up to 151
undefined residues that can be any amino acid

<220>

<221> UNSURE

<222> (10)..(14)

<223> Xaa at positions 10-14 can be any amino acid

<220>

<221> UNSURE

<222> (23)..(24)

<223> Xaa at positions 23-24 can be any amino acid

<220>

<221> UNSURE

<222> (33)..(71)

<223> Xaa at positions 33-71 can be any amino acid

<220>

<221> UNSURE

<222> (77)..(85)

<223> Xaa at positions 77-85 can be any amino acid

<220>

<221> UNSURE

<222> (91)..(127)

<223> Xaa at positions 91-127 can be any amino acid

<220>

<221> UNSURE

<222> (135)..(170)

<223> Xaa at positions 135-170 can be any amino acid

<220>

<221> UNSURE

<222> (178)..(193)

<223> Xaa at positions 178-193 can be any amino acid

<220>

<221> UNSURE

<222> (207)..(238)

<223> Xaa at positions 207-238 can be any amino acid

<220>

<221> UNSURE

<222> (247)..(248)

<223> Xaa at positions 247-248 can be any amino acid

<220>

<221> UNSURE

<222> (254)..(285)

<223> Xaa at positions 254-285 can be any amino acid

<220>

<221> UNSURE

<222> (295)..(299)

<223> Xaa at positions 295-299 can be any amino acid

<220>

<221> UNSURE

<222> (306)

<223> Xaa at position 306 represents from 15 up to 30
undefined residues that can be any amino acid

<400> 48

Xaa	Thr	Gly	Thr	Tyr	Arg	His	Thr	Thr	Xaa	Xaa	Xaa	Xaa	Xaa	Gly	Ala
1				5					10					15	

Arg	Val	Ala	Trp	Arg	Asn	Xaa	Xaa	Arg	Cys	Ile	Gly	Arg	Leu	Tyr	Trp
			20					25					30		

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			35					40					45		

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			50					55					60		

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Val	Phe	Ala	Pro	Asp	Xaa	Xaa	Xaa	Xaa
			65				70				75				80

Xaa	Xaa	Xaa	Xaa	Xaa	Trp	Asn	Glu	Gln	Leu	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
					85				90					95	

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
														100	110

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Pro
														115	125

Phe	Asp	Val	Leu	Pro	Leu	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

130 135 140
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 145 150 155 160
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Arg Trp His Ala Val
 165 170 175
 Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 180 185 190
 Xaa Ala Pro Phe Asn Gly Trp Tyr Met Gly Thr Glu Ile Gly Xaa Xaa
 195 200 205
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 210 215 220
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asp Arg
 225 230 235 240
 Ala Leu Val Glu Leu Asn Xaa Xaa Val Leu His Ser Phe Xaa Xaa Xaa
 245 250 255
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 260 265 270
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ala Asp Trp
 275 280 285
 Ser Trp Ile Val Pro Pro Xaa Xaa Xaa Xaa Xaa Thr Pro Val Phe His
 290 295 300
 Arg Xaa
 305

<210> 49

<211> 23

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<220>

<221> NON_TER

<222> (1)

<223> Xaa at position 1 represents from 29 up to 166

undefined residues that can be any amino acid

<220>

<221> UNSURE

<222> (2)

<223> Xaa at position 2 can be R or K

<220>

<221> UNSURE

<222> (3)

<223> Xaa at position 3 can be M or V

<220>

<221> UNSURE

<222> (8)

<223> Xaa at position 8 can be S or A

<220>

<221> UNSURE

<222> (9)

<223> Xaa at position 9 can be any amino acid

<220>

<221> UNSURE

<222> (12)

<223> Xaa at position 12 can be I or V

<220>

<221> UNSURE

<222> (15)

<223> Xaa at position 15 can be L or I

<220>

<221> UNSURE

<222> (16)

<223> Xaa at position 16 represents from 52 up to 58
undefined residues that can be any amino acid

<220>

<221> UNSURE

<222> (19)

<223> Xaa at position 19 can be I or V

<220>

<221> UNSURE

<222> (22)

<223> Xaa at position 22 represents from 220 up to 243
undefined residues that can be any amino acid

<400> 49
Xaa Xaa Xaa Ala Trp Arg Asn Xaa Xaa Arg Cys Xaa Gly Arg Xaa Xaa
1 5 10 15

Gln Leu Xaa Arg Tyr Ala Xaa
20

<210> 50

<211> 23

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<220>

<221> NON_TER

<222> (1)

<223> Xaa at position 1 represents from 37 up to 174
undefined residues that can be any amino acid

<220>

<221> UNSURE

<222> (7)

<223> Xaa at position 7 represents from 48 up to 67
undefined residues that can be any amino acid

<220>

<221> UNSURE

<222> (11)..(13)

<223> Xaa at positions 11-13 can be any amino acid

<220>

<221> UNSURE

<222> (15)

<223> Xaa at position 15 represents up to 83 undefined
residues that can be any amino acid

<220>

<221> UNSURE

<222> (19)..(21)

<223> Xaa at positions 19-21 can be any amino acid

<220>

<221> UNSURE

<222> (23)

<223> Xaa at position 23 represents from 20 up to 38
undefined residues that can be any amino acid

<400> 50

Xaa Arg Cys Ile Gly Arg Xaa Gly Trp Tyr Xaa Xaa Xaa Glu Xaa Trp
1 5 10 15

Ser Trp Xaa Xaa Xaa Pro Xaa
20

<210> 51

<211> 356

<212> PRT

<213> Bacillus anthracis

<400> 51

Met Ser Lys Thr Lys Gln Leu Ile Glu Glu Ala Ser His Phe Ile Thr
1 5 10 15

Ile Cys Tyr Lys Glu Leu Ser Lys Glu His Phe Ile Glu Glu Arg Met
20 25 30

Lys Glu Ile Gln Ala Glu Ile Glu Lys Thr Gly Thr Tyr Glu His Thr
35 40 45

Phe Glu Glu Leu Val His Gly Ser Arg Met Ala Trp Arg Asn Ser Asn
50 55 60

Arg Cys Ile Gly Arg Leu Phe Trp Ser Lys Met His Ile Leu Asp Ala
65 70 75 80

Arg Glu Val Asn Asp Glu Glu Gly Val Tyr His Ala Leu Ile His His
85 90 95

Ile Lys Tyr Ala Thr Asn Asp Gly Lys Val Lys Pro Thr Ile Thr Ile
100 105 110

Phe Lys Gln Tyr Gln Gly Glu Glu Asn Asn Ile Arg Ile Tyr Asn His
115 120 125

Gln Leu Ile Arg Tyr Ala Gly Tyr Lys Thr Glu Met Gly Val Thr Gly
130 135 140

Asp Ser His Ser Thr Ala Phe Thr Asp Phe Cys Gln Glu Leu Gly Trp
145 150 155 160

Gln Gly Glu Gly Thr Asn Phe Asp Val Leu Pro Leu Val Phe Ser Ile

31/45

165	170	175
Asp Gly Lys Ala Pro Ile Tyr Lys Glu Ile Pro Lys Glu Glu Val Lys		
180	185	190
Glu Val Pro Ile Glu His Pro Glu Tyr Pro Ile Ser Ser Leu Gly Ala		
195	200	205
Lys Trp Tyr Gly Val Pro Met Ile Ser Asp Met Arg Leu Glu Ile Gly		
210	215	220
Gly Ile Ser Tyr Thr Ala Ala Pro Phe Asn Gly Trp Tyr Met Gly Thr		
225	230	235 240
Glu Ile Gly Ala Arg Asn Leu Ala Asp His Asp Arg Tyr Asn Leu Leu		
245	250	255
Pro Ala Val Ala Glu Met Met Asp Leu Asp Thr Ser Arg Asn Gly Thr		
260	265	270
Leu Trp Lys Asp Lys Ala Leu Ile Glu Leu Asn Val Ala Val Leu His		
275	280	285
Ser Phe Lys Lys Gln Gly Val Ser Ile Val Asp His His Thr Ala Ala		
290	295	300
Gln Gln Phe Gln Gln Phe Glu Lys Gln Glu Ala Ala Cys Gly Arg Val		
305	310	315 320
Val Thr Gly Asn Trp Val Trp Leu Ile Pro Pro Leu Ser Pro Ala Thr		
325	330	335
Thr His Ile Tyr His Lys Pro Tyr Pro Asn Glu Ile Leu Lys Pro Asn		
340	345	350
Phe Phe His Lys		
355		
<210> 52		
<211> 356		
<212> PRT		
<213> Bacillus cereus		
<400> 52		
Met Ser Lys Thr Lys Gln Leu Ile Glu Glu Ala Ser Asn Phe Ile Thr		
1	5	10 15

32/45

Ile Cys Tyr Lys Glu Leu His Lys Glu Gln Leu Ile Glu Glu Arg Ile
 20 25 30

Lys Glu Ile Gln Ile Glu Ile Glu Lys Thr Gly Thr Tyr Glu His Thr
 35 40 45

Phe Glu Glu Leu Val His Gly Ser Arg Met Ala Trp Arg Asn Ser Asn
 50 55 60

Arg Cys Ile Gly Arg Leu Phe Trp Ser Lys Met His Ile Leu Asp Ala
 65 70 75 80

Arg Glu Val Asn Asp Glu Glu Gly Val Tyr Asn Ala Leu Ile His His
 85 90 95

Ile Lys Tyr Ala Thr Asn Asp Gly Lys Val Lys Pro Thr Ile Thr Ile
 100 105 110

Phe Lys Gln Tyr Gln Gly Glu Glu Asn Asn Ile Arg Ile Tyr Asn His
 115 120 125

Gln Leu Ile Arg Tyr Ala Gly Tyr Lys Thr Glu Thr Gly Val Ile Gly
 130 135 140

Asp Ser His Ser Ala Thr Phe Thr Asp Phe Cys Gln Gly Leu Gly Trp
 145 150 155 160

Gln Gly Glu Gly Thr Asn Phe Asp Val Leu Pro Leu Val Phe Ser Ile
 165 170 175

Asn Gly Lys Ala Pro Thr Tyr Lys Glu Ile Pro Arg Glu Glu Val Lys
 180 185 190

Glu Val Pro Ile Glu His Pro Glu Tyr Pro Ile Ser Ser Leu Gly Val
 195 200 205

Lys Trp Tyr Gly Val Pro Met Ile Ser Asp Met Arg Leu Glu Ile Gly
 210 215 220

Gly Ile Ser Tyr Thr Ala Ala Pro Phe Asn Gly Trp Tyr Met Gly Thr
 225 230 235 240

Glu Ile Gly Ala Arg Asn Leu Ala Asp His Asp Arg Tyr Asn Leu Leu
 245 250 255

Pro Ala Val Ala Glu Met Met Asp Leu Asp Thr Ser Arg Asn Gly Thr
 260 265 270

33/45

Leu Trp Lys Asp Lys Ala Leu Ile Glu Leu Asn Ile Ala Val Leu His
 275 280 285

Ser Phe Lys Lys Gln Gly Val Ser Ile Val Asp His His Thr Ala Ala
 290 295 300

Gln Gln Phe Gln Gln Phe Glu Lys Gln Glu Ala Ala Cys Gly Arg Val
 305 310 315 320

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

Thr His Ile Tyr His Lys Pro Tyr Pro Asn Glu Ile Leu Lys Pro Asn
 340 345 350

Phe Phe His Lys
 355

<210> 53

<211> 358

<212> PRT

<213> Staphylococcus aureus

<400> 53

Met Leu Phe Lys Glu Ala Gln Ala Phe Ile Glu Asn Met Tyr Lys Glu
 1 5 10 15

Cys His Tyr Glu Thr Gln Ile Ile Asn Lys Arg Leu His Asp Ile Glu
 20 25 30

Leu Glu Ile Lys Glu Thr Gly Thr Tyr Thr His Thr Glu Glu Glu Leu
 35 40 45

Ile Tyr Gly Ala Lys Met Ala Trp Arg Asn Ser Asn Arg Cys Ile Gly
 50 55 60

Arg Leu Phe Trp Asp Ser Leu Asn Val Ile Asp Ala Arg Asp Val Thr
 65 70 75 80

Asp Glu Ala Ser Phe Leu Ser Ser Ile Thr Tyr His Ile Thr Gln Ala
 85 90 95

Thr Asn Glu Gly Lys Leu Lys Pro Tyr Ile Thr Ile Tyr Ala Pro Lys
 100 105 110

Asp Gly Pro Lys Ile Phe Asn Asn Gln Leu Ile Arg Tyr Ala Gly Tyr
 115 120 125

Asp Asn Cys Gly Asp Pro Ala Glu Lys Glu Val Thr Arg Leu Ala Asn
 130 135 140
 His Leu Gly Trp Lys Gly Lys Gly Thr Asn Phe Asp Val Leu Pro Leu
 145 150 155 160
 Ile Tyr Gln Leu Pro Asn Glu Ser Val Lys Phe Tyr Glu Tyr Pro Thr
 165 170 175
 Ser Leu Ile Lys Glu Val Pro Ile Glu His Asn His Tyr Pro Lys Leu
 180 185 190
 Arg Lys Leu Asn Leu Lys Trp Tyr Ala Val Pro Ile Ile Ser Asn Met
 195 200 205
 Asp Leu Lys Ile Gly Gly Ile Val Tyr Pro Thr Ala Pro Phe Asn Gly
 210 215 220
 Trp Tyr Met Val Thr Glu Ile Gly Val Arg Asn Phe Ile Asp Asp Tyr
 225 230 235 240
 Arg Tyr Asn Leu Leu Glu Lys Val Ala Asp Ala Phe Glu Phe Asp Thr
 245 250 255
 Leu Lys Asn Asn Ser Phe Asn Lys Asp Arg Ala Leu Val Glu Leu Asn
 260 265 270
 Tyr Ala Val Tyr His Ser Phe Lys Lys Glu Gly Val Ser Ile Val Asp
 275 280 285
 His Leu Thr Ala Ala Lys Gln Phe Glu Leu Phe Glu Arg Asn Glu Ala
 290 295 300
 Gln Gln Gly Arg Gln Val Thr Gly Lys Trp Ser Trp Leu Ala Pro Pro
 305 310 315 320
 Leu Ser Pro Thr Leu Thr Ser Asn Tyr His His Gly Tyr Asp Asn Thr
 325 330 335
 Val Lys Asp Pro Asn Phe Phe Tyr Lys Lys Lys Glu Ser Asn Ala Asn
 340 345 350
 Gln Cys Pro Phe His His
 355

<210> 54

<211> 355

<212> PRT

<213> Staphylococcus epidermidis

<400> 54

```

Met Leu Ile Asp Lys Ala Arg Ser Phe Ile Gln Thr Met Tyr Ser Glu
  1             5             10             15

Leu Lys Tyr Asn Thr Asn Glu Ile Glu Asn Arg Met Lys Glu Ile Glu
      20             25             30

Gln Glu Ile Asn Leu Thr Gly Ser Tyr Thr His Thr Tyr Glu Glu Leu
      35             40             45

Ser Tyr Gly Ala Lys Met Ala Trp Arg Asn Ser Asn Arg Cys Ile Gly
      50             55             60

Arg Leu Phe Trp Asn Ser Leu Asn Val Lys Asp Ala Arg Asp Val Cys
      65             70             75             80

Asp Glu Lys Glu Phe Ile Lys Phe Ile His Thr His Ile Lys Glu Ala
      85             90             95

Thr Asn Gly Gly Lys Ile Lys Pro Tyr Ile Thr Ile Phe Ser Pro Glu
      100            105            110

Asp Thr Pro Lys Ile Tyr Asn Asn Gln Leu Ile Arg Tyr Ala Gly Tyr
      115            120            125

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

His Leu Gly Trp Lys Gly Lys Gly Ser Asn Phe Asp Ile Leu Pro Leu
      145            150            155            160

Ile Tyr Gln Leu Pro Asn Asp Thr Ile Lys Ile His Glu Leu Pro Asn
      165            170            175

Asp Ile Val Lys Glu Val Ser Ile His His Glu His Tyr Pro Lys Leu
      180            185            190

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

Asp Leu Lys Ile Gly Gly Ile Thr Tyr Pro Thr Ala Pro Phe Asn Gly
      210            215            220

Trp Tyr Met Val Thr Glu Ile Ala Val Arg Asn Phe Thr Asp Thr Tyr

```

[illegible]

```
<210> 55
<211> 366
<212> PRT
<213> Bacillus halodurans
```

<400> 55

Met	Glu	Glu	Lys	Glu	Arg	Leu	Gln	Leu	Glu	Ala	Glu	Ser	Phe	Leu	Thr
1				5					10					15	
Lys	Cys	Tyr	Glu	Glu	Leu	Gly	Ser	Thr	Gly	Glu	Leu	Ser	Lys	Arg	Leu
			20					25					30		
Glu	Glu	Val	Arg	Lys	Glu	Ile	Asp	Lys	Thr	Gly	Thr	Tyr	Val	His	Thr
		35					40					45			
Thr	Lys	Glu	Leu	Ala	His	Gly	Ala	Arg	Met	Ala	Trp	Arg	Asn	Ser	Asn
	50					55					60				
Arg	Cys	Ile	Gly	Arg	Leu	Phe	Trp	Glu	Ser	Leu	His	Val	Leu	Asp	Cys
65					70					75					80

Arg His Leu Gln Thr Glu Glu Glu Met Ala Glu Ala Leu Val Asp His
85 90 95

Ile Thr Tyr Ala Thr Asn Asp Gly Lys Ile Leu Pro Thr Ile Ser Val
100 105 110

Phe Arg Pro Arg His Pro Asn Lys Gly Asp Val Arg Ile Trp Asn Gln
115 120 125

Gln Leu Ile Arg Tyr Ala Gly Tyr Glu Glu Gly Asp Gln Val Ile Gly
130 135 140

Asp Pro Ile Ser Thr Lys Phe Thr Gln Ala Cys Glu Arg Leu Gly Trp
145 150 155 160

Ser Gly Glu Arg Thr Pro Phe Asp Val Leu Pro Leu Val Ile Gln Asp
165 170 175

Gly Ser Lys Pro Pro Lys Trp Phe Ala Val Pro Asn Glu Ser Val Lys
180 185 190

Glu Val Pro Leu Arg His Pro Glu Tyr Glu Trp Phe Ala Gly Phe Gln
195 200 205

Leu Lys Trp Tyr Ala Val Pro Ile Val Ser Asn Met Arg Leu Glu Ile
210 215 220

Gly Gly Ile His Tyr Pro Ala Ala Pro Phe Asn Gly Trp Tyr Met Gly
225 230 235 240

Thr Glu Ile Gly Ala Arg Asn Leu Ala Asp Glu Asp Arg Tyr Asn Ile
245 250 255

Leu Pro Lys Met Ala Glu Tyr Met Gly Leu Ser Thr Gly Lys Asp Ser
260 265 270

Thr Leu Trp Lys Asp Lys Ala Leu Val Glu Leu Asn Val Ala Ile Leu
275 280 285

Tyr Ser Tyr Lys Gln Glu Gly Val Ser Ile Val Asp His His Thr Ala
290 295 300

Ala Lys Gln Phe Ala Arg Phe Glu Gln Ala Glu Gln Ala Ala Asn Arg
305 310 315 320

Lys Val Thr Gly Arg Trp Ser Trp Leu Ile Pro Pro Val Ser Pro Ala
325 330 335

38/45

Thr Thr His Ile Phe His His Glu Tyr Glu Asp Glu Thr Val Leu Pro
 340 345 350

Asn Tyr Phe Tyr Gln Pro Ala Pro Tyr Glu Ser Asp Thr Phe
 355 360 365

<210> 56

<211> 344

<212> PRT

<213> Bacillus subtilis

<400> 56

Met Lys Asp Arg Leu Ala Asp Ile Lys Ser Glu Ile Asp Leu Thr Gly
 1 5 10 15

Ser Tyr Val His Thr Lys Glu Glu Leu Glu His Gly Ala Lys Met Ala
 20 25 30

Trp Arg Asn Ser Asn Arg Cys Ile Gly Arg Leu Phe Trp Asn Ser Leu
 35 40 45

Asn Val Ile Asp Arg Arg Asp Val Arg Thr Lys Glu Glu Val Arg Asp
 50 55 60

Ala Leu Phe His His Ile Glu Thr Ala Thr Asn Asn Gly Lys Ile Arg
 65 70 75 80

Pro Thr Ile Thr Ile Phe Pro Pro Glu Glu Lys Gly Glu Lys Gln Val
 85 90 95

Glu Ile Trp Asn His Gln Leu Ile Arg Tyr Ala Gly Tyr Glu Ser Asp
 100 105 110

Gly Glu Arg Ile Gly Asp Pro Ala Ser Cys Ser Leu Thr Ala Ala Cys
 115 120 125

Glu Glu Leu Gly Trp Arg Gly Glu Arg Thr Asp Phe Asp Leu Leu Pro
 130 135 140

Leu Ile Phe Arg Met Lys Gly Asp Glu Gln Pro Val Trp Tyr Glu Leu
 145 150 155 160

Pro Arg Ser Leu Val Ile Glu Val Pro Ile Thr His Pro Asp Ile Glu
 165 170 175

Ala Phe Ser Asp Leu Glu Leu Lys Trp Tyr Gly Val Pro Ile Ile Ser
 180 185 190

Asp Met Lys Leu Glu Val Gly Gly Ile His Tyr Asn Ala Ala Pro Phe
 195 200 205

Asn Gly Trp Tyr Met Gly Thr Glu Ile Gly Ala Arg Asn Leu Ala Asp
 210 215 220

Glu Lys Arg Tyr Asp Lys Leu Lys Lys Val Ala Ser Val Ile Gly Ile
 225 230 235 240

Ala Ala Asp Tyr Asn Thr Asp Leu Trp Lys Asp Gln Ala Leu Val Glu
 245 250 255

Leu Asn Lys Ala Val Leu His Ser Tyr Lys Lys Gln Gly Val Ser Ile
 260 265 270

Val Asp His His Thr Ala Ala Ser Gln Phe Lys Arg Phe Glu Glu Gln
 275 280 285

Glu Glu Glu Ala Gly Arg Lys Leu Thr Gly Asp Trp Thr Trp Leu Ile
 290 295 300

Pro Pro Ile Ser Pro Ala Ala Thr His Ile Phe His Arg Ser Tyr Asp
 305 310 315 320

Asn Ser Ile Val Lys Pro Asn Tyr Phe Tyr Gln Asp Lys Pro Tyr Glu
 325 330 335

Asx Thr Thr Met Phe Phe Arg Met
 340

<210> 57

<211> 356

<212> PRT

<213> Deinococcus radiodurans

<400> 57

Met Ser Cys Pro Ala Ala Ala Val Leu Thr Pro Asp Met Arg Ala Phe
 1 5 10 15

Leu Arg Arg Phe His Glu Glu Met Gly Glu Pro Gly Leu Pro Ala Arg
 20 25 30

Leu Arg Ala Val Glu Glu Ala Gly Leu Trp Trp Pro Thr Ser Ala Glu
 35 40 45

Leu Thr Trp Gly Ala Lys Val Ala Trp Arg Asn Ser Thr Arg Cys Val

50	55	60
Gly Arg Leu Tyr Trp Glu Ala Leu Ser Val Arg Asp Leu Arg Glu Leu		
65	70	75 80
Asn Thr Ala Gln Ala Val Tyr Glu Ala Leu Leu Gln His Leu Asp Asp		
	85	90 95
Ala Phe Cys Gly Gly His Ile Arg Pro Val Ile Ser Val Phe Gly Pro		
	100	105 110
Gly Val Arg Leu His Asn Pro Gln Leu Ile Arg Tyr Ala Asp Asp Pro		
	115	120 125
Ile Asn Ala Asp Phe Val Asp Lys Leu Arg Arg Phe Gly Trp Gln Pro		
	130	135 140
Arg Gly Glu Arg Phe Glu Val Leu Pro Leu Leu Ile Glu Val Asn Gly		
145	150	155 160
Arg Ala Glu Leu Phe Ser Leu Pro Pro Gln Ala Val Gln Glu Val Ala		
	165	170 175
Ile Thr His Pro Val Cys Leu Gly Ile Gly Glu Leu Gly Leu Arg Trp		
	180	185 190
His Ala Leu Pro Val Ile Ser Asp Met His Leu Asp Ile Gly Gly Leu		
	195	200 205
His Leu Pro Cys Ala Phe Ser Gly Trp Tyr Val Gln Thr Glu Ile Ala		
210	215	220
Ala Arg Asp Leu Ala Asp Val Gly Arg Tyr Asp Gln Leu Pro Ala Val		
225	230	235 240
Ala Arg Ala Leu Gly Leu Asp Thr Ser Arg Glu Arg Thr Leu Trp Arg		
	245	250 255
Asp Arg Ala Leu Val Glu Leu Asn Val Ala Val Leu His Ser Phe Asp		
	260	265 270
Ala Ala Gly Val Lys Leu Ala Asp His His Thr Val Thr Ala His His		
	275	280 285
Val Arg Phe Glu Glu Arg Glu Ala Arg Ala Gly Arg Glu Val Arg Gly		
290	295	300
Lys Trp Ser Trp Leu Val Pro Pro Leu Ser Pro Ala Thr Thr Pro Leu		

[illegible]

```
<210> 58
<211> 485
<212> PRT
<213> Streptomyces avermitilis
```

<400> 58
Leu Ala Asp Ala Pro Pro Glu Ser Ala Ala Ser Arg Pro Gly Gly Asp
1 5 10 15

Ala Arg Asp Arg His Gly Ala Pro Ser Arg Ala Arg Gly Trp Ser Thr
20 25 30

Glu Thr Trp Cys Ala Pro His Gly Arg Lys Asp Thr Ala Gly Glu Ser
35 40 45

Glu His Arg Pro Pro Gly Pro Thr Leu Pro Gly Ala Arg Gly Trp Gly
50 55 60

Ala Ala Arg Thr Arg Gly Asp Arg Pro Asp Gly His Pro Asp His Pro
65 70 75 80

Thr Gly Ser Pro Thr Thr Asp Gly Pro Ala Arg Pro Ala Arg Pro Asp
85 90 95

Arg Arg Glu Glu Gly Asp Gly His Asp Gly His Glu Gly Gln Asp Gly
100 105 110

His Asp Gly His Asp Leu Leu Cys Ala Ala Thr Ala Phe Leu Thr Leu
115 120 125

His His Thr Glu Glu Arg Leu Gly Asp Pro Ala Arg Arg Ile Ala Ala
130 135 140

Ala His Ala Glu Ile Ala Glu Thr Gly Thr Tyr Arg His Thr Thr Glu
145 150 155 160

42/45

Glu Leu Val Phe Gly Ala Arg Val Ala Trp Arg Asn Ala Asn Arg Cys
 165 170 175

Ile Gly Arg Leu Tyr Trp His Ser Leu Cys Val Arg Asp Arg Arg Asp
 180 185 190

Val Arg Asp Ala Lys Asp Val Ala Glu Ala Ser Ala Asp His Leu Arg
 195 200 205

Glu Ala Thr Arg Asp Gly Arg Ile Arg Ala Leu Ile Thr Val Phe Ala
 210 215 220

Pro Asp Ala Pro Gly Arg Pro Gly Pro Arg Ile Trp Asn Glu Gln Leu
 225 230 235 240

Ile Arg Tyr Ala Gly Tyr Ala Arg Pro Gly Gly Ala Val Thr Gly Asp
 245 250 255

Pro Arg Asn Ala Gly Leu Thr Ala Leu Ala Arg Arg Leu Gly Trp Pro
 260 265 270

Gly Gly Pro Gly Ser Pro Phe Asp Val Leu Pro Leu Ile Val Gln Ser
 275 280 285

Ala Gly Asp Arg Pro Arg Trp Phe Thr Leu Pro Glu Asp Ala Val Leu
 290 295 300

Glu Val Ala Leu Thr His Pro Glu Tyr Pro Trp Trp Arg Ser Leu Gly
 305 310 315 320

Leu Arg Trp His Ala Val Pro Ala Leu Ala Gly Met Cys Leu Glu Ser
 325 330 335

Gly Gly Ile Cys Tyr Pro Ala Ala Pro Phe Asn Gly Trp Tyr Met Gly
 340 345 350

Thr Glu Ile Gly Ala Arg Asn Leu Ala Asp Ala Asp Arg Tyr Asp Leu
 355 360 365

Leu Pro His Leu Ala Asp Arg Leu Gly Leu Asp Thr Arg Ser Asp Arg
 370 375 380

Ser Leu Trp Lys Asp Arg Ala Leu Val Glu Leu Asn Arg Ser Val Leu
 385 390 395 400

His Ser Phe Asp Arg Ala Gly Val Thr Val Thr Asp His His Thr Glu
 405 410 415

43/45

Ser Leu Arg Phe Leu Thr His Leu Asp Arg Glu Glu Arg Lys Gly Arg
 420 425 430

Arg Val Gly Ala Asp Trp Ser Trp Ile Val Pro Pro Ile Ser Gly Ser
 435 440 445

Ala Thr Pro Val Phe His Arg Thr Tyr Glu Thr Val Glu Arg His Pro
 450 455 460

Ala Tyr Val His His Pro Glu Ala Leu Ala Arg Ala Arg Gly Glu Ile
 465 470 475 480

Asp Glu Ile Leu Val
 485

<210> 59
 <211> 423
 <212> PRT
 <213> Mus musculus

<400> 59
 Gln Tyr Val Arg Ile Lys Asn Trp Gly Ser Gly Glu Ile Leu His Asp
 1 5 10 15

Thr Leu His His Lys Ala Thr Ser Asp Phe Thr Cys Lys Ser Lys Ser
 20 25 30

Cys Leu Gly Ser Ile Met Asn Pro Lys Ser Leu Thr Arg Gly Pro Arg
 35 40 45

Asp Lys Pro Thr Pro Leu Glu Glu Leu Leu Pro His Ala Ile Glu Phe
 50 55 60

Ile Asn Gln Tyr Tyr Gly Ser Phe Lys Glu Ala Lys Ile Glu Glu His
 65 70 75 80

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

Tyr Gln Leu Thr Leu Asp Glu Leu Ile Phe Ala Thr Lys Met Ala Trp
 100 105 110

Arg Asn Ala Pro Arg Cys Ile Gly Arg Ile Gln Trp Ser Asn Leu Gln
 115 120 125

Val Phe Asp Ala Arg Asn Cys Ser Thr Ala Gln Glu Met Phe Gln His
 130 135 140

Ile Cys Arg His Ile Leu Tyr Ala Thr Asn Asn Gly Asn Ile Arg Ser
 145 150 155 160
 Ala Ile Thr Val Phe Pro Gln Arg Ser Asp Gly Lys His Asp Phe Arg
 165 170 175
 Leu Trp Asn Ser Gln Leu Ile Arg Tyr Ala Gly Tyr Gln Met Pro Asp
 180 185 190
 Gly Thr Ile Arg Gly Asp Ala Ala Thr Leu Glu Phe Thr Gln Leu Cys
 195 200 205
 Ile Asp Leu Gly Trp Lys Pro Arg Tyr Gly Arg Phe Asp Val Leu Pro
 210 215 220
 Leu Val Leu Gln Ala Asp Gly Gln Asp Pro Glu Val Phe Glu Ile Pro
 225 230 235 240
 Pro Asp Leu Val Leu Glu Val Thr Met Glu His Pro Lys Tyr Glu Trp
 245 250 255
 Phe Gln Glu Leu Gly Leu Lys Trp Tyr Ala Leu Pro Ala Val Ala Asn
 260 265 270
 Met Leu Leu Glu Val Gly Gly Leu Glu Phe Pro Ala Cys Pro Phe Asn
 275 280 285
 Gly Trp Tyr Met Gly Thr Glu Ile Gly Val Arg Asp Phe Cys Asp Thr
 290 295 300
 Gln Arg Tyr Asn Ile Leu Glu Glu Val Gly Arg Arg Met Gly Leu Glu
 305 310 315 320
 Thr His Thr Leu Ala Ser Leu Trp Lys Asp Arg Ala Val Thr Glu Ile
 325 330 335
 Asn Val Ala Val Leu His Ser Phe Gln Lys Gln Asn Val Thr Ile Met
 340 345 350
 Asp His His Thr Ala Ser Glu Ser Phe Met Lys His Met Gln Asn Glu
 355 360 365
 Tyr Arg Ala Arg Gly Gly Cys Pro Ala Asp Trp Ile Trp Leu Val Pro
 370 375 380
 Pro Val Ser Gly Ser Ile Thr Pro Val Phe His Gln Glu Met Leu Asn
 385 390 395 400

Tyr Val Leu Ser Pro Phe Tyr Tyr Tyr Gln Ile Glu Pro Trp Lys Thr
405 410 415

His Ile Trp Gln Asn Glu Lys
420