(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)特許出願公表番号
特表2005-508151
(P2005-508151A)
(43)公表日　平成17年3月31日(2005.3.31)

(51)Int.Cl.[7]	F I		テーマコード（参考）
C12N 15/09	C12N 15/00	ZNAA	4B024
A61K 35/60	A61K 35/60		4B064
A61K 38/00	A61P 31/00		4B065
A61P 31/00	A61P 31/04		4C084
A61P 31/04	A61P 31/10		4C087

審査請求 未請求　予備審査請求 未請求　（全 100 頁）　最終頁に続く

(21)出願番号　特願2003-517222(P2003-517222)
(86)(22)出願日　平成14年7月30日(2002.7.30)
(85)翻訳文提出日　平成16年1月30日(2004.1.30)
(86)国際出願番号　PCT/US2002/024019
(87)国際公開番号　WO2003/012044
(87)国際公開日　平成15年2月13日(2003.2.13)
(31)優先権主張番号　60/308,652
(32)優先日　平成13年7月30日(2001.7.30)
(33)優先権主張国　米国(US)

(71)出願人　501097189
ジェネーラ・コーポレーション
アメリカ合衆国ペンシルバニア州１９４６２，プリマス・ミーティング，キャンパス・ドライブ　５１１０
(74)代理人　100078282
弁理士　山本　秀策
(74)代理人　100062409
弁理士　安村　高明
(74)代理人　100113413
弁理士　森下　夏樹
(72)発明者　ウゼル，　トーマス
アメリカ合衆国　ペンシルバニア　１９４６２，　プリマス　ミーティング，　キャンパス　ドライブ　５１１０
最終頁に続く

(54)【発明の名称】メクラウナギカテリン関連抗菌ペプチドおよび遺伝子

(57)【要約】
本発明は、Ｍｙｘｉｎｅ　ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドおよびこれらのペプチドをコードする遺伝子を包含する。本発明はまた、組成物およびこれらのペプチドを生成する方法ならびにこれらのペプチドを用いて微生物感染を予防および処置する方法を包含する。本発明の単離された核酸分子は、１つ以上の発現制御エレメントに作動可能に連結され得る。本発明はさらに、上記の単離された核酸分子のいずれかを含むベクターおよびこのベクターを含む宿主細胞を包含する。

【特許請求の範囲】
【請求項１】
以下：
（ａ）配列番号２または４を含むＭｙｘｉｎｅ　ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドのアミノ酸配列をコードする、単離された核酸分子；および
（ｂ）ストリンジェントな条件下で、配列番号２または４を含むＭｙｘｉｎｅ　ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドをコードするヌクレオチド配列を含む核酸分子の相補体にハイブリダイズする、単離された核酸分子、
からなる群より選択される、単離された核酸分子。
【請求項２】
前記核酸分子が、配列番号１および３からなる群より選択される、請求項１に記載の単離された核酸分子。
【請求項３】
前記核酸分子が、配列番号１のヌクレオチド４～５４０を含む、請求項２に記載の単離された核酸分子。
【請求項４】
前記核酸分子が、配列番号１のヌクレオチド４２７～５４０を含む、請求項２に記載の単離された核酸分子。
【請求項５】
前記核酸分子が、配列番号３のヌクレオチド２～５０５を含む、請求項１に記載の単離された核酸分子。
【請求項６】
前記核酸分子が、配列番号３のヌクレオチド４１６～５０５を含む、請求項１に記載の単離された核酸分子。
【請求項７】
以下：
（ａ）配列番号５または６を含むＭｙｘｉｎｅ　ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドのアミノ酸配列をコードする、単離された核酸分子；および
（ｂ）ストリンジェントな条件下で、配列番号５または６を含むＭｙｘｉｎｅ　ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドをコードするヌクレオチド配列を含む核酸分子の相補体にハイブリダイズする、単離された核酸分子、
からなる群より選択される、単離された核酸分子。
【請求項８】
前記核酸分子が、１つ以上の発現制御エレメントに作動可能に連結されている、請求項１～７のいずれか１項に記載の単離された核酸分子。
【請求項９】
請求項１～７のいずれか１項に記載の単離された核酸分子を含む、ベクター。
【請求項１０】
請求項９に記載のベクターを含む、宿主細胞。
【請求項１１】
請求項１～７のいずれか１項に記載の核酸分子を含むよう形質転換された、宿主細胞。
【請求項１２】
ペプチドまたはそのフラグメントを生成する方法であって、請求項１１に記載の宿主細胞を、前記核酸分子によりコードされるタンパク質またはタンパク質フラグメントが発現される条件下で培養する工程を包含する、方法。
【請求項１３】
前記宿主細胞が、原核生物宿主および真核生物宿主からなる群より選択される、請求項１２に記載の方法。
【請求項１４】
請求項１２に記載の方法により生成される、単離されたペプチドまたはそのフラグメント

。
【請求項１５】
以下：
（ａ）配列番号２、４、５、または６に示されるアミノ酸配列を含む、単離されたペプチド；
（ｂ）配列番号２、４、５、または６に示される配列のいずれかの少なくとも６アミノ酸を含む、単離されたペプチドフラグメント；
（ｃ）配列番号２、４、５、または６に示される配列のいずれかの保存的アミノ酸置換を含む、単離されたポリペプチド；および
（ｄ）配列番号２、４、５、または６に示されるアミノ酸配列のいずれかの天然に存在するアミノ酸配列ペプチド改変体、
からなる群より選択される、単離されたペプチドまたはそのフラグメント。
【請求項１６】
前記単離されたペプチドが、配列番号２、４、５、または６に示されるアミノ酸配列から本質的になる、請求項１５に記載の単離されたタンパク質フラグメント。
【請求項１７】
請求項１５に記載の１つ以上のペプチドを含む、抗菌組成物。
【請求項１８】
前記組成物が、局所投与または非経口投与に適している、請求項１７に記載の抗菌組成物。
【請求項１９】
哺乳動物における微生物感染を処置または予防する方法であって、該哺乳動物に、有効量のＭｙｘｉｎｅ　ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドを投与する工程を包含する、方法。
【請求項２０】
前記Ｍｙｘｉｎｅ　ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドが、配列番号２、４、５、または６からなる群より選択される、請求項１９に記載の方法。
【請求項２１】
前記哺乳動物が、ヒトである、請求項１９に記載の方法。
【請求項２２】
前記微生物感染が、細菌感染および真菌感染からなる群より選択される、請求項１９に記載の方法。
【請求項２３】
前記Ｍｙｘｉｎｅ　ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドが、経口経路、局所経路、および非経口経路からなる群より選択される経路により投与される、請求項１９に記載の方法。
【請求項２４】
前記局所投与が、吸入により達成される、請求項２３に記載の方法。
【請求項２５】
前記非経口投与が、静脈内投与、皮下投与、または筋内投与により達成される、請求項２４に記載の方法。
【発明の詳細な説明】
【技術分野】
【０００１】
（関連出願の相互参照）
本願は、２００２年７月３０日に出願された、米国仮出願６０／３０８，６５２の利益を主張する。
【０００２】
（発明の分野）
本発明は、抗生ペプチドに対する遺伝子およびこのペプチドを生成するためのそれらの使

用に関する。本発明はまた、抗菌ペプチドおよび抗生物質としてのそれらの使用に関する。

【背景技術】

【０００３】

（発明の背景）

カテリシジン遺伝子ファミリー（Ｚａｎｅｔｔｉら（１９９５）ＦＥＢＳ Ｌｅｔｔｅｒｓ ３７４，１－５；Ｚａｎｅｔｔｉら（１９９７）Ａｎｎ．Ｎｅｗ Ｙｏｒｋ．Ａｃａｄ．Ｓｃｉ．１４７－１６２；Ｇｅｎｎａｒｏら（２０００）Ｂｉｏｐｏｌｙｍｅｒｓ Ｐｅｐｔｉｄｅ Ｓｃｉｅｎｃｅ ５５，３１－４９）は、シグナル配列、カテリン様酸性スペーサー、および抗生ペプチドからなるプレプロペプチドをコードする（図１）。以下のような種々の哺乳動物由来のカテリシジンが、現在知られている：Ｂｏｓ ｔａｕｒｕｓ（ウシ）、Ｃａｐｒａ ｈｉｒｃｕｓ（ヤギ）、Ｃａｖｉａ ｐｏｒｃｅｌｌｕｓ（モルモット）、Ｅｑｕｕｓ ｃａｂａｌｌｕｓ（ウマ）、Ｈｏｍｏ ｓａｐｉｅｎｓ（ヒト）、Ｍｕｓ ｍｕｓｃｕｌｕｓ（マウス）、Ｏｒｙｃｔｏｌａｇｕｓ ｃｕｎｎｉｃｕｌｕｓ（ウサギ）、Ｏｖｉｓ ａｒｉｅｓ（ヒツジ）、およびＳｕｓ ｓｃｒｏｆａ（ブタ）。公知の抗生ペプチドは、哺乳動物種内および哺乳動物種間の両方で、大きく異なるが、ｃＤＮＡクローンの配列分析により明らかにされたように、関連するシグナル配列およびカテリン様酸性スペーサーは、抗生ペプチド自体が異なる種の間で大いに保存されている。

【０００４】

哺乳動物カテリシジンは、主に、骨髄性細胞において合成され、そして循環好中球に豊富である。２０数個のカテリシジン遺伝子の構造が公知であり、そして通常４つのエキソンを含み、そのうちの最後のエキソンは、カテリンのいくつかのＣ末端残基と共に抗生ペプチドをコードする。８、１１、およびそれ以上の番号を付されたカテリシジン遺伝子のクラスターが、それぞれ、Ｏｖｉｓ、Ｂｏｓ、およびＳｕｓにおいて報告されている（Ｇｅｎｎａｒｏら（２０００）Ｂｉｏｐｏｌｙｍｅｒｓ Ｐｅｐｔｉｄｅ Ｓｃｉｅｎｃｅ ５５，３１－４９）。

【０００５】

２つの抗生ペプチドが、Ｍｙｘｉｎｅ ｇｌｕｔｉｎｏｓａ（大西洋メクラウナギ）（哺乳動物にわずかにしか関連しない遠縁の脊椎動物脊索動物分類群）の腸組織から単離された（Ｋａｕｍａｙａら（１９９６）Ｐｅｐｔｉｄｅｓ：Ｃｈｅｍｉｓｔｒｙ ａｎｄ Ｂｉｏｌｏｇｙ，Ｍａｙｆｌｏｗｅｒ Ｓｃｉｅｎｔｉｆｉｃ，１８９～１９１頁）。これらのペプチドは、１つまたは２つの臭素化トリプトファンを含有するという点で、抗菌ペプチド間で独特である。本明細書中には、以前に単離された抗生ペプチドと同じファミリー由来の２つの新規抗生ペプチドをコードするメクラウナギｃＤＮＡ配列が記載される。これらの配列は、Ｍｙｘｉｎｅ抗生ペプチドをカテリシジンファミリーのメンバーとして同定する。

【発明の開示】

【課題を解決するための手段】

【０００６】

（発明の要旨）

本発明は、以下：配列番号２または４を含むＭｙｘｉｎｅ ｇｌｕｔｉｎｏｓａカテリン（ｃａｔｈｅｌｉｎｅ）関連抗菌ペプチドのアミノ酸配列をコードする単離された核酸分子；およびストリンジェントな条件下で、配列番号２または４を含むＭｙｘｉｎｅ ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドをコードするヌクレオチド配列を含む核酸分子の相補体にハイブリダイズする単離された核酸分子、からなる群より選択される単離された核酸分子を包含する。いくつかの実施形態において、この核酸分子は、配列番号１および３；配列番号１のヌクレオチド４～５４０；配列番号１のヌクレオチド４２７～５４０；配列番号３のヌクレオチド２～５０５；または配列番号３のヌクレオチド４１６～５０５を含む。

【０００７】
本発明はまた、以下：配列番号５または６を含むＭｙｘｉｎｅ ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドのアミノ酸配列をコードする単離された核酸分子；およびストリンジェントな条件下で、配列番号５または６を含むＭｙｘｉｎｅ ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドをコードするヌクレオチド配列を含む核酸分子の相補体にハイブリダイズする単離された核酸分子、からなる群より選択される、単離された核酸分子を包含する。
【０００８】
本発明の単離された核酸分子は、１つ以上の発現制御エレメントに作動可能に連結され得る。本発明はさらに、上記の単離された核酸分子のいずれかを含むベクターおよびこのベクターを含む宿主細胞を包含する。
【０００９】
本発明は、ペプチドまたはそのフラグメントを生成するための方法を包含する。この方法は、上記宿主細胞を、上記核酸分子によりコードされるタンパク質またはタンパク質フラグメントが発現される条件下で培養する工程を包含する。この宿主細胞は、原核生物または真核生物であり得る。本発明は、この方法によって生成される組換えペプチドおよびそのフラグメントを包含する。
【００１０】
本発明は、以下：配列番号２、４、５、または６に示されるアミノ酸配列を含む、単離されたペプチド；配列番号２、４、５、または６に示される配列のいずれかの少なくとも６アミノ酸を含む、単離されたペプチドフラグメント；配列番号２、４、５、または６に示される配列のいずれかの保存的アミノ酸置換を含む、単離されたペプチド；および配列番号２、４、５、または６に示されるアミノ酸配列のいずれかの天然に存在するアミノ酸配列ペプチド改変体、からなる群より選択される、単離されたペプチドまたはそのフラグメントを包含する。
【００１１】
本発明はさらに、本発明の１つ以上のペプチドを含む抗菌組成物を包含する。この抗菌組成物は、局所投与または非経口投与に適切であり得る。
【００１２】
本発明は、哺乳動物における微生物感染を処置または予防する方法を包含する。この方法は、この哺乳動物に、有効量のＭｙｘｉｎｅ ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドを投与する工程を包含する。いくつかの実施形態において、Ｍｙｘｉｎｅ ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチドは、配列番号２、４、５、または６からなる群より選択され、そして哺乳動物は、細菌感染または真菌感染の危険を有するか、それに罹患しているヒトである。Ｍｙｘｉｎｅ ｇｌｕｔｉｎｏｓａカテリン関連抗菌ペプチド含有組成物は、経口経路、局所経路、および非経口経路からなる群より選択される経路により投与され得る。局所投与が用いられる場合、投与は吸入により達成され得る。非経口経路が用いられる場合、投与は静脈内投与、皮下投与、または筋内投与により達成され得る。
【００１３】
（詳細な説明）
（概説）
出願人らは、大西洋メクラウナギ（Ａｔｌａｎｔｉｃ ｈａｇｆｉｓｈ）（Ｍｙｘｉｎｅ ｇｌｕｔｉｎｏｓａ）由来の２つの新規の遺伝子を単離し、そして配列決定した。これらの遺伝子（配列番号１および３、図１）は、２つのタンパク質（配列番号２および４、図１）をコードし、これらタンパク質の各々は、Ｎ末端のシグナル配列、カテリン（ｃａｔｈｅｌｉｎ）様配列、およびＣ末端の抗菌ペプチドをコードする。これらの遺伝子由来の抗細菌ペプチド配列（ＧＷＦＫＫＡＷＲＫＶＫＨＡＧＲＲＶＬＤＴＡＫＧＶＧＲＨＹＬＮＮＷＬＮＲＹＲＧ（配列番号５）、およびＧＷＦＫＫＡＷＲＫＶＫＮＡＧＲＶＬＫＧＶＧＩＨＹＧＶＧＬＩＧ（配列番号６））は、米国特許第５，７３４，０１５に記載されたメクラウナギのペプチドに高度に類似であるが異なっている（図４）。従って、本発明は

、単離されたメクラウナギ遺伝子およびそれらにコードされるカテリン関連抗菌性ペプチドを包含する。

【００１４】

本発明において、Ｃ末端グリシン残基が翻訳後に、ペプチジルグリシンα－アミド化モノオキシゲナーゼ（ＰＡＭ）（Ｐｒｉｇｇｅら（２０００）Ｃｅｌｌ．Ｍｏｌ．Ｌｉｆｅ Ｓｃｉ．５７、１２３６－１２５９）のような酵素によってＣ末端のアミノ残基に転換された抗菌性ペプチド（配列番号１１および１２）、そして／または１つ以上のＴｒｐ残基が翻訳後にブロモ化酵素によってＢｒ－Ｔｒｐ残基に転換された抗菌性ペプチド（Ｓｈｉｎｎａｒら、（２０００）ＦＡＳＥＢ Ｊｏｕｒｎａｌ １４、Ａ１４４８）が包含される。

【００１５】

また本発明において、メクラウナギカテリン関連遺伝子ファミリー、それらのコードするタンパク質、およびそれらの抗菌性ペプチドが包含される。別の実施形態は、これら遺伝子の発現およびこれらを含む細胞についての方法である。メクラウナギカテリン関連遺伝子によってコードされるタンパク質の組換え産生は、これが関連の抗菌性ペプチドの産生である場合、本発明に包含される。

【００１６】

（定義）

本明細書中で使用される場合、用語「抗菌」とは、ある化合物が微生物の増殖を阻害するかまたは不可逆的に防止する能力をいう。このような阻害または防止は、殺菌性活性または静菌性阻害を介し得る。従って、用語「殺菌性阻害」とは、本明細書中で使用される場合、ある化合物が標的の微生物を殺すかまたは微生物に不可逆的に損傷を与える能力をいう。用語「静菌性阻害」とは、本明細書中で使用される場合、抗菌化合物が、標的微生物を殺さずに増殖を阻害する能力をいう。微生物の増殖を示す現況（治療状態）または微生物の増殖を支援するリスクのある環境（予防）における微生物の殺菌性阻害または静菌性阻害は、この定義のうちに包含される。

【００１７】

本明細書中で使用される場合、「細胞株」は、インビトロで多世代にわたって安定に増殖し得る、初代細胞のクローンである。

【００１８】

ＤＮＡ「コード配列」は、適切な調節配列の制御下に配置される場合に、インビボで転写され、そしてポリペプチドへと翻訳される、二重鎖ＤＮＡ配列である。コード配列の境目は、５’（アミノ）末端の開始コドンおよび３’（カルボキシ）末端の翻訳終止コドンによって決定される。ポリアデニル化シグナルおよび転写終結配列は、通常、コード配列から３’側に位置付けられる。

【００１９】

本明細書中で使用される場合、「核酸分子」とは、デオキシリボヌクレオチド（アデニン、グアニン、チミン、および／もしくはシトシン）、またはリボヌクレオチド（アデニン、グアニン、ウラシル、および／もしくはシトシン）のポリマー形態をいい、そしてその一本鎖形態、または二重鎖ヘリックス形態およびＲＮＡのいずれかを含み得る。この用語は、分子の一次構造および二次構造のみをいい、そしていずれかの特定の三次形態に限定されない。特定の二重鎖ＤＮＡ分子の構造を考察する際、配列は、翻訳されないＤＮＡ鎖（例えば、ｍＲＮＡに相同な配列を有する鎖）に沿う、５’から３’方向での配列のみを提供する通常の慣習に従って、本明細書中に記載され得る。転写制御配列および翻訳制御配列は、宿主細胞中でコード配列の発現のために提供されるＤＮＡ調節配列（例えば、プロモーター、エンハンサー、ポリアデニル化シグナル、ターミネーターなど）である。

【００２０】

本明細書中で使用される場合、「プロモーター配列」は、細胞中でＲＮＡポリメラーゼを結合し得、そして下流（３’方向）のコード配列の転写を開始し得るＤＮＡ調節配列である。本発明の定義の目的のために、プロモーター配列は、転写開始部位の側のその配列の

３’末端で（包括的に）結合され、そして上流（５’方向）に延びて、バックグラウンドを超えて検出可能なレベルで転写を開始するための最小の塩基数またはそれに必須なエレメントを含む。プロモーター配列中には、転写開始部位、およびＲＮＡポリメラーゼの結合を担うタンパク質結合ドメインが見出される。真核生物のプロモーターは、しばしば、「ＴＡＴＡ」ボックスおよび「ＣＡＴ」ボックスを含むが、必ずしも含むとは限らない。

【００２１】

「シグナル配列」は、コード配列の前方に含まれ得るか、またはエンベロープタンパク質由来のネイティブなアミノ酸シグナル配列が使用され得る。この配列は、細胞表面にそのポリペプチドを向けるか、または培地中にそのポリペプチドを分泌するように宿主細胞に伝えるシグナルペプチド（ポリペプチドに対してＮ末端）をコードする。このシグナルペプチドは、タンパク質が細胞から離れる前に宿主細胞によって切り取られる。シグナル配列は、原核生物および真核生物にネイティブな、種々のタンパク質に関連して見出され得る。例えば、ネイティブな酵母タンパク質のα因子、は、酵母から分泌され、そしてそのシグナル配列が異種性タンパク質に結合されて媒地中に分泌され得る（例えば、米国特許第４，５４６，０８２号およびＥＰ０１１６２０１を参照のこと）。さらに、このα因子およびそのアナログは、種々の酵母（例えば、ＳａｃｃｈａｒｏｍｙｃｅｓおよびＫｌｕｙｖｅｒｏｍｙｃｅｓ（ＥＰ８８３１２３０６．９；ＥＰ０３２４２７４公開およびＥＰ０３０１６６９））から異種性タンパク質を分泌することを見出された。哺乳動物細胞における使用例は、第ＶＩＩＩｃ因子軽鎖を発現するために使用されるｔＰＡシグナルである。

【００２２】

外来性ＤＮＡまたは異種性ＤＮＡが細胞内に導入される場合、細胞は、このようなＤＮＡによって「形質転換」されている。形質転換ＤＮＡは、細胞ゲノムを構築する染色体ＤＮＡ中に組み込まれても（共有結合で連結されても）よいし、組み込まれなくても（共有結合で連結されなくても）よい。例えば、原核生物において、形質転換ＤＮＡは、プラスミドまたはウイルスベクターのようなエピソーム性エレメント上で維持され得る。真核生物に関して、安定に形質転換された細胞とは、その形質転換ＤＮＡが染色体中に組み込まれて、その結果染色体複製を介して娘細胞に遺伝される細胞である。この安定性は、真核生物細胞が、この形質転換ＤＮＡを含有する娘細胞の集団を含む細胞株または細胞クローンを構築する能力によって実証される。

【００２３】

コード配列は、細胞内で、ＲＮＡポリメラーゼがｍＲＮＡにコード配列を転写し、次いでコード配列によってコードされたタンパク質へと翻訳される場合に、転写制御配列および翻訳制御配列の「制御下」にある。

【００２４】

本明細書中で使用される場合、「ベクター」は、結合されたセグメントの複製を引き起こすよう別のＤＮＡセグメントが結合されたレプリコン（例えば、プラスミド、ウイルス、ファージまたはコスミド）である。

【００２５】

（核酸分子）

２つのｃＤＮＡ配列（図１）は、８４８ヌクレオチド（配列番号１）および８６３ヌクレオチド（配列番号２）からなる。配列番号１の配列は、開始コドンおよび５’側の３つのさらなるヌクレオチドを含む；配列番号３の配列は、配列番号１とのアラインメントにより判断すると、５’末端に８つのコードヌクレオチドの欠損を有する。図１のアラインメントにおいて、８１４ヌクレオチドが２つの配列間で相当しており、７８５ヌクレオチドが合致する（９６．４％）。これら２つの推定のシグナル配列およびカテリン配列をコードするヌクレオチドは、非常に類似しており、７０個のシグナルヌクレオチドのうち６９個が同一（９８．６％）であり、そして３４５個のカテリンヌクレオチドのうち３３２個（９６．２％）が合致し；全体で４１５個のうち４０１個（９６．６％）が合致する。ｃＤＮＡ配列のこれらの部分において、挿入も欠失も認められない。

【００２６】
対照的に、このペプチドおよび３’非翻訳配列は、いくつかの挿入／欠失の変化の証拠を示し、この証拠は、これら配列の類似性の比較を困難にする。ペプチドコード配列において、配列番号３は、配列番号１との比較において、３つのギャップ（３ヌクレオチド、９ヌクレオチドおよび４ヌクレオチド）を有し；最後のギャップは終結コドンの直前の１７ヌクレオチドの代わりにフレームシフトを生じる。配列番号１のペプチドコード配列は、配列番号３の配列に対して、２１ヌクレオチドのギャップを有し；この結果、配列に沿ってさらに２６ヌクレオチド先の新規の終結コドンが、配列番号１のペプチドの末端の印となる。このアラインメントにおいて、配列番号１についての最後の１０個のコードヌクレオチドは、配列番号３のｃＤＮＡの３’非翻訳領域における１０個のヌクレオチドに対応する。全体で、ペプチドをコードする９８ヌクレオチドは比較可能であり；これらのうちの８９個（９０．８％）が合致する。

【００２７】
両方のｃＤＮＡの３’非翻訳末端は、－ｔｇｇ－反復を有する（配列番号１については１０マー、そして配列番号３については１７マー）。この－ｔｇｇ－反復を除いて、３’末端にある２４８ヌクレオチドを比較し得る；これらのうち、２４３（９８．０％）がマッチする。

【００２８】
本発明の核酸分子は、配列番号２を有するタンパク質、配列番号４を有するタンパク質、および本明細書中に記載される関連のタンパク質をコードする核酸分子（好ましくは、単離された形態）を含む。本明細書中で使用される場合、「核酸」は、上記に定義されるようなタンパク質もしくはペプチドをコードするか、またはそのようなペプチドをコードする核酸配列に対して相補的であるか、または適切なストリンジェンシーの条件下でそのような核酸にハイブリダイズし、そしてそれらに結合したまま安定に残存するか、またはそれらのペプチド配列と、少なくとも５５％の配列同一性、８５％の配列同一性、好ましくは、少なくとも８５％、そしてより好ましくは少なくとも９０％、最も好ましくは、少なくとも９５～９９％の配列同一性を共有するポリペプチドをコードする、ＲＮＡもしくはＤＮＡとして規定される。特に意図されるのは、天然供給源に由来するものであれ合成であれ、ゲノムＤＮＡ、ｃＤＮＡ、ｍＲＮＡ、およびアンチセンス分子、ならびに、代替的な骨格に基づくかまたは代替的な塩基を含む核酸である。しかし、このようなハイブリダイズする核酸または相補的な核酸は、あらゆる先行技術の核酸（本発明に従うタンパク質をコードする核酸をコードするもの、そのような核酸に適切なストリンジェンシー条件下でハイブリダイズするもの、またはそのような核酸に対して相補的であるものを含む）に対して新規かつ非自明であるものとしてさらに規定される。

【００２９】
本発明の好ましい実施形態は、本発明のメクラウナギカテリン関連タンパク質をコードする核酸分子（例えば、配列番号１もしくは配列番号１のヌクレオチド４～５４０（または５４３）により規定されるオープンリーディングフレームを含む核酸分子、配列番号１もしくは配列番号１のヌクレオチド４～５４０（または５４３）により規定されるオープンリーディングフレームから本質的になる核酸分子、または配列番号１もしくは配列番号１のヌクレオチド４～５４０（または５４３）により規定されるオープンリーディングフレームからなる核酸分子；あるいは例えば、配列番号３もしくは配列番号３のヌクレオチド２～５０５（または５０８）により規定されるオープンリーディングフレームを含む核酸分子、配列番号３もしくは配列番号３のヌクレオチド２～５０５（または５０８）により規定されるオープンリーディングフレームから本質的になる核酸分子、または配列番号３もしくは配列番号３のヌクレオチド２～５０５（または５０８）により規定されるオープンリーディングフレームからなる核酸分子）である。本発明のさらに好ましい実施形態は、本発明のメクラウナギ抗菌ペプチドをコードする核酸分子（例えば、配列番号１のヌクレオチド４２７～５４０、配列番号２のアミノ酸１４２～１７９をコードするヌクレオチドもしくは配列番号５をコードするヌクレオチドを含む核酸分子、配列番号１のヌクレオ

チド４２７～５４０、配列番号２のアミノ酸１４２～１７９をコードするヌクレオチドもしくは配列番号５をコードするヌクレオチドから本質的になる核酸分子、または配列番号１のヌクレオチド４２７～５４０、配列番号２のアミノ酸１４２～１７９をコードするヌクレオチドもしくは配列番号５をコードするヌクレオチドからなる核酸分子；あるいは例えば、配列番号３のヌクレオチド４１６～５０５、配列番号４のアミノ酸１３９～１６８をコードするヌクレオチドもしくは配列番号６をコードするヌクレオチドを含む核酸分子、配列番号３のヌクレオチド４１６～５０５、配列番号４のアミノ酸１３９～１６８をコードするヌクレオチドもしくは配列番号６をコードするヌクレオチドから本質的になる核酸分子、または配列番号３のヌクレオチド４１６～５０５、配列番号４のアミノ酸１３９～１６８をコードするヌクレオチドもしくは配列番号６をコードするヌクレオチドからなる核酸分子）である。

【００３０】

ヌクレオチド配列レベルまたはアミノ酸配列レベルでの相同性または配列同一性は、配列類似性検索のために調整されたプログラムｂｌａｓｔｐ、ｂｌａｓｔｎ、ｂｌａｓｔｘ、ｔｂｌａｓｔｎおよびｔｂｌａｓｔｘによって使用されるアルゴリズム（Ａｌｔｓｃｈｕｌら（１９９７）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．２５、３３８９－３４０２およびＫａｒｌｉｎら（１９９０）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８７、２２６４－２２６８（両方とも全体的に参考として援用される））を使用して、ＢＬＡＳＴ（Ｂａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ）分析によって決定される。このＢＬＡＳＴプログラムによって使用されるアプローチは、照会（ｑｕｅｒｙ）配列とデータベース配列との間で、ギャップを有する（非連続性）およびギャップを有さない（連続性）類似のセグメントをまず考慮に入れ得、次いで、同定されたすべてのマッチの統計学的有意性を評価し得、そして最後に、予め選択された有意性閾値を満たすマッチのみをまとめ得る。配列データベースの類似性検索に関する基本的な問題の考察については、Ａｌｔｓｃｈｕｌら（１９９４）Ｎａｔｕｒｅ Ｇｅｎｅｔｉｃｓ ６、１１９－１２９（これは、全体的に参考として援用される）を参照のこと。ヒストグラム、記述、アラインメント、期待値（すなわち、データベース配列に対するマッチを報告するための統計学的有意性の閾値）、カットオフ、マトリクス、およびフィルター（低複雑性）の検索パラメーターは、デフォルト設定である。ｂｌａｓｔｐ、ｂｌａｓｔｘ、ｔｂｌａｓｔｎおよびｔｂｌａｓｔｘにより使用されるデフォルトのスコア付けマトリクスは、８５以上のヌクレオチド長またはアミノ酸長に対する照会配列について推奨される、ＢＬＯＳＵＭ６２マトリクス（Ｈｅｎｉｋｏｆｆら（１９９２）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８９、１０９１５－１０９１９（全体的に参考として援用される））である。

【００３１】

ｂｌａｓｔｎについて、スコア付けマトリクスは、Ｍ（すなわち、マッチ残基の対についての報酬（ｒｅｗａｒｄ）スコア）対Ｎ（すなわち、ミスマッチ残基についてのペナルティスコア）の比によって設定され、ここでＭおよびＮについてのデフォルト値は、それぞれ、＋５および－４である。４つのｂｌａｓｔｎパラメーターを、以下のように調整した：Ｑ＝１０（ギャップ作成ペナルティ）；Ｒ＝１０（ギャップ伸長ペナルティ）；ｗｉｎｋ＝１（照会物に沿って、ｗｉｎｋ番目の位置毎におけるワードヒットを作成する）；およびｇａｐｗ＝１６（ギャップ形成されたアラインメントを作成するウィンドウ幅を設定する）。等価なＢｌａｓｔｐパラメーター設定は、以下の通りであった：Ｑ＝９；Ｒ＝２；ｗｉｎｋ＝１；およびｇａｐｗ＝３２。ＧＣＧパッケージバージョン１０．０において利用可能である、配列間のＢｅｓｔｆｉｔ比較は、ＤＮＡパラメーターＧＡＰ＝５０（ギャップ作成ペナルティ）およびＬＥＮ＝３（ギャップ伸長ペナルティ）を使用し、そしてタンパク質比較における等価な設定は、ＧＡＰ＝８およびＬＥＮ＝２である。

【００３２】

「ストリンジェント条件」とは、（１）洗浄のために低イオン強度および高温を使用する条件（例えば、５０℃での、０．０１５Ｍ ＮａＣｌ／０．００１５クエン酸ナトリウム

／０．１％ＳＤＳ）、または（２）ハイブリダイゼーションの間に、ホルムアミドのような変性剤を使用する条件（例えば、４２℃での、０．１％ウシ血清アルブミンを伴う５０％（容量／容量）ホルムアミド／０．１％Ｆｉｃｏｌｌ／０．１％ポリビニルピロリドン／７５０ｍＭ　ＮａＣｌ、７５ｍＭクエン酸ナトリウムを伴う５０ｍＭリン酸ナトリウム緩衝剤（ｐＨ６．５））である。別の例は、４２℃での５０％ホルムアミド、５×ＳＳＣ（０．７５Ｍ　ＮａＣｌ、０．０７５Ｍクエン酸ナトリウム）、５０ｍＭリン酸ナトリウム（ｐＨ６．８）、０．１％ピロリン酸ナトリウム、５×デンハルト溶液、超音波処理済サケ精子ＤＮＡ（５０μｇ／ｍｌ）、０．１％ＳＤＳ、および１０％デキストラン硫酸中におけるハイブリダイゼーションと、０．２×ＳＳＣおよび０．１％ＳＤＳ中において４２℃での洗浄である。当業者は、明瞭かつ検出可能なハイブリダイゼーションシグナルを得るために適切にストリンジェンシー条件を容易に決定し得、そして変動させ得る。好ましい核酸分子は、配列番号１または３の相補体に上記条件下でハイブリダイズし、そして抗菌活性を有するポリペプチドをコードする核酸分子である。さらにより好ましい核酸分子は、配列番号１または３のオープンリーディングフレームの相補鎖に上記条件下でハイブリダイズする核酸分子である。

【００３３】

本明細書中で使用される場合、「単離された」核酸分子は、そのネイティブの環境から取り出されたか、または核酸分子がその核酸供給源由来の他のポリペプチドをコードする夾雑核酸から実質的に分離されている場合の、核酸分子、ＤＮＡまたはＲＮＡを意味する。例えば、ベクター中に含まれる組換えＤＮＡ分子は、本発明の目的のために、単離されたとみなされる。単離されたＤＮＡ分子のさらなる例としては、異種宿主細胞中に維持されている組換えＤＮＡ分子、または溶液中にある精製（部分的または実質的）されたＤＮＡ分子が挙げられる。単離されたＲＮＡ分子としては、本発明のＤＮＡ分子のインビボまたはインビトロでのＲＮＡ転写物が挙げられる。本発明に従う単離された核酸分子はさらに、合成的に生成されたこのような分子を含む。本発明の単離された核酸分子としては、配列番号１または３に示されるオープンリーディングフレーム（ＯＲＦ）を含むＤＮＡ分子；成熟抗菌ペプチドタンパク質のコード配列を含むＤＮＡ分子；および上記の配列とは実質的に異なる配列を含むが、遺伝暗号の縮重に起因して、なお抗菌ペプチドタンパク質をコードするＤＮＡ分子が挙げられる。遺伝暗号は当該分野で周知であるので、このような縮重改変体を生成することは当業者には慣用的である。別の局面では、本発明は、本発明の核酸分子のポリヌクレオチドの一部に対して、ストリンジェントなハイブリダイゼーション条件下でハイブリダイズするポリヌクレオチドを含む単離された核酸分子を提供する。本発明はさらに、抗菌ペプチドタンパク質の部分、アナログまたは誘導体をコードする、本発明の核酸分子の改変体に関する。改変体は、天然の対立遺伝子改変体のように、天然において生じ得る。「対立遺伝子改変体」とは、生物の染色体上にある所定の遺伝子座を占める遺伝子のいくつかの代替的形態のうちの１つを意図する。

【００３４】

本発明はさらに、コード核酸分子のフラグメントを提供する。本明細書中で使用される場合、コード核酸分子のフラグメントは、タンパク質コード配列全体のなかの小さな部分をいう。フラグメントのサイズは、意図される用途によって決定される。例えば、フラグメントが、タンパク質の活性部分をコードするように選択される場合、フラグメントは、そのタンパク質の機能的領域（１つまたは複数）をコードするのに十分な大きさである必要があるか、または図１（配列番号２および４）のメクラウナギタンパク質の間の相同性領域をコードし得る。本発明の１つの実施形態において、抗菌ペプチドフラグメントは、好ましくは少なくとも１５～２０アミノ酸、より好ましくは２０～３０アミノ酸、そして最も好ましくは３０～４０アミノ酸である。フラグメントが核酸プローブまたはＰＣＲプライマーとして使用される場合、フラグメントの長さは、プロービング／プライミングの間に比較的少数の偽陽性を得るように選択される。

【００３５】

ポリメラーゼ連鎖反応（ＰＣＲ）のためのプローブもしくは特異的プライマーとして使用

されるか、または本発明のタンパク質をコードする遺伝子配列を合成するために使用される、本発明のコード核酸分子のフラグメント（すなわち、合成オリゴヌクレオチド）は、化学技術（例えば、Ｍａｔｔｅｕｃｃｉら（１９８１）Ｊ．Ａｍ．Ｃｈｅｍ．Ｓｏｃ．１０３、３１８５－３１９１）のホスホトリエステル方法または自動化合成方法の使用）によって、容易に合成され得る。さらに、より大きなＤＮＡセグメントは、周知の方法（例えば、遺伝子の種々のモジュールセグメントを規定するオリゴヌクレオチド群を合成し、次いで、オリゴヌクレオチドを連結して完全な改変遺伝子を構築すること）によって、容易に調製され得る。

【００３６】

本発明のコード核酸分子はさらに、診断目的およびプローブ目的のために検出可能な標識を含むように改変され得る。種々のこのような標識が当該分野で公知であり、そして本明細書中に記載されるコード分子と共に容易に使用され得る。このような標識としては、ビオチン、放射性標識ヌクレオチドなどが挙げられるが、これらに限定されない。当業者は、当該分野で公知の任意の標識を使用して、標識されたコード核酸分子を獲得し得る。

【００３７】

翻訳の間にタンパク質配列に組み込まれるアミノ酸を欠失、付加または変更することよる、一次構造自体に対する改変が、そのタンパク質の活性を壊すことなくなされ得る。このような置換または他の変更は、本発明の意図される範囲内にある核酸によってコードされるアミノ酸配列を有するタンパク質を生じさせる。

【００３８】

（他の関連核酸分子の単離）

上記のように、配列番号１および３を有するメクラウナギ核酸分子の同定は、当業者が、本明細書中に記載される配列に加えて、カテリン関連タンパク質ファミリーの他のメンバーをコードする核酸分子を単離することを可能にする。

【００３９】

基本的に、当業者は容易に、配列番号２または４のアミノ酸配列を使用して、適切な細胞から調製される発現ライブラリーをスクリーニングするための抗体プローブを作製し得る。代表的には、精製タンパク質（以下に記載される）で免疫された哺乳動物（例えば、ウサギ）由来のポリクローナル抗血清またはモノクローナル抗体を使用して、哺乳動物のｃＤＮＡライブラリーまたはゲノム発現ライブラリー（例えば、λｇｔ１１ライブラリー）を探索し、このタンパク質ファミリーの他のメンバーの適切なコード配列を獲得し得る。クローン化されたｃＤＮＡ配列は、融合タンパク質として発現され得るか、それ自体の制御配列を使用して直接的に発現され得るか、またはタンパク質発現のために使用される特定の宿主に適切な制御配列を使用した構築物により発現され得る。

【００４０】

あるいは、本明細書中に記載されるコード配列の一部分を合成し得、そして任意の哺乳動物生物由来のメクラウナギカテリン関連タンパク質ファミリーのメンバーをコードするＤＮＡを回収するためのプローブとして使用し得る。約１８～２０ヌクレオチドを含むオリゴマー（約６～７アミノ酸のストレッチをコードする）を調製し、そしてストリンジェントな条件下または過度な偽陽性レベルを排除するに十分なストリンジェンシーの条件下においてハイブリダイゼーションを得るように、ゲノムＤＮＡライブラリーまたはｃＤＮＡライブラリーをスクリーニングするために使用し得る。

【００４１】

さらに、オリゴヌクレオチドプライマーの対が、コード核酸分子を選択的にクローン化するためのポリメラーゼ連鎖反応（ＰＣＲ）における使用のために調製され得る。このようなＰＣＲプライマーを使用するためのＰＣＲの変性／アニール／伸長サイクルは、当該分野で周知であり、そして他のコード核酸分子を単離する際の使用のために容易に適合され得る。

【００４２】

（組換え核酸分子）

本発明はさらに、メクラウナギカテリン関連タンパク質コード配列を含む組換えＤＮＡ分子（ｒＤＮＡ）を提供する。本明細書中で使用される場合、ｒＤＮＡ分子は、インビトロでの分子操作に供されたＤＮＡ分子である。ｒＤＮＡ分子を作製する方法は当該分野で周知である（例えば、Ｓａｍｂｒｏｏｋら（２００１）Ｍｏｌｅｃｕｌａｒ Ｃｌｏｎｉｎｇ：Ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ，Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｌａｂｏｒａｔｏｒｙ Ｐｒｅｓｓを参照のこと）。好ましいｒＤＮＡ分子では、コードＤＮＡ分子は、発現制御配列および／またはベクター配列に作動可能に連結される。

【００４３】

本発明のタンパク質ファミリーコード配列の１つが作動可能に連結されるベクターおよび／または発現制御配列の選択は、当該分野で周知のように、所望される機能的特性（例えば、タンパク質発現）および形質転換される宿主細胞に直接的に依存する。本発明により意図されるベクターは、少なくとも、複製または宿主染色体への挿入を指示し得、そして好ましくは、そのｒＤＮＡ分子中に含まれる構造遺伝子の発現をもまた指示し得る。

【００４４】

作動可能に連結されたタンパク質コード配列の発現を調節するために使用される発現制御エレメントは、当該分野で公知であり、そしてこれには、誘導性プロモーター、構成的プロモーター、分泌シグナル、および他の調節エレメントが挙げられるが、これらに限定されない。好ましくは、誘導性プロモーターは、容易に制御される（例えば、宿主細胞の培地中の栄養素に対して応答性である）。

【００４５】

１実施形態において、コード核酸分子を含むベクターは、原核生物レプリコン（すなわち、自律的な複製を指向する能力、および形質転換された原核生物宿主（例えば、細菌宿主細胞）内の染色体外の組換えＤＮＡ分子の維持力を有するＤＮＡ配列）を含む。このようなレプリコンは、当該分野で周知である。さらに、原核生物レプリコンを含むベクターはまた、その発現によって薬物耐性のような検出可能なマーカーを与える遺伝子を含み得る。代表的な細菌薬物耐性遺伝子は、アンピシリンまたはテトラサイクリンに対する耐性を与える遺伝子である。

【００４６】

原核生物レプリコンを含むベクターは、細菌宿主細胞（例えば、Ｅ．ｃｏｌｉ．）におけるコード遺伝子の発現（転写および翻訳）を指向し得る原核生物プロモーターまたはバクテリオファージプロモーターをさらに含み得る。プロモーターは、ＲＮＡポリメラーゼが結合しそして転写を生じることを可能にするＤＮＡ配列によって形成される発現制御エレメントである。細菌宿主と適合性のプロモーター配列は、代表的には、本発明のＤＮＡセグメントの挿入のために都合のよい制限部位を含むプラスミドベクター中に提供される。このようなベクタープラスミドの代表的なものは、ｐＵＣ８、ｐＵＣ９、ｐＢＲ３２２およびｐＢＲＢ２９（Ｂｉｏ－Ｒａｄ Ｌａｂｏｒａｔｏｒｉｅｓ）、ｐＰＬ２２３、およびｐＫＫ２２３（Ｐｈａｒｍａｃｉａ）である。

【００４７】

真核生物細胞に適合性である（好ましくは、脊椎動物細胞と適合性である）発現ベクターはまた、コード配列を含むｒＤＮＡ分子を形成するために使用され得る。真核生物細胞発現ベクターは当該分野において周知であり、そしていくつかの市販供給元から利用可能である。代表的には、このようなベクターは、所望のＤＮＡセグメントの挿入のために都合のよい制限部位を含んで提供され得る。このようなベクターの代表的なものは、ｐＳＶＬおよびｐＫＳＶ－１０（Ｐｈａｒｍａｃｉａ）、ｐＢＰＶ－１／ｐＭＬ２ｄ（Ｉｎｔｅｒｎａｔｉｏｎａｌ Ｂｉｏｔｅｈｎｏｌｏｇｉｅｓ）、ｐＴＤＴ１（ＡＴＣＣ）、本明細書中に記載のベクターｐＣＤＭ８ならびに同様の真核生物発現ベクターである。

【００４８】

本発明のｒＤＮＡ分子を構築するのに使用される真核生物発現ベクターは、真核生物中で有効な選択マーカー（好ましくは、薬物耐性選択マーカー）をさらに含み得る。さらに、好ましい薬物耐性マーカーは、その発現によってネオマイシン耐性（すなわち、ネオマイ

シンホスホトランスフェラーゼ（ｎｅｏ）遺伝子）を生じる遺伝子である（Ｓｏｕｔｈｅｒｎ，Ｐ．Ｌ．ら、１９８２．Ｊ Ｍｏｌ Ａｐｐｌ Ｇｅｎｅｔ １（４）：３２７－４１）。あるいは、この選択マーカーは、別個のプラスミド上に存在し得、そして、これらの２つのベクターは宿主細胞の同時トランスフェクトによって導入され、そして、この選択マーカーにとって適切な薬物の中で培養されることによって選択される。

【００４９】

（外因性供給コード核酸分子を含む宿主細胞）

本発明は、さらに、本発明のメクラウナギカテリン（ｃａｔｈｅｌｉｎ）結合タンパク質をコードする核酸分子（好ましくは、配列番号１または３）で形質転換された宿主細胞が提供される。この宿主細胞は、原核生物または真核生物のいずれかであり得る。本発明のタンパク質の発現にとって有用な真核生物細胞は、その細胞株が細胞培養方法に適合しそして発現ベクターの増殖およびその遺伝子産物の発現に適合する限り、限定されるものではない。好ましい宿主細胞としては、酵母細胞、昆虫細胞、および哺乳動物細胞（好ましくは、マウス細胞株、ラット細胞株、サル細胞株、またはヒト細胞株）が挙げられるがこれらに限定されない。好ましい真核生物宿主細胞としては、メクラウナギ、ＡＴＣＣからＣＣＬ６１として入手可能であるチャイニーズハムスター卵巣（ＣＨＯ）細胞、ＡＴＣＣからＣＲＬ １６５８として入手可能なＮＩＨ Ｓｗｉｓｓマウス胚性細胞ＮＩＨ／３Ｔ３、胎仔ハムスター腎細胞（ＢＨＫ）ならびに同様の真核生物組織培養細胞株から樹立される細胞株を含む。

【００５０】

任意の原核宿主細胞は、本発明のタンパク質をコードするｒＤＮＡ分子を発現するために使用され得る。好ましい原核生物宿主は、Ｅ．ｃｏｌｉである。

【００５１】

本発明のｒＤＮＡ分子を有する適切な細胞宿主の形質転換は、使用されるベクターの型および使用される宿主系に代表的には依存する周知の方法によって達成される。原核生物宿主細胞の形質転換に関して、エレクトロポレーションおよび塩処理法が代表的には使用される（Ｃｏｈｅｎら（１９７２）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ６９，２１１０～２１１２およびＳａｍｂｒｏｏｋら（２００１）Ｍｏｌｅｃｕｌａｒ Ｃｌｏｎｉｎｇ，Ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ，Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｌａｂｏｒａｔｏｒｙ Ｐｒｅｓｓ）。ｒＤＮＡを含むベクターを用いる脊椎動物細胞の形質転換に関して、エレクトロポレーション、カチオン性脂質および塩処理法が代表的には使用される（例えば、Ｇｒａｈａｍら（１９７３）Ｖｉｒｏｌ．５２、４５６～４５８；Ｗｉｇｌｅｒら（１９７９）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ７６，１３７３～１３７６）。

【００５２】

首尾よく形質転換された細胞（すなわち、本発明のｒＤＮＡ分子を含む細胞）は、選択マーカーについての選択を含む周知の技術によって同定され得る。例えば、本発明のｒＤＮＡの導入から得られる細胞は、単一のコロニーを産生するためにクローニングされ得る。これらのコロニー由来の細胞が回収、溶解され得、そして、それらのＤＮＡ含有量が、存在するｒＤＮＡについて、Ｓｏｕｔｈｅｒｎ（１９７５）Ｊ．Ｍｏｌ．Ｂｉｏｌ．９８，５０３～５０６またはＢｅｒｅｎｔら（１９８５）Ｂｉｏｔｅｃｈ．３，２０８～２０９に記載されるような方法を使用して調べられるか、またはその細胞から産生され得るタンパク質が、免疫学的方法を介してアッセイされ得る。

【００５３】

（組換えタンパク質の産生）

本発明はさらに、本明細書に記載の核酸分子を使用する、メクラウナギカテリン結合タンパク質を産生するための方法を提供する。一般的な関係において、タンパク質の組換え形態の産生は、代表的には、以下の工程を包含する：

第１に、本発明のメクラウナギカテリン結合タンパク質をコードする核酸分子（例えば、配列番号１を含むか、本質的にはそれからなるか、もしくはそれからなるか、または、配

列番号１のヌクレオチド４位～５４０（または５４３）位によって定義されるオープンリーディングフレーム；あるいは、配列番号３を含むか、本質的にはそれからなるか、もしくはそれからなるか、または、配列番号３のヌクレオチド２位～５０５（または５０８）位によって定義されるオープンリーディングフレーム）が得られる。あるいは、本発明のメクラウナギ抗微生物ペプチドをコードする核酸分子（例えば、配列番号１のヌクレオチド４２７～５４０を含むか、本質的にそれからなるか、もしくはそれからなるか、または、配列番号２のアミノ酸残基１４２～１７９をコードするかもしくは配列番号５をコードするか；あるいは、例えば、配列番号３のヌクレオチド４１６～５０５を含むか、本質的にそれからなるか、もしくはそれからなるか、配列番号４のアミノ酸１３９～１６８をコードするか、または配列番号６をコードする、核酸分子）が得られる。コード配列がイントロンによって中断されておらず、このオープンリーディングフレームがそのままである場合、これは、任意の宿主での発現に直接適切である。

【００５４】

次いで、この核酸分子は、好ましくは、上記のような適切な制御配列と作動可能に連結されて配置されて、タンパク質オープンリーディングフレームを含む発現単位を形成する。発現単位を使用して、適切な宿主を形質転換し、そしてこの形質転換した宿主を、組換えタンパク質の産生を可能にする条件下で培養する。必要に応じて、この組換えタンパク質を、培地または細胞から単離する；このタンパク質の回収および精製は、いくらかの不純物が許容され得るいくつかの場合には、必要でないかもしれない。

【００５５】

上記の工程の各々は、種々の方法で行われ得る。例えば、所望のコード配列は、ゲノムフラグメントから得られ得、そして適切な宿主において直接使用され得る。種々の宿主において作動可能な発現ベクターの構築は、上記のような適切なレプリコンおよび制御配列を使用して達成される。制御配列、発現ベクターおよび形質転換方法は、遺伝子を発現するために使用される宿主細胞の型に依存し、そして上記でより詳細に記載した。適切な制限部位が、通常では利用可能でない場合、これらのベクター中に挿入される切り出し可能な遺伝子を提供するために、コード配列の末端に付加され得る。当業者は、組換えタンパク質を産生するために本発明の核酸分子と共に使用するための、当該分野で公知の任意の宿主／発現系を容易に適合し得る。

【００５６】

別の実施形態において、ポリペプチドは、改変形態（例えば、融合タンパク質）で発現され得、そして分泌シグナルだけでなく、さらなる異種性の機能的領域もまた含み得る。例えば、さらなるアミノ酸（特に、荷電したアミノ酸）の領域が、ポリペプチドのＮ末端に付加されて、複製または引き続く取り扱いおよび保存の間の、宿主細胞における安定性および持続性を改善し得る。また、ペプチド部分が、精製を容易にするためにこのポリペプチドに付加され得る。このような領域は、ポリペプチドの最終的な調製の前に、除去され得る。とりわけ、分泌または排出を生じるため、安定性を改善するため、および精製を容易にするための、ポリペプチドに対するペプチド部分の付加は、当該分野でよく知られた慣用技術である。

【００５７】

（メクラウナギカテリン（ｃａｔｈｅｌｉｎ）関連タンパク質）

アミノ酸配列（配列番号２および配列番号４）（図１）は、抗菌活性に関連するペプチドの領域が異なるのに主に起因して、異なる。シグナル配列は２６残基で同一であり；そしてカテリンは、その１１５残基中９１．３％を共有する。対照的に、配列番号２のアミノ酸残基１４２～１７９からなる抗菌ペプチドＧＷＦＫＫＡＷＲＫＶＫＨＡＧＲＲＶＬＤＴＡＫＧＶＧＲＨＹＬＮＮＷＬＮＲＹＲＧ（配列番号５）は、配列番号４のアミノ酸残基１３９～１６８からなる抗菌ペプチドＧＷＦＫＫＡＷＲＫＶＫＮＡＧＲＶＬＫＧＶＧＩＨＹＧＶＧＬＩＧ（配列番号６）の３０残基に対して、３８残基を有する。しかし、３つのギャップが配列番号６に挿入される場合、配列番号５の配列に対して３０残基中２３残基（７６％）が同一である。配列番号５および配列番号６の両方は、米国特許第５，７３４，

０１５号に記載されるメクラウナギ抗菌ペプチドと同様に、おそらくＣ末端アミドが翻訳後に変更されるＣ末端グリシン残基を有する。

【００５８】

本明細書中で使用される場合、メクラウナギカテリン関連タンパク質に関連するタンパク質のファミリーとは、メクラウナギまたはメクラウナギに密接に関連する生物から単離され得る他のカテリン関連タンパク質をいう。メクラウナギカテリン関連タンパク質に関連するタンパク質の他のファミリーメンバーを同定および単離するために使用される方法を、以下に記載する。

【００５９】

本発明のタンパク質は、好ましくは単離形態である。本明細書中で使用される場合、物理的方法、機械的方法または化学的方法を使用して、タンパク質に通常結合する細胞構成成分からタンパク質を取り出す場合、タンパク質は単離されると言われる。当業者は、単離されたタンパク質を得るために標準的な精製方法を使用し得る。単離されるとは、本発明のメクラウナギカテリン関連タンパク質または抗菌ペプチドは、他の成分との混合物の一部であり、そして本発明のタンパク質またはペプチドプチドは、混合物中の総タンパク質の少なくとも約１０％、好ましくは混合物中の総タンパク質の少なくとも２０％、より好ましくは混合物中の総タンパク質の少なくとも３０％、なおより好ましくは混合物中の総タンパク質の少なくとも４０％、さらにより好ましくは混合物中の総タンパク質の少なくとも５０％、なおより好ましくは混合物中の総タンパク質の少なくとも６０％、さらにより好ましくは混合物中の総タンパク質の少なくとも７０％、なおより好ましくは混合物中の総タンパク質の少なくとも８０％、さらにより好ましくは混合物中の総タンパク質の少なくとも９０％、なおより好ましくは混合物中の総タンパク質の少なくとも９５％、そして最も好ましくは混合物中の総タンパク質の少なくとも９９％を含むことをさらに意味する。

【００６０】

本発明のタンパク質はさらに、本明細書中に記載されるタンパク質の保存的改変体を含む。本明細書中で使用される場合、保存的改変体とは、タンパク質の生物学的機能に有害な影響を与えないアミノ酸配列の変更をいう。置換、挿入または欠失は、変更された配列が、タンパク質に関連する生物学的機能を防御または破壊する場合、タンパク質に有害な影響を与えることをいう。例えば、タンパク質の全体的な変化、構造または疎水性／親水性の特性は、生物学的活性に有害な影響を与えることなく変更され得る。従って、アミノ酸配列は、例えば、タンパク質の生物学的活性に有害な影響を与えることなく、ペプチドをより疎水性または親水性にするように変更され得る。

【００６１】

通常、対立遺伝子改変体、保存的置換改変体およびタンパク質ファミリーのメンバーは、配列番号２、４、５または６に記載される配列と、少なくとも約８５％、より好ましくは約９０％、なおより好ましくは約９５％、最も好ましくは約９９％のアミノ酸配列同一性を有するアミノ酸配列を有する。このような配列に関する同一性または相同性は、必要に応じて、その配列を整列させギャップを導入して、最大％相同性を実現した後に、既知のペプチドと同一である候補配列におけるアミノ酸残基のパーセンテージとして本明細書中に規定され、そして保存的置換を配列同一性の一部分として認めない。Ｎ末端、Ｃ末端または内部伸長、欠失あるいはペプチド配列中への挿入は、相同性に影響すると解釈されない。

【００６２】

従って、本発明のタンパク質としては、以下が挙げられる：配列番号２、４、５および６に開示されるアミノ酸配列を含む、それらから本質的になる、あるいはそれらから構成される分子；メクラウナギカテリン関連タンパク質またはその抗菌ペプチドの少なくとも３、４、５、６、１０、１５、２０、２５、３０、３５、３８またはそれより多いアミノ酸残基の保存的配列を有するそれらのフラグメント；このような配列のアミノ酸配列改変体、ここで、アミノ酸残基は、開示される配列のＮ末端もしくはＣ末端に挿入されるか、ま

たは開示される配列中に挿入される；および開示される配列のアミノ酸配列改変体、または（別の残基によって置換された）、あるいは上に規定されるようなそれらのフラグメント。企図される改変体は、例えば、以下：相同組換え、部位指向性突然変異誘発またはＰＣＲ突然変異誘発、および他の動物種（ウサギ、ラット、ブタ、ウシ、ヒツジ、ウマおよび非ヒト霊長類種が挙げられるが、これらに限定されない）の対応タンパク質、および対立遺伝子またはタンパク質ファミリーの他の天然に存在する改変体；ならびに誘導体（このタンパク質は、天然に存在するアミノ酸以外の部分（例えば、酵素もしくはラジオアイソトープのような検出可能部分）で、置換、化学的手段、酵素的手段、または他の適切な手段によって、共有結合的に改変されている）。

【００６３】

（治療的適用および添加剤適用）

本発明の別の局面は、微生物の増殖を阻害するための方法である。本発明のメクラウナギカテリン関連抗菌ペプチド（配列番号５および６により例示される）を使用して、種々の状況下で微生物の増殖を阻害し得る。例えば、本発明のペプチドは、治療的に投与して、微生物感染から生じる個体における疾患を処置または予防し得る。種々の微生物感染が、本発明の抗菌ペプチドを用いて処置され得る。例えば、本発明の抗菌ペプチドは、真菌感染または細菌感染の処置のための抗菌剤として使用され得る。

【００６４】

好ましい実施形態において、微生物感染は、細菌感染である。特に、本発明のペプチドの治療的投与は、細菌感染を阻害し、そして逆転し得る。細菌の新しい耐性株の発生および蔓延は、ますます公衆衛生に脅威を与えている。本発明の抗菌ペプチドは、現在公衆衛生を脅かしている微生物のこれらの耐性株に罹患している患者を処置する際に有用である。多くの細菌感染は、感染した個体において細菌性敗血症を引き起こす。敗血症性細菌感染に罹患した個体へのこのペプチドの投与は、感染により生じる敗血症、および感染自体を低減することにより、有意な治療効果を生じる。

【００６５】

処置に適切な個体は、上記の微生物感染の１つ以上に罹患しているか、そうでなければ感受性である任意の動物（哺乳動物またはその他）である。好ましい実施形態において、この個体はヒトである。別の実施形態において、この個体は、家畜動物である。別の実施形態において、この動物は、鑑賞用（ｓｈｏｗ）動物または家庭用のペットである。

【００６６】

本発明のメクラウナギカテリン関連抗菌ペプチドの個体への投与は、全身または局所的であり、そして主として処置される特定の感染により決定される。全身投与は、いくつかの経路（静脈内投与、吸入、粘膜、および食物摂取を含むがこれらに限定されない）により達成され得る。局所的投与は、表面または内部であり得る。このような投与は、いくつかの経路（皮下、皮膚、皮内、口腔内、粘膜、腹腔内、膣、吸入、および食物摂取を含むがこれらに限定されない）により達成され得る。

【００６７】

薬学的に受容可能なキャリアを含む、本発明のペプチドの処方物を投与することが、しばしば利用される。治療的投与に可能な処方物は、種々の薬学的組成物を含み、これらの適切な使用は、処置に必要とみなされる投与経路に依存する。局所投与に有用ないくつかの処方物は、例えば、点眼剤、点耳剤、または歯肉適用（例えば、点滴剤、うがい薬、クリームまたはペースト）である。患者に対する治療に役立つ投与レジメン（例えば、経路、用量および過程）は、処置される患者（例えば、健康、体重、代謝）、感染部位、および感染している病原体によって変化する。治療レジメンは、実験的観察と組み合わせて、同様の治療剤を用いる処置からの推定により開発されるべきである。個体における微生物感染を予防するための本発明のペプチドの投与は、上記の方法に類似する。

【００６８】

一般的提案として、用量あたりの非経口的に投与される抗菌ペプチドの薬学的有効量の合計は、患者の体重の約０．００１ｍｇ／ｋｇ／日～１０ｍｇ／ｋｇ／日の範囲であるが、

上記のように、この範囲は、治療的裁量に供される。より好ましくは、この用量は、少なくとも０．０１ｍｇ／ｋｇ／日であり、このペプチドに関してヒトに対して最も好ましくは約０．０１ｍｇ／ｋｇ／日と１ｍｇ／ｋｇ／日との間である。連続的に投与される場合、この抗菌ペプチドは、１日あたり１回～４回の注射によるか、または例えばミニポンプを使用する連続的な皮下注入のいずれかにより、代表的には、約０．００１ｍｇ／ｋｇ／時間～約０．０５０ｍｇ／ｋｇ／時間の用量速度で投与される。静脈内バッグ液剤もまた使用され得る。

【００６９】

本発明のメクラウナギカテリン関連抗菌ペプチド（配列番号５および６により例示される）はまた、他の抗菌薬または抗菌剤の治療作用を増強するために有用である。本発明の抗菌ポリペプチドとの他の抗菌剤の同時投与は、相乗抗菌効果を生じる。抗菌剤との本発明のペプチドの同時投与は、より低い用量の抗菌剤を用いる患者の治療的処置を可能にする。より低い用量は、高価な薬物、または所望でない副作用を生じる薬物、またはインビボでのその短い半減期により、有効であるために必要な濃度よりも低い濃度に急速に減少する薬物を用いて処置する場合のような状況において好ましい。さらに、抗菌剤または抗菌薬との同時投与はまた、より短い治療期間および／または耐性表現型の逆転を可能にし得る。限定することなく、その体内薬物濃度を低減すること（例えば、その薬物の減少した膜透過性または増加した細胞輸送もしくは代謝）により抗菌薬に抵抗する微生物は、これらのペプチドの相乗作用に特に感受性であることが期待される。

【００７０】

本発明の抗菌薬およびメクラウナギカテリン関連抗菌ペプチドまたは機能的改変体の投与レジメンは、患者および特定の感染によって変化し、そして症例毎に当業者により決定され得る。ポリペプチドの処方は、投与レジメンに依存し、例が上に記載される。

【００７１】

本発明のポリペプチドは、薬学的薬剤および機能性食品用剤としてだけではなく、食品および医薬または非医薬製品のような任意の製品のための添加剤としても使用され得、これらの製品は、体内に摂取されるか、あるいは他の方法でヒトもしくは他の動物の体表面、またはヒトもしくは他の動物由来の体液、器官、および細胞上に塗布されるかまたは接触される。

【００７２】

本発明は、種々の製品を処理するのに有用である。生物学的製品（本明細書中では、生物学的な生物またはプロセスに由来する製品として規定される）は特に、生物による汚染の危険に曝されている。生物学的製品の例としては、限定することなく、食品、組織、生存細胞、生存細胞由来の産物、血液またはその成分、任意の他の体液、薬物または他の分子調製物が挙げられる。非生物学的製品（本明細書中では、生物学的な生物またはプロセスから直接的には由来しない製品として規定される（例えば、ガラス製品、手術器具、合成薬物または他の分子調製物））もまた処理され得る。本発明の有効な使用のために、処理される製品は、このようにして適用されるペプチドの量の全ての抗菌活性を完全に不活化する活性を有さないべきである。本発明のペプチドは、微生物の増殖による汚染を予防するかまたは阻害することが一般的に所望される任意の製品に添加され得るか、組み合わせられ得るか、噴霧され得るか、製品上にコーティングされ得るか、吸着され得るか、化学的に架橋され得るか、または製品中に含浸され得る。あるいは、本発明のペプチドは、表面上に固定化され得、この表面上に製品を通して、その製品における微生物増殖を除去するかまたは阻害する。抗菌ペプチドで処理されてその抗菌ペプチドを保持する製品は、接触する別の製品を処理するためにさらに使用され得る。

【００７３】

さらに記載しなくても、当業者は、上述の記載および以下の例示的実施例を使用して、本発明を作製および利用し、特許請求の範囲に記載される方法を実施し得る。従って、以下の実施例は、本発明の好ましい実施形態を具体的に示し、そして本開示の残りの部分をいかなるようにも限定すると解釈されるべきではない。

【実施例１】
【００７４】
（メクラウナギカテリン関連抗菌ペプチドの遺伝子の単離）
成体タイセイヨウメクラウナギ（Ｍｙｘｉｎｅ ｇｌｕｔｉｎｏｓａ Ｌｉｎｎａｅｕｓ，１７５８；Ｍｙｘｉｎｉｄａｅ）を、Ｈｕｎｔｓｍａｎ Ｍａｒｉｎｅ Ｓｃｉｅｎｃｅ Ｃｅｎｔｒｅ（Ｎｅｗ Ｂｒｕｎｓｗｉｃｋ，Ｃａｎａｄａ）から購入した。この魚を、氷のスラリー中で麻酔し、そして断頭した。腸を取り出し、そして内容物を洗い流し、液体N_2中で凍結しそして－８０℃で貯蔵した。
【００７５】
組織を液体窒素中で粉砕し、そしてグアニジニウムイソチオシアネート溶液（Ｃｈｉｒｇｗｉｎら（１９７９）Ｂｉｏｃｈｅｍｉｓｔｒｙ １８，５２９４－５２９９）に移した。この溶液は、細胞分裂の間ＲＮａｓｅ活性を阻害する。全ＲＮＡを、ＣｓＣｌ段階勾配による遠心分離により単離し；ポリアデニル化ＲＮＡをオリゴ（ｄＴ）セルロース（Ｓｔｒａｔａｇｅｎｅ）により回収した。逆転写したｃＤＮＡを、Ｓｔｒａｔａｇｅｎｅのプロトコルに従って、最初にアダプターアーム、次いで、ファージλＺａｐＩＩにライゲーションした。
【００７６】
２０マーオリゴヌクレオチドプローブ５’－ＡＡＲＡＡＲＧＣＮＴＧＧＭＧＮＡＡＲＧＴ－３’（配列番号１１：２５６倍縮重）を、メクラウナギ腸組織から以前に単離されたペプチドのアミノ酸残基４～１０に基づいて設計した。相同配列を含むファージ粒子を、少数の粒子をＸＬ１ Ｂｌｕｅ ＭＲＦ（Ｓｔｒａｔａｇｅｎｅ）の菌叢にプレーティングすることにより単離した。ｃＤＮＡ挿入物を、インビボ切断により、ファージミドｐＢｌｕｅｓｃｒｉｐｔにおいて単離した。２つの挿入物を、両方向で、サンガーのジデオキシ法（Ｂａｒｎｅｓら（１９８３）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．１１，３４９－３６８）を使用して配列決定した。
【００７７】
相同配列を、ＧｅｎＢａｎｋのＢＬＡＳＴＸ（Ａｌｔｓｃｈｕｌら（１９９７）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．２５，３３８９－３４０２）検索を使用して同定した。相同配列を、ＥＳＥＥ（Ｃａｂｏｔら（１９８９）Ｃｏｍｐｕ．Ａｐｐｌ．Ｂｉｏｓｃｉ．５，２３３－２３４）を使用して整列した。
【００７８】
インサイチュハイブリダイゼーションのために、配列番号１からの挿入物３１～５２を含むプラスミドＤＮＡを、Ｓｔｙｌを用いて線状化し、そしてセンスプローブおよびアンチセンスプローブの両方を酵素により^{32}Ｐで標識し、そしてオートラジオグラフィーにより検出した。
【実施例２】
【００７９】
（メクラウナギカテリン関連配列と他のカテリン関連配列との比較）
ＢＬＡＳＴＮを使用する、デフォルト値を用いたＧｅｎＢａｎｋの検索により、２つのペプチドのコード領域のみを参照として使用した場合も、コード領域全体を使用した場合のいずれの場合でもマッチが全く得られなかった。しかし、各タンパク質およびＢＬＡＳＴＸに利用可能なコード配列全体を使用する（同一アミノ酸および保存アミノ酸に基づいてより遠い関係について検索する）と、多数のマッチが得られた。正確なリーディングフレームのものは、特に、カテリン様酸性スペーサーを含む哺乳動物抗菌ペプチドのプレプロペプチドのｃＤＮＡ配列を含んでいた。
【００８０】
見出されたマッチの中で、Ｃａｐｒａ ｈｉｒｃｕｓ（ヤギ）の骨髄細胞から単離された、Ｃａｐｒａ ＭＡＰ２８の配列は、抗菌ペプチドと２つのＭｙｘｉｎｅペプチドのうちの一方との間で少なくとも弱いマッチを示す点において独特であった。Ｃａｐｒａ ＭＡＰ２８を含むメクラウナギ配列のうちの一つのコード領域の妥当と思われるアラインメン

トを、図２に示す。
【実施例３】
【００８１】
（カテリン関連抗菌ペプチドの抗菌活性）
カテリン関連抗菌ペプチドの抗菌活性を、Ｎａｔｉｏｎａｌ Ｃｏｍｍｉｔｔｅｅ ｆｏｒ Ｃｌｉｎｉｃａｌ Ｌａｂｏｒａｔｏｒｙ Ｓｔａｎｄａｒｄｓのガイドライン（Ｄｏｃｕｍｅｎｔ Ｍ７－Ｔ２（１９８８）第８巻，第８号）に基づく以下の抗菌アッセイにより決定する。
【００８２】
試験ペプチドのストック溶液を、滅菌脱イオン蒸留水中の５１２μｇ／ｍｌの濃度で調製し、そして－７０℃で貯蔵する。このストックペプチド溶液を、ウェル中のペプチドの最終濃度が、０．２５、０．５０、１、２、４、８、１６、３２、６４、１２８および２５６μｇ／ｍｌになるように、ミクロタイタープレートのウェルに沿って連続希釈（１：２）で希釈する。１．５´１０^{5}ＣＦＵ／ｍｌの試験微生物を、中対数期培養物からの完全強度のＭｕｅｌｌｅｒ Ｈｉｎｔｏｎブロス（ＢＢＬ １１４４３）でこれらのウェルに添加する。接種物を、６００ｎｍにおいて分光光度測定法により標準化し、そしてコロニー係数により確認する。これらのプレートを、３７℃で１６～２４時間インキュベートし、そして各ペプチドについての最少阻害濃度（ＭＩＣ）を決定する。最少阻害濃度は、ミクロタイタープレートにおいて清澄なウェルを生じるペプチドの最低濃度として規定される。
【００８３】
本発明を、種々の具体的な材料、手順および実施例を参照して記載し、そして例示してきたが、本発明は、その目的のために選択される材料および手順の特定の組合せに制限されないことが理解される。このような細部の多数の改変は、当業者により理解されると推定され得る。本明細書および実施例は、例示のみとみなされ、本発明の真の範囲および本質は、添付の特許請求の範囲により示されることが意図される。本明細書中で参照される全ての刊行物、特許および特許出願は、その全体が、本明細書中に参考として援用される。
【図面の簡単な説明】
【００８４】
【図１Ａ】図１Ａは、メクラウナギカテリン関連抗菌ペプチドのヌクレオチド配列（配列番号１および３）およびアミノ酸配列（配列番号２および４）の比較を示す。推定されるシグナル、カテリン、ペプチド、および３’非翻訳領域の開始点に印を付している。そのペプチドアミノ酸を、斜体の太字で示し；開始コドンおよび停止コドンを大文字で示し；ポリアデニル化部位に下線を付している。ドットは、同一のヌクレオチドまたはアミノ酸を示し、ダッシュは、挿入または欠失を示す。＊は、この抗生ペプチドにおける推定Ｃ末端アミド化部位を示す。
【図１Ｂ】図１Ｂは、メクラウナギカテリン関連抗菌ペプチドのヌクレオチド配列（配列番号１および３）とアミノ酸配列（配列番号２および４）の比較を示す。推定されるシグナル、カテリン、ペプチド、および３’非翻訳領域の開始点に印を付している。そのペプチドアミノ酸を、斜体の太字で示し；開始コドンおよび停止コドンを大文字で示し；ポリアデニル化部位に下線を付している。ドットは、同一のヌクレオチドまたはアミノ酸を示し、ダッシュは、挿入または欠失を示す。＊は、この抗生ペプチドにおける推定Ｃ末端アミド化部位を示す。
【図２】図２は、配列番号３および４と、Ｃａｐｒａ ｈｉｒｃｕｓ ＭＡＰ２８との、ヌクレオチドおよび暗示されるアミノ酸残基の比較を示す。哺乳動物遺伝子構造により示唆されるエキソン間の境界に垂線を付す。
【図３】図３は、メクラウナギ抗体ペプチドのアミノ酸配列（配列番号５および６）の比較を示す。ダッシュは、アラインメントを最大限にするために挿入されたギャップを示す。
【図４】図４は、メクラウナギカテリン関連抗菌ペプチドのアミノ酸配列（配列番号５お

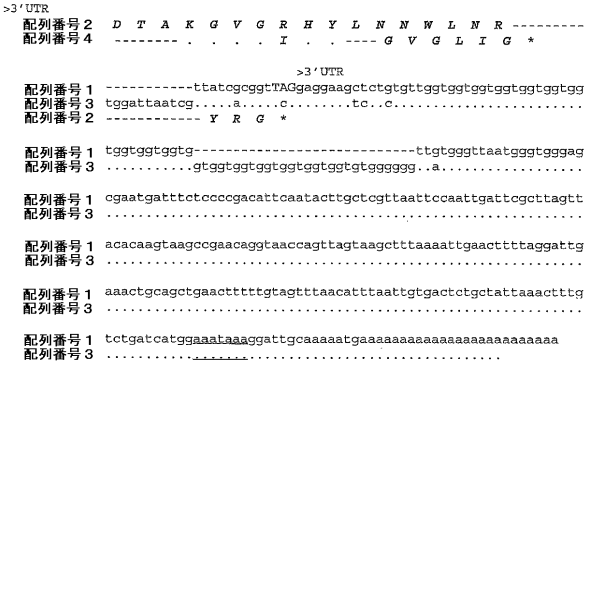
よび６）と、米国特許第５，７３４，０１５号に記載されるメクラウナギ腸組織から単離された抗菌ペプチド（配列番号９および１０）との比較を示す。ダッシュは、アラインメントを最大限にするために挿入されたギャップを示す。Ｘ＝Ｂｒ－ＴｒｐおよびＺは、Ｃ末端アミドを示す。

【図１Ａ】

Figure 1A

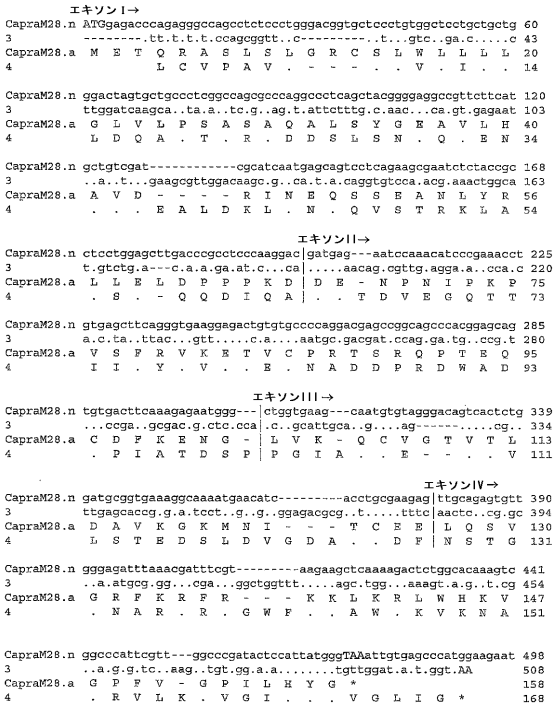
【図１Ｂ】

Figure 1B



【図２】

Figure 2



【図３】

Figure 3

配列番号５ GWFKKAWRKVKHAGRRVLDTAKGVGRHYLNNWLNRYRG
配列番号６ ..........N...-..---....I..-GVG.IG

【図４】

Figure 4

配列番号５ GWFKKAWRKVKHAGRRVLDTAKGVGRHYLNNWLNRYRG
配列番号９ .F....X.....V..X.....Z

配列番号６ GWFKKAWRKVKNAGR-VLKGVGIHYGVGLIG
配列番号１０ .X....X.....R.....Z

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 February 2003 (13.02.2003)

PCT

(10) International Publication Number
WO 03/012044 A2

(51) International Patent Classification⁷: C12N

(21) International Application Number: PCT/US02/24019

(22) International Filing Date: 30 July 2002 (30.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/308,652 30 July 2001 (30.07.2001) US

(71) Applicant *(for all designated States except US)*: **GENAERA CORPORATION** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **UZZELL, Thomas** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US). **STOLZENBERG, Ethan, D.** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US). **SHINNAR, Ann, E.** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US). **ZASLOFF, Michael, A.** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US).

(74) Agents: **SMYTH, Robert, J.** et al.; Morgan, Lewis & Bockius LLP, 1111 Pennsylvania Avenue, NW, Washington, DC 20004 (US).

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

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

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

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

WO 03/012044 A2

(54) Title: HAGFISH CATHELIN-ASSOCIATED ANTIMICROBIAL PEPTIDES AND GENES

(57) Abstract: The present invention includes Myxine glutinosa cathelin-associated antimicrobial peptides and genes encoding these peptides. The invention also includes compositions an methods for producing these peptides as well as method of preventing and treating microbial infections using these peptides.

Hagfish Cathelin-Associated Antimicrobial Peptides and Genes

Cross-Reference To Related Application

This application claims the benefit of U.S. provisional application 60/308,652 filed on July 30, 2002.

Field Of The Invention

This invention relates to genes for antibiotic peptides and their use for the production of said peptides. It also relates to the antimicrobial peptides and their use as antibiotics.

Background Of The Invention

The cathelicidin gene family (Zanetti *et al.* (1995) FEBS Letters 374, 1-5; Zanetti *et al.* (1997) Ann. New York. Acad. Sci. 147-162; Gennaro *et al.* (2000) Biopolymers Peptide Science 55, 31-49) encodes prepropeptides that consist of a signal sequence, a cathelin-like acidic spacer, and an antibiotic peptide (Figure 1). Cathelicidins are presently known from a variety of mammals: *Bos taurus* (cattle), *Capra hircus* (goat), *Cavia porcellus* (guinea pig), *Equus caballus* (horse), *Homo sapiens* (human), *Mus musculus* (mouse), *Oryctolagus cuniculus* (rabbit), *Ovis aries* (sheep), and *Sus scrofa* (pig). Although the known antibiotic peptides are remarkably diverse, both within and between mammalian species, the associated signal sequences and cathelin-like acidic spacers, as revealed by sequence analysis of cDNA clones, are as remarkably conserved within and among species as the antibiotic peptides themselves are diverse.

Mammalian cathelicidins are synthesized primarily in myeloid cells, and are abundant in circulating neutrophils. The structure of some 20 cathelicidin genes is known and regularly involves four exons, the last of which encodes the antibiotic peptide along with a few C-terminal residues of the cathelin. Clusters of cathelicidin genes numbering eight, eleven, and more have been reported in *Ovis*, *Bos* and *Sus*, respectively (Gennaro *et al.* (2000) Biopolymers Peptide Science 55, 31-49).

Two antibiotic peptides have been isolated (Kaumaya *et al.* (1996) Peptides: Chemistry and Biology, Mayflower Scientific, pages 189-191) from intestinal tissue of *Myxine glutinosa* (Atlantic hagfish), a craniate chordate taxon only distantly related to the Mammalia. These peptides are unique among antimicrobial peptides in containing one or two brominated tryptophans. Herein are described hagfish cDNA sequences encoding two new antibiotic peptides from the same family as those isolated earlier. These sequences identify the *Myxine* antibiotic peptides as members of the cathelicidin family.

Summary Of The Invention

The invention includes an isolated nucleic acid molecule selected from the group consisting of an isolated nucleic acid molecule that encodes the amino acid sequence of a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 2 or 4; and an isolated nucleic acid molecule, which hybridizes to the complement of a nucleic acid molecule comprising a nucleotide sequence encoding a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 2 or 4 under stringent conditions. In some embodiments, the nucleic acid molecule comprises SEQ ID NO: 1 and 3; nucleotides 4 to 540 of SEQ ID NO: 1; nucleotides 427 to 540 of SEQ ID NO: 1; nucleotides 2 to 505 of SEQ ID NO: 3; or nucleotides 416 to 505 of SEQ ID NO: 3.

The invention also includes an isolated nucleic acid molecule selected from the group consisting of an isolated nucleic acid molecule that encodes the amino acid sequence of a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 5 or 6; and an isolated nucleic acid molecule, which hybridizes to the complement of a nucleic acid molecule comprising a nucleotide sequence encoding a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 5 or 6 under stringent conditions.

The isolated nucleic acid molecules of the invention can be operably linked to one or more expression control elements. The invention further includes a vector comprising an any of the isolated nucleic acid molecules recited above and a host cell comprising the vector.

The invention includes a method for producing a peptide or fragment thereof comprising the step of culturing the aforementioned host cell under conditions in which the protein or protein fragment encoded by said nucleic acid molecule is expressed. The host cell can be prokaryotic or eukaryotic. The invention includes recombinant peptides and fragments thereof produced by this method.

The invention includes an isolated peptide or fragment thereof selected from the group consisting of an isolated peptide comprising the amino acid sequences depicted in SEQ ID NO: 2, 4, 5 or 6; an isolated peptide fragment comprising at least six amino acids of any of the sequences depicted in SEQ ID NO: 2, 4, 5 or 6; an isolated peptide comprising conservative amino acid substitutions of any of the sequences depicted in SEQ ID NO: 2, 4, 5 or 6; and naturally occurring amino acid sequence peptide variants of any of the amino acid sequences depicted in SEQ ID NO: 2, 4, 5 or 6.

The invention further includes an antimicrobial composition comprising one or more of the peptides of of the invention. The antimicrobial composition can be suitable for topical or parenteral administration.

The invention includes a method of treating or preventing a microbial infection in a mammal, comprising administering to the mammal an effective amount of a *Myxine glutinosa* cathelin-associated antimicrobial peptide. In some embodiments, the *Myxine glutinosa* cathelin-associated antimicrobial peptide is selected from the group consisting of SEQ ID NO: 2, 4, 5 or 6 and the mammal is a human at risk of, or suffering from, a bacterial infection or a fungal infection. The *Myxine glutinosa* cathelin-associated antimicrobial peptide containing compositions can be administered by a route selected from the group consisting of oral, topical and parenteral. When topical administration is employed, it can be accomplished by inhalation. When parenteral administration is employed, it can be accomplished by intravenous, subcutaneous or intramuscular administration.

Brief Description Of The Figures

Figure 1 depicts a comparison of nucleotide (SEQ ID NO: 1 and 3) and amino acid sequences (SEQ ID NO: 2 and 4) for hagfish cathelin-associated antimicrobial peptides. Beginnings of presumed signal, cathelin, peptide, and 3' untranslated regions are marked. The peptide amino acids are italicized and in bold; start and stop codons are in upper case; the polyadenylation site is underlined. Dots indicate identical nucleotides of amino acids while dashes indicate insertions or deletions. * indicates the putative C-terminal amidation site in the antibiotic peptides.

Figure 2 depicts a comparison of nucleotides and implied amino acid residues of SEQ ID NOS. 3 and 4 with *Capra hircus* MAP28. Borders between exons suggested by mammalian gene structures are marked by vertical bars.

Figure 3 depicts a comparison of the amino acid sequences of the hagfish antibiotic peptides (SEQ ID NO: 5 and 6). Dashes indicate gaps inserted to maximize the alignment.

Figure 4 depicts a comparison of amino acid sequences for hagfish cathelin-associated antimicrobial peptides (SEQ ID NO: 5 and 6) with antimicrobial peptides isolated from hagfish intestinal tissues (SEQ ID NO: 9 and 10) described in U.S. Patent 5,734,015. Dashes indicate gaps inserted to maximize the alignment. X=Br-Trp and Z indicates a C-terminal amide.

Detailed Description

General

Applicants have isolated and sequenced two new genes from the Atlantic hagfish (*Myxine glutinosa*). The genes (SEQ ID NO: 1 and 3, Figure 1) encode two proteins (SEQ ID NO: 2 and 4, Figure 1) which each comprise an N-terminal signal sequence, a cathelin-like sequence and a C-terminal antimicrobial peptide. The antimicrobial peptide sequences deduced

from these genes, GWFKKAWRKVKHAGRRVLDTAKGVGRHYLNNWLNRYRG (SEQ ID NO: 5) and GWFKKAWRKVKNAGRVLKGVGIHYGVGLIG (SEQ ID NO: 6), are highly similar yet distinct from the hagfish peptides described in U.S. Patent 5,734,015 (Figure 4). Thus, the invention includes the isolated hagfish genes and their encoded cathelin-associated antibiotic peptides.

Included in this invention is the antibiotic peptides wherein the C-terminal glycine residue has been post-translationally transformed into a C-terminal amide residue (SEQ ID NO: 11 and 12) by an enzyme such as peptidylglycine α-amidating monooxygenase (PAM) (Prigge *et al.* (2000) Cell. Mol. Life Sci. 57, 1236-1259) and/or one or more of the Trp residues have been post-translationally transformed into Br-Trp residues by a brominating enzyme (Shinnar *et al.* (2000) FASEB Journal 14, A1488).

Also included in the invention is the family of hagfish cathelin-associated genes, their encoded proteins and their antibiotic peptides. Another embodiment is methods for the expression of these genes and cells that contain them. The recombinant production of the proteins encoded by the hagfish cathelin-associated genes is included in this invention, as is the production of the associated antibiotic peptides.

Definitions

As used herein, the term "antimicrobial" refers to the ability of a compound to inhibit or irreversibly prevent the growth of a microorganism. Such inhibition or prevention can be through a microbicidal action or microbistatic inhibition. Therefore, the term "microbicidal inhibition" as used herein, refers to the ability of a compound to kill, or irreversibly damage the target organism. The term "microbistatic inhibition" as used herein, refers to the ability of the antimicrobial compound to inhibit the growth of the target microorganism without death. Microbicidal or microbistatic inhibition of microorganisms in an environment presently exhibiting microbial growth (therapeutic treatment) or an environment at risk of supporting microbial growth (prevention) is included within this definition.

As used herein, a "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many generations.

A DNA "coding sequence" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

As used herein, a "nucleic acid molecule" refers to the polymeric form of deoxyribonucleotides (adenine, guanine, thymine, and/or cytosine) or ribonucleotides (adenine, guanine, uracil, and/or cytosine) and may include in either its single stranded form, or in double-stranded helix as well as RNA. This term refers only to the primary and secondary structure of the molecule and is not limited to any particular tertiary form. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (*e.g.*, the strand having a sequence homologous to the mRNA). Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, polyadenylation signals, terminators, and the like, that provide for the expression of a coding sequence in a host cell.

As used herein, a "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded (inclusively) at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site, as well as protein binding domains responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes.

A "signal sequence" can be included before the coding sequence or the native amino acid signal sequence from the envelope protein may be used. This sequence encodes a signal peptide, N-terminal to the polypeptide, that communicates to the host cell to direct the polypeptide to the cell surface or secrete the polypeptide into the media. This signal peptide is clipped off by the host cell before the protein leaves the cell. Signal sequences can be found associated with a variety of proteins native to prokaryotes and eukaryotes. For instance, alpha-factor, a native yeast protein, is secreted from yeast, and its signal sequence can be attached to heterologous proteins to be secreted into the media (see *e.g.*, U.S. Patent 4,546,082 and EP 0116201). Further, the alpha-factor and its analogs have been found to secrete heterologous proteins from a variety of yeast, such as *Saccharomyces* and *Kluyveromyces* (EP 88312306.9; EP 0324274 publication, and EP 0301669). An example for use in mammalian cells is the tPA signal used for expressing Factor VIIIc light chain.

A cell has been "transformed" by exogenous or heterologous DNA when such DNA as been introduced inside the cell. The transforming DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes,

for example, the transforming DNA may be maintained on an episomal element such as a plasmid or viral vector. With respect to eukaryotic cells, a stably transformed cell is one in which the transforming DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA.

A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then translated into the protein encoded by the coding sequence.

As used herein, a "vector" is a replicon, such as plasmid, virus, phage or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment.

Nucleic Acid Molecules

The two cDNA sequences (Figure 1) consist of 848 (SEQ ID NO: 1) and 863 (SEQ ID NO: 3) nucleotides. The sequence for SEQ ID NO: 1 includes the initiation codon and 3 additional 5'-nucleotides; the sequence for SEQ ID NO: 3, judged by alignment with SEQ ID NO: 1, has eight coding nucleotides missing at the 5'-end. In the alignment of Figure 1, 814 nucleotides are comparable between the two sequences, with 785 matching (96.4%). The nucleotides encoding the putative signal and cathelin sequences of the two are very similar, with 69 of 70 signal nucleotides identical (98.6%) and 332 of 345 (96.2%) cathelin nucleotides matching; 401 of 415 (96.6%) match overall. Neither insertions nor deletions are evident in these parts of the cDNA sequences.

The peptide and 3'-untranslated sequences, in contrast, show evidence of several insertion/deletion changes, which makes comparison of their similarities difficult. In the peptide-encoding sequence, SEQ ID NO: 3 has three gaps (of 3, 9, and 4 nucleotides) in comparison to SEQ ID NO: 1; the last gap results in a frame-shift for 17 nucleotides just before the termination codon. The peptide-encoding sequence of SEQ ID NO: 1 has a gap of 21 nucleotides with respect to the sequence for SEQ ID NO: 3; as a result, a new termination codon 26 nucleotides farther along the sequence marks the end of the peptide for SEQ ID NO: 1. In this alignment, the last 10 coding nucleotides for SEQ ID NO: 1 correspond to 10 nucleotides in the 3'-untranslated region of the SEQ ID NO: 3 cDNA. Overall, 98 nucleotides encoding peptide can be compared; 89 of these (90.8%) match.

The 3'-untranslated ends of both cDNAs have -tgg- repeats, a 10-mer for SEQ ID NO: 1, and a 17-mer for SEQ ID NO: 3. Excluding the -tgg- repeats, 248 nucleotides in the 3'-end can be compared; of these, 243 (98.0%) match.

The nucleic acid molecules of the present invention include nucleic acid molecules that encode the proteins having SEQ ID NO: 2, SEQ ID NO: 4 and the related proteins herein described, preferably in isolated form. As used herein, "nucleic acid" is defined as RNA or DNA that encodes a protein or peptide as defined above, or is complementary to nucleic acid sequence encoding such peptides, or hybridizes to such nucleic acid and remains stably bound to it under appropriate stringency conditions, or encodes a polypeptide sharing at least 55% sequence identity, 85 % sequence identity, preferably at least 85%, and more preferably at least 90%, most preferably at least 95 to 99% sequence identity with the peptide sequences. Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as nucleic acids based on alternative backbones or including alternative bases whether derived from natural sources or synthesized. Such hybridizing or complementary nucleic acids, however, are defined further as being novel and unobvious over any prior art nucleic acid including that which encodes, hybridizes under appropriate stringency conditions, or is complementary to nucleic acid encoding a protein according to the present invention.

A preferred embodiment of the present invention is a nucleic acid molecule that encodes a hagfish cathelin-associated protein of the invention, such as a nucleic acid molecule comprising, consisting essentially of or consisting of SEQ ID NO: 1, or the open reading frame defined by nucleotides 4-540 (or 543) of SEQ ID NO: 1; or such as a nucleic acid molecule comprising, consisting essentially of or consisting of SEQ ID NO: 3, or the open reading frame defined by nucleotides 2-505 (or 508) of SEQ ID NO: 3. A further preferred embodiment of the present invention is a nucleic acid molecule that encodes a hagfish antimicrobial peptide of the invention, such as a nucleic acid molecule comprising, consisting essentially of or consisting of nucleotides 427-540 of SEQ ID NO: 1, encoding amino acids 142-179 of SEQ ID NO: 2, or encoding SEQ ID NO: 5; or such as a nucleic acid molecule comprising, consisting essentially of or consisting of nucleotides 416-505 of SEQ ID NO: 3, encoding amino acids 139-168 of SEQ ID NO: 4, or encoding SEQ ID NO: 6.

Homology or sequence identity at the nucleotide or amino acid sequence level is determined by **BLAST** (Basic Local Alignment Search Tool) analysis using the algorithm employed by the programs **blastp, blastn, blastx, tblastn** and **tblastx** (Altschul *et al.* (1997) Nucleic Acids Res. 25, 3389-3402 and Karlin *et al.* (1990) Proc. Natl. Acad. Sci. USA 87, 2264-2268, both fully incorporated by reference) which are tailored for sequence similarity searching. The approach used by the **BLAST** program is to first consider similar segments,

with gaps (non-contiguous) and without gaps (contiguous), between a query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only those matches which satisfy a preselected threshold of significance. For a discussion of basic issues in similarity searching of sequence databases, see Altschul *et al.* (1994) Nature Genetics 6, 119-129 which is fully incorporated by reference. The search parameters for histogram, descriptions, alignments, expect (*i.e.*, the statistical significance threshold for reporting matches against database sequences), cutoff, matrix and filter (low complexity) are at the default settings. The default scoring matrix used by blastp, blastx, tblastn, and tblastx is the **BLOSUM62** matrix (Henikoff *et al.* (1992) Proc. Natl. Acad. Sci. USA 89, 10915-10919, fully incorporated by reference), recommended for query sequences over 85 nucleotides or amino acids in length.

For **blastn**, the scoring matrix is set by the ratios of **M** (*i.e.*, the reward score for a pair of matching residues) to **N** (*i.e.*, the penalty score for mismatching residues), wherein the default values for **M** and **N** are +5 and -4, respectively. Four **blastn** parameters were adjusted as follows: Q=10 (gap creation penalty); R=10 (gap extension penalty); wink=1 (generates word hits at every winkth position along the query); and gapw=16 (sets the window width within which gapped alignments are generated). The equivalent **Blastp** parameter settings were Q=9; R=2; wink=1; and gapw=32. A **Bestfit** comparison between sequences, available in the GCG package version 10.0, uses DNA parameters GAP=50 (gap creation penalty) and LEN=3 (gap extension penalty) and the equivalent settings in protein comparisons are GAP=8 and LEN=2.

"Stringent conditions" are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50°C, or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer (pH 6.5) with 750 mM NaCl, 75 mM sodium citrate at 42°C. Another example is hybridization in 50% formamide, 5× SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5× Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2× SSC and 0.1% SDS. A skilled artisan can readily determine and vary the stringency conditions appropriately to obtain a clear and detectable hybridization signal. Preferred nucleic acid molecules are those that hybridize under the above conditions to the complement of SEQ ID NO: 1 or 3 and which encode a polypeptide with antimicrobial activity. Even more preferred nucleic acid molecules are those

that hybridize under the above conditions to the complement strand of the open reading frame of SEQ ID NO: 1 or 3.

As used herein, an "isolated" nucleic acid molecule(s) means a nucleic acid molecule, DNA or RNA, which has been removed from its native environment or when the nucleic acid molecule is substantially separated from contaminant nucleic acid encoding other polypeptides from the source of nucleic acid. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically. Isolated nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) shown in SEQ ID NO: 1 or 3; DNA molecules comprising the coding sequence for the mature antimicrobial peptide protein; and DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode the antimicrobial peptide protein. As the genetic code is well known in the art, it would be routine for one skilled in the art to generate such degenerate variants. In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention. The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the antimicrobial peptide protein. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism.

The present invention further provides fragments of the encoding nucleic acid molecule. As used herein, a fragment of an encoding nucleic acid molecule refers to a small portion of the entire protein encoding sequence. The size of the fragment will be determined by the intended use. For example, if the fragment is chosen so as to encode an active portion of the protein, the fragment will need to be large enough to encode the functional region(s) of the protein or may encode regions of homology between the hagfish proteins in Figure 1 (SEQ ID NO: 2 and 4). In one embodiment of the invention an antimicrobial peptide fragment will be preferably at least 15-20 amino acids, more preferably 20-30 amino acids and most preferably 30-40 amino acids. If the fragment is to be used as a nucleic acid probe or PCR primer, then the

fragment length is chosen so as to obtain a relatively small number of false positives during probing/priming.

Fragments of the encoding nucleic acid molecules of the present invention (*i.e.*, synthetic oligonucleotides) that are used as probes or specific primers for the polymerase chain reaction (PCR), or to synthesize gene sequences encoding proteins of the invention can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci *et al.* (1981) J. Am. Chem. Soc. 103, 3185-3191 or using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define various modular segments of the gene, followed by ligation of oligonucleotides to build the complete modified gene.

The encoding nucleic acid molecules of the present invention may further be modified so as to contain a detectable label for diagnostic and probe purposes. A variety of such labels are known in the art and can readily be employed with the encoding molecules herein described. Suitable labels include, but are not limited to, biotin, radiolabeled nucleotides and the like. A skilled artisan can employ any of the art known labels to obtain a labeled encoding nucleic acid molecule.

Modifications to the primary structure itself by deletion, addition, or alteration of the amino acids incorporated into the protein sequence during translation can be made without destroying the activity of the protein. Such substitutions or other alterations result in proteins having an amino acid sequence encoded by a nucleic acid falling within the contemplated scope of the present invention.

Isolation of Other Related Nucleic Acid Molecules

As described above, the identification of the hagfish nucleic acid molecules having SEQ ID NO: 1 and 3 allows a skilled artisan to isolate nucleic acid molecules that encode other members of the cathelin-associated protein family in addition to the sequences herein described.

Essentially, a skilled artisan can readily use the amino acid sequence of SEQ ID NO: 2 or 4 to generate antibody probes to screen expression libraries prepared from appropriate cells. Typically, polyclonal antiserum from mammals such as rabbits immunized with the purified protein (as described below) or monoclonal antibodies can be used to probe a mammalian cDNA or genomic expression library, such as a λgtll library, to obtain the appropriate coding sequence for other members of the protein family. The cloned cDNA sequence can be expressed as a fusion protein, expressed directly using its own control sequences, or expressed

by constructions using control sequences appropriate to the particular host used for expression of the protein.

Alternatively, a portion of the coding sequence herein described can be synthesized and used as a probe to retrieve DNA encoding a member of the hagfish cathelin-associated protein family from any mammalian organism. Oligomers containing approximately 18-20 nucleotides (encoding about a 6-7 amino acid stretch) are prepared and used to screen genomic DNA or cDNA libraries to obtain hybridization under stringent conditions or conditions of sufficient stringency to eliminate an undue level of false positives.

Additionally, pairs of oligonucleotide primers can be prepared for use in a polymerase chain reaction (PCR) to selectively clone an encoding nucleic acid molecule. A PCR denature/anneal/extend cycle for using such PCR primers is well known in the art and can readily be adapted for use in isolating other encoding nucleic acid molecules.

Recombinant Nucleic Acid Molecules

The present invention further provides recombinant DNA molecules (rDNA) that contain a hagfish cathelin-associated protein coding sequence. As used herein, an rDNA molecule is a DNA molecule that has been subjected to molecular manipulation *in vitro*. Methods for generating rDNA molecules are well known in the art (see, *e.g.*, Sambrook *et al.* (2001) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press). In the preferred rDNA molecules, a coding DNA sequence is operably linked to expression control sequences and/or vector sequences.

The choice of vector and/or expression control sequences to which one of the protein family encoding sequences of the present invention is operably linked depends directly, as is well known in the art, on the functional properties desired, *e.g.*, protein expression, and the host cell to be transformed. A vector contemplated by the present invention is at least capable of directing the replication or insertion into the host chromosome, and preferably also expression, of the structural gene included in the rDNA molecule.

Expression control elements that are used for regulating the expression of an operably linked protein encoding sequence are known in the art and include, but are not limited to, inducible promoters, constitutive promoters, secretion signals, and other regulatory elements. Preferably, the inducible promoter is readily controlled, such as being responsive to a nutrient in the host cell's medium.

In one embodiment, the vector containing a coding nucleic acid molecule will include a prokaryotic replicon, *i.e.*, a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extrachromosomally in a prokaryotic host

cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, vectors that include a prokaryotic replicon may also include a gene whose expression confers a detectable marker such as a drug resistance. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

Vectors that include a prokaryotic replicon can further include a prokaryotic or bacteriophage promoter capable of directing the expression (transcription and translation) of the coding gene sequences in a bacterial host cell, such as *E. coli*. A promoter is an expression control element formed by a DNA sequence that permits RNA polymerase to bind and transcription to occur. Promoter sequences compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 (Bio-Rad Laboratories), pPL and pKK223 (Pharmacia).

Expression vectors compatible with eukaryotic cells, preferably those compatible with vertebrate cells, can also be used to form an rDNA molecules that contains a coding sequence. Eukaryotic cell expression vectors are well known in the art and are available from several commercial sources. Typically, such vectors are provided containing convenient restriction sites for insertion of the desired DNA segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies), pTDT1 (ATCC), the vector pCDM8 described herein, and the like eukaryotic expression vectors.

Eukaryotic cell expression vectors used to construct the rDNA molecules of the present invention may further include a selectable marker that is effective in an eukaryotic cell, preferably a drug resistance selection marker. A preferred drug resistance marker is the gene whose expression results in neomycin resistance, i.e., the neomycin phosphotransferase (*neo*) gene (Southern, P.J. *et al*. 1982. J Mol Appl Genet 1(4): 327-41). Alternatively, the selectable marker can be present on a separate plasmid, and the two vectors are introduced by co-transfection of the host cell, and selected by culturing in the appropriate drug for the selectable marker.

Host Cells Containing an Exogenously Supplied Coding Nucleic Acid Molecule

The present invention further provides host cells transformed with a nucleic acid molecule that encodes a hagfish cathelin-associated protein, preferably SEQ ID NO: 1 or 3, of the present invention. The host cell can be either prokaryotic or eukaryotic. Eukaryotic cells useful for expression of a protein of the invention are not limited, so long as the cell line is compatible with cell culture methods and compatible with the propagation of the expression vector and expression of the gene product. Preferred eukaryotic host cells include, but are not

limited to, yeast, insect, and mammalian cells, preferably those from a mouse, rat, monkey or human cell line. Preferred eukaryotic host cells include cell lines established from hagfish cells, Chinese hamster ovary (CHO) cells available from the ATCC as CCL61, NIH Swiss mouse embryo cells NIH/3T3 available from the ATCC as CRL 1658, baby hamster kidney cells (BHK), and the like eukaryotic tissue culture cell lines.

Any prokaryotic host can be used to express an rDNA molecule encoding a protein of the invention. The preferred prokaryotic host is *E. coli*.

Transformation of appropriate cell hosts with an rDNA molecule of the present invention is accomplished by well known methods that typically depend on the type of vector used and host system employed. With regard to transformation of prokaryotic host cells, electroporation and salt treatment methods are typically employed (see, *e.g.*, Cohen *et al.* (1972) Proc. Natl. Acad. Sci. USA 69, 2110-2112 and Sambrook *et al.* (2001) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press). With regard to transformation of vertebrate cells with vectors containing rDNAs, electroporation, cationic, lipid or salt treatment methods are typically employed (see, *e.g.*, Graham *et al.* (1973) Virol. 52, 456-458; Wigler *et al.* (1979) Proc. Natl. Acad. Sci. USA 76, 1373-1376).

Successfully transformed cells (*i.e.*, cells that contain an rDNA molecule of the present invention) can be identified by well known techniques including the selection for a selectable marker. For example, cells resulting from the introduction of an rDNA of the present invention can be cloned to produce single colonies. Cells from those colonies can be harvested, lysed and their DNA content examined for the presence of the rDNA using a method such as that described by Southern (1975) J. Mol. Biol. 98, 503-506) or Berent *et al.* (1985) Biotech. 3, 208-209) or the proteins produced from the cell assayed via an immunological method.

Production of Recombinant Proteins

The present invention further provides methods for producing a hagfish cathelin-associated protein of the invention using nucleic acid molecules herein described. In general terms, the production of a recombinant form of a protein typically involves the following steps:

First, a nucleic acid molecule is obtained that encodes a hagfish cathelin-associated protein of the invention, such as a nucleic acid molecule comprising, consisting essentially of or consisting of SEQ ID NO: 1, or the open reading frame defined by nucleotides 4-540 (or 543) of SEQ ID NO: 1; or such as a nucleic acid molecule comprising, consisting essentially of or consisting of SEQ ID NO: 3, or the open reading frame defined by nucleotides 2-505 (or 508) of SEQ ID NO: 3. Alternatively, a nucleic acid molecule is obtained that encodes a hagfish antimicrobial peptide of the invention, such as a nucleic acid molecule comprising,

consisting essentially of or consisting of nucleotides 427-540 of SEQ ID NO: 1, encoding amino acid residues 142-179 of SEQ ID NO: 2, or encoding SEQ ID NO: 5; or such as a nucleic acid molecule comprising, consisting essentially of or consisting of nucleotides 416-505 of SEQ ID NO: 3, encoding amino acid residues 139-168 of SEQ ID NO: 4, or encoding SEQ ID NO: 6. If the encoding sequence is uninterrupted by introns, as is this open reading frame, it is directly suitable for expression in any host.

The nucleic acid molecule is then preferably placed in operable linkage with suitable control sequences, as described above, to form an expression unit containing the protein open reading frame. The expression unit is used to transform a suitable host and the transformed host is cultured under conditions that allow the production of the recombinant protein. Optionally the recombinant protein is isolated from the medium or from the cells; recovery and purification of the protein may not be necessary in some instances where some impurities may be tolerated.

Each of the foregoing steps can be done in a variety of ways. For example, the desired coding sequences may be obtained from genomic fragments and used directly in appropriate hosts. The construction of expression vectors that are operable in a variety of hosts is accomplished using appropriate replicons and control sequences, as set forth above. The control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene and were discussed in detail earlier. Suitable restriction sites can, if not normally available, be added to the ends of the coding sequence so as to provide an excisable gene to insert into these vectors. A skilled artisan can readily adapt any host/expression system known in the art for use with the nucleic acid molecules of the invention to produce recombinant protein.

In another embodiment, the polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art.

Hagfish cathelin-associated proteins

The amino acid sequences, SEQ ID NO: 2 and SEQ ID NO: 4 (Figure 1) differ mainly because of differences in the region of the peptide associated with antimicrobial activity. The signal sequences are identical, with 26 residues; and the cathelins share 91.3% of their 115 residues. In contrast, the antimicrobial peptide consisting of amino acid residues 142-179 of SEQ ID NO: 2, GWFKKAWRKVKHAGRRVLDTAKGVGRHYLNNWLNRYRG (SEQ ID NO: 5), has 38 residues compared to 30 for the antimicrobial peptide consisting of amino acid residues 139-168 from SEQ ID NO: 4, GWFKKAWRKVKNAGRVLKGVGIHYGVGLIG (SEQ ID NO: 6). When three gaps are inserted in SEQ ID NO: 6, however, 23 of the 30 residues (76%) are identical to those of SEQ ID NO: 5. Both SEQ ID NO: 5 and SEQ ID NO: 6, have a C-terminal glycine residue which is likely post-translationally converted to a C-terminal amide as is the case with the hagfish antimicrobial peptides described in U.S. Patent 5,734,015.

As used herein, the family of proteins related to the hagfish cathelin-associated proteins, refers to other cathelin-associated proteins that can be isolated from the hagfish or organisms closely related to the hagfish. The methods used to identify and isolate other members of the family of proteins related to the hagfish cathelin-associated proteins are described below.

The proteins of the present invention are preferably in isolated form. As used herein, a protein is said to be isolated when physical, mechanical or chemical methods are employed to remove the protein from cellular constituents that are normally associated with the protein. A skilled artisan can readily employ standard purification methods to obtain an isolated protein. By isolated, it is further meant that the hagfish cathelin-associated protein or the antimicrobial peptide of the present invention is part of a mixture with other components and wherein the protein or peptide of the present invention comprises at least about 10% of the total protein in the mixture, preferably at least about 20% of the total protein in the mixture, more preferably at least about 30% of the total protein in the mixture, even more preferably at least about 40% of the total protein in the mixture, still more preferably at least about 50% of the total protein in the mixture, yet more preferably at least about 60% of the total protein in the mixture, even still more preferably at least about 70% of the total protein in the mixture, yet still more preferably at least about 80% of the total protein in the mixture, much more preferably at least about 90% of the total protein in the mixture, still much more preferably at least about 95% of the total protein in the mixture, and most preferably at least about 99% of the total protein in the mixture.

The proteins of the present invention further include conservative variants of the proteins herein described. As used herein, a conservative variant refers to alterations in the amino acid sequence that do not adversely affect the biological functions of the protein. A substitution, insertion or deletion is said to adversely affect the protein when the altered sequence prevents or disrupts a biological function associated with the protein. For example, the overall charge, structure or hydrophobic/hydrophilic properties of the protein may be altered without adversely affecting a biological activity. Accordingly, the amino acid sequence can be altered, for example to render the peptide more hydrophobic or hydrophilic, without adversely affecting the biological activities of the protein.

Ordinarily, the allelic variants, the conservative substitution variants, and the members of the protein family, will have an amino acid sequence having at least about 85% amino acid sequence identity with the sequences set forth in SEQ ID NO: 2, 4, 5 or 6, more preferably at least 90%, even more preferably at least 95% and most preferably at least 99%. Identity or homology with respect to such sequences is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the known peptides, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity. N-terminal, C-terminal or internal extensions, deletions, or insertions into the peptide sequence shall not be construed as affecting homology.

Thus, the proteins of the present invention include molecules comprising, consisting essentially of, or consisting of the amino acid sequence disclosed in SEQ ID NO: 2, 4, 5 and 6; fragments thereof having a consecutive sequence of at least about 3, 4, 5, 6, 10, 15, 20, 25, 30, 35, 38 or more amino acid residues of the hagfish cathelin-associated protein or antimicrobial peptide thereof; amino acid sequence variants of such sequence wherein an amino acid residue has been inserted N- or C-terminal to, or within, the disclosed sequence; and amino acid sequence variants of the disclosed sequence, or their fragments as defined above, that have been substituted by another residue. Contemplated variants further include those containing predetermined mutations by, *e.g.*, homologous recombination, site-directed or PCR mutagenesis, and the corresponding proteins of other animal species, including but not limited to rabbit, rat, porcine, bovine, ovine, equine and non-human primate species, and the alleles or other naturally occurring variants of the family of proteins; and derivatives wherein the protein has been covalently modified by substitution, chemical, enzymatic, or other appropriate means with a moiety other than a naturally occurring amino acid (for example a detectable moiety such as an enzyme or radioisotope).

Therapeutic and additive applications

Another aspect of the present invention is a method for inhibiting microbial growth. Hagfish cathelin-associated antimicrobial peptides of the present invention, exemplified by SEQ ID NOs: 5 and 6, can be used to inhibit microbial growth under various circumstances. For example, the peptides of the present invention can be administered therapeutically to treat or prevent disease in an individual resulting from a microbial infection. A variety of microbial infections can be inhibited by treatment with the peptides of the present invention. For example, the antimicrobial peptides of the present invention can be used as an antimicrobial agent for the treatment of fungal or bacterial infections.

In preferred embodiments, the microbial infection is a bacterial infection. In particular, therapeutic administration of the peptides of the present invention inhibits and can reverse the progression of a bacterial infection. The development and spread of new resistant strains of bacteria are increasingly posing a public health threat. The antimicrobial peptides of the present invention are useful in treating patients suffering from these resistant strains of microbes which currently threaten public health. Many bacterial infections cause bacterial sepsis in the infected individual. Administration of the peptides to an individual suffering from a septic bacterial infection produces a significant therapeutic effect, by reducing both the sepsis caused by the infection, and the infection itself.

An individual suitable for treatment is any animal (mammal or otherwise) which is afflicted with or otherwise susceptible to one or more of the above described microbial infections. In a preferred embodiment, the individual is a human. In another embodiment, the individual is a livestock animal. In another embodiment, the animal is a show animal or a household pet.

Administration of the hagfish cathelin-associated antimicrobial peptides of the present invention to the individual is either systemic or localized, and is largely determined by the specific infection being treated. Systemic administration can be accomplished by several routes, including but not limited to intravenous administration, inhalation, mucosal, and ingestion. Localized administration can be topical or internal. Such administration can be accomplished by several routes, including, but not limited to, subcutaneous, dermal, intradermal, buccal, mucosal, intraperitoneal, vaginal, inhalation, and ingestion.

It will often be of use to administer a formulation of the peptides of the present invention which includes a pharmaceutically acceptable carrier. Possible formulations for therapeutic administration include a variety of pharmaceutical compositions, the appropriate use of which will depend upon the route of administration deemed necessary for treatment. Some useful formulations for topical administration are, for example, eyedrops, eardrops, or

gingival applications (*e.g.*, drops, mouthwash, cream, or paste). The regimen of administration (*e.g.*, route, dose, and course) which is therapeutic to the patient will vary with the individual to be treated (*e.g.*, health, weight, metabolism), the site of infection, and the infecting pathogen. A therapeutic regimen should be developed by extrapolation from treatment with similar therapeutics in combination with empirical observation. Administration of the peptides of the present invention to prevent a microbial infection in an individual parallel the above described methods.

As a general proposition, the total pharmaceutically effective amount of antimicrobial peptide administered parenterally per dose will be in the range of about 0.001 mg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the peptide. If given continuously, the antimicrobial peptide is typically administered at a dose rate of about 0.001 mg/kg/hour to about 0.050 mg/kg/hour, either by one to four injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed.

Hagfish cathelin-associated antimicrobial peptides of the present invention, exemplified by SEQ ID NO: 5 and 6, are also useful for potentiating the therapeutic action of other antimicrobial drugs or agents. Co-administration of other antimicrobial agents with the antimicrobial polypeptides of the present invention, produce a synergistic antibiotic effect. Co-administration of the peptides of the present invention with an antimicrobial agent enables therapeutic treatment of a patient with lower doses of the antimicrobial agent. Lower doses are preferable in situations such as when treating with an expensive drug, or one that produces undesired side effects, or one whose short half-life *in vivo* would otherwise rapidly reduce its concentration below that which is required for it to be efficacious. In addition, co-administration with antimicrobial agents or drugs may also allow for a shorter therapy period and/or the reversal of resistant phenotypes. Without limitation, microbes which resist an antimicrobial drug by decreasing their internal drug concentration (*e.g.*, with decreased membrane permeability or increased cellular export or metabolism of the drug) are expected to be especially susceptible to the potentiating activity of these peptides.

The regimen of administration of the antimicrobial drug and the hagfish cathelin-associated antimicrobial peptide or functional variant of the present invention varies with the patient and the particular infection, and can be determined by one of skill in the art on a case by case basis. Formulations of the polypeptide will depend upon the regimen of administration, examples described above.

The polypeptides of the present invention can be used not only as pharmaceutical and neutraceutical agents but also as additives for any products such as foods and medicinal or non-medicinal products which are taken into the bodies or otherwise applied onto or contacted with the body surface of humans or other animals or fluids, organs, and cells derived therefrom.

The present method is useful for treating a variety of products. Biological products, defined herein as products which are derived from biological organisms or processes, are particularly at risk for contamination with microorganisms. Examples of biological products include without limitation, food products, tissue, living cells, products derived from living cells, blood or components thereof, as any other bodily fluid, drugs or other molecular preparations. Non-biological products, defined herein as a product not directly derived from a biological organism or process, for example glassware, surgical equipment, synthetic drugs or other molecular preparations, can also be treated. For effective use of the present invention, the product which is to be treated should not possess an activity which completely inactivates the antimicrobial activity of all of the peptide quantity so applied. The peptides of the present invention can be added, assorted to, sprayed to, adhered to, coated onto, adsorbed to, chemically crosslinked to, or impregnated into any products which are generally desired to be prevented or inhibited from contamination by proliferating microorganisms. Alternatively, the peptide of the present invention can be immobilized on a surface over which a product is passed to remove or inhibit microbial growth in the product. A product which is treated with and retains antimicrobial peptide can further be used to treat another product with which it is contacted.

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the present invention and practice the claimed methods. The following working examples therefore, specifically point out the preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

Example 1

Isolation of genes for hagfish cathelin-associated antibiotic peptides

Adult Atlantic hagfish (*Myxine glutinosa* Linnaeus, 1758; Myxinidae) were purchased from Huntsman Marine Science Centre (New Brunswick, Canada). The fish were anesthetized in a slurry of ice and decapitated. Intestines were removed and cleaned of contents, frozen in liquid N_2, and stored at -80°C.

Tissue was ground in liquid nitrogen and transferred to a guanidinium isothiocyanate solution (Chirgwin *et al.* (1979) Biochemistry 18, 5294-5299), which inhibits RNase activity

while disrupting cells. Total RNA was isolated by centrifugation through a CsCl step gradient; polyadenylated RNA was recovered by means of oligo(dT) cellulose (Stratagene). Reverse transcribed cDNA was ligated first to adaptor arms and then into phage Lambda Zap II following Stratagene protocols.

An twenty-mer oligonucleotide probe 5'-AARAARGCNTGGMGNAARGT-3' (SEQ ID NO: 11; 256-fold degenerate) was designed based on amino acid residues 4 to 10 of the peptides previously isolated from hagfish intestinal tissue. Phage particles containing homologous sequences were isolated by plating a small number of particles on a lawn of XL1 Blue MRF (Stratagene). The cDNA inserts were isolated in phagmid pBluescript by *in vivo* excision. Two inserts were sequenced in both directions using Sanger's dideoxy method (Barnes *et al.* (1983) Nucleic Acids Res. 11, 349-368).

Homologous sequences were identified using BLASTX (Altschul *et al.* (1997) Nucleic Acids Res. 25, 3389-3402) searches of GenBank. Homologous sequences were aligned using ESEE (Cabot *et al.* (1989) Compu. Appl. Biosci. 5, 233-234).

For *in situ* hybridizations, plasmid DNA containing insert 31-52 from SEQ ID NO: 1 was linearized with *StyI*. Both sense and antisense probes were labeled enzymatically with ^{32}P and detected by autoradiography.

Example 2

Comparison of hagfish cathelin-associated sequences with other cathelin-associated sequences

Searches of GenBank using BLASTN with default values yielded no matches either when just the coding region for the two peptides were used as queries or when the entire coding regions were used. Using the entire coding sequence available for each protein and BLASTX, which searches for more distant relationships based on identical and conserved amino acids, however, yielded a large number of matches. Those in the correct reading frame included especially cDNA sequences for prepropeptides of mammalian antibiotic peptides containing a cathelin-like acidic spacer.

Among the matches found, the sequence for *Capra* MAP28, isolated from bone marrow cells of *Capra hircus* (goat) was unique in showing at least a weak match between the antibiotic peptide and one of the two *Myxine* peptides. A plausible alignment of the coding regions of one of the hagfish sequences with *Capra* MAP28 is shown in Figure 2.

Example 3

Antimicrobial activity of cathelin-associated antibiotic peptides

The antimicrobial activity of cathelin-associated antibiotic peptides is determined by the following antibacterial assay based upon the guidelines of the National Committee for Clinical Laboratory Standards (Document M7-T2, (1988) Volume 8, No. 8).

A stock solution of the test peptide is prepared at a concentration of 512 μg/ml in sterile deionized distilled water and stored at -70°C. The stock peptide solutions are diluted in serial dilutions (1:2) down the wells of a microtiter plate so that the final concentrations of peptide in the well are 0.25, 0.50, 1, 2, 4, 8, 16, 32, 64, 128 and 256 μg/ml. 1.5×10^5 CFU/ml of the test microbe is added to the wells in full strength Mueller Hinton broth (BBL 11443) from a mid-log culture. The inoculum is standardized spectrophotometrically at 600 nm and is verified by colony counts. The plates are incubated for 16-24 hours at 37°C and the minimal inhibitory concentration (MIC) for each peptide is determined. Minimal inhibitory concentration is defined as the lowest concentration of peptide which produces a clear well in the microtiter plate.

While the invention has been described and illustrated herein by references to various specific materials, procedures and examples, it is understood that the invention is not restricted to the particular combinations of material and procedures selected for that purpose. Numerous variations of such details can be implied as will be appreciated by those skilled in the art. It is intended that the specification and examples be considered as exemplary, only, with the true scope and spirit of the invention being indicated by the following claims. All publications, patents and patent applications referred to in this application are herein incorporated by reference in their entirety.

What is claimed:

1. An isolated nucleic acid molecule selected from the group consisting of:

(a) an isolated nucleic acid molecule that encodes the amino acid sequence of a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 2 or 4; and

(b) an isolated nucleic acid molecule, which hybridizes to the complement of a nucleic acid molecule comprising a nucleotide sequence encoding a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 2 or 4 under stringent conditions.

2. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule is selected from the group consisting of SEQ ID NO: 1 and 3.

3. The isolated nucleic acid molecule of claim 2, wherein the nucleic acid molecule comprises nucleotides 4 to 540 of SEQ ID NO: 1.

4. The isolated nucleic acid molecule of claim 2, wherein the nucleic acid molecule comprises nucleotides 427 to 540 of SEQ ID NO: 1.

5. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 2 to 505 of SEQ ID NO: 3.

6. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 416 to 505 of SEQ ID NO: 3.

7. An isolated nucleic acid molecule selected from the group consisting of:

(a) an isolated nucleic acid molecule that encodes the amino acid sequence of a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 5 or 6; and

(b) an isolated nucleic acid molecule, which hybridizes to the complement of a nucleic acid molecule comprising a nucleotide sequence encoding a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 5 or 6 under stringent conditions.

8. The isolated nucleic acid molecule of any one of claims 1-7, wherein said nucleic acid molecule is operably linked to one or more expression control elements.

9. A vector comprising an isolated nucleic acid molecule of any one of claims 1-7.

10. A host cell comprising a vector of claim 9.

11. A host cell transformed to contain the nucleic acid molecule of any one of claims 1-7.

12. A method for producing a peptide or fragment thereof comprising the step of culturing a host cell of claim 11 under conditions in which the protein or protein fragment encoded by said nucleic acid molecule is expressed.

14. The method of claim 12, wherein said host cell is selected from the group consisting of prokaryotic hosts and eukaryotic hosts.

15. An isolated peptide or fragment thereof produced by the method of claim 12.

15. An isolated peptide or fragment thereof selected from the group consisting of:

(a) an isolated peptide comprising the amino acid sequences depicted in SEQ ID NO: 2, 4; 5 or 6;

(b) an isolated peptide fragment comprising at least six amino acids of any of the sequences depicted in SEQ ID NO: 2, 4, 5 or 6;

(c) an isolated peptide comprising conservative amino acid substitutions of any of the sequences depicted in SEQ ID NO: 2, 4, 5 or 6; and

(d) naturally occurring amino acid sequence peptide variants of any of the amino acid sequences depicted in SEQ ID NO: 2, 4, 5 or 6.

16. The isolated protein fragment of claim 15 wherein the isolated peptide consists essentially of the amino acid sequences depicted in SEQ ID NO: 2, 4, 5 or 6.

17. A antimicrobial composition comprising one or more of the peptides of claim 15

18. The antimicrobial composition of claim 17, wherein the composition is suitable for topical or parenteral administration.

19. A method of treating or preventing a microbial infection in a mammal, comprising administering to the mammal an effective amount of a *Myxine glutinosa* cathelin-associated antimicrobial peptide.

20. The method of claim 19, wherein the *Myxine glutinosa* cathelin-associated antimicrobial peptide is selected from the group consisting of SEQ ID NO: 2, 4; 5 or 6.

21. The method of claim 19, wherein the mammal is a human.

22. The method of claim 19, wherein the microbial infection is selected from the group consisting of a bacterial infection and a fungal infection.

23. The method of claim 17, wherein the *Myxine glutinosa* cathelin-associated antimicrobial peptide is administered by a route selected from the group consisting of oral, topical and parenteral.

24. The method of claim 23, wherein the topical administration is accomplished by inhalation.

25. The method of claim 24, wherein the parenteral administration is accomplished by intravenous, subcutaneous or intramuscular administration.

Figure 1A

```
                >Signal sequence
SEQ ID NO. 1 gtcATGaagtccttgtgtgtccccgcggttctctctctggtcctgatcctgctcttggat
SEQ ID NO. 3           ..........a.....................................
SEQ ID NO. 2     M  K  S  L  C  V  P  A  V  L  S  L  V  L  I  L  L  L  D
SEQ ID NO. 4           .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

                                  >Cathelin
SEQ ID NO. 1 caagcacctacagctcgggcagatgattctttgtccaaagagcaggtggagaatgcagtt
SEQ ID NO. 3 .....................................c..................
SEQ ID NO. 2  Q  A  P  T  A  R  A  D  D  S  L  S  K  E  Q  V  E  N  A  V
SEQ ID NO. 4  .  .  .  .  .  .  .  .  .  .  .  N  .  .  .  .  .  .  .  .

SEQ ID NO. 1 gatgaagcgttggacaagctgaacaaacagcaggtgtccacacgtaaactggcattgtct
SEQ ID NO. 3 ........................t.............................
SEQ ID NO. 2  D  E  A  L  D  K  L  N  K  Q  Q  V  S  T  R  K  L  A  L  S
SEQ ID NO. 4  .  .  .  .  .  .  .  .  N  .  .  .  .  .  .  .  .  .  .  .

SEQ ID NO. 1 gaacaacaggatatccaggcagatgaaacagacgttgaaggacagttcaccatcaaattt
SEQ ID NO. 3 ..............................................ac........t....
SEQ ID NO. 2  E  Q  Q  D  I  Q  A  D  E  T  D  V  E  G  Q  F  T  I  K  F
SEQ ID NO. 4  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  T  .  .  I  .

SEQ ID NO. 1 gacgtggttgagaccgagtgcaatgcggacgatcccagggactgggccgattgtccgatc
SEQ ID NO. 3 t...........................................................
SEQ ID NO. 2  D  V  V  E  T  E  C  N  A  D  D  P  R  D  W  A  D  C  P  I
SEQ ID NO. 4  Y  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

SEQ ID NO. 1 gcgacggactcgacaccggtcgatgcacagtgtgaggtcaccgtgttgagcaccgaggac
SEQ ID NO. 3 ............c......g.at.....................................
SEQ ID NO. 2  A  T  D  S  T  P  V  D  A  Q  C  E  V  T  V  L  S  T  E  D
SEQ ID NO. 4  .  .  .  .  P  .  G  I  .  .  .  .  .  .  .  .  .  .  .  .

SEQ ID NO. 1 tccttggacgtcggagacgcgacttgcgatttcaacaggacggacggaaatgcgaggcga
SEQ ID NO. 3 ..................................tc.....g................
SEQ ID NO. 2  S  L  D  V  G  D  A  T  C  D  F  N  R  T  D  G  N  A  R  R
SEQ ID NO. 4  .  .  .  .  .  .  .  .  .  .  .  S  .  G  .  .  .  .  .  .

                >Antibiotic peptide
SEQ ID NO. 1 cgacgtggctggtttaagaaagcctggagaaaagtgaagcatgcgggacgacgagttctt
SEQ ID NO. 3 .....................................a.....---......
SEQ ID NO. 2  R  R  G  W  F  K  K  A  W  R  K  V  K  H  A  G  R  R  V  L
SEQ ID NO. 4  .  .  .  .  .  .  .  .  .  .  .  .  N  .  .  . --- .  .
```

Figure 1B

```
SEQ ID NO. 1 gataccgcgaagggtgtgggaagacattatttgaataattggcttaatcg----------
SEQ ID NO. 3 ---------.............t.....----ggtg....a.......gtTAAtcggt

>3'UTR
SEQ ID NO. 2  D  T  A  K  G  V  G  R  H  Y  L  N  N  W  L  N  R ---------
SEQ ID NO. 4  -------- .  .  .  .  I  .  . ----G  V  G  L  I  G  *

                                        >3'UTR
SEQ ID NO. 1 -----------ttatcgcggtTAGgaggaagctctgtgttggtggtggtggtggtggtgg
SEQ ID NO. 3 tggattaatcg.....a.....c.........tc..c.......................
SEQ ID NO. 2 ------------ Y  R  G  *

SEQ ID NO. 1 tggtggtggtg----------------------------ttgtgggttaatgggtgggag
SEQ ID NO. 3 ...........gtggtggtggtggtggtggtgtgggggg..a..................

SEQ ID NO. 1 cgaatgatttctccccgacattcaatacttgctcgttaattccaattgattcgcttagtt
SEQ ID NO. 3 ............................................................

SEQ ID NO. 1 acacaagtaagccgaacaggtaaccagttagtaagctttaaaattgaacttttaggattg
SEQ ID NO. 3 ............................................................

SEQ ID NO. 1 aaactgcagctgaacttttttgtagtttaacatttaattgtgactctgctattaaactttg
SEQ ID NO. 3 ............................................................

SEQ ID NO. 1 tctgatcatggaaataaaggattgcaaaaatgaaaaaaaaaaaaaaaaaaaaaaaaa
SEQ ID NO. 3 ...........aaaaaaa.....................................
```

Figure 2

```
          Exon I→
CapraM28.n ATGgagacccagagggccagcctctccctgggacggtgctccctgtggctcctgctgctg 60
3          ---------.tt.t.t.t.ccagcggtt..c---------..t...gtc..ga.c.....c 43
CapraM28.a  M  E  T  Q  R  A  S  L  S  L  G  R  C  S  L  W  L  L  L  L  20
4                 L  C  V  P  A  V  .  -  -  -  .  .  V  .  I  .  .  14

CapraM28.n ggactagtgctgccctcggccagcgcccaggccctcagctacggggaggccgttcttcat 120
3          ttggatcaagca..ta.a..tc.g..ag.t.attctttg.c.aac...ca.gt.gagaat 103
CapraM28.a  G  L  V  L  P  S  A  S  A  Q  A  L  S  Y  G  E  A  V  L  H  40
4           L  D  Q  A  .  T  .  R  .  D  D  S  L  S  N  .  Q  .  E  N  34

CapraM28.n gctgtcgat------------cgcatcaatgagcagtcctcagaagcgaatctctaccgc 168
3          ..a..t...gaagcgttggacaagc.g..ca.t.a.caggtgtcca.acg.aaactggca 163
CapraM28.a  A  V  D  -  -  -  -  R  I  N  E  Q  S  S  E  A  N  L  Y  R  56
4           .  .  .  E  A  L  D  K  L  .  N  .  Q  V  S  T  R  K  L  A  54

                                        Exon II→
CapraM28.n ctcctggagcttgacccgcctcccaaggac|gatgag---aatccaaacatcccgaaacct 225
3          t.gtctg.a---c.a.a.ga.at.c...ca|.....aacag.cgttg.agga.a..cca.c 220
CapraM28.a  L  L  E  L  D  P  P  P  K  D | D  E  -  N  P  N  I  P  K  P  75
4           .  S  .  -  Q  Q  D  I  Q  A |  .  .  T  D  V  E  G  Q  T  T  73

CapraM28.n gtgagcttcagggtgaaggagactgtgtgcccaggacgagccggcagcccacggagcag 285
3          a.c.ta..ttac...gtt.....c.a....aatgc.gacgat.ccag.ga.tg..ccg.t 280
CapraM28.a  V  S  F  R  V  K  E  T  V  C  P  R  T  S  R  Q  P  T  E  Q  95
4           I  I  .  Y  .  V  .  .  E  .  N  A  D  D  P  R  D  W  A  D  93

                                  Exon III→
CapraM28.n tgtgacttcaaagagaatggg---|ctggtgaag---caatgtgtagggacagtcactctg 339
3          ...ccga..gcgac.g.ctc.cca|.c..gcattgca..g....ag------.....cg.. 334
CapraM28.a  C  D  F  K  E  N  G  -  | L  V  K  -  Q  C  V  G  T  V  T  L  113
4           .  P  I  A  T  D  S  P |  P  G  I  A  .  .  E  -  -  .  .  V  111

                                                                Exon IV→
CapraM28.n gatgcggtgaaaggcaaaatgaacatc---------acctgcgaagag|ttgcagagtgtt 390
3          ttgagcaccg.g.a.tcct..g..g..ggagacgcg..t.....tttc|aactc..cg.gc 394
CapraM28.a  D  A  V  K  G  K  M  N  I  -  -  -  T  C  E  E | L  Q  S  V  130
4           L  S  T  E  D  S  L  D  V  G  D  A  .  .  D  F | N  S  T  G  131

CapraM28.n gggagatttaaacgatttcgt---------aagaagctcaaaagactctggcacaaagtc 441
3          ..a.atgcg.gg...cga...ggctggttt.....agc.tgg...aaagt.a.g..t.cg 454
CapraM28.a  G  R  F  K  R  F  R  -  -  -  K  K  L  K  R  L  W  H  K  V  147
4           .  N  A  R  .  R  .  G  W  F  .  .  A  W  .  K  V  K  N  A  151


CapraM28.n ggcccattcgtt---ggcccgatactccattatgggTAAattgtgagcccatggaagaat 498
3          ..a.g.g.tc..aag..tgt.gg.a.a........tgttggat.a.t.ggt.AA       508
CapraM28.a  G  P  F  V  -  G  P  I  L  H  Y  G  *                     158
4           .  R  V  L  K  .  V  G  I  .  .  .  V  G  L  I  G  *       168
```

Figure 3

SEQ ID NO. 5 GWFKKAWRKVKHAGRRVLDTAKGVGRHYLNNWLNRYRG
SEQ ID NO. 6 ..........N...-..--....I..-GVG.IG

Figure 4

SEQ ID NO. 5 GWFKKAWRKVKHAGRRVLDTAKGVGRHYLNNWLNRYRG
SEQ ID NO. 9 .F....X.....V..X.....Z

SEQ ID NO. 6 GWFKKAWRKVKNAGR-VLKGVGIHYGVGLIG
SEQ ID NO. 10 .X....X.....R.....Z

SEQUENCE LISTING

<110> GENAERA CORPORATION
UZZELL, Thomas
STOLZENBERG, Ethan D
SHINNAR, Ann E
ZASLOFF, Michael A

<120> HAGFISH CATHELIN-ASSOCIATED ANTIMICROBIAL PEPTIDES AND GENES

<130> 36870-5085

<150> US 60/308,652
<151> 2001-07-30

<160> 11

<170> PatentIn version 3.1

<210> 1
<211> 848
<212> DNA
<213> Myxine glutinosa

<220>
<221> CDS
<222> (4)..(543)
<223>

<400> 1

```
gtc atg aag tcc ttg tgt gtc ccc gcg gtt ctc tct ctg gtc ctg atc       48
    Met Lys Ser Leu Cys Val Pro Ala Val Leu Ser Leu Val Leu Ile
    1               5                  10                  15

ctg ctc ttg gat caa gca cct aca gct cgg gca gat gat tct ttg tcc       96
Leu Leu Leu Asp Gln Ala Pro Thr Ala Arg Ala Asp Asp Ser Leu Ser
                20                  25                  30

aaa gag cag gtg gag aat gca gtt gat gaa gcg ttg gac aag ctg aac      144
Lys Glu Gln Val Glu Asn Ala Val Asp Glu Ala Leu Asp Lys Leu Asn
            35                  40                  45

aaa cag cag gtg tcc aca cgt aaa ctg gca ttg tct gaa caa cag gat      192
Lys Gln Gln Val Ser Thr Arg Lys Leu Ala Leu Ser Glu Gln Gln Asp
        50                  55                  60

atc cag gca gat gaa aca gac gtt gaa gga cag ttc acc atc aaa ttt      240
Ile Gln Ala Asp Glu Thr Asp Val Glu Gly Gln Phe Thr Ile Lys Phe
    65                  70                  75

gac gtg gtt gag acc gag tgc aat gcg gac gat ccc agg gac tgg gcc      288
Asp Val Val Glu Thr Glu Cys Asn Ala Asp Asp Pro Arg Asp Trp Ala
80                  85                  90                  95

gat tgt ccg atc gcg acg gac tcg aca ccg gtc gat gca cag tgt gag      336
Asp Cys Pro Ile Ala Thr Asp Ser Thr Pro Val Asp Ala Gln Cys Glu
                100                 105                 110

gtc acc gtg ttg agc acc gag gac tcc ttg gac gtc gga gac gcg act      384
Val Thr Val Leu Ser Thr Glu Asp Ser Leu Asp Val Gly Asp Ala Thr
```

```
           115                 120                 125

tgc gat ttc aac agg acg gac gga aat gcg agg cga cga cgt ggc tgg      432
Cys Asp Phe Asn Arg Thr Asp Gly Asn Ala Arg Arg Arg Arg Gly Trp
        130                 135                 140

ttt aag aaa gcc tgg aga aaa gtg aag cat gcg gga cga cga gtt ctt      480
Phe Lys Lys Ala Trp Arg Lys Val Lys His Ala Gly Arg Arg Val Leu
    145                 150                 155

gat acc gcg aag ggt gtg gga aga cat tat ttg aat aat tgg ctt aat      528
Asp Thr Ala Lys Gly Val Gly Arg His Tyr Leu Asn Asn Trp Leu Asn
160                 165                 170                 175

cgt tat cgc ggt tag gaggaagctc tgtgttggtg gtggtggtgg tggtggtggt      583
Arg Tyr Arg Gly


ggtggtgttg tgggttaatg ggtgggagcg aatgatttct ccccgacatt caatacttgc    643

tcgttaattc caattgattc gcttagttac acaagtaagc cgaacaggta accagttagt    703

aagctttaaa attgaacttt taggattgaa actgcagctg aacttttttgt agtttaacat   763

ttaattgtga ctctgctatt aaactttgtc tgatcatgga aataaaggat tgcaaaaatg    823

aaaaaaaaaa aaaaaaaaaa aaaaa                                          848


<210>  2
<211>  179
<212>  PRT
<213>  Myxine glutinosa

<400>  2

Met Lys Ser Leu Cys Val Pro Ala Val Leu Ser Leu Val Leu Ile Leu
1               5                   10                  15


Leu Leu Asp Gln Ala Pro Thr Ala Arg Ala Asp Asp Ser Leu Ser Lys
            20                  25                  30


Glu Gln Val Glu Asn Ala Val Asp Glu Ala Leu Asp Lys Leu Asn Lys
        35                  40                  45


Gln Gln Val Ser Thr Arg Lys Leu Ala Leu Ser Glu Gln Gln Asp Ile
    50                  55                  60


Gln Ala Asp Glu Thr Asp Val Glu Gly Gln Phe Thr Ile Lys Phe Asp
65                  70                  75                  80


Val Val Glu Thr Glu Cys Asn Ala Asp Asp Pro Arg Asp Trp Ala Asp
                85                  90                  95


Cys Pro Ile Ala Thr Asp Ser Thr Pro Val Asp Ala Gln Cys Glu Val
```

```
           100                 105                 110

Thr Val Leu Ser Thr Glu Asp Ser Leu Asp Val Gly Asp Ala Thr Cys
        115                 120                 125

Asp Phe Asn Arg Thr Asp Gly Asn Ala Arg Arg Arg Arg Gly Trp Phe
    130                 135                 140

Lys Lys Ala Trp Arg Lys Val Lys His Ala Gly Arg Arg Val Leu Asp
145                 150                 155                 160

Thr Ala Lys Gly Val Gly Arg His Tyr Leu Asn Asn Trp Leu Asn Arg
                165                 170                 175

Tyr Arg Gly
```

<210> 3
<211> 863
<212> DNA
<213> Myxine glutinosa

<220>
<221> CDS
<222> (2)..(508)
<223>

<400> 3

```
c ttg tgt gtc cca gcg gtt ctc tct ctg gtc ctg atc ctg ctc ttg gat      49
  Leu Cys Val Pro Ala Val Leu Ser Leu Val Leu Ile Leu Leu Leu Asp
  1               5                  10                  15

caa gca cct aca gct cgg gca gat gat tct ttg tcc aac gag cag gtg       97
Gln Ala Pro Thr Ala Arg Ala Asp Asp Ser Leu Ser Asn Glu Gln Val
            20                  25                  30

gag aat gca gtt gat gaa gcg ttg gac aag ctg aac aat cag cag gtg      145
Glu Asn Ala Val Asp Glu Ala Leu Asp Lys Leu Asn Asn Gln Gln Val
        35                  40                  45

tcc aca cgt aaa ctg gca ttg tct gaa caa cag gat atc cag gca gat      193
Ser Thr Arg Lys Leu Ala Leu Ser Glu Gln Gln Asp Ile Gln Ala Asp
    50                  55                  60

gaa aca gac gtt gaa gga cag acc acc atc ata ttt tac gtg gtt gag      241
Glu Thr Asp Val Glu Gly Gln Thr Thr Ile Ile Phe Tyr Val Val Glu
65                  70                  75                  80

acc gag tgc aat gcg gac gat ccc agg gac tgg gcc gat tgt ccg atc      289
Thr Glu Cys Asn Ala Asp Asp Pro Arg Asp Trp Ala Asp Cys Pro Ile
                85                  90                  95

gcg acg gac tcg cca ccg ggc att gca cag tgt gag gtc acc gtg ttg      337
Ala Thr Asp Ser Pro Pro Gly Ile Ala Gln Cys Glu Val Thr Val Leu
```

```
           100                 105                 110

agc acc gag gac tcc ttg gac gtc gga gac gcg act tgc gat ttc aac      385
Ser Thr Glu Asp Ser Leu Asp Val Gly Asp Ala Thr Cys Asp Phe Asn
        115                 120                 125

tcg acg ggc gga aat gcg agg cga cga cgt ggc tgg ttt aag aaa gcc      433
Ser Thr Gly Gly Asn Ala Arg Arg Arg Arg Gly Trp Phe Lys Lys Ala
    130                 135                 140

tgg aga aaa gtg aag aat gcg gga cga gtt ctt aag ggt gtg gga ata      481
Trp Arg Lys Val Lys Asn Ala Gly Arg Val Leu Lys Gly Val Gly Ile
145                 150                 155                 160

cat tat ggt gtt gga tta atc ggt taa tcggttggat taatcgttat            528
His Tyr Gly Val Gly Leu Ile Gly
                165

cacggttcgg aggaagtcct ctgttggtgg tggtggtggt ggtggtggtg gtggtggtgg    588

tggtggtggt ggtggtgtgg ggggttatgg gttaatgggt gggagcgaat gatttctccc    648

cgacattcaa tacttgctcg ttaattccaa ttgattcgct tagttacaca agtaagccga    708

acaggtaacc agttagtaag ctttaaaatt gaacttttag gattgaaact gcagctgaac    768

tttttgtagt ttaacattta attgtgactc tgctattaaa ctttgtctga tcatggaaat    828

aaaggattgc aaaaatgaaa aaaaaaaaaa aaaaa                               863


<210>  4
<211>  168
<212>  PRT
<213>  Myxine glutinosa

<400>  4

Leu Cys Val Pro Ala Val Leu Ser Leu Val Leu Ile Leu Leu Leu Asp
1               5                   10                  15

Gln Ala Pro Thr Ala Arg Ala Asp Asp Ser Leu Ser Asn Glu Gln Val
            20                  25                  30

Glu Asn Ala Val Asp Glu Ala Leu Asp Lys Leu Asn Asn Gln Gln Val
        35                  40                  45

Ser Thr Arg Lys Leu Ala Leu Ser Glu Gln Gln Asp Ile Gln Ala Asp
    50                  55                  60

Glu Thr Asp Val Glu Gly Gln Thr Thr Ile Ile Phe Tyr Val Val Glu
65                  70                  75                  80

Thr Glu Cys Asn Ala Asp Asp Pro Arg Asp Trp Ala Asp Cys Pro Ile
                85                  90                  95
```

Ala Thr Asp Ser Pro Pro Gly Ile Ala Gln Cys Glu Val Thr Val Leu
100 105 110

Ser Thr Glu Asp Ser Leu Asp Val Gly Asp Ala Thr Cys Asp Phe Asn
115 120 125

Ser Thr Gly Gly Asn Ala Arg Arg Arg Arg Gly Trp Phe Lys Lys Ala
130 135 140

Trp Arg Lys Val Lys Asn Ala Gly Arg Val Leu Lys Gly Val Gly Ile
145 150 155 160

His Tyr Gly Val Gly Leu Ile Gly
165

<210> 5
<211> 38
<212> PRT
<213> Myxine glutinosa

<400> 5

Gly Trp Phe Lys Lys Ala Trp Arg Lys Val Lys His Ala Gly Arg Arg
1 5 10 15

Val Leu Asp Thr Ala Lys Gly Val Gly Arg His Tyr Leu Asn Asn Trp
20 25 30

Leu Asn Arg Tyr Arg Gly
35

<210> 6
<211> 30
<212> PRT
<213> Myxine glutinosa

<400> 6

Gly Trp Phe Lys Lys Ala Trp Arg Lys Val Lys Asn Ala Gly Arg Val
1 5 10 15

Leu Lys Gly Val Gly Ile His Tyr Gly Val Gly Leu Ile Gly
20 25 30

<210> 7
<211> 498
<212> DNA
<213> Capra hircus

<220>

<221> CDS
<222> (1)..(477)
<223>

<400> 7
atg gag acc cag agg gcc agc ctc tcc ctg gga cgg tgc tcc ctg tgg 48
Met Glu Thr Gln Arg Ala Ser Leu Ser Leu Gly Arg Cys Ser Leu Trp
1 5 10 15

ctc ctg ctg ctg gga cta gtg ctg ccc tcg gcc agc gcc cag gcc ctc 96
Leu Leu Leu Leu Gly Leu Val Leu Pro Ser Ala Ser Ala Gln Ala Leu
20 25 30

agc tac ggg gag gcc gtt ctt cat gct gtc gat cgc atc aat gag cag 144
Ser Tyr Gly Glu Ala Val Leu His Ala Val Asp Arg Ile Asn Glu Gln
35 40 45

tcc tca gaa gcg aat ctc tac cgc ctc ctg gag ctt gac ccg cct ccc 192
Ser Ser Glu Ala Asn Leu Tyr Arg Leu Leu Glu Leu Asp Pro Pro Pro
50 55 60

aag gac gat gag aat cca aac atc ccg aaa cct gtg agc ttc agg gtg 240
Lys Asp Asp Glu Asn Pro Asn Ile Pro Lys Pro Val Ser Phe Arg Val
65 70 75 80

aag gag act gtg tgc ccc agg acg agc cgg cag ccc acg gag cag tgt 288
Lys Glu Thr Val Cys Pro Arg Thr Ser Arg Gln Pro Thr Glu Gln Cys
85 90 95

gac ttc aaa gag aat ggg ctg gtg aag caa tgt gta ggg aca gtc act 336
Asp Phe Lys Glu Asn Gly Leu Val Lys Gln Cys Val Gly Thr Val Thr
100 105 110

ctg gat gcg gtg aaa ggc aaa atg aac atc acc tgc gaa gag ttg cag 384
Leu Asp Ala Val Lys Gly Lys Met Asn Ile Thr Cys Glu Glu Leu Gln
115 120 125

agt gtt ggg aga ttt aaa cga ttt cgt aag aag ctc aaa aga ctc tgg 432
Ser Val Gly Arg Phe Lys Arg Phe Arg Lys Lys Leu Lys Arg Leu Trp
130 135 140

cac aaa gtc ggc cca ttc gtt ggc ccg ata ctc cat tat ggg taa 477
His Lys Val Gly Pro Phe Val Gly Pro Ile Leu His Tyr Gly
145 150 155

attgtgagcc catggaagaa t 498

<210> 8
<211> 158
<212> PRT
<213> Capra hircus

<400> 8

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

Leu Leu Leu Leu Gly Leu Val Leu Pro Ser Ala Ser Ala Gln Ala Leu

20 25 30

Ser Tyr Gly Glu Ala Val Leu His Ala Val Asp Arg Ile Asn Glu Gln
35 40 45

Ser Ser Glu Ala Asn Leu Tyr Arg Leu Leu Glu Leu Asp Pro Pro Pro
50 55 60

Lys Asp Asp Glu Asn Pro Asn Ile Pro Lys Pro Val Ser Phe Arg Val
65 70 75 80

Lys Glu Thr Val Cys Pro Arg Thr Ser Arg Gln Pro Thr Glu Gln Cys
85 90 95

Asp Phe Lys Glu Asn Gly Leu Val Lys Gln Cys Val Gly Thr Val Thr
100 105 110

Leu Asp Ala Val Lys Gly Lys Met Asn Ile Thr Cys Glu Glu Leu Gln
115 120 125

Ser Val Gly Arg Phe Lys Arg Phe Arg Lys Lys Leu Lys Arg Leu Trp
130 135 140

His Lys Val Gly Pro Phe Val Gly Pro Ile Leu His Tyr Gly
145 150 155

<210> 9
<211> 38
<212> PRT
<213> Myxine glutinosa

<220>
<221> MISC_FEATURE
<222> (1)..(38)
<223> Xaa is Brominated Tryptophan

<220>
<221> MISC_FEATURE
<222> (38)..(38)
<223> Glx is C-terminal amide residue

<400> 9

Gly Phe Phe Lys Lys Ala Xaa Arg Lys Val Lys His Ala Gly Arg Arg
1 5 10 15

Val Leu Asp Thr Ala Lys Gly Val Gly Arg His Tyr Val Asn Asn Xaa
20 25 30

Leu Asn Arg Tyr Arg Glx
35

<210> 10
<211> 31
<212> PRT
<213> Myxine glutinosa

<220>
<221> MISC_FEATURE
<222> (1)..(31)
<223> Xaa is Brominated Tryptophan

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> Glx is C-terminal amide residue

<400> 10

Gly Xaa Phe Lys Lys Ala Xaa Arg Lys Val Lys Asn Ala Gly Arg Arg
1 5 10 15

Val Leu Lys Gly Val Gly Ile His Tyr Gly Val Gly Leu Ile Glx
20 25 30

<210> 11
<211> 20
<212> DNA
<213> Myxine glutinosa

<220>
<221> misc_feature
<222> (1)..(20)
<223> r = g or a

<220>
<221> misc_feature
<222> (1)..(20)
<223> n = a or t or c or g

<400> 11
aaraargcnt ggmgnaargt 20

【国際公開パンフレット（コレクトバージョン）】

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

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 February 2003 (13.02.2003)

PCT

(10) International Publication Number
WO 2003/012044 A2

(51) International Patent Classification⁷: C12N

(21) International Application Number: PCT/US2002/024019

(22) International Filing Date: 30 July 2002 (30.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/308,652 30 July 2001 (30.07.2001) US

(71) Applicant (*for all designated States except US*): **GENAERA CORPORATION** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **UZZELL, Thomas** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US). **STOLZENBERG, Ethan, D.** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US). **SHINNAR, Ann, E.** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US). **ZASLOFF, Michael, A.** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US).

(74) Agents: **SMYTH, Robert, J.** et al.; Morgan, Lewis & Bockius LLP, 1111 Pennsylvania Avenue, NW, Washington, DC 20004 (US).

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

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

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

(48) Date of publication of this corrected version:
1 April 2004

(15) Information about Correction:
see PCT Gazette No. 14/2004 of 1 April 2004, Section II

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

WO 2003/012044 A2

(54) Title: HAGFISH CATHELIN-ASSOCIATED ANTIMICROBIAL PEPTIDES AND GENES

(57) Abstract: The present invention includes *Myxine glutinosa* cathelin-associated antimicrobial peptides and genes encoding these peptides. The invention also includes compositions an methods for producing these peptides as well as method of preventing and treating microbial infections using these peptides.

Hagfish Cathelin-Associated Antimicrobial Peptides and Genes

Cross-Reference To Related Application

This application claims the benefit of U.S. provisional application 60/308,652 filed on July 30, 2002.

Field Of The Invention

This invention relates to genes for antibiotic peptides and their use for the production of said peptides. It also relates to the antimicrobial peptides and their use as antibiotics.

Background Of The Invention

The cathelicidin gene family (Zanetti *et al.* (1995) FEBS Letters 374, 1-5; Zanetti *et al.* (1997) Ann. New York. Acad. Sci. 147-162; Gennaro *et al.* (2000) Biopolymers Peptide Science 55, 31-49) encodes prepropeptides that consist of a signal sequence, a cathelin-like acidic spacer, and an antibiotic peptide (Figure 1). Cathelicidins are presently known from a variety of mammals: *Bos taurus* (cattle), *Capra hircus* (goat), *Cavia porcellus* (guinea pig), *Equus caballus* (horse), *Homo sapiens* (human), *Mus musculus* (mouse), *Oryctolagus cunniculus* (rabbit), *Ovis aries* (sheep), and *Sus scrofa* (pig). Although the known antibiotic peptides are remarkably diverse, both within and between mammalian species, the associated signal sequences and cathelin-like acidic spacers, as revealed by sequence analysis of cDNA clones, are as remarkably conserved within and among species as the antibiotic peptides themselves are diverse.

Mammalian cathelicidins are synthesized primarily in myeloid cells, and are abundant in circulating neutrophils. The structure of some 20 cathelicidin genes is known and regularly involves four exons, the last of which encodes the antibiotic peptide along with a few C-terminal residues of the cathelin. Clusters of cathelicidin genes numbering eight, eleven, and more have been reported in *Ovis*, *Bos* and *Sus*, respectively (Gennaro *et al.* (2000) Biopolymers Peptide Science 55, 31-49).

Two antibiotic peptides have been isolated (Kaumaya *et al.* (1996) Peptides: Chemistry and Biology, Mayflower Scientific, pages 189-191) from intestinal tissue of *Myxine glutinosa* (Atlantic hagfish), a craniate chordate taxon only distantly related to the Mammalia. These peptides are unique among antimicrobial peptides in containing one or two brominated tryptophans. Herein are described hagfish cDNA sequences encoding two new antibiotic peptides from the same family as those isolated earlier. These sequences identify the *Myxine* antibiotic peptides as members of the cathelicidin family.

Summary Of The Invention

The invention includes an isolated nucleic acid molecule selected from the group consisting of an isolated nucleic acid molecule that encodes the amino acid sequence of a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 2 or 4; and an isolated nucleic acid molecule, which hybridizes to the complement of a nucleic acid molecule comprising a nucleotide sequence encoding a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 2 or 4 under stringent conditions. In some embodiments, the nucleic acid molecule comprises SEQ ID NO: 1 and 3; nucleotides 4 to 540 of SEQ ID NO: 1; nucleotides 427 to 540 of SEQ ID NO: 1; nucleotides 2 to 505 of SEQ ID NO: 3; or nucleotides 416 to 505 of SEQ ID NO: 3.

The invention also includes an isolated nucleic acid molecule selected from the group consisting of an isolated nucleic acid molecule that encodes the amino acid sequence of a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 5 or 6; and an isolated nucleic acid molecule, which hybridizes to the complement of a nucleic acid molecule comprising a nucleotide sequence encoding a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 5 or 6 under stringent conditions.

The isolated nucleic acid molecules of the invention can be operably linked to one or more expression control elements. The invention further includes a vector comprising an any of the isolated nucleic acid molecules recited above and a host cell comprising the vector.

The invention includes a method for producing a peptide or fragment thereof comprising the step of culturing the aforementioned host cell under conditions in which the protein or protein fragment encoded by said nucleic acid molecule is expressed. The host cell can be prokaryotic or eukaryotic. The invention includes recombinant peptides and fragments thereof produced by this method.

The invention includes an isolated peptide or fragment thereof selected from the group consisting of an isolated peptide comprising the amino acid sequences depicted in SEQ ID NO: 2, 4, 5 or 6; an isolated peptide fragment comprising at least six amino acids of any of the sequences depicted in SEQ ID NO: 2, 4, 5 or 6; an isolated peptide comprising conservative amino acid substitutions of any of the sequences depicted in SEQ ID NO: 2, 4, 5 or 6; and naturally occurring amino acid sequence peptide variants of any of the amino acid sequences depicted in SEQ ID NO: 2, 4, 5 or 6.

The invention further includes an antimicrobial composition comprising one or more of the peptides of of the invention. The antimicrobial composition can be suitable for topical or parenteral administration.

The invention includes a method of treating or preventing a microbial infection in a mammal, comprising administering to the mammal an effective amount of a *Myxine glutinosa* cathelin-associated antimicrobial peptide. In some embodiments, the *Myxine glutinosa* cathelin-associated antimicrobial peptide is selected from the group consisting of SEQ ID NO: 2, 4, 5 or 6 and the mammal is a human at risk of, or suffering from, a bacterial infection or a fungal infection. The *Myxine glutinosa* cathelin-associated antimicrobial peptide containing compositions can be administered by a route selected from the group consisting of oral, topical and parenteral. When topical administration is employed, it can be accomplished by inhalation. When parenteral administration is employed, it can be accomplished by intravenous, subcutaneous or intramuscular administration.

Brief Description Of The Figures

Figure 1 depicts a comparison of nucleotide (SEQ ID NO: 1 and 3) and amino acid sequences (SEQ ID NO: 2 and 4) for hagfish cathelin-associated antimicrobial peptides. Beginnings of presumed signal, cathelin, peptide, and 3' untranslated regions are marked. The peptide amino acids are italicized and in bold; start and stop codons are in upper case; the polyadenylation site is underlined. Dots indicate identical nucleotides of amino acids while dashes indicate insertions or deletions. * indicates the putative C-terminal amidation site in the antibiotic peptides.

Figure 2 depicts a comparison of nucleotides and implied amino acid residues of SEQ ID NOS. 3 and 4 with *Capra hircus* MAP28. Borders between exons suggested by mammalian gene structures are marked by vertical bars.

Figure 3 depicts a comparison of the amino acid sequences of the hagfish antibiotic peptides (SEQ ID NO: 5 and 6). Dashes indicate gaps inserted to maximize the alignment.

Figure 4 depicts a comparison of amino acid sequences for hagfish cathelin-associated antimicrobial peptides (SEQ ID NO: 5 and 6) with antimicrobial peptides isolated from hagfish intestinal tissues (SEQ ID NO: 9 and 10) described in U.S. Patent 5,734,015. Dashes indicate gaps inserted to maximize the alignment. X=Br-Trp and Z indicates a C-terminal amide.

Detailed Description

General

Applicants have isolated and sequenced two new genes from the Atlantic hagfish (*Myxine glutinosa*). The genes (SEQ ID NO: 1 and 3, Figure 1) encode two proteins (SEQ ID NO: 2 and 4, Figure 1) which each comprise an N-terminal signal sequence, a cathelin-like sequence and a C-terminal antimicrobial peptide. The antimicrobial peptide sequences deduced

from these genes, GWFKKAWRKVKHAGRRVLDTAKGVGRHYLNNWLNRYRG (SEQ ID NO: 5) and GWFKKAWRKVKNAGRVLKGVGIHYGVGLIG (SEQ ID NO: 6), are highly similar yet distinct from the hagfish peptides described in U.S. Patent 5,734,015 (Figure 4). Thus, the invention includes the isolated hagfish genes and their encoded cathelin-associated antibiotic peptides.

Included in this invention is the antibiotic peptides wherein the C-terminal glycine residue has been post-translationally transformed into a C-terminal amide residue (SEQ ID NO: 11 and 12) by an enzyme such as peptidylglycine α-amidating monooxygenase (PAM) (Prigge *et al.* (2000) Cell. Mol. Life Sci. 57, 1236-1259) and/or one or more of the Trp residues have been post-translationally transformed into Br-Trp residues by a brominating enzyme (Shinnar *et al.* (2000) FASEB Journal 14, A1488).

Also included in the invention is the family of hagfish cathelin-associated genes, their encoded proteins and their antibiotic peptides. Another embodiment is methods for the expression of these genes and cells that contain them. The recombinant production of the proteins encoded by the hagfish cathelin-associated genes is included in this invention, as is the production of the associated antibiotic peptides.

Definitions

As used herein, the term "antimicrobial" refers to the ability of a compound to inhibit or irreversibly prevent the growth of a microorganism. Such inhibition or prevention can be through a microbicidal action or microbistatic inhibition. Therefore, the term "microbicidal inhibition" as used herein, refers to the ability of a compound to kill, or irreversibly damage the target organism. The term "microbistatic inhibition" as used herein, refers to the ability of the antimicrobial compound to inhibit the growth of the target microorganism without death. Microbicidal or microbistatic inhibition of microorganisms in an environment presently exhibiting microbial growth (therapeutic treatment) or an environment at risk of supporting microbial growth (prevention) is included within this definition.

As used herein, a "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many generations.

A DNA "coding sequence" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

As used herein, a "nucleic acid molecule" refers to the polymeric form of deoxyribonucleotides (adenine, guanine, thymine, and/or cytosine) or ribonucleotides (adenine, guanine, uracil, and/or cytosine) and may include in either its single stranded form, or in double-stranded helix as well as RNA. This term refers only to the primary and secondary structure of the molecule and is not limited to any particular tertiary form. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (*e.g.*, the strand having a sequence homologous to the mRNA). Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, polyadenylation signals, terminators, and the like, that provide for the expression of a coding sequence in a host cell.

As used herein, a "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded (inclusively) at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site, as well as protein binding domains responsible for the binding of RNA polymerase. Eukaryotic promoters will often, but not always, contain "TATA" boxes and "CAT" boxes.

A "signal sequence" can be included before the coding sequence or the native amino acid signal sequence from the envelope protein may be used. This sequence encodes a signal peptide, N-terminal to the polypeptide, that communicates to the host cell to direct the polypeptide to the cell surface or secrete the polypeptide into the media. This signal peptide is clipped off by the host cell before the protein leaves the cell. Signal sequences can be found associated with a variety of proteins native to prokaryotes and eukaryotes. For instance, alpha-factor, a native yeast protein, is secreted from yeast, and its signal sequence can be attached to heterologous proteins to be secreted into the media (see *e.g.*, U.S. Patent 4,546,082 and EP 0116201). Further, the alpha-factor and its analogs have been found to secrete heterologous proteins from a variety of yeast, such as *Saccharomyces* and *Kluyveromyces* (EP 88312306.9; EP 0324274 publication, and EP 0301669). An example for use in mammalian cells is the tPA signal used for expressing Factor VIIIc light chain.

A cell has been "transformed" by exogenous or heterologous DNA when such DNA as been introduced inside the cell. The transforming DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes,

for example, the transforming DNA may be maintained on an episomal element such as a plasmid or viral vector. With respect to eukaryotic cells, a stably transformed cell is one in which the transforming DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA.

A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then translated into the protein encoded by the coding sequence.

As used herein, a "vector" is a replicon, such as plasmid, virus, phage or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment.

Nucleic Acid Molecules

The two cDNA sequences (Figure 1) consist of 848 (SEQ ID NO: 1) and 863 (SEQ ID NO: 3) nucleotides. The sequence for SEQ ID NO: 1 includes the initiation codon and 3 additional 5'-nucleotides; the sequence for SEQ ID NO: 3, judged by alignment with SEQ ID NO: 1, has eight coding nucleotides missing at the 5'-end. In the alignment of Figure 1, 814 nucleotides are comparable between the two sequences, with 785 matching (96.4%). The nucleotides encoding the putative signal and cathelin sequences of the two are very similar, with 69 of 70 signal nucleotides identical (98.6%) and 332 of 345 (96.2%) cathelin nucleotides matching; 401 of 415 (96.6%) match overall. Neither insertions nor deletions are evident in these parts of the cDNA sequences.

The peptide and 3'-untranslated sequences, in contrast, show evidence of several insertion/deletion changes, which makes comparison of their similarities difficult. In the peptide-encoding sequence, SEQ ID NO: 3 has three gaps (of 3, 9, and 4 nucleotides) in comparison to SEQ ID NO: 1; the last gap results in a frame-shift for 17 nucleotides just before the termination codon. The peptide-encoding sequence of SEQ ID NO: 1 has a gap of 21 nucleotides with respect to the sequence for SEQ ID NO: 3; as a result, a new termination codon 26 nucleotides farther along the sequence marks the end of the peptide for SEQ ID NO: 1. In this alignment, the last 10 coding nucleotides for SEQ ID NO: 1 correspond to 10 nucleotides in the 3'-untranslated region of the SEQ ID NO: 3 cDNA. Overall, 98 nucleotides encoding peptide can be compared; 89 of these (90.8%) match.

The 3'-untranslated ends of both cDNAs have -tgg- repeats, a 10-mer for SEQ ID NO: 1, and a 17-mer for SEQ ID NO: 3. Excluding the -tgg- repeats, 248 nucleotides in the 3'-end can be compared; of these, 243 (98.0%) match.

The nucleic acid molecules of the present invention include nucleic acid molecules that encode the proteins having SEQ ID NO: 2, SEQ ID NO: 4 and the related proteins herein described, preferably in isolated form. As used herein, "nucleic acid" is defined as RNA or DNA that encodes a protein or peptide as defined above, or is complementary to nucleic acid sequence encoding such peptides, or hybridizes to such nucleic acid and remains stably bound to it under appropriate stringency conditions, or encodes a polypeptide sharing at least 55% sequence identity, 85 % sequence identity, preferably at least 85%, and more preferably at least 90%, most preferably at least 95 to 99% sequence identity with the peptide sequences. Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as nucleic acids based on alternative backbones or including alternative bases whether derived from natural sources or synthesized. Such hybridizing or complementary nucleic acids, however, are defined further as being novel and unobvious over any prior art nucleic acid including that which encodes, hybridizes under appropriate stringency conditions, or is complementary to nucleic acid encoding a protein according to the present invention.

A preferred embodiment of the present invention is a nucleic acid molecule that encodes a hagfish cathelin-associated protein of the invention, such as a nucleic acid molecule comprising, consisting essentially of or consisting of SEQ ID NO: 1, or the open reading frame defined by nucleotides 4-540 (or 543) of SEQ ID NO: 1; or such as a nucleic acid molecule comprising, consisting essentially of or consisting of SEQ ID NO: 3, or the open reading frame defined by nucleotides 2-505 (or 508) of SEQ ID NO: 3. A further preferred embodiment of the present invention is a nucleic acid molecule that encodes a hagfish antimicrobial peptide of the invention, such as a nucleic acid molecule comprising, consisting essentially of or consisting of nucleotides 427-540 of SEQ ID NO: 1, encoding amino acids 142-179 of SEQ ID NO: 2, or encoding SEQ ID NO: 5; or such as a nucleic acid molecule comprising, consisting essentially of or consisting of nucleotides 416-505 of SEQ ID NO: 3, encoding amino acids 139-168 of SEQ ID NO: 4, or encoding SEQ ID NO: 6.

Homology or sequence identity at the nucleotide or amino acid sequence level is determined by **BLAST** (**B**asic **L**ocal **A**lignment **S**earch **T**ool) analysis using the algorithm employed by the programs **blastp, blastn, blastx, tblastn** and **tblastx** (Altschul *et al.* (1997) Nucleic Acids Res. 25, 3389-3402 and Karlin *et al.* (1990) Proc. Natl. Acad. Sci. USA 87, 2264-2268, both fully incorporated by reference) which are tailored for sequence similarity searching. The approach used by the **BLAST** program is to first consider similar segments,

with gaps (non-contiguous) and without gaps (contiguous), between a query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only those matches which satisfy a preselected threshold of significance. For a discussion of basic issues in similarity searching of sequence databases, see Altschul *et al.* (1994) Nature Genetics 6, 119-129 which is fully incorporated by reference. The search parameters for histogram, descriptions, alignments, expect (*i.e.*, the statistical significance threshold for reporting matches against database sequences), cutoff, matrix and filter (low complexity) are at the default settings. The default scoring matrix used by blastp, blastx, tblastn, and tblastx is the **BLOSUM62** matrix (Henikoff *et al.* (1992) Proc. Natl. Acad. Sci. USA 89, 10915-10919, fully incorporated by reference), recommended for query sequences over 85 nucleotides or amino acids in length.

For **blastn**, the scoring matrix is set by the ratios of **M** (*i.e.*, the reward score for a pair of matching residues) to **N** (*i.e.*, the penalty score for mismatching residues), wherein the default values for **M** and **N** are +5 and -4, respectively. Four **blastn** parameters were adjusted as follows: Q=10 (gap creation penalty); R=10 (gap extension penalty); wink=1 (generates word hits at every winkth position along the query); and gapw=16 (sets the window width within which gapped alignments are generated). The equivalent **Blastp** parameter settings were Q=9; R=2; wink=1; and gapw=32. A **Bestfit** comparison between sequences, available in the GCG package version 10.0, uses DNA parameters GAP=50 (gap creation penalty) and LEN=3 (gap extension penalty) and the equivalent settings in protein comparisons are GAP=8 and LEN=2.

"Stringent conditions" are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50°C, or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer (pH 6.5) with 750 mM NaCl, 75 mM sodium citrate at 42°C. Another example is hybridization in 50% formamide, 5× SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5× Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2× SSC and 0.1% SDS. A skilled artisan can readily determine and vary the stringency conditions appropriately to obtain a clear and detectable hybridization signal. Preferred nucleic acid molecules are those that hybridize under the above conditions to the complement of SEQ ID NO: 1 or 3 and which encode a polypeptide with antimicrobial activity. Even more preferred nucleic acid molecules are those

that hybridize under the above conditions to the complement strand of the open reading frame of SEQ ID NO: 1 or 3.

As used herein, an "isolated" nucleic acid molecule(s) means a nucleic acid molecule, DNA or RNA, which has been removed from its native environment or when the nucleic acid molecule is substantially separated from contaminant nucleic acid encoding other polypeptides from the source of nucleic acid. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically. Isolated nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) shown in SEQ ID NO: 1 or 3; DNA molecules comprising the coding sequence for the mature antimicrobial peptide protein; and DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode the antimicrobial peptide protein. As the genetic code is well known in the art, it would be routine for one skilled in the art to generate such degenerate variants. In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention. The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the antimicrobial peptide protein. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism.

The present invention further provides fragments of the encoding nucleic acid molecule. As used herein, a fragment of an encoding nucleic acid molecule refers to a small portion of the entire protein encoding sequence. The size of the fragment will be determined by the intended use. For example, if the fragment is chosen so as to encode an active portion of the protein, the fragment will need to be large enough to encode the functional region(s) of the protein or may encode regions of homology between the hagfish proteins in Figure 1 (SEQ ID NO: 2 and 4). In one embodiment of the invention an antimicrobial peptide fragment will be preferably at least 15-20 amino acids, more preferably 20-30 amino acids and most preferably 30-40 amino acids. If the fragment is to be used as a nucleic acid probe or PCR primer, then the

fragment length is chosen so as to obtain a relatively small number of false positives during probing/priming.

Fragments of the encoding nucleic acid molecules of the present invention (*i.e.*, synthetic oligonucleotides) that are used as probes or specific primers for the polymerase chain reaction (PCR), or to synthesize gene sequences encoding proteins of the invention can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci *et al.* (1981) J. Am. Chem. Soc. 103, 3185-3191 or using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define various modular segments of the gene, followed by ligation of oligonucleotides to build the complete modified gene.

The encoding nucleic acid molecules of the present invention may further be modified so as to contain a detectable label for diagnostic and probe purposes. A variety of such labels are known in the art and can readily be employed with the encoding molecules herein described. Suitable labels include, but are not limited to, biotin, radiolabeled nucleotides and the like. A skilled artisan can employ any of the art known labels to obtain a labeled encoding nucleic acid molecule.

Modifications to the primary structure itself by deletion, addition, or alteration of the amino acids incorporated into the protein sequence during translation can be made without destroying the activity of the protein. Such substitutions or other alterations result in proteins having an amino acid sequence encoded by a nucleic acid falling within the contemplated scope of the present invention.

Isolation of Other Related Nucleic Acid Molecules

As described above, the identification of the hagfish nucleic acid molecules having SEQ ID NO: 1 and 3 allows a skilled artisan to isolate nucleic acid molecules that encode other members of the cathelin-associated protein family in addition to the sequences herein described.

Essentially, a skilled artisan can readily use the amino acid sequence of SEQ ID NO: 2 or 4 to generate antibody probes to screen expression libraries prepared from appropriate cells. Typically, polyclonal antiserum from mammals such as rabbits immunized with the purified protein (as described below) or monoclonal antibodies can be used to probe a mammalian cDNA or genomic expression library, such as a λgtll library, to obtain the appropriate coding sequence for other members of the protein family. The cloned cDNA sequence can be expressed as a fusion protein, expressed directly using its own control sequences, or expressed

by constructions using control sequences appropriate to the particular host used for expression of the protein.

Alternatively, a portion of the coding sequence herein described can be synthesized and used as a probe to retrieve DNA encoding a member of the hagfish cathelin-associated protein family from any mammalian organism. Oligomers containing approximately 18-20 nucleotides (encoding about a 6-7 amino acid stretch) are prepared and used to screen genomic DNA or cDNA libraries to obtain hybridization under stringent conditions or conditions of sufficient stringency to eliminate an undue level of false positives.

Additionally, pairs of oligonucleotide primers can be prepared for use in a polymerase chain reaction (PCR) to selectively clone an encoding nucleic acid molecule. A PCR denature/anneal/extend cycle for using such PCR primers is well known in the art and can readily be adapted for use in isolating other encoding nucleic acid molecules.

Recombinant Nucleic Acid Molecules

The present invention further provides recombinant DNA molecules (rDNA) that contain a hagfish cathelin-associated protein coding sequence. As used herein, an rDNA molecule is a DNA molecule that has been subjected to molecular manipulation *in vitro*. Methods for generating rDNA molecules are well known in the art (see, *e.g.*, Sambrook *et al.* (2001) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press). In the preferred rDNA molecules, a coding DNA sequence is operably linked to expression control sequences and/or vector sequences.

The choice of vector and/or expression control sequences to which one of the protein family encoding sequences of the present invention is operably linked depends directly, as is well known in the art, on the functional properties desired, *e.g.*, protein expression, and the host cell to be transformed. A vector contemplated by the present invention is at least capable of directing the replication or insertion into the host chromosome, and preferably also expression, of the structural gene included in the rDNA molecule.

Expression control elements that are used for regulating the expression of an operably linked protein encoding sequence are known in the art and include, but are not limited to, inducible promoters, constitutive promoters, secretion signals, and other regulatory elements. Preferably, the inducible promoter is readily controlled, such as being responsive to a nutrient in the host cell's medium.

In one embodiment, the vector containing a coding nucleic acid molecule will include a prokaryotic replicon, *i.e.*, a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extrachromosomally in a prokaryotic host

cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, vectors that include a prokaryotic replicon may also include a gene whose expression confers a detectable marker such as a drug resistance. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

Vectors that include a prokaryotic replicon can further include a prokaryotic or bacteriophage promoter capable of directing the expression (transcription and translation) of the coding gene sequences in a bacterial host cell, such as *E. coli*. A promoter is an expression control element formed by a DNA sequence that permits RNA polymerase to bind and transcription to occur. Promoter sequences compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 (Bio-Rad Laboratories), pPL and pKK223 (Pharmacia).

Expression vectors compatible with eukaryotic cells, preferably those compatible with vertebrate cells, can also be used to form an rDNA molecules that contains a coding sequence. Eukaryotic cell expression vectors are well known in the art and are available from several commercial sources. Typically, such vectors are provided containing convenient restriction sites for insertion of the desired DNA segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies), pTDT1 (ATCC), the vector pCDM8 described herein, and the like eukaryotic expression vectors.

Eukaryotic cell expression vectors used to construct the rDNA molecules of the present invention may further include a selectable marker that is effective in an eukaryotic cell, preferably a drug resistance selection marker. A preferred drug resistance marker is the gene whose expression results in neomycin resistance, i.e., the neomycin phosphotransferase (*neo*) gene (Southern, P.J. *et al.* 1982. J Mol Appl Genet 1(4): 327-41). Alternatively, the selectable marker can be present on a separate plasmid, and the two vectors are introduced by co-transfection of the host cell, and selected by culturing in the appropriate drug for the selectable marker.

Host Cells Containing an Exogenously Supplied Coding Nucleic Acid Molecule

The present invention further provides host cells transformed with a nucleic acid molecule that encodes a hagfish cathelin-associated protein, preferably SEQ ID NO: 1 or 3, of the present invention. The host cell can be either prokaryotic or eukaryotic. Eukaryotic cells useful for expression of a protein of the invention are not limited, so long as the cell line is compatible with cell culture methods and compatible with the propagation of the expression vector and expression of the gene product. Preferred eukaryotic host cells include, but are not

limited to, yeast, insect, and mammalian cells, preferably those from a mouse, rat, monkey or human cell line. Preferred eukaryotic host cells include cell lines established from hagfish cells, Chinese hamster ovary (CHO) cells available from the ATCC as CCL61, NIH Swiss mouse embryo cells NIH/3T3 available from the ATCC as CRL 1658, baby hamster kidney cells (BHK), and the like eukaryotic tissue culture cell lines.

Any prokaryotic host can be used to express an rDNA molecule encoding a protein of the invention. The preferred prokaryotic host is *E. coli*.

Transformation of appropriate cell hosts with an rDNA molecule of the present invention is accomplished by well known methods that typically depend on the type of vector used and host system employed. With regard to transformation of prokaryotic host cells, electroporation and salt treatment methods are typically employed (see, *e.g.*, Cohen *et al.* (1972) Proc. Natl. Acad. Sci. USA 69, 2110-2112 and Sambrook *et al.* (2001) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press). With regard to transformation of vertebrate cells with vectors containing rDNAs, electroporation, cationic, lipid or salt treatment methods are typically employed (see, *e.g.*, Graham *et al.* (1973) Virol. 52, 456-458; Wigler *et al.* (1979) Proc. Natl. Acad. Sci. USA 76, 1373-1376).

Successfully transformed cells (*i.e.*, cells that contain an rDNA molecule of the present invention) can be identified by well known techniques including the selection for a selectable marker. For example, cells resulting from the introduction of an rDNA of the present invention can be cloned to produce single colonies. Cells from those colonies can be harvested, lysed and their DNA content examined for the presence of the rDNA using a method such as that described by Southern (1975) J. Mol. Biol. 98, 503-506) or Berent *et al.* (1985) Biotech. 3, 208-209) or the proteins produced from the cell assayed via an immunological method.

Production of Recombinant Proteins

The present invention further provides methods for producing a hagfish cathelin-associated protein of the invention using nucleic acid molecules herein described. In general terms, the production of a recombinant form of a protein typically involves the following steps:

First, a nucleic acid molecule is obtained that encodes a hagfish cathelin-associated protein of the invention, such as a nucleic acid molecule comprising, consisting essentially of or consisting of SEQ ID NO: 1, or the open reading frame defined by nucleotides 4-540 (or 543) of SEQ ID NO: 1; or such as a nucleic acid molecule comprising, consisting essentially of or consisting of SEQ ID NO: 3, or the open reading frame defined by nucleotides 2-505 (or 508) of SEQ ID NO: 3. Alternatively, a nucleic acid molecule is obtained that encodes a hagfish antimicrobial peptide of the invention, such as a nucleic acid molecule comprising,

consisting essentially of or consisting of nucleotides 427-540 of SEQ ID NO: 1, encoding amino acid residues 142-179 of SEQ ID NO: 2, or encoding SEQ ID NO: 5; or such as a nucleic acid molecule comprising, consisting essentially of or consisting of nucleotides 416-505 of SEQ ID NO: 3, encoding amino acid residues 139-168 of SEQ ID NO: 4, or encoding SEQ ID NO: 6. If the encoding sequence is uninterrupted by introns, as is this open reading frame, it is directly suitable for expression in any host.

The nucleic acid molecule is then preferably placed in operable linkage with suitable control sequences, as described above, to form an expression unit containing the protein open reading frame. The expression unit is used to transform a suitable host and the transformed host is cultured under conditions that allow the production of the recombinant protein. Optionally the recombinant protein is isolated from the medium or from the cells; recovery and purification of the protein may not be necessary in some instances where some impurities may be tolerated.

Each of the foregoing steps can be done in a variety of ways. For example, the desired coding sequences may be obtained from genomic fragments and used directly in appropriate hosts. The construction of expression vectors that are operable in a variety of hosts is accomplished using appropriate replicons and control sequences, as set forth above. The control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene and were discussed in detail earlier. Suitable restriction sites can, if not normally available, be added to the ends of the coding sequence so as to provide an excisable gene to insert into these vectors. A skilled artisan can readily adapt any host/expression system known in the art for use with the nucleic acid molecules of the invention to produce recombinant protein.

In another embodiment, the polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art.

Hagfish cathelin-associated proteins

The amino acid sequences, SEQ ID NO: 2 and SEQ ID NO: 4 (Figure 1) differ mainly because of differences in the region of the peptide associated with antimicrobial activity. The signal sequences are identical, with 26 residues; and the cathelins share 91.3% of their 115 residues. In contrast, the antimicrobial peptide consisting of amino acid residues 142-179 of SEQ ID NO: 2, GWFKKAWRKVKHAGRRVLDTAKGVGRHYLNNWLNRYRG (SEQ ID NO: 5), has 38 residues compared to 30 for the antimicrobial peptide consisting of amino acid residues 139-168 from SEQ ID NO: 4, GWFKKAWRKVKNAGRVLKGVGIHYGVGLIG (SEQ ID NO: 6). When three gaps are inserted in SEQ ID NO: 6, however, 23 of the 30 residues (76%) are identical to those of SEQ ID NO: 5. Both SEQ ID NO: 5 and SEQ ID NO: 6, have a C-terminal glycine residue which is likely post-translationally converted to a C-terminal amide as is the case with the hagfish antimicrobial peptides described in U.S. Patent 5,734,015.

As used herein, the family of proteins related to the hagfish cathelin-associated proteins, refers to other cathelin-associated proteins that can be isolated from the hagfish or organisms closely related to the hagfish. The methods used to identify and isolate other members of the family of proteins related to the hagfish cathelin-associated proteins are described below.

The proteins of the present invention are preferably in isolated form. As used herein, a protein is said to be isolated when physical, mechanical or chemical methods are employed to remove the protein from cellular constituents that are normally associated with the protein. A skilled artisan can readily employ standard purification methods to obtain an isolated protein. By isolated, it is further meant that the hagfish cathelin-associated protein or the antimicrobial peptide of the present invention is part of a mixture with other components and wherein the protein or peptide of the present invention comprises at least about 10% of the total protein in the mixture, preferably at least about 20% of the total protein in the mixture, more preferably at least about 30% of the total protein in the mixture, even more preferably at least about 40% of the total protein in the mixture, still more preferably at least about 50% of the total protein in the mixture, yet more preferably at least about 60% of the total protein in the mixture, even still more preferably at least about 70% of the total protein in the mixture, yet still more preferably at least about 80% of the total protein in the mixture, much more preferably at least about 90% of the total protein in the mixture, still much more preferably at least about 95% of the total protein in the mixture, and most preferably at least about 99% of the total protein in the mixture.

The proteins of the present invention further include conservative variants of the proteins herein described. As used herein, a conservative variant refers to alterations in the amino acid sequence that do not adversely affect the biological functions of the protein. A substitution, insertion or deletion is said to adversely affect the protein when the altered sequence prevents or disrupts a biological function associated with the protein. For example, the overall charge, structure or hydrophobic/hydrophilic properties of the protein may be altered without adversely affecting a biological activity. Accordingly, the amino acid sequence can be altered, for example to render the peptide more hydrophobic or hydrophilic, without adversely affecting the biological activities of the protein.

Ordinarily, the allelic variants, the conservative substitution variants, and the members of the protein family, will have an amino acid sequence having at least about 85% amino acid sequence identity with the sequences set forth in SEQ ID NO: 2, 4, 5 or 6, more preferably at least 90%, even more preferably at least 95% and most preferably at least 99%. Identity or homology with respect to such sequences is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the known peptides, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity. N-terminal, C-terminal or internal extensions, deletions, or insertions into the peptide sequence shall not be construed as affecting homology.

Thus, the proteins of the present invention include molecules comprising, consisting essentially of, or consisting of the amino acid sequence disclosed in SEQ ID NO: 2, 4, 5 and 6; fragments thereof having a consecutive sequence of at least about 3, 4, 5, 6, 10, 15, 20, 25, 30, 35, 38 or more amino acid residues of the hagfish cathelin-associated protein or antimicrobial peptide thereof; amino acid sequence variants of such sequence wherein an amino acid residue has been inserted N- or C-terminal to, or within, the disclosed sequence; and amino acid sequence variants of the disclosed sequence, or their fragments as defined above, that have been substituted by another residue. Contemplated variants further include those containing predetermined mutations by, *e.g.*, homologous recombination, site-directed or PCR mutagenesis, and the corresponding proteins of other animal species, including but not limited to rabbit, rat, porcine, bovine, ovine, equine and non-human primate species, and the alleles or other naturally occurring variants of the family of proteins; and derivatives wherein the protein has been covalently modified by substitution, chemical, enzymatic, or other appropriate means with a moiety other than a naturally occurring amino acid (for example a detectable moiety such as an enzyme or radioisotope).

Therapeutic and additive applications

Another aspect of the present invention is a method for inhibiting microbial growth. Hagfish cathelin-associated antimicrobial peptides of the present invention, exemplified by SEQ ID NOs: 5 and 6, can be used to inhibit microbial growth under various circumstances. For example, the peptides of the present invention can be administered therapeutically to treat or prevent disease in an individual resulting from a microbial infection. A variety of microbial infections can be inhibited by treatment with the peptides of the present invention. For example, the antimicrobial peptides of the present invention can be used as an antimicrobial agent for the treatment of fungal or bacterial infections.

In preferred embodiments, the microbial infection is a bacterial infection. In particular, therapeutic administration of the peptides of the present invention inhibits and can reverse the progression of a bacterial infection. The development and spread of new resistant strains of bacteria are increasingly posing a public health threat. The antimicrobial peptides of the present invention are useful in treating patients suffering from these resistant strains of microbes which currently threaten public health. Many bacterial infections cause bacterial sepsis in the infected individual. Administration of the peptides to an individual suffering from a septic bacterial infection produces a significant therapeutic effect, by reducing both the sepsis caused by the infection, and the infection itself.

An individual suitable for treatment is any animal (mammal or otherwise) which is afflicted with or otherwise susceptible to one or more of the above described microbial infections. In a preferred embodiment, the individual is a human. In another embodiment, the individual is a livestock animal. In another embodiment, the animal is a show animal or a household pet.

Administration of the hagfish cathelin-associated antimicrobial peptides of the present invention to the individual is either systemic or localized, and is largely determined by the specific infection being treated. Systemic administration can be accomplished by several routes, including but not limited to intravenous administration, inhalation, mucosal, and ingestion. Localized administration can be topical or internal. Such administration can be accomplished by several routes, including, but not limited to, subcutaneous, dermal, intradermal, buccal, mucosal, intraperitoneal, vaginal, inhalation, and ingestion.

It will often be of use to administer a formulation of the peptides of the present invention which includes a pharmaceutically acceptable carrier. Possible formulations for therapeutic administration include a variety of pharmaceutical compositions, the appropriate use of which will depend upon the route of administration deemed necessary for treatment. Some useful formulations for topical administration are, for example, eyedrops, eardrops, or

gingival applications (*e.g.*, drops, mouthwash, cream, or paste). The regimen of administration (*e.g.*, route, dose, and course) which is therapeutic to the patient will vary with the individual to be treated (*e.g.*, health, weight, metabolism), the site of infection, and the infecting pathogen. A therapeutic regimen should be developed by extrapolation from treatment with similar therapeutics in combination with empirical observation. Administration of the peptides of the present invention to prevent a microbial infection in an individual parallel the above described methods.

As a general proposition, the total pharmaceutically effective amount of antimicrobial peptide administered parenterally per dose will be in the range of about 0.001 mg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the peptide. If given continuously, the antimicrobial peptide is typically administered at a dose rate of about 0.001 mg/kg/hour to about 0.050 mg/kg/hour, either by one to four injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed.

Hagfish cathelin-associated antimicrobial peptides of the present invention, exemplified by SEQ ID NO: 5 and 6, are also useful for potentiating the therapeutic action of other antimicrobial drugs or agents. Co-administration of other antimicrobial agents with the antimicrobial polypeptides of the present invention, produce a synergistic antibiotic effect. Co-administration of the peptides of the present invention with an antimicrobial agent enables therapeutic treatment of a patient with lower doses of the antimicrobial agent. Lower doses are preferable in situations such as when treating with an expensive drug, or one that produces undesired side effects, or one whose short half-life *in vivo* would otherwise rapidly reduce its concentration below that which is required for it to be efficacious. In addition, co-administration with antimicrobial agents or drugs may also allow for a shorter therapy period and/or the reversal of resistant phenotypes. Without limitation, microbes which resist an antimicrobial drug by decreasing their internal drug concentration (*e.g.*, with decreased membrane permeability or increased cellular export or metabolism of the drug) are expected to be especially susceptible to the potentiating activity of these peptides.

The regimen of administration of the antimicrobial drug and the hagfish cathelin-associated antimicrobial peptide or functional variant of the present invention varies with the patient and the particular infection, and can be determined by one of skill in the art on a case by case basis. Formulations of the polypeptide will depend upon the regimen of administration, examples described above.

The polypeptides of the present invention can be used not only as pharmaceutical and neutraceutical agents but also as additives for any products such as foods and medicinal or non-medicinal products which are taken into the bodies or otherwise applied onto or contacted with the body surface of humans or other animals or fluids, organs, and cells derived therefrom.

The present method is useful for treating a variety of products. Biological products, defined herein as products which are derived from biological organisms or processes, are particularly at risk for contamination with microorganisms. Examples of biological products include without limitation, food products, tissue, living cells, products derived from living cells, blood or components thereof, as any other bodily fluid, drugs or other molecular preparations. Non-biological products, defined herein as a product not directly derived from a biological organism or process, for example glassware, surgical equipment, synthetic drugs or other molecular preparations, can also be treated. For effective use of the present invention, the product which is to be treated should not possess an activity which completely inactivates the antimicrobial activity of all of the peptide quantity so applied. The peptides of the present invention can be added, assorted to, sprayed to, adhered to, coated onto, adsorbed to, chemically crosslinked to, or impregnated into any products which are generally desired to be prevented or inhibited from contamination by proliferating microorganisms. Alternatively, the peptide of the present invention can be immobilized on a surface over which a product is passed to remove or inhibit microbial growth in the product. A product which is treated with and retains antimicrobial peptide can further be used to treat another product with which it is contacted.

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the present invention and practice the claimed methods. The following working examples therefore, specifically point out the preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

Example 1

Isolation of genes for hagfish cathelin-associated antibiotic peptides

Adult Atlantic hagfish (*Myxine glutinosa* Linnaeus, 1758; Myxinidae) were purchased from Huntsman Marine Science Centre (New Brunswick, Canada). The fish were anesthetized in a slurry of ice and decapitated. Intestines were removed and cleaned of contents, frozen in liquid N_2, and stored at -80°C.

Tissue was ground in liquid nitrogen and transferred to a guanidinium isothiocyanate solution (Chirgwin *et al.* (1979) Biochemistry 18, 5294-5299), which inhibits RNase activity

while disrupting cells. Total RNA was isolated by centrifugation through a CsCl step gradient; polyadenylated RNA was recovered by means of oligo(dT) cellulose (Stratagene). Reverse transcribed cDNA was ligated first to adaptor arms and then into phage Lambda Zap II following Stratagene protocols.

An twenty-mer oligonucleotide probe 5'-AARAARGCNTGGMGNAARGT-3' (SEQ ID NO: 11; 256-fold degenerate) was designed based on amino acid residues 4 to 10 of the peptides previously isolated from hagfish intestinal tissue. Phage particles containing homologous sequences were isolated by plating a small number of particles on a lawn of XL1 Blue MRF (Stratagene). The cDNA inserts were isolated in phagmid pBluescript by *in vivo* excision. Two inserts were sequenced in both directions using Sanger's dideoxy method (Barnes *et al.* (1983) Nucleic Acids Res. 11, 349-368).

Homologous sequences were identified using BLASTX (Altschul *et al.* (1997) Nucleic Acids Res. 25, 3389-3402) searches of GenBank. Homologous sequences were aligned using ESEE (Cabot *et al.* (1989) Compu. Appl. Biosci. 5, 233-234).

For *in situ* hybridizations, plasmid DNA containing insert 31-52 from SEQ ID NO: 1 was linearized with *StyI*. Both sense and antisense probes were labeled enzymatically with ^{32}P and detected by autoradiography.

Example 2

Comparison of hagfish cathelin-associated sequences with other cathelin-associated sequences

Searches of GenBank using BLASTN with default values yielded no matches either when just the coding region for the two peptides were used as queries or when the entire coding regions were used. Using the entire coding sequence available for each protein and BLASTX, which searches for more distant relationships based on identical and conserved amino acids, however, yielded a large number of matches. Those in the correct reading frame included especially cDNA sequences for prepropeptides of mammalian antibiotic peptides containing a cathelin-like acidic spacer.

Among the matches found, the sequence for *Capra* MAP28, isolated from bone marrow cells of *Capra hircus* (goat) was unique in showing at least a weak match between the antibiotic peptide and one of the two *Myxine* peptides. A plausible alignment of the coding regions of one of the hagfish sequences with *Capra* MAP28 is shown in Figure 2.

Example 3

Antimicrobial activity of cathelin-associated antibiotic peptides

The antimicrobial activity of cathelin-associated antibiotic peptides is determined by the following antibacterial assay based upon the guidelines of the National Committee for Clinical Laboratory Standards (Document M7-T2, (1988) Volume 8, No. 8).

A stock solution of the test peptide is prepared at a concentration of 512 μg/ml in sterile deionized distilled water and stored at -70°C. The stock peptide solutions are diluted in serial dilutions (1:2) down the wells of a microtiter plate so that the final concentrations of peptide in the well are 0.25, 0.50, 1, 2, 4, 8, 16, 32, 64, 128 and 256 μg/ml. 1.5×10^5 CFU/ml of the test microbe is added to the wells in full strength Mueller Hinton broth (BBL 11443) from a mid-log culture. The inoculum is standardized spectrophotometrically at 600 nm and is verified by colony counts. The plates are incubated for 16-24 hours at 37°C and the minimal inhibitory concentration (MIC) for each peptide is determined. Minimal inhibitory concentration is defined as the lowest concentration of peptide which produces a clear well in the microtiter plate.

While the invention has been described and illustrated herein by references to various specific materials, procedures and examples, it is understood that the invention is not restricted to the particular combinations of material and procedures selected for that purpose. Numerous variations of such details can be implied as will be appreciated by those skilled in the art. It is intended that the specification and examples be considered as exemplary, only, with the true scope and spirit of the invention being indicated by the following claims. All publications, patents and patent applications referred to in this application are herein incorporated by reference in their entirety.

What is claimed:

1. An isolated nucleic acid molecule selected from the group consisting of:

(a) an isolated nucleic acid molecule that encodes the amino acid sequence of a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 2 or 4; and

(b) an isolated nucleic acid molecule, which hybridizes to the complement of a nucleic acid molecule comprising a nucleotide sequence encoding a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 2 or 4 under stringent conditions.

2. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule is selected from the group consisting of SEQ ID NO: 1 and 3.

3. The isolated nucleic acid molecule of claim 2, wherein the nucleic acid molecule comprises nucleotides 4 to 540 of SEQ ID NO: 1.

4. The isolated nucleic acid molecule of claim 2, wherein the nucleic acid molecule comprises nucleotides 427 to 540 of SEQ ID NO: 1.

5. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 2 to 505 of SEQ ID NO: 3.

6. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule comprises nucleotides 416 to 505 of SEQ ID NO: 3.

7. An isolated nucleic acid molecule selected from the group consisting of:

(a) an isolated nucleic acid molecule that encodes the amino acid sequence of a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 5 or 6; and

(b) an isolated nucleic acid molecule, which hybridizes to the complement of a nucleic acid molecule comprising a nucleotide sequence encoding a *Myxine glutinosa* cathelin-associated antimicrobial peptide comprising SEQ ID NO: 5 or 6 under stringent conditions.

8. The isolated nucleic acid molecule of any one of claims 1-7, wherein said nucleic acid molecule is operably linked to one or more expression control elements.

9. A vector comprising an isolated nucleic acid molecule of any one of claims 1-7.

10. A host cell comprising a vector of claim 9.

11. A host cell transformed to contain the nucleic acid molecule of any one of claims 1-7.

12. A method for producing a peptide or fragment thereof comprising the step of culturing a host cell of claim 11 under conditions in which the protein or protein fragment encoded by said nucleic acid molecule is expressed.

13. The method of claim 12, wherein said host cell is selected from the group consisting of prokaryotic hosts and eukaryotic hosts.

14. An isolated peptide or fragment thereof produced by the method of claim 12.

15. An isolated peptide or fragment thereof selected from the group consisting of :

(a) an isolated peptide comprising the amino acid sequences depicted in SEQ ID NO: 2, 4; 5 or 6;

(b) an isolated peptide fragment comprising at least six amino acids of any of the sequences depicted in SEQ ID NO: 2, 4, 5 or 6;

(c) an isolated peptide comprising conservative amino acid substitutions of any of the sequences depicted in SEQ ID NO: 2, 4, 5 or 6; and

(d) naturally occurring amino acid sequence peptide variants of any of the amino acid sequences depicted in SEQ ID NO: 2, 4, 5 or 6.

16. The isolated protein fragment of claim 15 wherein the isolated peptide consists essentially of the amino acid sequences depicted in SEQ ID NO: 2, 4, 5 or 6.

17. A antimicrobial composition comprising one or more of the peptides of claim 15

18. The antimicrobial composition of claim 17, wherein the composition is suitable for topical or parenteral administration.

19. A method of treating or preventing a microbial infection in a mammal, comprising administering to the mammal an effective amount of a *Myxine glutinosa* cathelin-associated antimicrobial peptide.

20. The method of claim 19, wherein the *Myxine glutinosa* cathelin-associated antimicrobial peptide is selected from the group consisting of SEQ ID NO: 2, 4; 5 or 6.

21. The method of claim 19, wherein the mammal is a human.

22. The method of claim 19, wherein the microbial infection is selected from the group consisting of a bacterial infection and a fungal infection.

23. The method of claim 17, wherein the *Myxine glutinosa* cathelin-associated antimicrobial peptide is administered by a route selected from the group consisting of oral, topical and parenteral.

24. The method of claim 23, wherein the topical administration is accomplished by inhalation.

25. The method of claim 24, wherein the parenteral administration is accomplished by intravenous, subcutaneous or intramuscular administration.

Figure 1A

```
                >Signal sequence
SEQ ID NO. 1 gtcATGaagtccttgtgtgtccccgcggttctctctctggtcctgatcctgctcttggat
SEQ ID NO. 3           ...........a.....................................
SEQ ID NO. 2     M  K  S  L  C  V  P  A  V  L  S  L  V  L  I  L  L  L  D
SEQ ID NO. 4           .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

                              >Cathelin
SEQ ID NO. 1 caagcacctacagctcgggcagatgattctttgtccaaagagcaggtggagaatgcagtt
SEQ ID NO. 3 .....................................c.....................
SEQ ID NO. 2  Q  A  P  T  A  R  A  D  D  S  L  S  K  E  Q  V  E  N  A  V
SEQ ID NO. 4  .  .  .  .  .  .  .  .  .  .  .  N  .  .  .  .  .  .  .  .

SEQ ID NO. 1 gatgaagcgttggacaagctgaacaaacagcaggtgtccacacgtaaactggcattgtct
SEQ ID NO. 3 .........................t..............................
SEQ ID NO. 2  D  E  A  L  D  K  L  N  K  Q  Q  V  S  T  R  K  L  A  L  S
SEQ ID NO. 4  .  .  .  .  .  .  .  .  N  .  .  .  .  .  .  .  .  .  .  .

SEQ ID NO. 1 gaacaacaggatatccaggcagatgaaacagacgttgaaggacagttcaccatcaaatttt
SEQ ID NO. 3 ...........................................ac........t....
SEQ ID NO. 2  E  Q  Q  D  I  Q  A  D  E  T  D  V  E  G  Q  F  T  I  K  F
SEQ ID NO. 4  .  .  .  .  .  .  .  .  .  .  .  .  .  .  T  .  .  I  .

SEQ ID NO. 1 gacgtggttgagaccgagtgcaatgcggacgatcccagggactgggccgattgtccgatc
SEQ ID NO. 3 t...........................................................
SEQ ID NO. 2  D  V  V  E  T  E  C  N  A  D  D  P  R  D  W  A  D  C  P  I
SEQ ID NO. 4  Y  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

SEQ ID NO. 1 gcgacggactcgacaccggtcgatgcacagtgtgaggtcaccgtgttgagcaccgaggac
SEQ ID NO. 3 ............c......g.at.....................................
SEQ ID NO. 2  A  T  D  S  T  P  V  D  A  Q  C  E  V  T  V  L  S  T  E  D
SEQ ID NO. 4  .  .  .  .  P  .  G  I  .  .  .  .  .  .  .  .  .  .  .  .

SEQ ID NO. 1 tccttggacgtcggagacgcgacttgcgatttcaacaggacggacggaaatgcgaggcga
SEQ ID NO. 3 ..................................tc.....g.................
SEQ ID NO. 2  S  L  D  V  G  D  A  T  C  D  F  N  R  T  D  G  N  A  R  R
SEQ ID NO. 4  .  .  .  .  .  .  .  .  .  .  .  S  .  G  .  .  .  .  .  .

                >Antibiotic peptide
SEQ ID NO. 1 cgacgtggctggtttaagaaagcctggagaaaagtgaagcatgcgggacgacgagttctt
SEQ ID NO. 3 ...........................................a.....---......
SEQ ID NO. 2  R  R  G  W  F  K  K  A  W  R  K  V  K  H  A  G  R  R  V  L
SEQ ID NO. 4  .  .  .  .  .  .  .  .  .  .  .  .  .  N  .  .  . --- .  .
```

Figure 1B

```
SEQ ID NO. 1 gataccgcgaagggtgtgggaagacattatttgaataattggcttaatcg----------
SEQ ID NO. 3 ---------.............t.......----ggtg....a.......gtTAAtcggt

>3'UTR
SEQ ID NO. 2  D  T  A  K  G  V  G  R  H  Y  L  N  N  W  L  N  R ---------
SEQ ID NO. 4  -------- .  .  .  .  I  .  . ---- G  V  G  L  I  G  *

                                        >3'UTR
SEQ ID NO. 1 -----------ttatcgcggtTAGgaggaagctctgtgttggtggtggtggtggtggtgg
SEQ ID NO. 3' tggattaatcg.....a.....c........tc..c.......................
SEQ ID NO. 2 ------------ Y  R  G  *

SEQ ID NO. 1 tggtggtggtg----------------------------ttgtgggttaatgggtgggag
SEQ ID NO. 3 ...........gtggtggtggtggtggtgtgggggg..a..................

SEQ ID NO. 1 cgaatgatttctccccgacattcaatacttgctcgttaattccaattgattcgcttagtt
SEQ ID NO. 3 ............................................................

SEQ ID NO. 1 acacaagtaagccgaacaggtaaccagttagtaagctttaaaattgaacttttaggattg
SEQ ID NO. 3 ............................................................

SEQ ID NO. 1 aaactgcagctgaacttttgtagtttaacatttaattgtgactctgctattaaactttg
SEQ ID NO. 3 ............................................................

SEQ ID NO. 1 tctgatcatggaaataaaggattgcaaaaatgaaaaaaaaaaaaaaaaaaaaaaaaa
SEQ ID NO. 3 ..................................................
```

Figure 2

```
          Exon I→
CapraM28.n ATGgagacccagagggccagcctctccctgggacggtgctccctgtggctcctgctgctg 60
3          ~~~~~~~.tt.t.t.t.ccagcggtt..c---------..t...gtc..ga.c.....c 43
CapraM28.a  M  E  T  Q  R  A  S  L  S  L  G  R  C  S  L  W  L  L  L  L  20
4                L  C  V  P  A  V  .  -  -  -  .  .  V  .  I  .  .  14

CapraM28.n ggactagtgctgccctcggccagcgcccaggccctcagctacgggggaggccgttcttcat 120
3          ttggatcaagca..ta.a..tc.g..ag.t.attctttg.c.aac...ca.gt.gagaat 103
CapraM28.a  G  L  V  L  P  S  A  S  A  Q  A  L  S  Y  G  E  A  V  L  H  40
4           L  D  Q  A  .  T  .  R  .  D  D  S  L  S  N  .  Q  .  E  N  34

CapraM28.n gctgtcgat------------cgcatcaatgagcagtcctcagaagcgaatctctaccgc 168
3          ..a..t...gaagcgttggacaagc.g..ca.t.a.caggtgtcca.acg.aaactggca 163
CapraM28.a  A  V  D  -  -  -  -  R  I  N  E  Q  S  S  E  A  N  L  Y  R  56
4           .  .  .  E  A  L  D  K  L  .  N  .  Q  V  S  T  R  K  L  A  54

                                          Exon II→
CapraM28.n ctcctggagcttgacccgcctcccaaggac|gatgag---aatccaaacatcccgaaacct 225
3          t.gtctg.a---c.a.a.ga.at.c...ca|.....aacag.cgttg.agga.a..cca.c 220
CapraM28.a  L  L  E  L  D  P  P  P  K  D |D  E  -  N  P  N  I  P  K  P  75
4           .  S  .  -  Q  Q  D  I  Q  A |.  .  T  D  V  E  G  Q  T  T  73

CapraM28.n gtgagcttcagggtgaaggagactgtgtgccccaggacgagccggcagcccacggagcag 285
3          a.c.ta..ttac...gtt.....c.a....aatgc.gacgat.ccag.ga.tg..ccg.t 280
CapraM28.a  V  S  F  R  V  K  E  T  V  C  P  R  T  S  R  Q  P  T  E  Q  95
4           I  I  .  Y  .  V  .  .  E  .  N  A  D  D  P  R  D  W  A  D  93

                                    Exon III→
CapraM28.n tgtgacttcaaagagaatggg---|ctggtgaag---caatgtgtagggacagtcactctg 339
3          ...ccga..gcgac.g.ctc.cca|.c..gcattgca..g....ag-----.....cg.. 334
CapraM28.a  C  D  F  K  E  N  G  -  |L  V  K  -  Q  C  V  G  T  V  T  L  113
4           .  P  I  A  T  D  S  P |P  G  I  A  .  .  E  -  -  .  .  V  111

                                                                Exon IV→
CapraM28.n gatgcggtgaaaggcaaaatgaacatc---------acctgcgaagag|ttgcagagtgtt 390
3          ttgagcaccg.g.a.tcct..g..g..ggagacgcg..t.....tttc|aactc..cg.gc 394
CapraM28.a  D  A  V  K  G  K  M  N  I  -  -  -  T  C  E  E |L  Q  S  V  130
4           L  S  T  E  D  S  L  D  V  G  D  A  .  .  D  F |N  S  T  G  131

CapraM28.n gggagatttaaacgatttcgt---------aagaagctcaaaagactctggcacaaagtc 441
3          ..a.atgcg.gg...cga...ggctggttt.....agc.tgg...aaagt.a.g..t.cg 454
CapraM28.a  G  R  F  K  R  F  R  -  -  -  K  K  L  K  R  L  W  H  K  V  147
4           .  N  A  R  .  R  .  G  W  F  .  .  A  W  .  K  V  K  N  A  151


CapraM28.n ggcccattcgtt---ggcccgatactccattatgggTAAattgtgagcccatggaagaat 498
3          ..a.g.g.tc..aag..tgt.gg.a.a........tgttggat.a.t.ggt.AA       508
CapraM28.a  G  P  F  V  -  G  P  I  L  H  Y  G  *                     158
4           .  R  V  L  K  .  V  G  I  .  .  .  V  G  L  I  G  *        168
```

Figure 3

SEQ ID NO. 5 GWFKKAWRKVKHAGRRVLDTAKGVGRHYLNNWLNRYRG
SEQ ID NO. 6 ..........N...-..---....I..-GVG.IG

Figure 4

SEQ ID NO. 5 GWFKKAWRKVKHAGRRVLDTAKGVGRHYLNNWLNRYRG
SEQ ID NO. 9 .F....X.....V..X.....Z

SEQ ID NO. 6 GWFKKAWRKVKNAGR-VLKGVGIHYGVGLIG
SEQ ID NO. 10 .X....X.....R.....Z

SEQUENCE LISTING

<110> GENAERA CORPORATION
UZZELL, Thomas
STOLZENBERG, Ethan D
SHINNAR, Ann E
ZASLOFF, Michael A

<120> HAGFISH CATHELIN-ASSOCIATED ANTIMICROBIAL PEPTIDES AND GENES

<130> 36870-5085

<150> US 60/308,652
<151> 2001-07-30

<160> 11

<170> PatentIn version 3.1

<210> 1
<211> 848
<212> DNA
<213> Myxine glutinosa

<220>
<221> CDS
<222> (4)..(543)
<223>

<400> 1

```
gtc atg aag tcc ttg tgt gtc ccc gcg gtt ctc tct ctg gtc ctg atc       48
    Met Lys Ser Leu Cys Val Pro Ala Val Leu Ser Leu Val Leu Ile
    1               5                  10                  15

ctg ctc ttg gat caa gca cct aca gct cgg gca gat gat tct ttg tcc       96
Leu Leu Leu Asp Gln Ala Pro Thr Ala Arg Ala Asp Asp Ser Leu Ser
                20                  25                  30

aaa gag cag gtg gag aat gca gtt gat gaa gcg ttg gac aag ctg aac      144
Lys Glu Gln Val Glu Asn Ala Val Asp Glu Ala Leu Asp Lys Leu Asn
            35                  40                  45

aaa cag cag gtg tcc aca cgt aaa ctg gca ttg tct gaa caa cag gat      192
Lys Gln Gln Val Ser Thr Arg Lys Leu Ala Leu Ser Glu Gln Gln Asp
        50                  55                  60

atc cag gca gat gaa aca gac gtt gaa gga cag ttc acc atc aaa ttt      240
Ile Gln Ala Asp Glu Thr Asp Val Glu Gly Gln Phe Thr Ile Lys Phe
    65                  70                  75

gac gtg gtt gag acc gag tgc aat gcg gac gat ccc agg gac tgg gcc      288
Asp Val Val Glu Thr Glu Cys Asn Ala Asp Asp Pro Arg Asp Trp Ala
80                  85                  90                  95

gat tgt ccg atc gcg acg gac tcg aca ccg gtc gat gca cag tgt gag      336
Asp Cys Pro Ile Ala Thr Asp Ser Thr Pro Val Asp Ala Gln Cys Glu
                100                 105                 110

gtc acc gtg ttg agc acc gag gac tcc ttg gac gtc gga gac gcg act      384
Val Thr Val Leu Ser Thr Glu Asp Ser Leu Asp Val Gly Asp Ala Thr
```

```
            115                 120                 125

tgc gat ttc aac agg acg gac gga aat gcg agg cga cga cgt ggc tgg      432
Cys Asp Phe Asn Arg Thr Asp Gly Asn Ala Arg Arg Arg Arg Gly Trp
        130                 135                 140

ttt aag aaa gcc tgg aga aaa gtg aag cat gcg gga cga cga gtt ctt      480
Phe Lys Lys Ala Trp Arg Lys Val Lys His Ala Gly Arg Arg Val Leu
    145                 150                 155

gat acc gcg aag ggt gtg gga aga cat tat ttg aat aat tgg ctt aat      528
Asp Thr Ala Lys Gly Val Gly Arg His Tyr Leu Asn Asn Trp Leu Asn
160                 165                 170                 175

cgt tat cgc ggt tag gaggaagctc tgtgttggtg gtggtggtgg tggtggtggt      583
Arg Tyr Arg Gly


ggtggtgttg tgggttaatg ggtgggagcg aatgatttct ccccgacatt caatacttgc    643

tcgttaattc caattgattc gcttagttac acaagtaagc cgaacaggta accagttagt    703

aagctttaaa attgaacttt taggattgaa actgcagctg aacttttttgt agtttaacat   763

ttaattgtga ctctgctatt aaactttgtc tgatcatgga aataaaggat tgcaaaaatg    823

aaaaaaaaaa aaaaaaaaaa aaaaa                                          848


<210>  2
<211>  179
<212>  PRT
<213>  Myxine glutinosa

<400>  2

Met Lys Ser Leu Cys Val Pro Ala Val Leu Ser Leu Val Leu Ile Leu
1               5                   10                  15


Leu Leu Asp Gln Ala Pro Thr Ala Arg Ala Asp Asp Ser Leu Ser Lys
            20                  25                  30


Glu Gln Val Glu Asn Ala Val Asp Glu Ala Leu Asp Lys Leu Asn Lys
        35                  40                  45


Gln Gln Val Ser Thr Arg Lys Leu Ala Leu Ser Glu Gln Gln Asp Ile
    50                  55                  60


Gln Ala Asp Glu Thr Asp Val Glu Gly Gln Phe Thr Ile Lys Phe Asp
65                  70                  75                  80


Val Val Glu Thr Glu Cys Asn Ala Asp Asp Pro Arg Asp Trp Ala Asp
                85                  90                  95


Cys Pro Ile Ala Thr Asp Ser Thr Pro Val Asp Ala Gln Cys Glu Val
```

```
            100                 105                 110

Thr Val Leu Ser Thr Glu Asp Ser Leu Asp Val Gly Asp Ala Thr Cys
        115                 120                 125

Asp Phe Asn Arg Thr Asp Gly Asn Ala Arg Arg Arg Arg Gly Trp Phe
    130                 135                 140

Lys Lys Ala Trp Arg Lys Val Lys His Ala Gly Arg Arg Val Leu Asp
145                 150                 155                 160

Thr Ala Lys Gly Val Gly Arg His Tyr Leu Asn Asn Trp Leu Asn Arg
                165                 170                 175

Tyr Arg Gly
```

<210> 3
<211> 863
<212> DNA
<213> Myxine glutinosa

<220>
<221> CDS
<222> (2)..(508)
<223>

<400> 3

```
c ttg tgt gtc cca gcg gtt ctc tct ctg gtc ctg atc ctg ctc ttg gat      49
  Leu Cys Val Pro Ala Val Leu Ser Leu Val Leu Ile Leu Leu Leu Asp
  1               5                  10                  15

caa gca cct aca gct cgg gca gat gat tct ttg tcc aac gag cag gtg       97
Gln Ala Pro Thr Ala Arg Ala Asp Asp Ser Leu Ser Asn Glu Gln Val
            20                  25                  30

gag aat gca gtt gat gaa gcg ttg gac aag ctg aac aat cag cag gtg      145
Glu Asn Ala Val Asp Glu Ala Leu Asp Lys Leu Asn Asn Gln Gln Val
        35                  40                  45

tcc aca cgt aaa ctg gca ttg tct gaa caa cag gat atc cag gca gat      193
Ser Thr Arg Lys Leu Ala Leu Ser Glu Gln Gln Asp Ile Gln Ala Asp
    50                  55                  60

gaa aca gac gtt gaa gga cag acc acc atc ata ttt tac gtg gtt gag      241
Glu Thr Asp Val Glu Gly Gln Thr Thr Ile Ile Phe Tyr Val Val Glu
65                  70                  75                  80

acc gag tgc aat gcg gac gat ccc agg gac tgg gcc gat tgt ccg atc      289
Thr Glu Cys Asn Ala Asp Asp Pro Arg Asp Trp Ala Asp Cys Pro Ile
                85                  90                  95

gcg acg gac tcg cca ccg ggc att gca cag tgt gag gtc acc gtg ttg      337
Ala Thr Asp Ser Pro Pro Gly Ile Ala Gln Cys Glu Val Thr Val Leu
```

```
            100                 105                 110

agc acc gag gac tcc ttg gac gtc gga gac gcg act tgc gat ttc aac      385
Ser Thr Glu Asp Ser Leu Asp Val Gly Asp Ala Thr Cys Asp Phe Asn
        115                 120                 125

tcg acg ggc gga aat gcg agg cga cga cgt ggc tgg ttt aag aaa gcc      433
Ser Thr Gly Gly Asn Ala Arg Arg Arg Arg Gly Trp Phe Lys Lys Ala
    130                 135                 140

tgg aga aaa gtg aag aat gcg gga cga gtt ctt aag ggt gtg gga ata      481
Trp Arg Lys Val Lys Asn Ala Gly Arg Val Leu Lys Gly Val Gly Ile
145                 150                 155                 160

cat tat ggt gtt gga tta atc ggt taa tcggttggat taatcgttat            528
His Tyr Gly Val Gly Leu Ile Gly
                165

cacggttcgg aggaagtcct ctgttggtgg tggtggtggt ggtggtggtg gtggtggtgg    588

tggtggtggt ggtggtgtgg ggggttatgg gttaatgggt gggagcgaat gatttctccc    648

cgacattcaa tacttgctcg ttaattccaa ttgattcgct tagttacaca agtaagccga    708

acaggtaacc agttagtaag ctttaaaatt gaacttttag gattgaaact gcagctgaac    768

tttttgtagt ttaacattta attgtgactc tgctattaaa ctttgtctga tcatggaaat    828

aaaggattgc aaaaatgaaa aaaaaaaaaa aaaaa                               863
```

<210> 4
<211> 168
<212> PRT
<213> Myxine glutinosa

<400> 4

```
Leu Cys Val Pro Ala Val Leu Ser Leu Val Leu Ile Leu Leu Leu Asp
1               5                   10                  15

Gln Ala Pro Thr Ala Arg Ala Asp Asp Ser Leu Ser Asn Glu Gln Val
            20                  25                  30

Glu Asn Ala Val Asp Glu Ala Leu Asp Lys Leu Asn Asn Gln Gln Val
        35                  40                  45

Ser Thr Arg Lys Leu Ala Leu Ser Glu Gln Gln Asp Ile Gln Ala Asp
    50                  55                  60

Glu Thr Asp Val Glu Gly Gln Thr Thr Ile Ile Phe Tyr Val Val Glu
65                  70                  75                  80

Thr Glu Cys Asn Ala Asp Asp Pro Arg Asp Trp Ala Asp Cys Pro Ile
                85                  90                  95
```

Ala Thr Asp Ser Pro Pro Gly Ile Ala Gln Cys Glu Val Thr Val Leu
100 105 110

Ser Thr Glu Asp Ser Leu Asp Val Gly Asp Ala Thr Cys Asp Phe Asn
115 120 125

Ser Thr Gly Gly Asn Ala Arg Arg Arg Arg Gly Trp Phe Lys Lys Ala
130 135 140

Trp Arg Lys Val Lys Asn Ala Gly Arg Val Leu Lys Gly Val Gly Ile
145 150 155 160

His Tyr Gly Val Gly Leu Ile Gly
165

<210> 5
<211> 38
<212> PRT
<213> Myxine glutinosa

<400> 5

Gly Trp Phe Lys Lys Ala Trp Arg Lys Val Lys His Ala Gly Arg Arg
1 5 10 15

Val Leu Asp Thr Ala Lys Gly Val Gly Arg His Tyr Leu Asn Asn Trp
20 25 30

Leu Asn Arg Tyr Arg Gly
35

<210> 6
<211> 30
<212> PRT
<213> Myxine glutinosa

<400> 6

Gly Trp Phe Lys Lys Ala Trp Arg Lys Val Lys Asn Ala Gly Arg Val
1 5 10 15

Leu Lys Gly Val Gly Ile His Tyr Gly Val Gly Leu Ile Gly
20 25 30

<210> 7
<211> 498
<212> DNA
<213> Capra hircus

<220>

<221> CDS
<222> (1)..(477)
<223>

<400> 7

```
atg gag acc cag agg gcc agc ctc tcc ctg gga cgg tgc tcc ctg tgg       48
Met Glu Thr Gln Arg Ala Ser Leu Ser Leu Gly Arg Cys Ser Leu Trp
1               5                   10                  15

ctc ctg ctg ctg gga cta gtg ctg ccc tcg gcc agc gcc cag gcc ctc       96
Leu Leu Leu Leu Gly Leu Val Leu Pro Ser Ala Ser Ala Gln Ala Leu
            20                  25                  30

agc tac ggg gag gcc gtt ctt cat gct gtc gat cgc atc aat gag cag      144
Ser Tyr Gly Glu Ala Val Leu His Ala Val Asp Arg Ile Asn Glu Gln
        35                  40                  45

tcc tca gaa gcg aat ctc tac cgc ctc ctg gag ctt gac ccg cct ccc      192
Ser Ser Glu Ala Asn Leu Tyr Arg Leu Leu Glu Leu Asp Pro Pro Pro
    50                  55                  60

aag gac gat gag aat cca aac atc ccg aaa cct gtg agc ttc agg gtg      240
Lys Asp Asp Glu Asn Pro Asn Ile Pro Lys Pro Val Ser Phe Arg Val
65                  70                  75                  80

aag gag act gtg tgc ccc agg acg agc cgg cag ccc acg gag cag tgt      288
Lys Glu Thr Val Cys Pro Arg Thr Ser Arg Gln Pro Thr Glu Gln Cys
                85                  90                  95

gac ttc aaa gag aat ggg ctg gtg aag caa tgt gta ggg aca gtc act      336
Asp Phe Lys Glu Asn Gly Leu Val Lys Gln Cys Val Gly Thr Val Thr
            100                 105                 110

ctg gat gcg gtg aaa ggc aaa atg aac atc acc tgc gaa gag ttg cag      384
Leu Asp Ala Val Lys Gly Lys Met Asn Ile Thr Cys Glu Glu Leu Gln
        115                 120                 125

agt gtt ggg aga ttt aaa cga ttt cgt aag aag ctc aaa aga ctc tgg      432
Ser Val Gly Arg Phe Lys Arg Phe Arg Lys Lys Leu Lys Arg Leu Trp
    130                 135                 140

cac aaa gtc ggc cca ttc gtt ggc ccg ata ctc cat tat ggg taa          477
His Lys Val Gly Pro Phe Val Gly Pro Ile Leu His Tyr Gly
145                 150                 155

attgtgagcc catggaagaa t                                             498
```

<210> 8
<211> 158
<212> PRT
<213> Capra hircus

<400> 8

```
Met Glu Thr Gln Arg Ala Ser Leu Ser Leu Gly Arg Cys Ser Leu Trp
1               5                   10                  15

Leu Leu Leu Leu Gly Leu Val Leu Pro Ser Ala Ser Ala Gln Ala Leu
```

20 25 30

Ser Tyr Gly Glu Ala Val Leu His Ala Val Asp Arg Ile Asn Glu Gln
35 40 45

Ser Ser Glu Ala Asn Leu Tyr Arg Leu Leu Glu Leu Asp Pro Pro Pro
50 55 60

Lys Asp Asp Glu Asn Pro Asn Ile Pro Lys Pro Val Ser Phe Arg Val
65 70 75 80

Lys Glu Thr Val Cys Pro Arg Thr Ser Arg Gln Pro Thr Glu Gln Cys
85 90 95

Asp Phe Lys Glu Asn Gly Leu Val Lys Gln Cys Val Gly Thr Val Thr
100 105 110

Leu Asp Ala Val Lys Gly Lys Met Asn Ile Thr Cys Glu Glu Leu Gln
115 120 125

Ser Val Gly Arg Phe Lys Arg Phe Arg Lys Lys Leu Lys Arg Leu Trp
130 135 140

His Lys Val Gly Pro Phe Val Gly Pro Ile Leu His Tyr Gly
145 150 155

<210> 9
<211> 38
<212> PRT
<213> Myxine glutinosa

<220>
<221> MISC_FEATURE
<222> (1)..(38)
<223> Xaa is Brominated Tryptophan

<220>
<221> MISC_FEATURE
<222> (38)..(38)
<223> Glx is C-terminal amide residue

<400> 9

Gly Phe Phe Lys Lys Ala Xaa Arg Lys Val Lys His Ala Gly Arg Arg
1 5 10 15

Val Leu Asp Thr Ala Lys Gly Val Gly Arg His Tyr Val Asn Asn Xaa
20 25 30

Leu Asn Arg Tyr Arg Glx
35

<210> 10
<211> 31
<212> PRT
<213> Myxine glutinosa

<220>
<221> MISC_FEATURE
<222> (1)..(31)
<223> Xaa is Brominated Tryptophan

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> Glx is C-terminal amide residue

<400> 10

Gly Xaa Phe Lys Lys Ala Xaa Arg Lys Val Lys Asn Ala Gly Arg Arg
1 5 10 15

Val Leu Lys Gly Val Gly Ile His Tyr Gly Val Gly Leu Ile Glx
20 25 30

<210> 11
<211> 20
<212> DNA
<213> Myxine glutinosa

<220>
<221> misc_feature
<222> (1)..(20)
<223> r = g or a

<220>
<221> misc_feature
<222> (1)..(20)
<223> n = a or t or c or g

<400> 11
aaraargcnt ggmgnaargt 20

【国際公開パンフレット（コレクトバージョン）】

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 February 2003 (13.02.2003)

PCT

(10) International Publication Number
WO 2003/012044 A3

(51) International Patent Classification⁷: C07K 14/00

(21) International Application Number:
PCT/US2002/024019

(22) International Filing Date: 30 July 2002 (30.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/308,652 30 July 2001 (30.07.2001) US

(71) Applicant (*for all designated States except US*): **GENAERA CORPORATION** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **UZZELL, Thomas** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US). **STOLZENBERG, Ethan, D.** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US). **SHINNAR, Ann, E.** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US). **ZASLOFF, Michael, A.** [US/US]; 5110 Campus Drive, Plymouth Meeting, PA 19462 (US).

(74) Agents: **SMYTH, Robert, J.** et al.; Morgan, Lewis & Bockius LLP, 1111 Pennsylvania Avenue, NW, Washington, DC 20004 (US).

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

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

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:
11 November 2004

(15) Information about Correction:
Previous Correction:
see PCT Gazette No. 14/2004 of 1 April 2004, Section II

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:** HAGFISH CATHELIN-ASSOCIATED ANTIMICROBIAL PEPTIDES AND GENES

(57) **Abstract:** The present invention includes *Myxine glutinosa* cathelin-associated antimicrobial peptides and genes encoding these peptides. The invention also includes compositions an methods for producing these peptides as well as method of preventing and treating microbial infections using these peptides.

WO 2003/012044 A3

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/24019

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C07K 14/00
US CL : 530/350
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
U.S. : 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
SEQ ID NO: 1 and 2 against Data bases GenBank, EMBL, SwissProt, Gen_Seq, Published Patents, PGPub.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SHINNAR et al. Cathelicidin family of antimicrobial peptides: proteolytic processing and protease resistance. Bioorg Chem. December 2003 Vol. 31. No. 6, pages 425-436	1-4, 7-12, and 14.

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
30 August 2004 (30.08.2004)

Date of mailing of the international search report
08 OCT 2004

Name and mailing address of the ISA/US
Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
Facsimile No. (703) 305-3230

Authorized officer
Ponnathapura Achutamurthy
Telephone No. 703-308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

PCT/US02/24019

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

Group I claim(s) 1-14 drawn to a nucleic acid molecule.

Group II claim(s) 15-18 drawn to a peptide

Group III claim(s) 19-25 drawn to a method of treatment using a peptide.

For each of inventions I-III above, restriction to one of the following is also required under 35 USC 121 and 327. Therefore, election is required of one of inventions I-III and one of inventions (A)-(B).

(A) The polypeptide of SEQ ID NO: 2, which comprises SEQ ID NO: 5, or a polypeptide encoding said polypeptide.

(B) The polypeptide of SEQ ID NO: 4, which comprises SEQ ID NO: 6, or a polypeptide encoding said polypeptide.

The inventions listed as Groups I-III and (A)-(B) do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical feature for the following reasons: The technical feature linking Groups I-III and (A)-(B) appears to be that they all relate to antimicrobial peptides. However, Shinnar et al, 1997 (A_GeneSeq Data Base Acc#Aaw10359; Alignment with SEQ ID NO: 5) teach a polypeptide comprising 24 contiguous residues of SEQ ID NO: 5, as recited in Claim 15. Therefore Groups I-III and (A)-(B) share no special technical feature as defined by PCT Rule 13.2, as it does not define a contribution over the prior art. Furthermore, the products of Groups I and II (A)-(B) do not share a special common structural or functional feature while, the methods of Group III do not comprise all of the methods for making or using the products of Groups I and II and (A)-(B). Accordingly, Groups I-III and (A)-(B) are not so linked by the same or a corresponding special technical feature as to form a single general inventive concept.

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/24019

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claim Nos.: 13
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
Claim 13 is missing.

3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1, 2, 7-12, and 14, in part, 3 and 4.

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet(1)) (July 1998)

フロントページの続き

(51)Int.Cl.[7]	FI		テーマコード（参考）
Ａ６１Ｐ 31/10	Ｃ０７Ｋ 14/46		４Ｈ０４５
Ｃ０７Ｋ 14/46	Ｃ１２Ｎ 1/15		
Ｃ１２Ｎ 1/15	Ｃ１２Ｎ 1/19		
Ｃ１２Ｎ 1/19	Ｃ１２Ｎ 1/21		
Ｃ１２Ｎ 1/21	Ｃ１２Ｐ 21/02	Ｃ	
Ｃ１２Ｎ 5/10	Ｃ１２Ｎ 5/00	Ａ	
Ｃ１２Ｐ 21/02	Ａ６１Ｋ 37/02		

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,OM,PH,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZM,ZW

(72)発明者 ストルゼンバーグ, イーサン ディー.
アメリカ合衆国 ペンシルバニア １９４６２, プリマス ミーティング, キャンパス ドライブ ５１１０
(72)発明者 シナー, アン イー.
アメリカ合衆国 ペンシルバニア １９４６２, プリマス ミーティング, キャンパス ドライブ ５１１０
(72)発明者 ザスロフ, マイケル エー.
アメリカ合衆国 ペンシルバニア １９４６２, プリマス ミーティング, キャンパス ドライブ ５１１０

Ｆターム(参考) 4B024 AA01 AA03 AA10 BA80 CA04 CA07 CA09 CA20 DA06 EA03 GA11 HA13 HA14
4B064 AG01 CA01 CA02 CA19 CC01 CC24 DA02
4B065 AA01X AA58X AA72X AA87X AA90Y AB01 AC14 BA02 CA24 CA43 CA44
4C084 AA01 AA06 BA01 BA08 BA22 BA23 CA45 DA41 MA13 MA17 MA55 MA59 MA66 NA14 ZB32 ZB35
4C087 AA01 BB29 CA16 MA13 MA17 MA55 MA59 MA66 NA14 ZB32 ZB35
4H045 AA10 AA20 AA30 BA10 BA41 CA52 EA29 FA74