

(19) 日本国特許庁(JP)

## (12) 公開特許公報(A)

(11) 特許出願公開番号

特開2010-263909  
(P2010-263909A)

(43) 公開日 平成22年11月25日(2010.11.25)

(51) Int.Cl.	F 1	テーマコード (参考)
<b>C 12 N 15/09</b> (2006.01)	C 12 N 15/00	Z N A A 4 B 0 2 4
<b>C 07 K 14/705</b> (2006.01)	C 07 K 14/705	4 B 0 6 5
<b>C 12 N 1/15</b> (2006.01)	C 12 N 1/15	4 H 0 4 5
<b>C 12 N 1/19</b> (2006.01)	C 12 N 1/19	
<b>C 12 N 1/21</b> (2006.01)	C 12 N 1/21	

審査請求 有 請求項の数 1 O L (全 162 頁) 最終頁に続く

(21) 出願番号	特願2010-161113 (P2010-161113)	(71) 出願人	500478097 アーナ ファーマシューティカルズ, インコーポレイテッド アメリカ合衆国カリフォルニア州9212 1サンディエゴ・ナンシーリツジドライブ 6166
(22) 出願日	平成22年7月15日 (2010.7.15)	(74) 代理人	100078282 弁理士 山本 秀策
(62) 分割の表示	特願2010-87 (P2010-87) の分割 原出願日 平成11年10月13日 (1999.10.13)	(74) 代理人	100062409 弁理士 安村 高明
(31) 優先権主張番号	09/170,496	(74) 代理人	100113413 弁理士 森下 夏樹
(32) 優先日	平成10年10月13日 (1998.10.13)		
(33) 優先権主張国	米国(US)		
(31) 優先権主張番号	60/108,029		
(32) 優先日	平成10年11月12日 (1998.11.12)		
(33) 優先権主張国	米国(US)		
(31) 優先権主張番号	60/109,213		
(32) 優先日	平成10年11月20日 (1998.11.20)		
(33) 優先権主張国	米国(US)		
(31) 優先権主張番号	60/110,060		
(32) 優先日	平成10年11月27日 (1998.11.27)		
(33) 優先権主張国	米国(US)		

最終頁に続く

(54) 【発明の名称】非内在性の構成的に活性化されるヒトGタンパク質共役型受容体

## (57) 【要約】

【課題】受容体は、内在性リガンドもしくは薬物のような化合物により活性状態で安定化させることができる。独占的に限定されるものでないが受容体のアミノ酸配列に対する改変を挙げることができる最近の発見は、活性状態のコンホーメーションにある受容体を助長かつ安定化するための内在性リガンドもしくは薬物以外の手段を提供する。これらの手段は、受容体への内在性リガンドの結合の効果を刺激することにより、受容体を活性状態で効果的に安定化する。こうしたリガンドに依存しない手段による安定化を「構成的受容体活性化」と命名する。構成的受容体活性化した受容体を提供することを本発明の課題とする。

【解決手段】上記課題は、明細書に記載の受容体を提供することによって、解決された。

【選択図】なし

## 【特許請求の範囲】

## 【請求項 1】

明細書に記載の c D N A。

## 【発明の詳細な説明】

## 【技術分野】

## 【0 0 0 1】

本特許出願は、1998年10月13日に米国特許・商標庁に出願された米国第09/170,496号の一部継続出願であり、かつそれからの優先権を主張する。本出願は、以下の仮出願（全部は示された日付に米国特許・商標庁に米国エクスプレスメール（U.S. Express Mail）を介して出願された）、すなわち1998年11月27日に出願された米国仮出願第60/110,060号；1999年2月16日に出願された米国仮出願第60/120,416号；1998年11月20日に出願された米国仮出願第60/109,213号の利益を主張する1999年2月26日に出願された米国仮出願第60/121,852号；1999年3月12日に出願された米国仮出願第60/123,944号；1999年3月12日に出願された米国仮出願第60/123,945号；1999年3月12日に出願された米国仮出願第60/123,948号；1999年3月12日に出願された米国仮出願第60/123,951号；1999年3月12日に出願された米国仮出願第60/123,946号；1999年3月12日に出願された米国仮出願第60/123,949号；1999年8月27日に出願された米国仮出願第60/151,114号および1998年11月12日に出願された米国仮出願第60/108,029号の利益を主張する1999年9月3日に出願された米国仮出願第60/152,524号；1999年5月28日に出願された米国仮出願第60/136,436号；1999年5月28日に出願された米国仮出願第60/136,567号；1999年5月28日に出願された米国仮出願第60/137,127号；1999年5月28日に出願された米国仮出願第60/137,131号；1999年5月28日に出願された米国仮出願第60/136,437号の利益を主張する1999年6月29日に出願された米国仮出願第60/141,448号；1999年9月29日に出願された米国仮出願第60/156,633号；1999年9月29日に出願された米国仮出願第60/156,555号；1999年9月29日に出願された米国仮出願第60/156,634号；1999年9月29日に出願された米国仮出願第\_\_\_\_\_号（アリーナ ファーマシューティカルズ インク（Arena Pharmaceuticals, Inc.）処理予定表番号：CHN10-1）；1999年10月1日に出願された米国仮出願第\_\_\_\_\_号（アリーナ ファーマシューティカルズ インク（Arena Pharmaceuticals, Inc.）処理予定表番号：RUP6-1）；1999年10月1日に出願された米国仮出願第\_\_\_\_\_号（アリーナ ファーマシューティカルズ インク（Arena Pharmaceuticals, Inc.）処理予定表番号：RUP7-1）；1999年10月1日に出願された米国仮出願第\_\_\_\_\_号（アリーナ ファーマシューティカルズ インク（Arena Pharmaceuticals, Inc.）処理予定表番号：CHN6-1）；1999年10月1日に出願された米国仮出願第\_\_\_\_\_号（アリーナ ファーマシューティカルズ インク（Arena Pharmaceuticals, Inc.）処理予定表番号：RUP5-1）；ならびに1999年10月1日に出願された米国仮出願第\_\_\_\_\_号（アリーナ ファーマシューティカルズ インク（Arena Pharmaceuticals, Inc.）処理予定表番号：CHN9-1）からの優先権の利益もまた主張する。本出願は、1999年10月12日に（米国エクスプレスメール（U.S. Express Mail）を

10

20

30

40

50

介して)出願された同時係属中の米国第\_\_\_\_\_号(ウッドコック(Woodcock)、ウォッシュバーン(Washburn)、クルツ(Kurtz)、マキエヴィツ( Makiewicz)とノリス(Norris) )、LLP処理予定表番号AREN-0050)および1999年7月30日に出願された米国第09/364,425号(双方は引用により本明細書に組み込まれる)にもまた関する。本出願は、1999年10月12日に(米国エクスプレスメール(U.S.Express Mail)を介して)出願された米国第\_\_\_\_\_号(ウッドコック(Woodcock)、ウォッシュバーン(Washburn)、クルツ(Kurtz)、マキエヴィツ( Makiewicz)とノリス(Norris) )、LLP処理予定表番号AREN-0054)(そっくりそのまま引用により本明細書に組み込まれる)に対する優先権もまた主張する。前述の出願のそれぞれはそっくりそのまま引用により本明細書に組み込まれる。

10

## 【0002】

## (発明の分野)

本特許明細書に開示される発明は、膜貫通受容体、およびより具体的にはヒトGタンパク質共役型受容体、そしてとりわけ該受容体の構成的活性を確立するもしくは高めるよう改変されているGPCRに関する。好ましくは、改変されたGPCRは、治療薬として潜在的応用性を有する受容体のアゴニスト、反作用薬もしくは部分的アゴニストとしての候補化合物の直接の同定に使用される。

20

## 【背景技術】

## 【0003】

## (発明の背景)

ヒトでは多数の受容体のクラスが存在するが、はるかに最も豊富かつ治療に関係するものは、Gタンパク質共役型受容体(GPCRもしくは複数GPCR)のクラスにより代表される。ヒトゲノム内には数十万個の遺伝子が存在することが推定されており、そしてこれらのなかでおよそ2%もしくは2,000個の遺伝子がGPCRをコードすると推定されている。内在性リガンドが同定されている、GPCRを包含する受容体は「既知」受容体と称される一方、内在性リガンドが同定されていない受容体は「オーファン」受容体と称される。GPCRは製薬学的製品の開発に重要な一領域を代表する。すなわち、100種の既知のGPCRのおよそ20種から全処方薬の60%が開発されている。

30

## 【0004】

GPCRは1つの共通な構造モチーフを共有する。全部のこれらの受容体は、7個のヘリックスを形成する22ないし24個の間の疎水性アミノ酸の7個の連なりを有し、そのそれぞれは膜にまたがる(各広がり(span)は数字により同定されている。すなわち膜貫通-1(TM-1)、膜貫通-2(TM-2)など)。膜貫通ヘリックスは、細胞膜の外もしくは「細胞外」側で、膜貫通-2と膜貫通-3、膜貫通-4と膜貫通-5、および膜貫通-6と膜貫通-7の間でアミノ酸の鎖により結合されている(これらはそれぞれ「細胞外」領域1、2および3(EC-1、EC-2およびEC-3)と称される)。膜貫通ヘリックスは、細胞膜の内もしくは「細胞内」側で、膜貫通-1と膜貫通-2、膜貫通-3と膜貫通-4、および膜貫通-5と膜貫通-6の間でもまたアミノ酸の鎖により結合されている(これらはそれぞれ「細胞内」領域1、2および3(IC-1、IC-2およびIC-3)と称される)。受容体の「カルボキシ」(「C」)末端は細胞内の細胞内空隙中に存し、また、受容体の「アミノ」(「N」)末端は細胞の外側の細胞外空隙中に存する。

40

## 【0005】

一般に、内在性リガンドが受容体と結合する場合(しばしば受容体の「活性化」と称される)、細胞内領域のコンホメーションの変化が存在し、それは細胞内領域と細胞内の「Gタンパク質」との間の共役(coupling)を見込む。GPCRは

50

Gタンパク質に関して「混雑して」いる、すなわちGPCRは1種以上のGタンパク質と相互作用する可能性があることが報告されている。ケナキン (Kena kin, T.)、43 Life Sciences 1095 (1988) を参照されたい。他のGタンパク質が存在するが、現在のところGq、Gs、Gi、GzおよびGoが同定されたGタンパク質である。Gタンパク質との内在性リガンドで活性化されたGPCRの共役がシグナル伝達カスケード過程（「シグナル伝達」と称される）を開始する。正常な条件下では、シグナル伝達は最終的に細胞の活性化もしくは細胞の阻害をもたらす。受容体のIC-3ループならびにカルボキシ末端がGタンパク質と相互作用すると考えられている。

## 【0006】

10

GPCRは、生理学的条件下では2種の異なるコンホメーション、すなわち「不活性」状態と「活性状態」との間の平衡で細胞膜中に存在する。不活性状態の受容体は、細胞内のシグナル伝達経路に連結して生物学的応答を生じさせることができない。受容体のコンホメーションの活性状態への変化は（Gタンパク質を介する）伝達経路への連鎖を可能にし、そして生物学的応答を生じさせる。

## 【0007】

20

受容体は、内在性リガンドもしくは薬物のような化合物により活性状態で安定化させることができる。独占的に限定されるものでないが受容体のアミノ酸配列に対する改変を挙げることができる最近の発見は、活性状態のコンホメーションにある受容体を助長かつ安定化するための内在性リガンドもしくは薬物以外の手段を提供する。これらの手段は、受容体への内在性リガンドの結合の効果を刺激することにより、受容体を活性状態で効果的に安定化する。こうしたリガンドに依存しない手段による安定化を「構成的受容体活性化」と命名する。

## 【発明の概要】

## 【課題を解決するための手段】

## 【0008】

## (発明の要約)

内在性ヒトGPCRの非内在性変種およびそれらの用途を本明細書で開示する。

本発明は、例えば、以下を提供する：

30

(項目1) hARE-3 (F313K) を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目2) 項目1記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目3) ベクターおよび項目1記載のcDNAを含んで成るプラスミド。

(項目4) 項目3記載のプラスミドを含んで成る宿主細胞。

(項目5) hARE-4 (V233K) を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目6) 項目5記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

40

(項目7) ベクターおよび項目5記載のcDNAを含んで成るプラスミド。

(項目8) 項目7記載のプラスミドを含んで成る宿主細胞。

(項目9) hARE-5 (A240K) を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目10) 項目9記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目11) ベクターおよび項目5記載のcDNAを含んで成るプラスミド。

(項目12) 項目11記載のプラスミドを含んで成る宿主細胞。

50

(項目13) hGPCR14(L257K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目14) 項目13記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目15) ベクターおよび項目13記載のcDNAを含んで成るプラスミド。

(項目16) 項目15記載のプラスミドを含んで成る宿主細胞。

(項目17) hGPCR27(C283K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

10

(項目18) 項目17記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目19) ベクターおよび項目17記載のcDNAを含んで成るプラスミド。

(項目20) 項目19記載のプラスミドを含んで成る宿主細胞。

(項目21) hARE-1(E232K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目22) 項目21記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

20

(項目23) ベクターおよび項目21記載のcDNAを含んで成るプラスミド。

(項目24) 項目23記載のプラスミドを含んで成る宿主細胞。

(項目25) hARE-2(G285K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目26) 項目25記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目27) ベクターおよび項目25記載のcDNAを含んで成るプラスミド。

(項目28) 項目27記載のプラスミドを含んで成る宿主細胞。

30

(項目29) hPPR1(L239K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目30) 項目29記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目31) ベクターおよび項目29記載のcDNAを含んで成るプラスミド。

(項目32) 項目31記載のプラスミドを含んで成る宿主細胞。

(項目33) hG2A(K232A)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

40

(項目34) 項目33記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目35) ベクターおよび項目33記載のcDNAを含んで成るプラスミド。

(項目36) 項目35記載のプラスミドを含んで成る宿主細胞。

(項目37) hRUP3(L224K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目38) 項目37記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目39) ベクターおよび項目37記載のcDNAを含んで成るプラスミド。

50

(項目40) 項目39記載のプラスミドを含んで成る宿主細胞。  
 (項目41) hRUP5(A236K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目42) 項目41記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目43) ベクターおよび項目41記載のcDNAを含んで成るプラスミド。

(項目44) 項目42記載のプラスミドを含んで成る宿主細胞。  
 (項目45) hRUP6(N267K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目46) 項目45記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目47) ベクターおよび項目45記載のcDNAを含んで成るプラスミド。

(項目48) 項目47記載のプラスミドを含んで成る宿主細胞。  
 (項目49) hRUP7(A302K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目50) 項目49記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目51) ベクターおよび項目49記載のcDNAを含んで成るプラスミド。

(項目52) 項目51記載のプラスミドを含んで成る宿主細胞。  
 (項目53) hCHN4(V236K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目54) 項目53記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目55) ベクターおよび項目53記載のcDNAを含んで成るプラスミド。

(項目56) 項目55記載のプラスミドを含んで成る宿主細胞。  
 (項目57) hM C4(A244K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目58) 項目57記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目59) ベクターおよび項目57記載のcDNAを含んで成るプラスミド。

(項目60) 項目60記載のプラスミドを含んで成る宿主細胞。  
 (項目61) hCHN3(S284K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目62) 項目61記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目63) ベクターおよび項目61記載のcDNAを含んで成るプラスミド。

(項目64) 項目63記載のプラスミドを含んで成る宿主細胞。  
 (項目65) hCHN6(L352K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目66) 項目65記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目67) ベクターおよび項目65記載のcDNAを含んで成るプラスミド。

(項目68) 項目67記載のプラスミドを含んで成る宿主細胞。

10

20

30

40

50

(項目69) hCHN8(N235K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目70) 項目69記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目71) ベクターおよび項目69記載のcDNAを含んで成るプラスミド。

(項目72) 項目71記載のプラスミドを含んで成る宿主細胞。

(項目73) hH9(F236K)を含んで成る、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型受容体をコードするcDNA。

(項目74) 項目73記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目75) ベクターおよび項目73記載のcDNAを含んで成るプラスミド。

(項目76) 項目74記載のプラスミドを含んで成る宿主細胞。

(項目77) hAT1(F239K)；hAT1(N111A)；hAT1(AT2K255IC3)；およびhAT1(A243+)より成る群から選択される、非内在性で、構成的に活性化される変種ヒトGタンパク質共役型AT1受容体をコードするcDNA。

(項目78) 項目77記載のcDNAによりコードされる、非内在性の変種ヒトGタンパク質共役型受容体。

(項目79) ベクターおよび項目77記載のcDNAを含んで成るプラスミド。

(項目80) 項目79記載のプラスミドを含んで成る宿主細胞。

【図面の簡単な説明】

【0009】

【図1】発明にかかる、8XCRE-Lucレポータープラスミドの表示である(実施例4(c)3を参照されたい)。

【図2】発明にかかる、内在性TDAg8へのATPおよびADPの結合の結果のグラフ表示(2A)ならびに血清および血清を含まない培地中での比較(2B)である。

【図3】発明にかかる、GPCR融合タンパク質H9(F236K)：Gsに対するCMVの比較的なシグナル伝達の結果のグラフ表示である。

【発明を実施するための形態】

【0010】

(詳細な記述)

受容体を中心として発展してきた学術文献は、受容体に対する多様な影響を有するリガンドを指す多数の用語を採用している。明瞭化および一貫性のため、本特許明細書を通じて以下の定義を使用するであろう。これらの定義がこれらの用語の他の定義と矛盾する限り、以下の定義が支配する：

アゴニストは、それらが受容体に結合する場合に細胞内応答を活性化する、もしくは膜へのGTP結合を高める物質(例えばリガンド、候補化合物)を意味する。

【0011】

本明細書で使用されるアミノ酸の略語を表Aに示す：

【0012】

10

20

30

40

## 【表1】

表A		
アラニン	ALA	A
アルギニン	ARG	R
アスパラキシン	ASN	N
アスパラギン酸	ASP	D
システイン	CYS	C
グルタミン酸	GLU	E
グルタミン	GLN	Q
グリシン	GLY	G
ヒスチジン	HIS	H
イソロイシン	ILE	I
ロイシン	LEU	L
リシン	LYS	K
メオニン	MET	M
フェニルアラニン	PHE	F
プロリン	PRO	P
セリン	SER	S
トレオニン	THR	T
トリプトファン	TRP	W
チロシン	TYR	Y
バリン	VAL	V

10

20

## 【0013】

部分的アゴニストは、それらが受容体に結合する場合にアゴニストが活性化するより小さい度合(degree) / 程度(extent)まで細胞内応答を活性化する、もしくはアゴニストが高めるより小さい度合(degree) / 程度(extent)まで膜へのGTP結合を高める物質(例えばリガンド、候補化合物)を意味する。

## 【0014】

アンタゴニストは、アゴニストと同一の部位で受容体に競争的に結合するがしかし活性型の受容体により開始される細胞内応答を活性化せず、そしてそれによりアゴニストもしくは部分的アゴニストによる細胞内応答を阻害する可能性のある物質(例えばリガンド、候補化合物)を意味する。アンタゴニストは、アゴニストもしくは部分的アゴニストの非存在下で基線の細胞内応答を減少させない。

30

## 【0015】

候補化合物は、スクリーニング技術に基づいて分析できる分子(例えば、そして制限でなく化合物)を意味する。好ましくは、「候補化合物」という句は、間接的同定方法により既に決定されたような、受容体に対する反作用薬、アゴニストもしくはアンタゴニストより成る群から選択される化合物であることが公知であった化合物(「間接的同定された化合物」)を包含せず;より好ましくは、最低1種の哺乳動物で治療上の効能を有することが既に決定されている間接的同定された化合物を包含せず;そして最も好ましくは、ヒトで治療上の有用性を有することが既に決定されている間接的同定された化合物を包含しない。

40

## 【0016】

組成物は最低1種の成分を含んで成る物質を意味し、「製薬学的組成物」は組成物の一例である。

## 【0017】

化合物の効能は、受容体結合親和性と対照的に、受容体の機能性を阻害もしくは刺激する化合物の能力の大きさを意味する。化合物の効能の検出の例示的手段を本特許明細書の実施例のセクションに開示する。

## 【0018】

コドンは、リン酸基に結合されたヌクレオシド(アデノシン(A)、グアノシ

50

ン ( G ) 、シチジン ( C ) 、ウリジン ( U ) およびチミジン ( T ) ) を一般に含んで成りかつ翻訳される場合に 1 個のアミノ酸をコードするひとそろいの 3 個のヌクレオチド ( もしくはヌクレオチドに対する同等物 ) を意味する。

【 0019 】

構成的に活性化される受容体は構成的受容体活性化にさらされる受容体を意味する。構成的に活性化される受容体は内在性もしくは非内在性であることができる。

【 0020 】

構成的な受容体の活性化は、その内在性リガンドもしくはその化学的同等物との受容体の結合以外の手段による、活性状態での受容体の安定化を意味する。

10

【 0021 】

接触させるもしくは接触することは、インビトロの系であろうとインビボの系であろうと最低 2 つの部分を一緒にすることを意味する。

【 0022 】

直接同定するもしくは直接同定されるは、「候補化合物」という句に関係して、構成的に活性化される受容体、好ましくは構成的に活性化されるオーファン受容体、そして最も好ましくは構成的に活性化される G タンパク質共役型の細胞表面のオーファン受容体に対する候補化合物のスクリーニング、およびこうした化合物の化合物の効能の評価を意味する。この句は、どの状況下でも、「間接的に同定する」もしくは「間接的に同定される」という句により包含されるもしくはそれを包含すると解釈もしくは理解されるべきでない。

20

【 0023 】

内在性は哺乳動物が天然に産生する物質を意味する。例えば、そして制限でなく「受容体」という用語に関しての内在性は、哺乳動物 ( 例えば、そして制限でなくヒト ) もしくはウイルスにより天然に産生されるものを意味する。対照的に、これに関して非内在性という用語は、哺乳動物 ( 例えば、そして制限でなくヒト ) もしくはウイルスにより天然に産生されないものを意味する。例えば、そして制限でなく、その内在性の形態で構成的に活性でないがしかし操作される場合に構成的に活性になる受容体は、本明細書で「非内在性の構成的に活性化される受容体」と最も好ましく称される。双方の用語は「インビボ」および「インビトロ」双方の系を記述するのに利用することができる。例えば、そして制限でなく、スクリーニングのアプローチにおいて、内在性もしくは非内在性の受容体はインビトロのスクリーニング系に関してであってよい。さらなる一例として、そして制限としてでなく、非内在性の構成的に活性化される受容体を包含するように哺乳動物のゲノムが操作されている場合には、インビボ系による候補化合物のスクリーニングが実現可能である。

30

【 0024 】

G タンパク質共役型受容体融合タンパク質および G P C R 融合タンパク質は、本明細書に開示される本発明に関して、それぞれ、最低 1 種の G タンパク質、最も好ましくはこうした G タンパク質の サブユニット ( これは G T P を結合するサブユニットである ) に融合された、内在性の構成的に活性化される G P C R もしくは非内在性の構成的に活性化される G P C R を含んで成る非内在性タンパク質を意味し、G タンパク質は好ましくは内在性のオーファン G P C R と天然に共役する G タンパク質と同一の型のものである。例えば、そして制限でなく、内在性の状態において、G タンパク質「 G s 」が G P C R と共に主な G タンパク質である場合、特定の G P C R に基づく G P C R 融合タンパク質は G s に融合された G P C R を含んで成る非内在性タンパク質であることができ ; 下に示されるであろうようないくつかの状況では、主なものでない G タンパク質を G P C R に融合することができる。G タンパク質は構成的に活性な G P C R の C 末端に直接融合することができるか、もしくは 2 者の間にスペーサーが存在してよい。

40

50

## 【0025】

宿主細胞は、その中に組み込まれたプラスミドおよび／もしくはベクターを有することが可能な細胞を意味する。原核生物宿主細胞の場合には、プラスミドは宿主細胞の複製物のような自律的分子として典型的に複製され（一般に、プラスミドはその後真核生物宿主細胞への導入のため単離される）；真核生物宿主細胞の場合には、真核生物宿主細胞が複製する場合にプラスミドが複製するように、プラスミドが宿主細胞の細胞DNAに組み込まれる。好ましくは、本明細書に開示される本発明の目的上、宿主細胞は真核生物、より好ましくは哺乳動物であり、そして最も好ましくは293、293TおよびCOS-7細胞より成る群から選択される。

10

## 【0026】

間接的に同定するもしくは間接的に同定されるは、内在性受容体に特異的な内在性リガンドの同定、リガンド-受容体の相互作用を妨害および／もしくはそれと競争するものの決定のための受容体に対する候補化合物のスクリーニング、ならびに活性化された受容体に関連する最低1種のセカンドメッセンジャー経路に影響を及ぼす化合物の効能の評価を必要とする、薬物発見過程への伝統的アプローチを意味する。

## 【0027】

「応答」という用語に関係した阻害するもしくは阻害は、化合物の非存在下と対照的に、化合物の存在下で応答が低下もしくは予防されることを意味する。

20

## 【0028】

反作用薬は、内在性の形態の受容体もしくは構成的に活性化された形態の受容体のいずれかに結合し、かつ、アゴニストもしくは部分的アゴニストの非存在下で観察される正常の基準レベルの活性より低い、活性の形態の受容体により開始される基線の細胞内応答を阻害するか、もしくは膜へのGTP結合を減少させる物質（例えばリガンド、候補化合物）を意味する。好ましくは、基線の細胞内応答は、反作用薬の非存在下の基礎の応答に比較して、反作用薬の存在下で最低30%、より好ましくは最低50%、そして最も好ましくは最低75%阻害される。

## 【0029】

既知の受容体は、その受容体に特異的な内在性リガンドが同定されている内在性受容体を意味する。

30

## 【0030】

リガンドは、内在性の天然に存在する受容体に特異的な内在性の天然に存在する分子を意味する。

## 【0031】

内在性受容体の核酸および／もしくはアミノ酸の配列に関しての突然変異体もしくは突然変異は、突然変異された形態の内在性の構成的に活性化されない受容体が該受容体の構成的活性化を明示するような、こうした内在性の配列に対する指定された1個の変化もしくは複数の変化を意味する。特定の配列に対する同等物に関して、（a）その後の突然変異された形態のヒト受容体の構成的活性化のレベルが該受容体の第一の突然変異により明示されるものと実質的に同一であり；そして（b）その後の突然変異された形態の受容体と該受容体の第一の突然変異との間の配列（アミノ酸および／もしくは核酸）の相同性のパーセントが最低約80%、より好ましくは最低約90%、そして最も好ましくは最低95%である場合、その後の突然変異された形態のヒト受容体は該ヒト受容体の第一の突然変異に同等であるとみなされる。理想的には、また、構成的活性化を達成するための本明細書に開示される最も好ましいカセットはGPCRの内在性の形態と非内在性の形態との間の单一のアミノ酸および／もしくはコドンの変化を包含するという事実のために、配列の相同性のパーセントは最低98%であるべきである。

40

50

## 【0032】

非オーファン受容体は、内在性の天然に存在するリガンドに特異的な内在性の天然に存在する分子を意味し、ここで、受容体へのリガンドの結合が細胞内シグナル伝達経路を活性化する。

## 【0033】

オーファン受容体は、その受容体に特異的な内在性リガンドが同定されていないかもしくは未知である内在性受容体を意味する。

## 【0034】

製薬学的組成物は最低1種の有効成分を含んで成る組成物を意味し、それにより該組成物は哺乳動物（例えば、そして制限でなくヒト）における特定の効能のある結果についての検討に基づいて分析できる。当業者は、ある有効成分が当業者のニーズに基づき所望の効能のある結果を有するかどうかを決定するのに適切な技術を理解かつ認識するであろう。

10

## 【0035】

プラスミドはベクターおよびcDNAの組み合わせを意味する。一般に、プラスミドは、cDNAの複製および／もしくはそのタンパク質としての発現の目的上、宿主細胞に導入される。

## 【0036】

「応答」という用語に関係した刺激するもしくは刺激することは、化合物の非存在下と対照的に化合物の存在下で応答が増大されることを意味する。

20

## 【0037】

cDNAに関してのベクターは、最低1個のcDNAを組み込むことが可能かつ宿主細胞への組み込みが可能な環状DNAを意味する。

## 【0038】

以下のセクションの階層は表象的な効能について示され、また、後に続く開示もしくは請求の範囲に対する制限として意図されず、またそのように解釈されるべきでもない。

## A. 緒言

受容体の伝統的な研究は、発見が受容体に影響を及ぼす可能性のあるアンタゴニストおよび他の分子を見出すよう進めることができる前に、内在性リガンドを最初に同定しなければならないという（歴史的に基づく）演繹的仮定から常に進められてきた。アンタゴニストを最初に知ることができた場合であっても、内在性リガンドを探すように研究が直ちに拡大されてきた。この思考様式は、構成的に活性化される受容体の発見の後でさえ、受容体の研究で持続してきた。これまで認識されていなかったことは、受容体のアゴニスト、部分的アゴニストおよび反作用薬の発見に最も有用であるのは活性状態の受容体であるということである。過度に活性の受容体もしくは過小活性の受容体から生じる疾患に対して、治療薬で望まれるものは、必ずしも内在性リガンドに対するアンタゴニストである薬物でなく、それぞれ、受容体の活性状態を低下させるもしくは受容体の活性を高めるよう作用する化合物である。これは、活性の受容体状態の活性を低下させるもしくは高める化合物は内在性リガンドと同一の部位で結合する必要がないためである。従って、本発明の方法により教示されるとおり、治療的化合物についてのいかなる研究も、リガンドに依存しない活性状態に対して化合物をスクリーニングすることにより開始すべきである。

30

## B. ヒトGPCRの同定

ヒトゲノムプロジェクトの努力の成果は、ヒトゲノム内に配置されている核酸配列に関する夥しい量の情報の同定につながり；それは、いずれかの特定のゲノム配列がヒトタンパク質を翻訳する読取り枠情報を含有するもしくは含有するかも知れないかどうかに関する理解もしくは認識を伴わずに、遺伝子配列情報が利用可能にされているというこの努力において真実である。ヒトゲノム内の核酸配

40

50

列のいくつかの同定方法は当業者の範囲内にある。例えば、そして制限でなく、本明細書で開示される多様なヒトG P C Rは、ジェンバンク [ G e n B a n k ] (商標)データベースを再検討することにより発見された一方、他のG P C Rは、既に配列決定されたG P C Rの核酸配列を利用してE S TデータベースのB L A S T (商標)検索を実施することにより発見された。下の表Bは、G P C Rのそれぞれの相同受容体と一緒に、われわれが発見した数種の内在性G P C Rを列挙する。

【0039】

【表2】

表B

10

開示されるヒトオーファンGPCRs	固定された受託番号	読み取り枠(塩基対)	指定されたGPCRとの相同性パーセント	相同GPCRに対する言及(受託番号)
<b>hARE-3</b>	AL033379	1,260 bp	52.3% LPA-R	U92642
<b>hARE-4</b>	AC006087	1,119 bp	36% P2Y5	AF000546
<b>hARE-5</b>	AC006255	1,104 bp	32% オリジアスラティペス ( <i>Oryzias latipes</i> )	D43633
<b>hGPR27</b>	AA775870	1,128 bp		
<b>hARE-1</b>	AJ090920	999 bp	43% KIAA0001	D13626
<b>hARE-2</b>	AA359504	1,122 bp	53% GPR27	
<b>hPPR1</b>	H67224	1,053 bp	39% EBI1	L31581
<b>hG2A</b>	AA754702	1,113 bp	31% GPR4	L36148
<b>hRUP3</b>	AL035423	1,005 bp	30% キイロショウジョウバエ ( <i>Drosophila melanogaster</i> )	2133653
<b>hRUP4</b>	AI307658	1,296 bp	32% pNPGPR 28% および 29% それぞれセブライッシュ ( <i>Zebrafish</i> ) Ya および Yb	NP_004876 AAC41276 AAB94616
<b>hRUP5</b>	AC005849	1,413 bp	25% DEZ 23% FMLPR	Q99788 P21462
<b>hRUP6</b>	AC005871	1,245 bp	48% GPR66	NP_006047
<b>hRUP7</b>	AC007922	1,173 bp	43% H3R	AF140538
<b>hCHN3</b>	EST 36581	1,113 bp	53% GPR27	
<b>hCHN4</b>	AA804531	1,077 bp	32% トロンビン	4503637
<b>hCHN6</b>	EST 2134670	1,503 bp	36% edg-1	NP_001391
<b>hCHN8</b>	EST 764455	1,029 bp	47% KIAA0001	D13626
<b>hCHN9</b>	EST 1541536	1,077 bp	41% LTB4R	NM_000752
<b>hCHN10</b>	EST 1365839	1,055 bp	35% P2Y	NM_002563

20

30

40

50

【0040】

受容体の相同性は、人体内での受容体の役割の正しい認識を得ることに関して有用である。本特許明細書が進めるように、われわれは、これらの受容体の非内在性で構成的に活性化される変種を確立するようにこれらの受容体を突然変異するための技術を開示するであろう。

## 【0041】

本明細書に開示される技術は、本特許明細書が進行するに従って明らかになるであろうように、技術既知の他のヒトのオーファン GPCR にもまた適用されている。

## C . 受容体スクリーニング

本明細書に開示される、非内在性で、構成的に活性化される変種ヒト GPCR に対する候補化合物のスクリーニングは、受容体の内在性リガンドの使用を必要とすることなく、この細胞表面受容体で作用する候補化合物の直接の同定を許す。本明細書に開示される内在性の変種ヒト GPCR が発現および／もしくは過剰発現されている身体内の領域を決定することにより、受容体の発現および／もしくは過剰発現を伴う関連疾患／障害状態を決定することが可能であり；こうしたアプローチを本特許明細書で開示する。

## 【0042】

本明細書に開示されるヒト GPCR の構成的活性化を明示することができる突然変異の創製に関して、 GPCR の TM6 内に配置されると想定されているプロリン残基からの距離に基づき；このアルゴリズム技術は、引用により本明細書に組み込まれる同時係属中かつ共通に譲渡される特許明細書米国第 09/170,496 号に開示されている。該アルゴリズム技術は、伝統的な配列の「整列」ではなく、しかしむしろ前述の TM6 のプロリン残基からの指定された距離に基づく。（思うに受容体の ICB 領域中に配置される）この残基から 16 アミノ酸残基に配置されるアミノ酸残基を最も好ましくはリシン残基に突然変異することにより、こうした活性化を得てよい。この目的を達成するためにはこの位置の突然変異で他のアミノ酸残基が有用であるかも知れない。

## D . 疾患／障害の同定および／もしくは選択

下により詳細に示されるであろうとおり、最も好ましくは非内在性で、構成的に活性化される GPCR に対する反作用薬を本発明の方法論により同定することができる。こうした反作用薬は、本受容体に関連する疾患治療のための薬物発見プログラムでのリード化合物として理想的な候補である。 GPCR に対する反作用薬を直接同定してそれにより製薬学的組成物の開発を可能にする能力のため、 GPCR に関連する疾患および障害の研究が適切である。例えば、 GPCR の存在について疾患に罹った組織サンプルおよび正常な組織サンプルの双方を走査することは、今や、学究的修練、もしくは特定の GPCR に対する内在性リガンドを同定する途に沿って追跡することができるもの以上となっている。組織の走査は広範な健康なおよび疾患に罹った組織にわたって実施することができる。こうした組織の走査は、特定の受容体のある疾患および／もしくは障害との関連付けの好ましい第一段階を提供する。例えば、本明細書に開示される GPCR のいくつかの例示的ドットプロットおよび RT-PCR の結果については、同時係属中の出願（処理予定表番号 A R E - 0050 ）を参照されたい。

## 【0043】

好ましくは、（a）組織 mRNA に対するドットプロット分析、および／もしくは（b）組織サンプルでの受容体の発現の RT-PCR 同定のためのプローブを作成するのにヒト GPCR の DNA 配列を使用する。組織供給源、もしくは疾患に罹った組織中での受容体の存在、または正常組織に比較して疾患に罹った組織中での上昇された濃度での受容体の存在は、限定されるものでないがその疾患に関連する疾患を挙げができる治療レジメンとの相関を同定するのに好ましく利用することができる。受容体はこの技術により器官の領域に等しく十分に局在化する可能性がある。受容体が局在化されている特定の組織の既知の機能に基づき、受容体の推定の機能上の役割を推定することができる。

## E . 候補化合物のスクリーニング

## 1 . 包括的 GPCR スクリーニングアッセイの技術

10

20

30

40

50

Gタンパク質受容体が構成的に活性になった場合、それはGタンパク質（例えばG<sub>q</sub>、G<sub>s</sub>、G<sub>i</sub>、G<sub>z</sub>、G<sub>o</sub>）に結合しかつGタンパク質へのGTPの結合を刺激する。その後、Gタンパク質がGTPアーゼとして作用し、そしてGTPをGDPにゆっくりと加水分解し、それにより受容体は正常条件下で失活したようになる。しかしながら、構成的に活性化された受容体はGDPをGTPに交換し続ける。構成的に活性化される受容体を発現する膜への高められた結合をモニターするのに、GTPの加水分解不可能な類似物 [<sup>35</sup>S]GTP-Sを使用することができる。 [<sup>35</sup>S]GTP-Sはリガンドの非存在下および存在下での膜へのGタンパク質の共役をモニターするのに使用することができることが報告されている。このモニタリングの一例、なかんずく当業者に公知かつ利用可能な例は、1995年にトレイノル（T raynor）とナホルスキ（Nahorski）により報告された。本アッセイ系の好ましい使用は候補化合物の初期スクリーニングのためである。なぜなら、該系は、受容体の細胞内ドメインと相互作用する特定のGタンパク質に関係なく、全部のGタンパク質共役型受容体に包括的に応用可能であるからである。

10

## 2. 特異的GPCRスクリーニングアッセイの技術

「包括的」Gタンパク質共役型受容体アッセイ（すなわちアゴニスト、部分的アゴニストもしくは反作用薬である化合物を選択するためのアッセイ）を使用して候補化合物が同定されれば、該化合物が受容体部位で相互作用していることを確認するためのさらなるスクリーニングが好ましい。例えば、「包括的」アッセイにより同定された化合物は受容体に結合しないかも知れないが、しかし代わりに細胞内ドメインからGタンパク質を単に「分離する」かも知れない。

20

### a. G<sub>s</sub>、G<sub>z</sub>およびG<sub>i</sub>

G<sub>s</sub>は酵素アデニリルシクラーゼを刺激する。他方、G<sub>i</sub>（ならびにG<sub>z</sub>およびG<sub>o</sub>）はこの酵素を阻害する。アデニリルシクラーゼはcAMPへの転化を触媒し；従って、G<sub>s</sub>タンパク質を共役する構成的に活性化されたGPCRは、cAMPの増大された細胞レベルと関連する。他方、G<sub>i</sub>（もしくはG<sub>z</sub>、G<sub>o</sub>）タンパク質を共役する構成的に活性化されたGPCRは、cAMPの低下された細胞レベルと関連する。一般に、“Indirect Mechanisms of Synaptic Transmission,” 第8章、From Neuron To Brain（第3版）ニコルス（Nicholls, J. G.）ら編 サイナウア アソシエーツ インク（Sinclair Associates, Inc.）（1992）を参照されたい。従って、cAMPを検出するアッセイを利用して、ある候補化合物が例えば受容体に対する反作用薬（すなわちこうした化合物はcAMPのレベルを低下させることができる）であるかどうかを決定することができる。cAMPを測定するための当該技術分野で既知の多様なアプローチを利用することができます；最も好ましいアプローチはELISAに基づく形式での抗cAMP抗体の使用に頼る。利用することができる別の型のアッセイは全細胞セカンドメッセンジャーレポーター系アッセイである。遺伝子のプロモーターは特定の遺伝子がコードするタンパク質の発現を司る。環状AMPは、cAMP応答性のDNA結合タンパク質もしくは転写因子（CREB）（その後cAMP応答要素と呼ばれる特定の部位でプロモーターに結合しそして遺伝子の発現を司る）の結合を促進することにより遺伝子発現を司る。レポーター遺伝子（例えば-ガラクトシダーゼもしくはルシフェラーゼ）の前に複数のcAMP応答要素を含有するプロモーターを有するレポーター系を構築することができる。従って、構成的に活性化されたG<sub>s</sub>に結合された受容体は、cAMPの蓄積を引き起こし、これがその後遺伝子およびレポータータンパク質の発現を活性化する。その後、標準的な生化学的アッセイを使用して、-ガラクトシダーゼもしくはルシフェラーゼのようなレポータータンパク質を検出することができる（チェン（Chen）ら 1995）。

30

40

50

b . G o および G q

G<sub>q</sub> および G<sub>o</sub> は酵素ホスホリパーゼ C の活性化に関連し、この酵素は順にリン脂質 PIP<sub>2</sub> を加水分解して、2種の細胞内メッセンジャー、すなわちジアシルグリセロール (DAG) およびイノシトール 1,4,5-三リン酸 (IP<sub>3</sub>) を放出する。IP<sub>3</sub> の増大された蓄積は G<sub>q</sub> および G<sub>o</sub> 会合型受容体の活性化と関連する。一般に、“*Indirect Mechanisms of Synaptic Transmission*,” 第8章、From Neuron To Brain (第3版) ニコルス (Nichols, J. G.) ら編 サイナウア アソシエーツ インク (Sinauer Associates, Inc.) (1992) を参照されたい。IP<sub>3</sub> の蓄積を検出するアッセイは、候補化合物が例えれば G<sub>q</sub> もしくは G<sub>o</sub> 会合型受容体に対する反作用薬である (すなわちこうした化合物は IP<sub>3</sub> のレベルを低下させることができる) かどうかを決定するのに利用することができる。G<sub>q</sub> 会合型受容体は AP1 レポーターアッセイ (ここで G<sub>q</sub> 依存性のホスホリパーゼ C が AP1 要素を含有する遺伝子の活性化を引き起こす) を使用してもまた検査することができ；従って、活性化された G<sub>q</sub> 会合型受容体は、こうした遺伝子の発現の増大を明示することができ、それにより、それに対する反作用薬はこうした発現の減少を明示することができ、また、アゴニストはこうした発現の増大を明示することができる。こうした検出のための商業的に入手可能なアッセイが利用可能である。

### 3. GPCR 融合タンパク質

反作用薬、アゴニストおよび部分的アゴニストの直接の同定のための候補化合物のスクリーニングでの使用のための内在性で、構成的に活性化されるオーファンG P C Rもしくは非内在性で、構成的に活性化されるオーファンG P C Rの使用は興味深いスクリーニングの挑戦を提供し、ここでは当然、それに結合される内在性リガンドの非存在下でさえ該受容体は活性である。従って、こうした化合物が反作用薬、アゴニスト、部分的アゴニストであることができるかどうか、もしくはこうした受容体に対する影響を有することができるかどうかに関しての理解を許すこうした識別の目的をもって、例えば候補化合物の存在下の非内在性受容体とその化合物の非存在下の非内在性受容体とを識別するために、こうした識別を高める可能性のあるアプローチを利用することが好ましい。好ましいアプローチはG P C R融合タンパク質の使用である。

【 0 0 4 4 】

一般に、非内在性のオーファン G P C R が上で示されたアッセイ技術（ならびに他者）を使用して構成的に活性化されていることを決定すれば、内在性 G P C R と共に役する優勢な G タンパク質を決定することが可能である。G P C R への G タンパク質の共役はシグナル伝達経路を提供し、これは評価することが可能である。哺乳動物発現系の使用によりスクリーニングが行われることが最も好ましいため、こうした系はその中に内在性 G タンパク質を有することが期待されよう。従って、当然、こうした系においては、非内在性の構成的に活性化されたオーファン G P C R が連続的にシグナルを発するであろう。この点に関して、例えば受容体に対する反作用薬の存在下で、とりわけスクリーニングに関してそれが反作用薬と接触される場合にそれが受容体をより容易に識別することが可能であろうことがよりありそうであるように、このシグナルを高めることが好ましい。

【 0 0 4 5 】

G P C R 融合タンパク質は、非内在性 G P C R と共に作用する G タンパク質の効能を高めることを意図している。G P C R 融合タンパク質は非内在性の構成的に活性化された G P C R を用いるスクリーニングに好ましい。なぜなら、こうしたアプローチは、こうしたスクリーニング技術で最も好ましく利用されるシグナルを増大させるからである。これは大きな「S/N」比の助長で重要であり；こうした大きな比は本明細書で開示されるような候補化合物のスクリーニングに重要な

つ好ましい。

【0046】

G P C R 融合タンパク質の発現に有用な構築物の構築は当業者の範囲内にある。商業的に入手可能な発現ベクターおよび系が、研究者の特定のニーズに合う可能性のある多様なアプローチを提供する。こうした G P C R 融合タンパク質構築物に対する重要な基準は、内在性の G P C R 配列および G タンパク質配列の双方が同じ読み枠にある（好ましくは内在性 G P C R の配列は G タンパク質配列の上流である）こと、および、 G P C R の発現に際して G タンパク質もまた発現することが可能であるように G P C R の「終止」コドンを欠失もしくは置き換えてはならないことである。 G P C R は G タンパク質に直接連結することができるか、もしくは 2 者の間にスペーサー残基（好ましくは約 12 を越えないが、この数字は当業者により容易に確かめることが可能である）が存在することができる。われわれは、使用されないいくつかの制限部位が発現に際して効果的にスペーサーとなるであろうことに、スペーサーの使用が（便宜性に基づくと）好ましい。最も好ましくは、 G P C R 融合タンパク質構築物の創製に先立ち非内在性 G P C R に共役する G タンパク質を同定することができる。同定された数種の G タンパク質が存在するにすぎないため、その中の内在性の G P C R 配列の挿入に G タンパク質の配列を含んで成る構築物（すなわち普遍的 G タンパク質構築物）が利用可能であることが好ましく；これは、異なる配列を有する多様な異なった内在性 G P C R の大スケールのスクリーニングにおいて効率を提供する。

10

20

30

【0047】

上に示されたとおり、 G i 、 G z および G o に共役する構成的に活性化される G P C R は、これらの型の G P C R に基づくアッセイを挑戦的にする c A M P の形成を阻害すると期待される（すなわち、 c A M P シグナルは活性化に際して減少し、従って例えば（このシグナルをさらに減少させることができる）反作用薬の直接の同定を興味深いものとする）。本明細書に開示されるであろうとおり、われわれは、実現可能なシクラーゼに基づくアッセイを確立する努力において、これらの型の受容体について内在性の G P C R の内在性の G タンパク質に基づかない G P C R 融合タンパク質を創製することが可能であることを確かめた。従って、例えば、 H 9 のような G z 共役型受容体、 G s 融合タンパク質を利用する G P C R 融合タンパク質を確立することが可能である。われわれは、こうした融合構築物は、発現に際して非内在性の G P C R が例え「天然の」 G z タンパク質よりもむしろ G s と共にすることを「誘導」もしくは「強要」し、その結果シクラーゼに基づくアッセイを確立することが可能であると考える。従って、 G i 、 G z および G o 共役型受容体について、われわれは、 G P C R 融合タンパク質を使用しあつアッセイがアデニルシクラーゼ活性の検出に基づく場合は、 G s （もしくは酵素アデニリルシクラーゼの形成を刺激する同等の G タンパク質）を用いて融合構築物を確立することを好む。

F. 医薬品化学

一般に（しかし常にでなく）、候補化合物の直接の同定はコンビナトリアル化学の技術を介して生成される化合物に関連して好ましく実施され、それにより何千もの化合物がこうした分析のため無作為に調製される。一般に、こうしたスクリーニングの結果は独特のコア構造を有する化合物であることができ；その後、これらの化合物は、その医薬特性をさらに高めるため、好ましいコア構造（1種もしくは複数）を取り巻く付加的な化学的改变に好ましくかけられる。こうした技術は当業者に既知でありかつ本特許明細書で詳細に取り扱わないであろう。

40

G. 製薬学的組成物

さらなる開発に選択された候補化合物は、当業者に公知の技術を使用して製薬学的組成物に処方することができる。適する製薬学的に許容できる担体は当業者に利用可能であり；例えば、 Remington's Pharmaceut

50

ical Sciences、第16版、1980、マック パブリッシング  
カンパニー (Mack Publishing Co.) (オスロ (Oslo))  
ら編) を参照されたい。

#### H. 他の利用性

本明細書に開示される非内在性の変種ヒトGPCRの好ましい使用は、(好ましくは製薬学的作用物質としての使用のための)反作用薬、アゴニストもしくは部分的アゴニストとしての候補化合物の直接の同定のためのものであることができるが、これらの変種ヒトGPCRはまた研究の設定でも利用することができる。例えば、GPCRを組み込むインビトロおよびインビボの系は、正常のおよび疾患に罹った双方のヒトの状態でこれらの受容体が演じる役割をさらに解明かつ理解するため、ならびに構成的活性化の役割の理解(それがシグナル伝達カスケードの理解に適用されるため)に利用することができる。非内在性のヒトGPCRの価値は、そのための内在性リガンドが同定される前に人体でのこれらの受容体の役割を理解するのに、それらの独特的の特徴のため非内在性のヒトGPCRを使用することができるることにおいて、研究ツールとしてのそれらの利用性が高められることである。開示される受容体の他の用途は、とりわけ本特許明細書の検討に基づき当業者に明らかとなるであろう。

10

#### 【実施例】

#### 【0048】

以下の実施例は本発明の解明の目的上(そして制限でなく)提示する。本明細書で特定の核酸およびアミノ酸の配列が開示される一方、当業者は、下に報告される同一のもしくは実質的に類似の結果を達成しつつこれらの配列に対する小さな改変を行う能力があると信じられる。1配列から別のものまで(例えばラット受容体からヒト受容体まで、もしくはヒト受容体Aからヒト受容体Bまで)の配列カセットの応用もしくは理解への伝統的アプローチは、配列整列の技術(それにより配列は共通の領域を決定する活動で整列される)に基づく。本明細書に開示される突然変異のアプローチはこのアプローチに頼らないが、しかし、代わりにアルゴリズムのアプローチ、およびヒトGPCRのTM6領域内に配置される保存されたプロリン残基からの位置上の距離に基づく。このアプローチが確実にされれば、当業者は、それに対して小さな変更を行って本明細書に開示される実質的に同一の結果(すなわち構成的活性化)を達成する能力があると信じられる。こうした改変されたアプローチは本開示の範囲内と考えられる。

20

#### 実施例1

#### 内在性のヒトGPCR

##### 1. ヒトGPCRの同定

ジェンバンク [GenBank] (商標)データベース情報の再検討に基づき、開示された内在性ヒトGPCRのあるものを同定した。データベースを検索する間に以下のcDNAクローンを下に明示されるとおり同定した(表C)。

30

#### 【0049】

## 【表3】

表C

開示された ヒトオーファン GPCRs	受託番号	完全なcDNA配列 (塩基対)	読み取り枠 (塩基対)	核酸配列 番号	アミノ酸配列 番号	
hARE-3	AL033379	111,389 bp	1,260 bp	1	2	
hARE-4	AC006087	226,925 bp	1,119 bp	3	4	10
hARE-5	AC006255	127,605 bp	1,104 bp	5	6	
hRUP3	AL035423	140,094 bp	1,005 bp	7	8	
hRUP5	AC005849	169,144 bp	1,413 bp	9	10	
hRUP6	AC005871	218,807 bp	1,245 bp	11	12	
hRUP7	AC007922	158,858 bp	1,173 bp	13	14	

20

## 【0050】

他の開示された内在性ヒトGPCRは、以下のESTクローンをクエリ配列として使用するESTデータベース(dbest)のBLAST(商標)検索を実施することにより同定した。その後、同定された以下のESTクローンをプロープとして使用して、ヒトゲノムライブラリーをスクリーニングした(表D)。

## 【0051】

【表4】

表D

開示されたヒトオーファンGPCRs	ケリ(配列)	同定されたESTクローン/受託番号	読み取り枠(塩基対)	核酸配列番号	アミノ酸配列番号	
hGPCR27	マウスGPCR27	AA775870	1,125 bp	17	18	
hARE-1	TDAG	1689643 AI090920	999 bp	19	20	10
hARE-2	GPCR27	68530 AA359504	1,122 bp	21	22	
hPPR1	ウシPPR1	238667 H67224	1,053 bp	23	24	
hG2A	マウス 1179426	下の実施例2(a)を 参照されたい	1,113 bp	25	26	
hCHN3	N.A.	EST 36581 (完全長)	1,113 bp	27	28	
hCHN4	TDAG	1184934 AA804531	1,077 bp	29	30	
hCHN6	N.A.	EST 2134670 (完全長)	1,503 bp	31	32	20
hCHN8	KIAA0001	EST 764455	1,029 bp	33	34	
hCHN9	1365839	EST 1541536	1,077 bp	35	36	
hCHN10	マウスEST 1365839	ヒ1365839	1,005 bp	37	38	
hRUP4	N.A.	AI307658	1,296 bp	39	40	
		"N.A. = 当てはまらない"				

【0052】

2. 完全長のクローニング

A. ヒトG2A

マウスESTクローン1179426を使用して、3種のアミノ酸G2Aコーディング配列を除く全部を含有するヒトゲノムクローンを得た。このコーディング配列の5'は5' RACEを使用することにより得、また、PCRのための鑄型はクロンテック(Clonitech)のヒト脾マラソン-レディ[Marat hon-Ready](商標)cDNAであった。開示されるヒトG2Aは、以下：

5' - C T G T G T A C A G C A G T T C G C A G A G T G - 3' (配列番号41; 1回目のPCR)

5' - G A G T G C C A G G C A G A G C A G G T A G A C - 3' (配列番号42; 2回目のPCR)

のような配列番号41および配列番号42に示されるような第一回および第二回のPCRのためのG2AのcDNAに特異的なプライマーを使用するPCRにより増幅した。PCRは、94 30秒間、次いで94 5秒間および72 4分間の5周期；ならびに94°5秒間および70°4分間の30周期で、アドバンテージ(Advantage)GCポリメラーゼキット(クロンテック(Clonitech)；製造説明書に従うことができる)を使用して実施した。およそ1.3kbのPCRフラグメントをアガロースゲルから精製し、HindIIIおよびXbaIで消化し、そして発現ベクターpRC/CMV2(インヴィトロジエン(Invitrogen))にクローン化した。T7シークエナーゼ[Sequase](商標)キット(USBアマーシャム(USB Amers

30

40

50

ham) ; 製造元の説明書が従われた) を使用してクローン化された挿入物を配列決定し、そして提示された配列と配列を比較した。P<sup>32</sup>標識されたフラグメントを用いてRNAドットプロット(クロントック(Clontech) ; 製造元の説明書が従われた)をプロービングすることによりヒトG2Aの発現を検出した。

#### b. CHN9

ESTクローン1541536の配列決定は、CHN9が1個の開始コドンのみを有する(すなわち終止コドンは欠けていた)部分的cDNAクローンであることを示した。CHN9をデータベース(nr)に対してBLAST検索する(blast)のに使用した場合、CHN9の3'の配列はロイコトリエンB4受容体のcDNAの5'非翻訳領域(CHN9のコーディング配列と同じ読み枠に終止コドンを含有した)に100%相同であった。LTB4RのcDNAの5'非翻訳領域がCHN9の3'配列であったかどうかを決定するために、CHN9に見出される開始コドンに隣接する5'配列およびLTB4Rの5'非翻訳領域に見出される終止コドンを取り巻く3'配列に基づくプライマーを使用してPCRを実施した。利用された5'プライマー配列は以下のとおりであった：

5' - C C C G A A T T C C T G C T T G C T C C C A G C T T G G C C C - 3' (配列番号43; センス) および

5' - T G T G G A T C C T G C T G T C A A A G G T C C C A T T C C G G - 3' (配列番号44; アンチセンス)

製造元により供給される緩衝液系、0.25μMの各プライマーおよび0.2mMの各4種のスクレオチドとともに鑄型としての胸腺cDNAおよびrTthポリメラーゼ(パーキンエルマー(Perkin Elmer))を使用してPCRを実施した。周期条件は、94 1分間、65 1分間ならびに72 1分および10秒間の30周期であった。予測された大きさと一致する1.1kbのフラグメントをPCRから得た。本PCRフラグメントをpCMVにサブクローンングし(下を参照されたい)そして配列決定した(配列番号35を参照されたい)。

#### c. RUP4

鑄型としてヒト脳cDNA(クロントック(Clontech))を用いるRT-PCRにより完全長のRUP4をクローン化した：

5' - T C A C A A T G C T A G G T G T G G T C - 3' (配列番号45; センス) および

5' - T G C A T A G A C A A T G G G A T T A C A G - 3' (配列番号46; アンチセンス)。

以下の周期、すなわち94 2分間；94 30秒間；55 30秒間、72 45秒間および72 10分間により、タックプラス プレシジョン[Taq Plus Precision](商標)ポリメラーゼ(ストラタジーン(Stratagene)；製造説明書が従われた)を使用してPCRを実施した。周期2から4を30回反復した。

#### 【0053】

PCR産物を1%アガロースゲルで分離し、そして500bpのPCRフラグメントを単離しあつPCR-II-TOPO(商標)ベクター(インヴィトロジエン(Invitrogen))にクローン化し、そしてT7 DNAシークェナーゼ[Sequenase](商標)キット(アマーシャム(Amersham))およびSP6/T7プライマー(ストラタジーン(Stratagene))を使用して配列決定した。PCRフラグメントから示される配列分析は、実際に、他のGPCRとの類似性をもつ1個の連続的読み取り枠を有する、代替スプライシングされた形態のAI307658であった。このPCRフラグメントの完了された配列は以下のとおりであった：

10

20

30

40

50

【0054】

【表5】

5'-TCACAATGCTAGGTGTGGCTGGCTGGCAGTCATCGTAGGATCACCCATGTGGCAC  
 GTGCAACAACCTGAGATCAAATATGACTTCCTATATGAAAAGGAACACATCTGCTGCTTAAGA  
 GTGGACCAGCCCTGTGACCAAGAGATCTACACCACCTTCATCCTGTATCCTCTTCCCTGC  
 CTCTTATGGTGTGCTTATTCTGTACGTAAAATTGGTTATGAACCTTGGATAAAGAAAAGAGTT  
 GGGGATGGTCAGTGCTTCGAACATATTGAAAGAAATGTCCAAAATAGCCAGGAAGAAG  
 AAACGAGCTGTGATTATGATGGTACAGTGGTGGCTCTTGTGTGCTGGGACCCATTCC  
 ATGTTGTCCATATGATGATTGAAACAGTAATTGGAAAAGGAATATGATGATGTACAATCAA  
 GATGATTTTGCTATCGTCAAATTATTGGATTTCACACTCTGTAATCCATTGTCTATGCA-  
 3' (配列番号:47)

10

【0055】

上の配列に基づき、2種のセンスオリゴヌクレオチドプライマーの組：

5' - C T G C T T A G A A G A G T G G A C C A G - 3' (配列番号48；オリゴ1)、

5' - C T G T G C A C C A G A A G A T C T A C A C - 3' (配列番号49；オリゴ2) および

2種のアンチセンスオリゴヌクレオチドプライマーの組：

5' - C A A G G A T G A A G G T G G T G T A G A - 3' (配列番号50；オリゴ3)

5' - G T G T A G A T C T T C T G G T G C A C A G G - 3' (配列番号51；オリゴ4)

を、製造元の説明書に従って鋳型としてヒト脳マラソン-レディ [Marathon-Ready] (商標) cDNA (クロンテック (Clontech)、カタログ番号7400-1) を用いる3' - および5' - RACE PCR に使用した。RACE PCR により生成されたDNAフラグメントをpCRII-TOP (商標) ベクター (インヴィトロジエン (Invitrogen)) にクローン化し、そしてSP6/T7プライマー (ストラタジーン (Stratagene)) および数種の内的プライマーを使用して配列決定した。3' RACE 産物はポリ(A)テール (tail) およびTAA終止コドンで終了する1個の完了される読み取り棒を含有した。5' RACE 産物は不完全な5'末端を含有した (すなわちATG開始コドンが存在しなかった)。

20

30

30

【0056】

新たな5'配列に基づき、オリゴ3および以下のプライマー：

5' - G C A A T G C A G G T C A T A G T G A G C - 3' (配列番号52；オリゴ5)

を第2回の5'race PCRに使用し、そしてPCR産物を上のとおり分析した。第3回の5'race PCRは、アンチセンスプライマー：

5' - T G G A G C A T G G T G A C G G G A A T G C A G A A G - 3' (配列番号53；オリゴ6) および

5' - G T G A T G A G C A G G T C A C T G A G C G C C A A G - 3' (配列番号54；オリゴ7)

40

を利用して実施した。5' RACE PCRの産物の配列は開始コドンATGの存在を示し、また、さらなる回の5'race PCRはいかなるそれ以上の5'配列も生成しなかった。プライマーとして、センスプライマー

5' - G C A A T G C A G G C G T T A A C A T T A C - 3' (配列番号55；オリゴ8)

およびオリゴ4を使用するRT-PCR、ならびにヒト脳および心のcDNA鋳型 (クロンテック (Clontech)、カタログ番号7404-1) から生成された650bpのPCR産物の配列分析により、完了された5'配列を確認し

50

た。オリゴ2および以下のアンチセンスプライマー：

5' - T T G G G T T A C A A T C T G A A G G G C A - 3' (配列番号56;  
オリゴ9)

を使用するR T - P C R、ならびにヒト脳および心のc D N A 鑄型(クロンテック(C l o n t e c h)、カタログ番号7404-1)から生成された670b pのP C R産物の配列分析により、完了された3'配列を確認した。

d. R U P 5

以下の配列：

5' - A C T C C G T G T C C A G C A G G A C T C T G - 3' (配列番号57)  
)

5' - T G C G T G T T C C T G G A C C C T C A C G T G - 3' (配列番号58)

を有した、A T G 開始コドンから上流のセンスプライマー(配列番号57)、および終止コドンとしてT C Aを含有するアンチセンスプライマー(配列番号58)、ならびに鑄型としてヒト末梢白血球c D N A(クロンテック(C l o n t e c h))を使用するR T - P C Rにより完全長のR U P 5をクローニングした。段階2から段階4が30回反復された以下の周期、すなわち94 30秒間; 94

15秒間; 69 40秒間; 72 3分間; および72 6分間による50 μ 1の反応中での増幅にアドバンテージ[Ad v a n t a g e](商標)c D N Aポリメラーゼ(クロンテック(C l o n t e c h))を使用した。1.4k bのP C Rフラグメントを単離し、そしてp C R I I - T O P O(商標)ベクター(インヴィトロジエン(Invitrogen))を用いてクローニングし、そしてT 7 D N A シーケナーゼ[Se quenase](商標)キット(アマーシャム(A m e r s h a m))を使用して完全に配列決定した。配列番号9を参照されたい。

e. R U P 6

プライマー：

5' - C A G G C C T T G G A T T T A A T G T C A G G G A T G G - 3' (配列番号59) および

5' - G G A G A G T C A G C T C T G A A A G A A T T C A G G - 3' (配列番号60)

ならびに鑄型としてヒト胸腺マラソン-レディ[Marathon-Ready](商標)c D N A(クロンテック(C l o n t e c h))を使用するR T - P C Rにより完全長のR U P 6をクローニングした。以下の周期、すなわち94 30秒間; 94 5秒間; 66 40秒間; 72 2.5秒間および72 7分間による50 μ 1反応中での増幅にアドバンテージ(Ad v a n t a g e)c D N Aポリメラーゼ(クロンテック(C l o n t e c h))、製造元の説明書に従う)を使用した。周期2から4を30回反復した。1.3k bのP C Rフラグメントを単離し、そしてp C R I I - T O P O[商標]ベクター(インヴィトロジエン(Invitrogen))にクローニングし、そしてA B I ビッグダイターミネーター[Big Dye Terminator](商標)キット(P. E.バイオシステム(P. E. B i o s y s t e m))を使用して完全に配列決定した(配列番号11を参照されたい)。

f. R U P 7

プライマー：

5' - T G A T G T G A T G C C A G A T A C T A A T A G C A C - 3' (配列番号61; センス) および

5' - C C T G A T T C A T T T A G G T G A G A T T G A G A C - 3' (配列番号62; アンチセンス)

ならびに鑄型としてヒト末梢白血球c D N A(クロンテック(C l o n t e c h))

10

20

30

40

50

) ) を使用する R T - P C R により完全長の R U P 7 をクローン化した。段階 2 ないし段階 4 が 30 回反復された以下の周期、すなわち 94 2 分間； 94 15 秒間； 60 20 秒間； 72 2 分間； 72 10 分間による 50  $\mu$  l 反応中の増幅にアドバンテージ [ A d v a n t a g e ] (商標) c D N A ポリメラーゼ (クロントック ( C l o n t e c h )) を使用した。1.25 kb の P C R フラグメントを単離し、そして p C R I I - T O P O (商標) ベクター (インヴィトロジエン ( I n v i t r o g e n )) にクローン化し、そして A B I ビッグダイターミネーター [ B i g D y e T e r m i n a t o r ] (商標) キット ( P . E . バイオシステム ( P . E . B i o s y s t e m )) を使用して完全に配列決定した。配列番号 13 を参照されたい。

### 3. アンジオテンシン II タイプ 1 受容体 (「 A T 1 」)

製造元により供給される緩衝液系、0.25  $\mu$  M の各プライマーおよび 0.2 mM の各 4 種のヌクレオチドとともに錆型としてのゲノム D N A および r T t h ポリメラーゼ (パークリン エルマー ( P e r k i n E l m e r )) を使用する P C R により、内在性のヒトアンジオテンシン II タイプ 1 受容体 (「 A T 1 」) を得た。周期条件は、94 1 分間、55 1 分間および 72 1.5 分間の 30 周期であった。5' P C R プライマーは、配列：

5' - C C C A A G C T T C C C C A G G T G T A T T T G A T - 3' (配列番号 63)

とともに H i n d I I I 部位を含有し、そして 3' プライマーは以下の配列：

5' - G T T G G A T C C A C A T A A T G C A T T T T C T C - 3' (配列番号 64)

とともに B a m H I 部位を含有する。生じる 1.3 kb の P C R フラグメントを H i n d I I I および B a m H I で消化し、そして p C M V 発現ベクターの H i n d I I I - B a m H I 部位にクローン化した。c D N A クローンを完全に配列決定した。その後、ヒト A T 1 の核酸 (配列番号 65) およびアミノ酸 (配列番号 66) の配列を決定しつつ確かめた。

### 4. G P R 38

G P R 38 を得るため、製造元により供給される緩衝液系、0.25  $\mu$  M の各プライマーおよび 0.2 mM の各 4 種のヌクレオチドとともに錆型としてのヒトゲノム c D N A および r T t h ポリメラーゼ (パークリン エルマー ( P e r k i n E l m e r )) を使用して、2 種の P C R フラグメントを組み合わせることにより P C R を実施した。各 P C R 反応の周期条件は 94 1 分間、62 1 分間および 72 2 分間の 30 周期であった。

### 【 0057 】

第一のフラグメントは、以下の配列：

5' - A C C A T G G G C A G C C C C T G G A A C G G C A G C - 3' (配列番号 67)

を伴う端部位を含有した 5' P C R プライマー、および以下の配列：

5' - A G A A C C A C C A C C A G C A G G A C G C G G A C G G T C T G C C G G T G G - 3' (配列番号 68)

を有する 3' プライマーを用いて増幅した。第二の P C R フラグメントは、以下の配列：

5' - G T C C G C G T C C T G C T G G T G G T G G T T C T G G C A T T T A T A A T T - 3' (配列番号 69)

を有する 5' プライマー、ならびに B a m H I 部位を含有しつつ以下の配列：

5' - C C T G G A T C C T T A T C C C A T C G T C T T C A C G T T A G C - 3' (配列番号 70)

を有する 3' プライマーを用いて増幅した。配列番号 67 および配列番号 70 をプライマーとして使用して (上に示された周期条件を使用する) 、2 種のフラグ

10

20

30

40

50

メントを鋳型として使用して G P R 3 8 を増幅した。生じる 1 . 4 4 k b の P C R フラグメントを B a m H I で消化し、そして p C M V 発現ベクターの平滑 B a m H I 部位にクローン化した。

#### 5 . M C 4

M C 4 を得るため、製造元により供給される緩衝液系、0 . 2 5  $\mu$  M の各プライマーおよび 0 . 2 mM の各 4 種のヌクレオチドとともに鋳型としてのヒトゲノム c D N A および r T t h ポリメラーゼ ( パーキン エルマー ( P e r k i n E l m e r ) ) を使用して P C R を実施した。各 P C R 反応の周期条件は 9 4 1 分間、5 4 1 分間および 7 2 1 . 5 分間の 3 0 周期であった。

#### 【 0 0 5 8 】

5 ' P C R は配列 :

5 ' - C T G G A A T T C T C C T G C C A G C A T G G T G A - 3 ' ( 配列番号 7 1 )

とともに E c o R I 部位を含有し、そして 3 ' プライマーは配列 :

5 ' - G C A G G A T C C T A T A T T G C G T G C T C T G T C C C C - 3 ' ( 配列番号 7 2 )

とともに B a m H I 部位を含有した。1 . 0 k b の P C R フラグメントを E c o R I および B a m H I で消化し、そして p C M V 発現ベクターの E c o R I - B a m H I 部位にクローン化した。その後、ヒト M C 4 の核酸 ( 配列番号 7 3 ) およびアミノ酸 ( 配列番号 7 4 ) の配列を決定した。

#### 6 . C C K B

C C K B を得るため、製造元により供給される緩衝液系、0 . 2 5  $\mu$  M の各プライマーおよび 0 . 2 mM の各 4 種のヌクレオチドとともに鋳型としてのヒト胃 c D N A および r T t h ポリメラーゼ ( パーキン エルマー ( P e r k i n E l m e r ) ) を使用して P C R を実施した。各 P C R 反応の周期条件は 9 4 1 分間、6 5 1 分間ならびに 7 2 1 分および 3 0 秒間の 3 0 周期であった。

#### 【 0 0 5 9 】

5 ' P C R は配列 :

5 ' - C C G A A G C T T C G A G C T G A G T A A G G C G G C G G G C T - 3 ' ( 配列番号 7 5 )

とともに H i n d I I I 部位を含有し、また、3 ' プライマーは配列 :

5 ' - G T G G A A T T C A T T T G C C C T G C C T C A A C C C C C A - 3 ' ( 配列番号 7 6 )

とともに E c o R I 部位を含有した。生じる 1 . 4 4 k b の P C R フラグメントを H i n d I I I および E c o R I で消化し、そして p C M V 発現ベクターの H i n d I I I - E c o R I 部位にクローン化した。その後、ヒト C C K B の核酸 ( 配列番号 7 7 ) およびアミノ酸 ( 配列番号 7 8 ) の配列を決定した。

#### 7 . T D A G 8

T D A G 8 を得るため、製造元により供給される緩衝液系、0 . 2 5  $\mu$  M の各プライマーおよび 0 . 2 mM の各 4 種のヌクレオチドとともに鋳型としてのゲノム D N A および r T t h ポリメラーゼ ( パーキン エルマー ( P e r k i n E l m e r ) ) を使用して P C R を実施した。周期条件は 9 4 1 分間、5 6 1 分間ならびに 7 2 1 分および 2 0 秒間の 3 0 周期であった。5 ' P C R プライマーは以下の配列 :

5 ' - T G C A A G C T T A A A A A G G A A A A A T G A A C A G C - 3 ' ( 配列番号 7 9 )

とともに H i n d I I I 部位を含有し、また、3 ' プライマーは以下の配列 :

5 ' - T A A G G A T C C C T T C C C T T C A A A A C A T C C T T G - 3 ' ( 配列番号 8 0 )

とともに B a m H I 部位を含有した。生じる 1 . 1 k b の P C R フラグメントを

10

20

30

40

50

Hind III および BamHI で消化し、そして pCMV 発現ベクターの Hind III - BamHI 部位にクローン化した。配列決定された 3 種の生じるクローンは、Pro から Ala へのアミノ酸 43、Lys から Asn へのアミノ酸 97、および Ile から Phe へのアミノ酸 130 の変化を伴う 3 種の潜在的多形を含有した。その後、ヒト T D A G 8 の核酸（配列番号 81）およびアミノ酸（配列番号 82）の配列を決定した。

#### 8. H9

H9を得るため、製造元により供給される緩衝液系、0.25 μM の各プライマーおよび 0.2 mM の各 4 種のヌクレオチドとともに鑄型としての下垂体 cDNA および rTth ポリメラーゼ（パーキン エルマー（Perkin Elmer））を使用して PCR を実施した。周期条件は 94 1 分間、62 1 分間および 72 2 分間の 30 周期であった。5' PCR プライマーは以下の配列：5' - G G A A A G C T T A A C G A T C C C C A G G A G C A A C A T - 3'（配列番号 15）

とともに Hind III 部位を含有し、また、3' プライマーは以下の配列：5' - C T G G G A T C C T A C G A G A G C A T T T T T C A C A C A G - 3'（配列番号 16）

とともに BamHI 部位を含有した。生じる 1.9 kb の PCR フラグメントを Hind III および BamHI で消化し、そして pCMV 発現ベクターの Hind III - BamHI 部位にクローン化した。H9 は、アミノ酸 P320S、S493N およびアミノ酸 G448A の変化を伴う 3 種の潜在的多形を含有した。その後、ヒト H9 の核酸（配列番号 139）およびアミノ酸（配列番号 140）の配列を決定しつつ確かめた。

#### 実施例 2

##### 非内在性の構成的に活性化される GPCR の調製

当業者は、核酸配列の突然変異のための技術を選択する能力があると信じられる。上に開示された非内在性の変種数種のヒト GPCR を創製するのに利用されたアプローチを下に提示する。下に開示される突然変異はアルゴリズムのアプローチに基づき、それにより（TM6 / IC3 の界面近くの GPCR の TM6 領域に配置される）保存されるプロリン残基からの（GPCR の IC3 領域に配置される）16 番目のアミノ酸が、最も好ましくはリシンアミノ酸残基に突然変異される。

##### 1. トランスフォーマー部位特異的 [Transformer Site-Directed]（商標）突然変異誘発

非内在性のヒト GPCR の調製は、製造元の説明書に従い、トランスフォーマー部位特異的 [Transformer Site-Directed]（商標）突然変異誘発キット（クロンテック（Clontech））を使用してヒト GPCR で達成することができる。2 種の突然変異誘発プライマー、最も好ましくはリシン突然変異を創製するリシン突然変異誘発オリゴヌクレオチド、および選択マーカーオリゴヌクレオチドを利用する。便宜上、ヒト GPCR に組み込まれるべきコドン突然変異を標準的形態でもまた示す（表 E）：

【0060】

10

20

30

40

【表6】

表E

レセプター識別名	コトン突然変異	
hARE-3	F313K	
hARE-4	V233K	
hARE-5	A240K	
hGPCR14	L257K	
hGPCR27	C283K	
hARE-1	E232K	
hARE-2	G285K	
hPPR1	L239K	
hG2A	K232A	
hRUP3	L224K	
hRUP5	A236K	
hRUP6	N267K	
hRUP7	A302K	
hCHN4	V236K	
hMC4	A244K	
hCHN3	S284K	
hCHN6	L352K	
hCHN8	N235K	
hCHN9	G223K	
hCHN10	L231K	
hH9	F236K	
		10
		20

【0061】

指定される配列プライマーを使用して、上の方法に従い、以下の G P C R を突然変異した（表F）。

【0062】

【表7】

表F

30

レセプター 識別名	コトン突然 変異	リン突然変異 (配列番号) 5'-3'の向き 突然変異配列に 下線をつけられる	選択マークー (配列番号) 5'-3'の向き
hRUP4	V272K	CAGGAAGAAG <u>AAAC</u> CGAGC TGTCA <u>TTATGATGGTGACA</u> GTG (83)	CACTGTCACCATCATAATG ACAGCTCGTTCTTCTTCC TG (84)
hAT1 hGPR38	下を参照 V297K	代替アプローチ；下を参照 GGCCACCGGCAGAC <u>CCAAAC</u> GCGTCCTGCTG (85)	代替アプローチ；下を参照 CTCCTTCGGTCCTCCTATC GTTGTCAGAAGT (86)
hCCKB hTDAG8	V332K I225K	代替アプローチ；下を参照 GGAAAAGAAGAGAATCAA <u>AAA</u> ACTACTTGTCA <u>GCACATC</u> (87)	代替アプローチ；下を参照 CTCCTTCGGTCCTCCTATC GTTGTCAGAAGT (88)
hH9	F236K	GCTGAGGTT <u>CGCAATAAAC</u> TAACCATGTTGTG (143)	CTCCTTCGGTCCTCCTATC GTTGTCAGAAGT (144)
hMC4	A244K	GCCAATATGAAGGGAAAA ATTACCTTGACCATC (137)	CTCCTTCGGTCCTCCTATC GTTGTCAGAAGT (138)
			40

【0063】

その後、非内在性のヒト G P C R を配列決定し、また、形成されかつ確かめられた核酸およびアミノ酸の配列を、下に表Gに要約されるとおり、本特許明細書への付随する「配列表」付録に列挙する：

50

【0064】

【表8】

表G

非内在的ヒト GPCR hRUP4 (V272K) hAT1 (下の代替アプローチを 参照) hGPR38 (V297K) hCCKB (V332K) HTDAG8 (I225K) hH9 (F236K) hMC4 (A244K)	核酸配列表 配列番号：127 (下の代替アプローチを 参照) 配列番号：129 配列番号：131 配列番号：133 配列番号：141 配列番号：135	アミノ酸配列表 配列番号：128 (下の代替アプローチを 参照) 配列番号：130 配列番号：132 配列番号：134 配列番号：142 配列番号：136	10 20

【0065】

2. 非内在性ヒトG P C Rの創製のための代替アプローチ

a. A T 1

1. F 2 3 9 K 突然変異

F 2 3 9 突然変異（核酸配列について配列番号89、およびアミノ酸配列について配列番号90を参照されたい）を創製することにより、非内在性の構成的に活性化されるヒトA T 1受容体の調製を達成した。突然変異誘発は、製造元の説明書に従い、トランスフォーマー部位特異的突然変異誘発 [T r a n s f o r m e r S i t e - D i r e c t e d M u t a g e n e s i s]（商標）キット（クロンテック（C l o n t e c h））を使用して実施した。2種の突然変異誘発プライマー、リシン突然変異誘発オリゴヌクレオチド（配列番号91）および選択マーカーオリゴヌクレオチド（配列番号92）を使用した。これらはそれぞれ以下の配列を有した：

5' - C C A A G A A A T G A T G A T A T T A A A A A G A T A A T T A T G  
G C - 3' (配列番号91)

5' - C T C C T T C G G T C C T C C T A T C G T T G T C A G A A G T - 3  
' (配列番号92)

2. N 1 1 1 A 突然変異

非内在性のヒトA T 1受容体の調製もまた、N 1 1 1 A 突然変異（核酸配列について配列番号93、およびアミノ酸配列について配列番号94を参照されたい）を創製することにより達成した。10% D M S O、0.25 μMの各プライマー、および0.5 mMの各4種のヌクレオチドを補充された、製造元により提供される緩衝液系を用いてp f uポリメラーゼ（ストラタジーン（S t r a t a g e n e））を使用して、2回のP C R反応を実施した。使用された5' P C Rのセンスプライマーは以下の配列：

5' - C C C A A G C T T C C C C A G G T G T A T T T G A T - 3' (配列番号95)

を有し、また、アンチセンスプライマーは以下の配列：

5' - C C T G C A G G C G A A A C T G A C T C T G G C T G A A G - 3' (配列番号96)

30

40

50

を有した。生じる 400 bp の PCR フラグメントを Hind III 部位で消化し、そして pCMV ベクターの Hind III - Sma I 部位にサブクローニングした (5' 構築物)。使用された 3' PCR のセンスプライマーは以下の配列：

5' - C T G T A C G C T A G T G T G T T C T A C T C A C G T G T C T C  
A G C A T T G A T - 3' (配列番号 97)

を有し、また、アンチセンスプライマーは以下の配列：

5' - G T T G G A T C C A C A T A A T G C A T T T C T C - 3' (配列番号 98)

を有した。生じる 880 bp の PCR フラグメントを Bam HI で消化し、そして 5' 構築物の Pst (T4 ポリメラーゼにより平滑化された) および Bam HI 部位に挿入して完全長の N111A 構築物を生成した。周期条件は、94 1 分間、60 1 分間および 72 1 分間 (5' PCR) もしくは 1.5 分間 (3' PCR) の 25 周期であった。

### 3. AT2K255IC3 突然変異

AT2K255IC3 の「ドメインスワップ(domain swap)」突然変異 (核酸配列について配列番号 99 を、およびアミノ酸配列について配列番号 100 を参照されたい) を創製することにより、非内在性の構成的に活性化されるヒト AT1 の調製を達成した。AT1 の IC3 に隣接する制限部位を生成させて、アンジオテンシン II タイプ 2 受容体 (AT2) からの対応する IC3 での IC3 の置き換えを助長した。これは 2 回の PCR 反応を実施することにより達成した。センスプライマーとして配列番号 63 を、およびアンチセンスプライマーとして以下の配列：

5' - T C C G A A T T C C A A A A T A A C T T G T A A G A A T G A T C A  
G A A A - 3' (配列番号 101)

を利用することにより、5' 非翻訳領域から IC3 の開始までをコードする 5' PCR フラグメント (フラグメント A) を生成させた。センスプライマーとして以下の配列：

5' - A G A T C T T A A G A A G A T A A T T A T G G C A A T T G T G C T  
- 3' (配列番号 102)

およびアンチセンスプライマーとして配列番号 64 を使用することにより、IC3 の終了から 3' 非翻訳領域までをコードする 3' PCR フラグメント (フラグメント B) を生成させた。PCR 条件は、10% DMSO、0.25 μM の各プライマーおよび 0.5 mM の各 4 種のヌクレオチドを補充された、製造元により供給される緩衝液系を用いて、鋳型としての内在性の AT1 cDNA クローンおよび pfu ポリメラーゼ (ストラタジーン (Stratagene)) を使用して、94 1 分間、55 1 分間および 72 1.5 分間の 30 周期であった。フラグメント A (720 bp) を Hind III および Eco RI で消化しそしてサブクローニングした。フラグメント B は Bam HI で消化し、そしてクローニングされた PCR フラグメントの 5' に Eco RI 部位をもつ pCMV ベクターにサブクローニングした。

### 【0066】

以下の配列：

5' A A T T C G A A A A C A C T T A C T G A A G A C G A A T A G C T A T  
G G G A A G A A C A G G A T A A C C C G T G A C C A A G - 3' (センス；  
配列番号 103)

5' T T A A C T T G G T C A C G G G T T A T C C T G T T C T T C C C A T  
A G C T A T T C G T C T T C A G T A A G T G T T T C G - 3' (アンチセンス；配列番号 104)

を有する 2 種の合成オリゴヌクレオチドをアニーリングすることにより、L25

10

20

20

30

40

50

5'K突然変異を伴いAT2のIC3をコードしそして5'にEcoRI付着端および3'にAfeIII付着端を含有するDNAフラグメント(フラグメントC)を生成させた。

【0067】

フラグメントCを、EcoRIおよびAfeIII部位によりフラグメントBの前に挿入した。その後、生じるクローンをEcoRI部位によりフラグメントAと連結して、AT2K255IC3をもつAT1を生成させた。

4. A243+突然変異

A243+突然変異(核酸配列について配列番号105、およびアミノ酸配列について配列番号106を参照されたい)を創製することにより、非内在性のヒトAT1受容体の調製もまた達成した。A243+突然変異は以下のPCRに基づく戦略を使用して構築した。すなわち、10%DMSO、0.25μMの各プライマーおよび0.5mMの各4種のスクレオチドを補充された、製造元により提供される緩衝液系とともにpfuポリメラーゼ(ストラタジーン(Stratagene))を使用して2回のPCR反応を実施した。利用された5'PCRのセンスプライマーは以下の配列：

5' - C C C A A G C T T C C C C A G G T G T A T T T G A T - 3' (配列番号107)

を有し、また、アンチセンスプライマーは以下の配列：

5' - A A G C A C A A T T G C T G C A T A A T T A T C T T A A A A A T A T C A T C - 3' (配列番号108)

を有した。利用された3'PCRのセンスプライマーは、Ala挿入を含有する以下の配列：

5' - A A G A T A A T T A T G G C A G C A A T T G T G C T T T C T T T T C T T T - 3' (配列番号109)

そしてアンチセンスプライマー：

5' - G T T G G A T C C A C A T A A T G C A T T T T C T C - 3' (配列番号110)

を有した。周期条件は、94 1分間、54 1分間および72 1.5分間の25周期であった。その後、5'および3'のPCRのアリコートを共鑄型(co-template)として使用して、5'PCRのセンスプライマーおよび3'PCRのアンチセンスプライマーを使用する二次的PCRを実施した。PCR条件は、伸長時間が2.5分であったことを除き一次PCRと同一であった。生じるPCRフラグメントをHindIIIおよびBamHIで消化し、そしてpCMVベクターにサブクローニングした(配列番号105を参照されたい)。

4. CCKB

V322K突然変異(核酸配列について配列番号111、およびアミノ酸配列について配列番号112を参照されたい)を創製することにより、非内在性の構成的に活性化されるヒトCCKB受容体の調製を達成した。突然変異誘発は、実施例1からの野性型CCKBを使用する増幅を介してPCRにより実施した。

【0068】

配列番号75、およびV322K突然変異を含んで成るアンチセンスプライマー：

5' - C A G C A G C A T G C G C T T C A C G C G C T T C T T A G C C C A G - 3' (配列番号113)

を使用することにより第一のPCRフラグメント(1kb)を増幅した。V322K突然変異を含んで成るセンスプライマー：

5' - A G A A G C G C G T G A A G C G C A T G C T G C T G G T G A T C G T T - 3' (配列番号114)および配列番号76を使用することにより第二のPCRフラグメント(0.44kb)を増幅した。その後、配列番号75および

10

20

30

40

50

配列番号 7 6、ならびに上で示された系および条件を使用して、V 3 3 2 K を含んで成る C C K B を増幅するための鑄型として 2 種の生じる P C R フラグメントを使用した。V 3 3 2 K 突然変異を含有する生じる 1 . 4 4 k b の P C R フラグメントを H i n d I I I および E c o R I で消化し、そして p C M V 発現ベクターの H i n d I I I - E c o R I 部位にクローン化した（配列番号 1 1 1 を参照されたい）。

3 . クイックチェンジ [ Q u i k C h a n g e ] ( 商標 ) 部位特異的 [ S i t e - D i r e c t e d ] ( 商標 ) 突然変異誘発

クイックチェンジ [ Q u i k C h a n g e ] ( 商標 ) 部位特異的 [ S i t e - D i r e c t e d ] ( 商標 ) 突然変異誘発キット（ストラタジーン（ S t r a t a g e n e ）、製造元の説明書に従う）を使用することにより、非内在性のヒト G P C R の調製もまた達成することができる。内在性 G P C R を鑄型として好ましく使用し、また、2 種の突然変異誘発プライマー、ならびに最も好ましくはリシン突然変異誘発オリゴヌクレオチドおよび選択マーカーオリゴヌクレオチド（キットに包含される）を利用する。便宜上、ヒト G P C R に組み込まれたコドンの突然変異およびそれぞれのオリゴヌクレオチドを標準的形態で示す（表 H ）：

【 0 0 6 9 】

【 表 9 】

表 H

10

20

受容体 識別名	コドン突然 変異	リシン突然変異誘発 (配列番号) 5'-3'の向き、 突然変異が 下線をつけられる	選択マーカー (配列番号) 5'-3'の向き
hCHN3	S284K	ATGGAGAAAAGAAT <u>CAAAAGAA</u> TGTTCTATATA (115)	TATATAGAACATTCTTTT GATTCTTTCTCCAT (116)
hCHN6	L352K	CGCTCTCTGGCCTT <u>GAAGCGCAC</u> GCTCAGC (117)	GCTGAGCGTGCGCTTCA AGGCCAGAGAGCG (118)
hCHN8	N235K	CCCAGGAAAAGGT <u>GAAGTCA</u> AAGTTTTC (119)	GAAAACTTGACTTCAC CTTTTCCCTGGG (120)
hCHN9	G223K	GGGGCGCGGT <u>GAAACGGCTGG</u> TGAGC (121)	GCTCACCAAGCCGTTCA CCCGCGCCCC (122)
hCHN10	L231K	CCCCTT <u>GAAAGCCTAAGAACTT</u> GGTCATC (123)	GATGACCAAGTTCTTAG GCTTTCAAGGGG (124)

30

【 0 0 7 0 】

実施例 3

受容体の発現

タンパク質の発現のために多様な細胞が当該技術に使用可能であるが、哺乳動物細胞を利用することが最も好ましい。これの主要な理由は実地的なことに基づく。すなわち、例えば G P C R の発現のための酵母細胞の利用が可能な一方で、受容体共役型の遺伝子的機構を包含しなくてもよい（事実、酵母の場合は包含しない）非哺乳動物細胞および哺乳動物の系について発展した分泌経路をプロトコルに導入し、従って、非哺乳動物細胞で得られる結果は、潜在的に有用な一方で哺乳動物細胞から得られるものと同じくらい好ましくはない。哺乳動物細胞のうち、 C O S - 7 、 2 9 3 および 2 9 3 T 細胞がとりわけ好ましいが、利用される特定の哺乳動物細胞は当業者の特定のニーズに基づくことが可能である。

40

【 0 0 7 1 】

第 1 日に、 1 5 0 m m プレートあたり  $1 \times 1 0^7$  個の 2 9 3 T 細胞をプレート培養した。第 2 日に 2 本の反応チューブを準備した（各チューブについて後に続

50

く比率はプレート1枚あたりである）。すなわち、チューブAは、1.2mlの血清を含まないD M E M（アーヴィンサイエンティフィック（I r v i n e S c i e n t i f i c）、カリフォルニア州アーヴィン）中に20μgのDNA（例えばp C M Vベクター；受容体c D N Aを含むp C M Vベクター、など）を混合することにより準備し；チューブBは1.2mlの血清を含まないD M E M中に120μlのリポフェクタミン（ギブコ（G i b c o）B R L）を混合することにより準備した。チューブAおよびBは反転（数回）により混合し、次いで室温で30～45分間インキュベートした。混合状態を「トランスフェクション混合物」と称する。プレート培養された293T細胞を1×P B Sで洗浄し、次いで10mlの血清を含まないD M E Mを添加した。細胞に2.4mlのトランスフェクション混合物を添加し、次いで37/5%CO<sub>2</sub>で4時間インキュベートした。トランスフェクション混合物を吸引により除去し、次いで25mlのD M E M/10%ウシ胎児血清を添加した。細胞を37/5%CO<sub>2</sub>でインキュベートした。72時間のインキュベーション後に細胞を収穫し、そして分析に利用した。

#### 実施例4

##### 非内在性G P C Rの構成的活性の測定のためのアッセイ

非内在性のヒトG P C Rの構成的活性の評価には多様なアプローチが利用可能である。以下は具体的説明であり；当業者は、当業者のニーズに優先的に利益をもたらす技術を決定する能力があると信じられる。

##### 1. 膜結合アッセイ：[<sup>35</sup>S]G T P Sアッセイ

Gタンパク質共役型受容体がリガンド結合もしくは構成的活性化のいずれかの結果としてその活性状態にある場合、受容体はGタンパク質に共役しそしてG D Pの放出およびG T PのGタンパク質へのその後の結合を刺激する。Gタンパク質-受容体複合体のサブユニットがG T Pアーゼとして作用し、そしてG T PをG D Pにゆっくりと加水分解し、この点で受容体が通常非活性化される。構成的に活性化された受容体はG D PをG T Pと交換し続ける。構成的に活性化される受容体を発現する膜への[<sup>35</sup>S]G T P Sの高められた結合を立証するのに、加水分解不可能なG T P類似物[<sup>35</sup>S]G T P Sを利用することができます。構成的活性化を測定するための[<sup>35</sup>S]G T P S結合の使用の利点は：(a)それが全部のGタンパク質共役型受容体に包括的に応用可能である；(b)それは膜表面の近接であり、細胞内カスケードに影響を及ぼす分子を拾い上げることをより少なくありそうにすることである。

##### 【0072】

該アッセイは、直接的に関連する受容体を発現する膜への[<sup>35</sup>S]G T P Sの結合を刺激するGタンパク質共役型受容体の能力を利用する。従って、該アッセイは、既知の、オーファンのおよび構成的に活性化されるGタンパク質共役型受容体に対する候補化合物をスクリーニングするための直接同定法で使用することができる。該アッセイは包括的であり、また、全部のGタンパク質共役型受容体での薬物発見に対する応用性を有する。

##### 【0073】

[<sup>35</sup>S]G T P Sアッセイは、20mM H E P E S、および1mMと約20mMとの間のM g C l<sub>2</sub>（この量は結果の至適化のため調節することができるが、20mMが好ましい）p H 7.4、約0.3nMと約1.2nMとの間の[<sup>35</sup>S]G T P S（この量は結果の至適化のため調節することができるが、1.2が好ましい）および12.5ないし75μgの膜タンパク質（例えば受容体を発現するC O S - 7細胞；この量は至適化のため調節することができるが、75μgが好ましい）および1μM G D P（この量は至適化のため変更することができる）を含む結合緩衝液中で1時間インキュベートすることが可能である。その後、コムギ胚芽アグルチニンビーズ（25μl；アマーシャム（A m e r s h

10

20

30

40

50

a m ) ) を添加し、そして混合物を室温で別の 30 分間インキュベートする。その後、チューブを  $1500 \times g$  で室温で 5 分間遠心分離し、そしてその後シンチレーション計数器で計数する。

【 0074 】

より少なく高価なしかし同等に応用可能な代替物が同定されており、これはまた大スケールのスクリーニングのニーズにも合致する。フラッシュプレート [ F l a s h p l a t e ] (商標) およびワラック [ W a l l a c ] (商標) シンチストリップを利用して、高スループットの [ $^{35}$  S] G T P S 結合アッセイの全体構成を定めることができる。さらに、この技術を使用すれば、 [ $^{35}$  S] G T P S 結合を介して効力をモニターするのと同一時点で受容体へのトリチウム化されたりガンドの結合を同時にモニターするために、該アッセイを既知の G P C R に利用することができる。トリチウム標識プローブおよび  $^{35}$  S 標識プローブの双方を見るようにワラック ( W a l l a c ) ベータ計数器がエネルギーインドウを切り替えることができるために、これが可能である。このアッセイはまた、受容体の活性化をもたらす他の型の膜活性化事象を検出するのにも使用してよい。例えば、該アッセイは多様な受容体 ( G タンパク質共役型受容体およびチロシンキナーゼ受容体の双方 ) の  $^{32}$  P ホスホリル化をモニターするのに使用してよい。膜がウェルの底部に遠心分離されている場合、結合された [ $^{35}$  S] G T P S もしくは  $^{32}$  P - ホスホリル化された受容体が、ウェルの被覆されているシンチラントを活性化することができる。この原理を立証するのにシンチ [ S c i n t i ] (商標) ストリップ (ワラック ( W a l l a c ) ) が使用されている。加えて、該アッセイはまた、放射活性に標識されたリガンドを使用して受容体へのリガンド結合を測定するための有用性も有する。類似の様式で、放射標識された結合されたリガンドがウェルの底部に遠心分離されている場合、シンチストリップの標識は放射標識されたリガンドに接近して活性化および検出をもたらす。

2. アデニリルシクラーゼ

細胞に基づくアッセイのため設計されたフラッシュプレート [ F l a s h P l a t e ] (商標) アデニリルシクラーゼキット ( ニュー イングランド ニューカリア ( N e w E n g l a n d N u c l e a r ) ; カタログ番号 S M P 0 0 4 A ) を、粗原形質膜との使用のため改変することができる。フラッシュプレート ( F l a s h P l a t e ) のウェルはシンチラントのコーティングを含有し、このコーティングは c A M P を認識する特異的抗体もまた含有する。 c A M P 抗体への放射活性の c A M P トレーサーの結合についての直接の競争により、ウェル中で生成された c A M P を定量した。以下は、受容体を発現する膜中の c A M P レベルの変化の測定のための簡潔なプロトコルとしてはたらく。

【 0075 】

トランスフェクションされた細胞をトランスフェクションのおよそ 3 日後に収穫する。 20 mM H E P E S 、 pH 7.4 および 10 mM M g C l <sub>2</sub> を含有する緩衝液に懸濁された細胞の均質化により膜を調製した。均質化は、ブリンクマン ( B r i n k m a n ) ポリトロン [ P o l y t r o n ] (商標) を使用し氷上でおよそ 10 秒間実施する。生じるホモジエネートを 4 で  $49,000 \times g$  で 15 分間遠心分離する。その後、生じるペレットを 20 mM H E P E S 、 pH 7.4 および 0.1 mM E D T A を含有する緩衝液に再懸濁し、 10 秒間均質化し、次いで 49,000  $\times g$  で 15 分間遠心分離する。生じるペレットは利用されるまで -80 で保存することができる。測定の日に膜ペレットを室温でゆっくりと融解させ、 20 mM H E P E S 、 pH 7.4 および 10 mM M g C l <sub>2</sub> ( これらの量は至適化することができるが、本明細書に列挙される値が好ましい ) を含有する緩衝液に再懸濁して 0.60 mg / ml の最終タンパク質濃度を生じる ( 再懸濁された膜は使用まで氷上に置いた ) 。

【 0076 】

10

20

30

30

40

50

cAMP 標準および検出緩衝液 (11 ml の検出緩衝液に対し 2  $\mu$ Ci のトレーサー [ $^{125}$ I] cAMP (100  $\mu$ l) を含んで成る) を調製し、そして製造元の説明書に従って維持する。アッセイ緩衝液はスクリーニングのため新たに調製し、そして 20 mM HEPES、pH 7.4、10 mM MgCl<sub>2</sub>、20 mM (シグマ (Sigma))、0.1 単位/ml のクレアチニンホスホキナーゼ (シグマ (Sigma))、50  $\mu$ M GTP (シグマ (Sigma)) および 0.2 mM ATP (シグマ (Sigma)) を含有し；アッセイ緩衝液は利用されるまで氷上で保存することができる。NEN フラッシュプレート (Flash Plate) への 50  $\mu$ l のアッセイ緩衝液の添加、次いで 50  $\mu$ l の膜懸濁液の添加によりアッセイを開始する。結果として生じるアッセイ混合物を室温で 60 分間インキュベートし、次いで 100  $\mu$ l の検出緩衝液を添加する。その後、プレートを追加の 2 ~ 4 時間インキュベートし、次いでワラック (Wallac) マイクロベータ [MicroBeta] (商標) シンチレーション計数器で計数する。各アッセイプレート内に含有される標準 cAMP 曲線からウェルあたりの cAMP の値を外挿する。

### C. レポーターに基づくアッセイ

#### 1. CREB レポーターアッセイ (Gs 会合型受容体)

Gs 刺激を検出する方法は cAMP 依存性の様式で活性化される転写因子 CREB の既知の特性に依存する。293 もしくは 293T 細胞中での Gs 共役型の活性についてアッセイするのに、パスディテクト [Path Detect] (商標) CREB トランスレポーティング (CREB trans-Reporting) 系 (ストラタジーン (Stratagene)、カタログ番号 219010) を利用することができる。細胞は、哺乳動物トランスフェクションキット (ストラタジーン (Stratagene)、カタログ番号 200285) を製造元の説明書に従って使用して、この上の系のプラスミド成分、および内在性もしくは突然変異体の受容体をコードする指定された発現プラスミドでトランスフェクションする。簡潔には、400 ng の pFR-Luc (Gal4 認識配列を含有するルシフェラーゼレポータープラスミド)、40 ng の FA2-CREB (Gal4 の DNA 結合ドメインを含有する Gal4-CREB 融合タンパク質)、80 ng の pCMV-受容体発現プラスミド (受容体を含んで成る) および 20 ng の CMV-SEAP (分泌型アルカリホスファターゼ発現プラスミド；トランスフェクションされた細胞の培地でアルカリホスファターゼ活性を測定して、サンプル間のトランスフェクション効率の変動について制御する) を、キットの説明書に従ってリン酸カルシウム沈殿中に組み合わせる。沈殿物の半分を 96 穴プレートの 3 個のウェルに同等に配分し、細胞上で一夜保ち、そして翌朝に新鮮培地で置き換える。トランスフェクションの開始 48 時間後に細胞を処理し、そして例えばルシフェラーゼ活性についてアッセイする。

#### 2. AP1 レポーターアッセイ (Gq 会合型受容体)

Gq 刺激を検出する方法は、それらのプロモーター中に AP1 要素を含有する遺伝子の活性化を引き起こす Gq 依存性ホスホリパーゼ C の既知の特性に依存する。リン酸カルシウム沈殿物の成分が 410 ng の pAP1-Luc、80 ng の pCMV-受容体発現プラスミドおよび 20 ng の CMV-SEAP であったことを除いて、CREB レポーターアッセイに関して上に示されたプロトコルに従い、パスディテクト [Path detect] (商標) AP-1 cis-Reporting 系 (ストラタジーン (Stratagene)、カタログ番号 219073) を利用することができる。

#### 3. CRE-Luc レポーターアッセイ

293 および 293T 細胞をウェルあたり細胞  $2 \times 10^4$  個の密度で 96 穴プレート上でプレート培養し、そして翌日、製造元の説明書に従ってリポフェクタミン試薬 (BRL) を使用してトランスフェクションした。DNA / 脂質混合物

10

20

30

40

50

を以下のとおり各 6 個のウェルのトランスフェクションのため調製する。すなわち、100 μl の D M E M 中の 260 ng のプラスミド D N A を、100 μl の D M E M 中 2 μl の脂質と穏やかに混合した (260 ng のプラスミド D N A は、200 ng の 8 × C R E - L u c レポータープラスミド (プラスミドの一部の表示については下および図 1 を参照されたい)、50 ng の内在性受容体もしくは非内在性受容体を含んで成る p C M V または p C M V 単独、ならびに 10 ng の G P R S 発現プラスミド (p c D N A 3 (インヴィトロジェン (I n v i t r o g e n) 中の G P R S) より成った)。8 × C R E - L u c レポータープラスミドは以下のとおり調製した。すなわち、p gal - 基本ベクター (B a s i c V e c t o r) (クロントック (C l o n t e c h)) の B g l V - H i n d I I I 部位にラットソマトスタチンプロモーター (-71/+51) をクローン化することにより、ベクター S R I F - - g a l を得た。アデノウイルス鑄型 A d p C F 1 2 6 C C R E 8 (7 Human Gene Therapy 1883 (1996) を参照されたい) からの P C R により、8 コピーの c A M P 応答要素を得、そして K p n - B g l V 部位で S R I F - - g a l ベクターにクローン化して 8 × C R E - - g a l レポーターべクターをもたらした。

8 × C R E - L u c レポータープラスミドは、8 × C R E - - g a l レポーターべクター中の - ガラクトシダーゼ遺伝子を、H i n d I I I - B a m H I 部位で p G L 3 - 基本ベクター (プロメガ (P r o m e g a)) から得られたルシフェラーゼ遺伝子で置き換えることにより生成した。室温で 30 分のインキュベーション後に、400 μl の D M E M で D N A / 脂質混合物を希釈し、そして 100 μl の希釈された混合物を各ウェルに添加した。細胞培養インキュベーター中の 4 時間のインキュベーション後に、10% F C S を含む 100 μl の D M E M を各ウェルに添加した。翌日、トランスフェクションされた細胞を、10% F C S を含むウェルあたり 200 μl の D M E M で変更した。8 時間後、P B S での 1 回の洗浄後に、ウェルをウェルあたり 100 μl のフェノールレッドを含まない D M E M に変更した。翌日、製造元の説明書に従って、ルックライト [L u c L i t e] (商標) レポーター遺伝子アッセイキット (パッカード (P a c k a r d)) を使用してルシフェラーゼ活性を測定し、そして 1450 マイクロベータ [M i c r o B e t a] (商標) シンチレーションおよび発光計数器 (ワラック (W a l l a c)) で読み取った。

#### 4. S R F - L u c レポーターアッセイ

G q 刺激を検出するための一方法は、それらのプロモーター中に血清応答因子を含有する遺伝子の活性化を引き起こす G q 依存性のホスホリバーゼ C の既知の特性に依存する。例えば C O S 7 細胞での G q 共役型の活性についてアッセイするのに、パスディテクト [P a t h d e t e c t] (商標) S R F - L u c レポーティング系 (ストラタジーン (S t r a t a g e n e)) を利用することができる。製造元の説明書に従って哺乳動物トランスフェクション [M a m m a l i a n T r a n s f e c t i o n] (商標) キット (ストラタジーン (S t r a t a g e n e))、カタログ番号 200285) を使用して、系のプラスミド成分および内在性もしくは非内在性の G P C R をコードする指定された発現プラスミドで細胞をトランスフェクションする。簡潔には、410 ng の S R F - L u c 、80 ng の p C M V 受容体発現プラスミドおよび 20 ng の C M V - S E A P (分泌型アルカリホスファターゼ発現プラスミド；トランスフェクションされた細胞の培地内でアルカリホスファターゼ活性を測定して、サンプル間のトランスフェクションの効率の変動について制御する) を、製造元の説明書に従ってリン酸カルシウム沈殿物中で組み合わせる。沈殿物の半分を 96 穴プレートの 3 個のウェルに同等に配分し、血清を含まない培地内で細胞上で 24 時間保つ。指定された場合は、最後の 5 時間、細胞を 1 μM のアンジオテンシンとともにインキュベートする。その後細胞を溶解し、そして、製造元の説明書に従ってルックライ

10

20

30

40

50

ト [Lucelite] (商標) キット (パッカード (Packard) 、カタログ番号 6016911) および「トライルックス (Tri lux) 1450マイクロベータ (MicroBeta)」液体シンチレーションおよび発光計数器 (ワラック (Wallac)) を使用してルシフェラーゼ活性についてアッセイする。データはグラフパッド プリズム [Graph Pad Prism] (商標) 2.0a (グラフパッド ソフトウェア インク (Graph Pad Software Inc.)) を使用して解析することができる。

### 5. 細胞内 $IP_3$ 蓄積アッセイ

第1日に、受容体 (内在性および/もしくは非内在性) を含んで成る細胞を24穴プレート上でプレート培養することができる (通常、ウェルあたり細胞  $1 \times 10^5$  個、とは言えこの数字は至適化することができる)。第2日に、ウェルあたり  $50 \mu l$  の血清を含まないD MEM 中の  $0.25 \mu g$  のDNAおよびウェルあたり  $50 \mu l$  の血清を含まないD MEM 中の  $2 \mu l$  のリポフェクタミンを最初に混合することにより、細胞をトランスフェクションすることができる。溶液は穏やかに混合し、そして室温で15~30分間インキュベートする。  $0.5 \text{ ml}$  のPBSで細胞を洗浄し、そして  $400 \mu l$  の血清を含まない培地をトランスフェクション培地と混合しあつ細胞に添加する。その後、細胞を  $37/5\%CO_2$  で3~4時間インキュベートし、そしてその後、トランスフェクション培地を除去しあつウェルあたり  $1 \text{ ml}$  の通常の成長培地で置き換える。第3日に、細胞を  $^3\text{H}$ -ミオイノシトールで標識する。簡潔には、培地を除去し、そして細胞を  $0.5 \text{ ml}$  のPBSで洗浄する。その後、  $0.5 \text{ ml}$  のイノシトールを含まない/血清を含まない培地 (ギブコ (GIBCO) BRL) を、ウェルあたり  $0.25 \mu \text{Ci}$  の  $^3\text{H}$ -ミオイノシトールとともにウェルあたりに添加し、そして細胞を  $37/5\%CO_2$  で16~18時間一夜インキュベートする。第4日に、細胞を  $0.5 \text{ ml}$  のPBSで洗浄し、そして、イノシトールを含まない/血清を含まない培地、  $10 \mu \text{M}$  パージリン、  $10 \text{ mM}$  塩化リチウムを含有する  $0.45 \text{ ml}$  のアッセイ培地、もしくは  $0.4 \text{ ml}$  のアッセイ培地および  $10 \mu \text{M}$  の最終濃度までの  $50 \mu l$  の  $10 \times$  ケタンセリン (ket) を添加する。その後細胞を  $37$  で30分間インキュベートする。その後細胞を  $0.5 \text{ ml}$  のPBSで洗浄し、そしてウェルあたり  $200 \mu l$  の新鮮/氷冷停止溶液 (1M KOH;  $18 \text{ mM}$  ホウ酸ナトリウム;  $3.8 \text{ mM}$  EDTA) を添加する。溶液を5~10分間もしくは細胞が溶解されるまで氷上で保ち、そしてその後、  $200 \mu l$  の新鮮/氷冷中和溶液 (7.5% HCl) により中和する。その後、ライセートを  $1.5 \text{ ml}$  のエッペンドルフチューブに移し、そしてチューブあたり  $1 \text{ ml}$  のクロロホルム/メタノール (1:2) を添加する。溶液を15秒間ボルテックス攪拌し、そして上層をバイオラッド (BioRad) AG1-X8 (商標) 陰イオン交換樹脂 (100~200メッシュ) に適用する。最初に樹脂を  $1:1.25 \text{ W/V}$  の水で洗浄し、そして  $0.9 \text{ ml}$  の上層をカラムに負荷する。カラムを  $10 \text{ ml}$  の  $5 \text{ mM}$  ミオイノシトールおよび  $10 \text{ ml}$  の  $5 \text{ mM}$  ホウ酸ナトリウム/ $60 \text{ mM}$  ギ酸ナトリウムで洗浄する。  $2 \text{ ml}$  の  $0.1 \text{ M}$  ギ酸/ $1 \text{ M}$  ギ酸アンモニウムを用い、  $10 \text{ ml}$  のシンチレーションカクテルを含有するシンチレーションバイアル中にイノシトールトリスリン酸を溶出する。  $10 \text{ ml}$  の  $0.1 \text{ M}$  ギ酸/ $3 \text{ M}$  ギ酸アンモニウムで洗浄することによりカラムを再生し、そして  $ddH_2O$  で2回すぎ、そして水中で4で保存する。

### 【0077】

例示的結果を下の表Iに提示する:

### 【0078】

## 【表10】

表I

受容体	突然変異	利用されたアッセイ	生成されたシグナル: 内在性バージョン (相対的光単位)	生成されたシグナル: 非内在性バージョン (相対的光単位)	差異のパーセント	
hAT1	F239K	SRF-LUC	34	137	75% <sup>1</sup>	10
	AT2K255IC3	SRF-LUC	34	127	73% <sup>1</sup>	
hTDAG8	I225K	CRE-LUC (293 細胞)	2,715	14,440	81% <sup>1</sup>	
	I225K	CRE-LUC (293T 細胞)	65,681	185,636	65% <sup>1</sup>	
hH9 hCCKB	F236K V332K	CRE-LUC CRE-LUC	1,887 785	6,096 3,223	69% <sup>1</sup> 76% <sup>1</sup>	

20

## 【0079】

## C. 細胞に基づく検出アッセイ(例 - T D A G 8 )

293細胞をプレートあたり細胞  $1.3 \times 10^7$  個の密度で  $150\text{mm}$  プレート上でプレート培養し、そしてプレートあたり  $12\mu\text{g}$  のそれぞれのDNAおよび  $60\mu\text{l}$  のリポフェクタミン試薬(BRL)を使用してトランスフェクションした。トランスフェクションされた細胞は、トランスフェクション後24時間に実施されるアッセイのため血清を含有する培地内で成長させた。トランスフェクション後48時間に実施される検出アッセイ(血清および血清を含まない培地を比較するアッセイ; 図3を参照されたい)のため、初期培地を血清もしくは血清を含まない培地のいずれかに変更した。血清を含まない培地は、単にダルベッコの変形イーグル(DME)高グルコース培地(アーヴィンサイエンティフィック(Irvine Scientific) #9024)から構成された。上のDME培地に加えて、血清を含む培地は以下、すなわち10%ウシ胎児血清(ハイクロン(Hyclone) #SH30071.03)、1%の  $100\text{mM}$  ピルビン酸ナトリウム(アーヴィンサイエンティフィック(Irvine Scientific) #9334)、1%の  $20\text{mM}$  L-グルタミン(アーヴィンサイエンティフィック(Irvine Scientific) #9317)および1%のペニシリン-ストレプトマイシン溶液(アーヴィンサイエンティフィック(Irvine Scientific) #9366)を含有した。

## 【0080】

96穴のアデニリルシクラーゼ活性化フラッシュプレート[Flashplate](商標)を使用した(NEN: #SMP004A)。最初に、アッセイのための標準  $50\mu\text{l}$  を、ウェルあたり  $50\text{pmol}$  から  $0\text{pmol}$  までのcAMP濃度の範囲にわたるプレート(二重(in duplicate))に添加した。標準cAMP(NEN: #SMP004A)を水で再構成し、そして  $1 \times$  PBS(アーヴィンサイエンティフィック(Irvine Scientific) #9240)を使用して連続的希釈物を作成した。次に、 $50\mu\text{l}$  の刺激緩衝液(NEN: #SMP004A)を全部のウェルに添加した。cAMPの活性化もしくは不活

30

40

50

性化を測定するのに化合物を使用する場合は、水で希釈された各化合物 10  $\mu$  l を三重でそのそれぞれのウェルに添加した。使用された多様な最終濃度は 1  $\mu$  M から 1 mM までの範囲にわたる。アデノシン 5' - 三リン酸、ATP (リサーチ

バイオケミカルズ インターナショナル (Research Biochemicals International : # A - 141) およびアデノシン 5' - ニリン酸、ADP (シグマ (Sigma) : # A 2754) をアッセイで使用した。次に、それぞれの cDNA (CMV もしくは TDAG8) でトランスフェクションされた 293 細胞を、トランスフェクション後 24 (血清培地中でのアッセイの検出) もしくは 48 (血清および血清を含まない培地を比較するアッセイの検出) 時間で収穫した。培地を吸引し、そして細胞を 1  $\times$  PBS で 1 回洗浄した。その後、5 ml の 1  $\times$  PBS を 3 ml の細胞解離緩衝液 (シグマ (Sigma) : # C - 1544) とともに細胞に添加した。解離された細胞を遠心分離チューブに移し、そして室温で 5 分間遠心分離した。上清を除去し、そして、1 ミリリットルあたり細胞  $2 \times 10^6$  個の最終濃度を得るように、細胞ペレットを適切な量の 1  $\times$  PBS に再懸濁した。化合物を含有するウェルに、1  $\times$  PBS 中の細胞 50  $\mu$  l (ウェルあたり細胞  $1 \times 10^5$  個) を添加した。プレートを室温で 15 分間振とう機でインキュベートした。トレーサー cAMP を含有する検出緩衝液を調製した。11 ml の検出緩衝液 (NEN : # SMP004A) 中で 50  $\mu$  l (1  $\mu$  Ci に等しい) の [ $^{125}$ I] cAMP (NEN : # SMP004A) を添加した。インキュベーション後に、トレーサー cAMP を含有するこの検出緩衝液 50  $\mu$  l を各ウェルに添加した。プレートを振とう機に設置し、そして室温で 2 時間インキュベートした。最後に、プレートのウェルからの溶液を吸引し、そしてワラック (Wallac) マイクロベータ [MicroBeta] (商標) シンチレーション計数器を使用してフラッシュプレートを計数した。

#### 【0081】

図 2A において、ATP および ADP は内在性の TDAG8 に結合し、それぞれ約 59% および約 55% の cAMP の増大をもたらす。図 2B は、内在性の TDAG8 への ATP および ADP の結合を明示し、ここでは内在性 TDAG8 がトランスフェクションされ、そして血清および血清を含まない培地中で成長された。血清培地中で成長された内在性の TDAG8 への ATP の結合は、化合物を伴わない内在性の TDAG8 に比較して、約 65% の cAMP の増大を明示し；血清を含まない培地中では約 68% の増大が存在した。血清中の内在性 TDAG8 への ADP の結合は約 61% の増大を明示する一方、血清を含まない培地中では ADP 結合が約 62% 増大の増大を明示する。ATP および ADP は、それぞれ 139.8  $\mu$ M および 120.5  $\mu$ M の EC50 値で内在性の TDAG8 に結合する (データは示されない)。

#### 【0082】

図 2B に提示される結果は、血清および血清を含まない培地を比較した場合に実質的に同一の結果を示すが、われわれの選択は血清に基づく培地を使用することである。とは言え血清を含まない培地もまた利用することが可能である。

#### 実施例 6

##### GPCR 融合タンパク質の調製

構成的に活性化される GPCR - G タンパク質融合構築物の設計を以下のとおり達成した。すなわち、ラット G タンパク質 Gs (長形態；伊藤 (Itoh, H.) ら、83 PNAS 3776 (1986)) の 5' および 3' 双方の端をそれに HindIII (5' - AAGCTT - 3') 配列を包含するように工作した。正しい配列 (隣接する HindIII 配列を包含する) の確認後に、そのベクターの HindIII 制限部位を使用してサブクローニングすることにより、配列全体を pCDNA3.1 (-) (インヴィトロジェン (Invitrogen)、カタログ番号 V795-20) に snuited した。pCDNA3

10

20

30

40

50

. 1 ( - )へのサブクローニング後に G s の配列について正しい向きを決定した。 H i n d I I I 配列でラット G s 遺伝子を含有する改変された p c D N A 3 . 1 ( - )をその後確認し；このベクターは今や「普遍的」 G s タンパク質ベクターとして利用可能となった。 p c D N A 3 . 1 ( - )ベクターは、 H i n d I I I 部位の上流に多様な公知の制限部位を含有し、従って G s タンパク質の上流に内在性の構成的に活性な G P C R のコーディング配列を挿入する能力を有益に提供する。他の「普遍的」 G タンパク質ベクターを創製するのにこの同一のアプローチを利用することができます、そして、もちろん、当業者に既知の他の商業的に入手可能なもしくは占有のベクターを利用することができる。重要な基準は、 G P C R の配列が G タンパク質の配列の上流にかつ同じ読み枠で存在することである。

10

## 【 0 0 8 3 】

T D A G 8 は G s を介して共役する一方、 H 9 は G z を介して共役する。以下の例示的 G P C R 融合タンパク質について G s への融合を達成した。

## 【 0 0 8 4 】

T D A G 8 ( I 2 2 5 K ) - G s 融合タンパク質構築物は以下のとおり作成した。すなわち、プライマーを以下のとおり設計した：

5 ' - g a t c T C T A G A A T G A A C A G C A C A T G T A T T G A A G -  
3 ' ( 配列番号 1 2 5 ; センス )

5 ' - c t a g G G T A C C C G C T C A A G G A C C T C T A A T T C C A T  
A G - 3 ' ( 配列番号 1 2 6 ; アンチセンス )。

20

## 【 0 0 8 5 】

小文字のヌクレオチドは G タンパク質と T D A G 8 との間の制限部位中のスペーサーとして包含される。該センスおよびアンチセンスプライマーはそれぞれ X b a I および K p n I の制限部位を包含した。

## 【 0 0 8 6 】

その後、それぞれについて以下のプロトコルを使用し、 P C R を利用して、上に開示された G s 普遍的ベクター内の融合についてそれぞれの受容体配列を保証した。すなわち、 2  $\mu$  l の各プライマー（センスおよびアンチセンス）、 3  $\mu$  L の 1 0 m M d N T P 、 1 0  $\mu$  L の 1 0  $\times$  タックプラス [ T a q P l u s ] (商標) プレシジョン ( P r e c i s i o n ) 緩衝液、 1  $\mu$  L のタックプラス [ T a q P l u s ] (商標) プレシジョン ( P r e c i s i o n ) ポリメラーゼ (ストラタジーン ( S t r a t a g e n e ) : # 6 0 0 2 1 1 ) ならびに 8 0  $\mu$  L の水を含有する別個のチューブに、 1 0 0 n g の T D A G 8 の c D N A を添加した。 T D A G 8 についての反応温度および周期時間は以下のとおりであった。すなわち、最初の変性段階はそれを 9 4 度 5 分間、そして、 9 4 3 0 秒間； 5 5

30

3 0 秒間； 7 2 2 分間の周期を行った。最後の伸長時間は 7 2 度 1 0 分間行った。順向きの P C R 産物 ( P C R product for ) を 1 % アガロースゲル上で泳動し、そしてその後精製した（データは示されない）。精製された産物を X b a I および K p n I ( ニュー イングランド バイオラブス ( N e w E n g l a n d B i o l a b s ) ) で消化し、そして所望の挿入物を精製しつつそれぞれの制限部位で G s 普遍的ベクターに連結した。形質転換後に陽性のクローニを単離し、そして制限酵素消化により決定し； 2 9 3 細胞を使用する発現は下に示されるプロトコルに従って達成した。 T D A G 8 : G s 融合タンパク質についてのそれぞれの陽性のクローニを配列決定して正しさを確かめた。

40

## 【 0 0 8 7 】

非内在性の構成的に活性化される T D A G 8 ( I 2 2 5 K ) を含んで成る G P C R 融合タンパク質を上のとおり分析し、そして構成的活性化について確かめた。

## 【 0 0 8 8 】

H 9 ( F 2 3 6 K ) - G s 融合タンパク質構築物は以下のとおり作成した。

50

すなわち、プライマーを以下のとおり設計した：

5' - T T A g a t a t c G G G G C C C A C C C T A G C G G T - 3' (配列番号 145 ; センス)

5' - g g t a c c C C C A C A G C C A T T T C A T C A G G A T C - 3' (配列番号 146 ; アンチセンス)。

【0089】

小文字のヌクレオチドは G タンパク質と H 9 との間の制限部位中のスペーサーとして包含される。センスおよびアンチセンスプライマーはそれぞれ E c o R V および K p n I の制限部位を包含し、その結果、スペーサー（制限部位に帰される）が G タンパク質と H 9 との間に存在する。

10

【0090】

その後、それについて以下のプロトコルを使用し、P C R を利用して、上に開示された G s 普遍的ベクター内の融合についてそれぞれの受容体配列を保証した。すなわち、1 0 0 n g の各プライマー（センスおよびアンチセンス）ならびに 4 5  $\mu$  L の P C R スーパーミックス [ S u p e r m i x ] (商標) (ギブコ (G i b c o) - B r 1 、ライフテック (L i f e T e c h) ) を含有する別個のチューブに、8 0 n g の H 9 の c D N A を添加した (5 0  $\mu$  L の総反応体積)。H 9 についての反応温度および周期時間は以下のとおりであった。すなわち、最初の変性段階はそれを 9 4 で 1 、そして、9 4 3 0 秒間 ; 5 5 3 0 秒間、7 2 2 分間の周期を行った。最後の伸長時間は 7 2 で 7 分間行った。順向きの P C R 産物を 1 % アガロースゲル上で泳動し、そしてその後精製した（データは示されない）。精製された産物を p C R I I - T O P O (商標) 系にクローン化し、次いで陽性のクローンを同定した。陽性のクローンを単離し、E c o R V および K p n I (ニュー イングランド バイオラブス (N e w E n g l a n d B i o l a b s) ) で消化し、そして所望の挿入物を単離し、精製しきつそれぞれの制限部位で G s 普遍的ベクターに連結した。形質転換後に陽性のクローンを単離し、そして制限酵素消化により決定し；2 9 3 細胞を使用する発現は下に示されたプロトコルに従って達成した。H 9 (F 2 3 6 K) : G s 融合タンパク質についての各陽性のクローンを配列決定して正しさを確かめた。膜は利用されるまで凍結 (- 8 0 ) した。

20

【0091】

(H 9 は G z と共に役するのであるが) G s タンパク質により媒介される c A M P 応答を測定する能力を確かめるため、N E N アデニルシクラーゼ活性化フラッシュプレート [ F l a s h p l a t e ] (商標) アッセイキット (9 6 穴の形式) に基づき、以下の c A M P 膜アッセイを利用した。「結合緩衝液」は 1 0 m M H E P E S 、 1 0 0 m M N a C l および 1 0 m M M g C l (p H 7 . 4) より成了た。「再生緩衝液」は結合緩衝液中で調製し、そして 2 0 m M ホスホクレアチニン、2 0 U のクレアチニンホスホキナーゼ、2 0  $\mu$  M G T P 、0 . 2 m M A T P および 0 . 6 m M I B M X より成了た。「c A M P 標準」は以下のとおり結合緩衝液中で調製した：

30

【0092】

40

【表11】

cAMPストック (2mlH <sub>2</sub> O中5,000 pmol/ml) ul	指定された量の 結合緩衝液に 添加	指定されたpmol/ウェルを 達成するための 最終アッセイ濃度 (100ul中50ul)	10
A 250	1ml	50	
B 500 of A	500ul	25	
C 500 of B	500ul	12.5	
D 500 of C	750ul	5.0	
E 500 of D	500ul	2.5	
F 500 of E	500ul	1.25	
G 500 of F	750ul	0.5	

【0093】

凍結された膜（対照としてのpCMVおよび非内在性のH（-Gs融合タンパク質の双方）を（溶液中まで室温で氷上で）融解した。膜を懸濁液中までポリトロンで均質化した（2×15秒）。ブラッドフォード（Bradford）アッセイプロトコル（下を参照されたい）を使用して膜タンパク質濃度を測定した。再生緩衝液中で膜の濃度を0.5mg/mlに希釈した（最終アッセイ濃度-ウェルあたり25μg）。その後、50μlの結合緩衝液を各ウェルに添加した。対照には、ウェルあたり50μlのcAMP標準をウェル11および12A-Gに、結合緩衝液単独を12Hに添加した（96穴の形式で）。その後、ウェルあたり50μlのタンパク質をウェルに添加し、そして室温で（振とう機上で）60分間インキュベートした。検出緩衝液（下を参照されたい）中100μlの[<sup>125</sup>I]cAMPを各ウェルに添加した（最終-11mlの検出緩衝液中に50μlの[<sup>125</sup>I]cAMP）。これらを室温で2時間インキュベートした。プレートを8チャンネルのマニホールドで吸引し、そしてプレートの蓋で封止した。結果（結合されたcAMPのpmol）を「プロット#15」でワラック[Wal11ac]（商標）1450で読み取った。結果を図3に提示する。

【0094】

図3に提示される結果は、Gs共役型融合がシクラーゼ反応を「駆動する」ことが可能であり、その結果H9(F236K)の構成的活性化の測定が実現可能であったことを示す。これらの結果に基づけば、反作用薬、アゴニストおよび部分的アゴニストである候補化合物の直接同定がシクラーゼに基づくアッセイを使用して可能である。

#### 実施例6

プロトコル：[<sup>35</sup>S]GTP-Sを使用する反作用薬およびアゴニストの直接同定

われわれは、全く理解されていない理由上、例えば反作用薬のような候補化合物の直接同定に内在性の構成的に活性のGPCRを利用したが、アッセイ間の変動が悪化されたようになる可能性がある。その場合、好ましくは、上に開示されたようなGPCR融合タンパク質を非内在性の構成的に活性化されるGPCRとともにまた利用する。われわれは、こうしたタンパク質を使用する場合にアッセイ間の変動が実質的に安定化されるようであり、それにより有効なS/N比が得られることを決定した。これは候補化合物のより確固たる同定を見込むという有益な結果を有する。従って、直接同定のためにはGPCR融合タンパク質を使用すること、また、利用される場合は以下のアッセイプロトコルを利用することが好ましい。

#### 膜の調製

目的の、および反作用薬、アゴニストもしくは部分的アゴニストとしての候補

10

20

20

30

40

50

化合物の直接同定での使用のための非内在性の構成的に活性のオーファン G P C R 融合タンパク質を含んで成る膜は、以下のように好ましく調製する：

a. 材料

「膜掻き取り緩衝液」は 20 mM HEPES および 10 mM EDTA、pH 7.4 より構成され；「膜洗浄緩衝液」は 20 mM HEPES および 0.1 mM EDTA、pH 7.4 から構成され；「結合緩衝液」は 20 mM HEPES、100 mM NaCl および 10 mM MgCl<sub>2</sub>、pH 7.4 から構成される。

b. 手順

全部の材料は処置の間中氷上に保つ。最初に、細胞のコンフルエントな単層から培地を吸引し、次いで 10 mL の冷 PBS ですすぎ、次いで吸引する。その後、5 mL の膜掻き取り緩衝液を添加して細胞を掻き取り；これに次いで細胞抽出物を 50 mL 遠心管に移す (20,000 rpm で 4 分で 17 分間遠心分離される)。その後、上清を吸引し、そしてペレットを 30 mL の膜洗浄緩衝液に再懸濁し、次いで 20,000 rpm で 4 分で 17 分間遠心分離する。上清をその後吸引し、そしてペレットを結合緩衝液に再懸濁する。その後、ブリンクマン (B r i n k m a n) ポリトロン [p o l y t r o n] (商標) ホモジエナイザーを使用してこれを均質化する (全部の物質が懸濁液中にあるまで 15 ~ 20 秒破裂させる)。これを本明細書で「膜タンパク質」と称する。

ブラッドフォード (B r a d f o r d) タンパク質アッセイ

均質化の後に、ブラッドフォード (B r a d f o r d) タンパク質アッセイを使用して膜のタンパク質濃度を測定する (タンパク質を約 1.5 mg / mL に希釈し、等分し、そして後の使用のため凍結する (-80 °C) ことができ；凍結される場合、使用のためのプロトコルは以下のとおりである。すなわち、アッセイの日に、凍結された膜タンパク質を室温で融解し、次いでボルテックス攪拌し、そしてその後、約 5 ~ 10 秒間約 12 × 1,000 rpm でポリトロンで均質化し；複数の調製のためには、異なる調製物の均質化の間にホモジエナイザーを徹底的に洗浄すべきであることが言及される)。

a. 材料

結合緩衝液 (上に従う)；ブラッドフォード (B r a d f o r d) 色素試薬；ブラッドフォード (B r a d f o r d) タンパク質標準を製造元の説明書に従って (バイオラッド (B i o R a d) カタログ番号 500-0006) 利用する。

b. 手順

二重のチューブを準備する (1 本は膜を包含し、そして 1 本は対照の「ブランク」として)。それぞれは 800 μL の結合緩衝液を含有した。その後、10 μL のブラッドフォード (B r a d f o r d) タンパク質標準 (1 mg / mL) を各チューブに添加し、そしてその後 10 μL の膜タンパク質を 1 本のチューブにだけ添加する (ブランクでない)。その後、200 μL のブラッドフォード (B r a d f o r d) 色素試薬を各チューブに添加し、次いでそれをボルテックス攪拌する。5 分後にチューブを再度ボルテックス攪拌し、そして、その中の物質をキュベットに移す。その後、波長 595 で C E C I L 3041 分光光度計を使用してキュベットを読み取る。

直接同定アッセイ

a. 材料

GDP 緩衝液は 37.5 mL の結合緩衝液および 2 mg の GDP (シグマ (S i g m a)、カタログ番号 G-7127) より成り、次いで 0.2 μM GDP を得るため結合緩衝液で一連の希釈を行い (各ウェル中での GDP の最終濃度は 0.1 μM GDP であった)；候補化合物を含んで成る各ウェルは、100 μL の GDP 緩衝液 (最終濃度、0.1 μM GDP)、50 μL の結合緩衝液中の膜タンパク質、および 50 μL の結合緩衝液中の [<sup>35</sup>S] GTP S (0.6

10

20

30

40

50

nM) (10mlの結合緩衝液あたり2.5μlの[<sup>35</sup>S]GTP-S)より成る200μlの最終体積を有する。

#### b. 手順

候補化合物は96穴プレートの形式(これらは-80で凍結させることができる)を使用して好ましくスクリーニングする。膜タンパク質(もしくは、対照として、GPCR融合タンパク質を除外する発現ベクターを含む膜)を懸濁液中まで短く均質化する。その後、上に示されたブラッドフォード(Bradford)タンパク質アッセイを使用してタンパク質濃度を測定する。その後、膜タンパク質(および対照)を結合緩衝液で0.25mg/mlに希釈する(最終アッセイ濃度、ウェルあたり12.5μg)。その後、100μlのGDP緩衝液をワラック(Wallac)シンチストリップ[Scintistrip](商標)(ワラック(Wallac))の各ウェルに添加する。その後、5μlのピンツール(pin-tool)を使用して、こうしたウェルに5μlの候補化合物を移す(すなわち、200μlの総アッセイ体積中の5μlは1:40の比であり、その結果候補化合物の最終スクリーニング濃度は10μMである)。再度、汚染を回避するため、各移送段階の後にはピンツールを水(1×)、エタノール(1×)および水(2×)を含んで成る3個の水槽ですすぐべきであり、過剰の液体は各すすぎの後にツールから振りそして紙およびキムワイプでぬぐって乾かすべきである。その後、50μlの膜タンパク質を各ウェルに添加し(GPCR融合タンパク質を含まない膜を含んで成る対照ウェルもまた利用する)、そして室温で5~10分間前インキュベートする。その後、50μlの結合緩衝液中の[<sup>35</sup>S]GTP-S(0.6nM)を各ウェルに添加し、次いで室温で振とう機で60分間インキュベートする(再度、この実施例において、プレートをホイルで覆った)。その後、22で4000 RPMで15分間のプレートの回転によりアッセイを停止する。その後、プレートを8チャンネルのマニホールドで吸引し、そしてプレートの蓋で封止する。その後、(製造元の説明書に従って)「Prot. #37」の設定を使用してワラック(Wallacc)1450でプレートを読み取る。

#### 実施例7

##### プロトコル：確認アッセイ

上に示されたような直接同定された候補化合物の確認を提供するための独立のアッセイのアプローチを使用して、その後確認アッセイを利用することが好ましい。この場合、好ましい確認アッセイはシクラーゼに基づくアッセイである。

##### 【0095】

改変されたフラッシュプレート[Flash Plate](商標)アデニリルシクラーゼキット(ニューイングランドニュークリア(New England Nuclear)；カタログ番号SMP004A)を、以下のプロトコルに従って、非内在性の構成的に活性化されるオーファンGPCRに対する反作用薬およびアゴニストとして直接同定された候補化合物の確認に、好ましく利用する。

##### 【0096】

トランスフェクションされた細胞をトランスフェクション後およそ3日に収穫する。20mM HEPES、pH7.4および10mM MgCl<sub>2</sub>を含有する緩衝液中に懸濁された細胞の均質化により膜を調製する。均質化は、ブリンクマン(Brinkman)ポリトロン[Polytron](商標)を使用して氷上でおよそ10秒間実施する。生じるホモジエネートは49,000×gで4で15分間遠心分離する。その後、生じるペレットを、20mM HEPES、pH7.4および0.1mM EDTAを含有する緩衝液に再懸濁し、10秒間均質化し、次いで49,000×gで4で15分間遠心分離する。生じるペレットは利用されるまで-80で保存することができる。直接同定スクリーニ

10

20

30

40

50

ングの日に、膜ペレットを室温でゆっくりと融解し、20 mM HEPES、pH 7.4 および 10 mM MgCl<sub>2</sub>を含有する緩衝液に再懸濁して、0.60 mg / ml の最終タンパク質濃度を生じる（再懸濁された膜は使用まで氷上に置く）。

【0097】

cAMP 標準および検出緩衝液（11 ml の検出緩衝液に対して 2 μCi のトレーサー [<sup>125</sup>I] cAMP (100 μl) を含んで成る）は、製造元の説明書に従って調製しつつ維持する。アッセイ緩衝液はスクリーニングのために新たに調製し、そして 20 mM HEPES、pH 7.4、10 mM MgCl<sub>2</sub>、20 mM ホスホクレアチニン（シグマ（Sigma））、0.1 単位 / ml のクレアチニンホスホキナーゼ（シグマ（Sigma））、50 μM GTP（シグマ（Sigma））および 0.2 mM ATP（シグマ（Sigma））を含有し；アッセイ緩衝液は利用されるまで氷上で保存することができる。

10

【0098】

上に従って同定された候補化合物（凍結されている場合は室温で融解させる）を、40 μM の膜タンパク質（ウェルあたり 30 μg）および 50 μl のアッセイ緩衝液と一緒に、好ましくは 96 穴プレートのウェルに添加する（ウェルあたり 3 μl；12 μM の最終アッセイ濃度）。その後、この混合状態を穏やかに振とうしながら室温で 30 分間インキュベートする。

20

【0099】

インキュベーション後に、各ウェルに 100 μl の検出緩衝液を添加し、次いで 2 ~ 24 時間インキュベートする。その後、プレートを、（製造元の説明書に従い）「Prot. # 31」を使用してワラック（Wallac）マイクロベータ [MicroBeta]（商標）プレートリーダーで計数する。

【0100】

本特許文書で挙げられる特許、出願および印刷された刊行物のそれぞれはこれによりそっくりそのまま引用により組み込まれることが意図される。

30

【0101】

当業者は、多数の変更および改変を、本発明の技術思想から離れることなく本発明の好ましい態様に対し行ってよいことを認識するであろう。全部のこうした変形物は本発明の範囲内にあることが意図される。

【0102】

多様な発現ベクターが当業者に利用可能であるが、内在性および非内在性双方のヒト GPCR に対する利用の目的上、利用されるベクターは pCMV であることが最も好ましい。このベクターは、特許手続上の微生物の寄託の国際的承認に関するブダペスト条約 (the Budapest Treaty for the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure) の規定の下に、1998 年 10 月 13 日にアメリカン タイプ カルチャー コレクション (American Type Culture Collection) (ATCC) (10801 University Blvd., 米国バージニア州マナサス 20110-2209) に寄託された。該 DNA は ATCC により試験され、そしてそうであることが決定された。ATCC は pCMV に対し以下の寄託番号：ATCC # 203351 を割り当てている。

40

（配列表）

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: Behan, Dominic P.  
5 Lehmann-Bruinsma, Karin  
Chalmers, Derek T.  
Lowitz, Kevin P.  
Lin, I-Lin  
Dang, Huong T.  
10 Chen, Ruoping  
Liau, Chen W.  
Gore, Martin J.  
White, Carol

(ii) TITLE OF INVENTION: Non-Endogenous, Constitutively Activated Human G  
15 Protein-Coupled Receptors

(iii) NUMBER OF SEQUENCES: 146

## (iv) CORRESPONDENCE ADDRESS:

20 (A) ADDRESSEE: Arena Pharmaceuticals, Inc.  
(B) STREET: 6166 Nancy Ridge Drive  
(C) CITY: San Diego  
(D) STATE: CA  
(E) COUNTRY: USA  
(F) ZIP: 92121

10

20

25 (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

30 (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: US  
(B) FILING DATE:  
(C) CLASSIFICATION:

## (vii) ATTORNEY/AGENT INFORMATION:

35 (A) NAME: Burgoon, Richard P.  
(B) REGISTRATION NUMBER: 34,787

30

(ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: (858) 453-7200  
(B) TELEFAX: (858) 453-7210

## 40 (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1260 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

## (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGGTCTTCT	CGGCAGTGTT	GAUTGCGTTC	CATACCGGGA	CATCCAACAC	AACATTTGTC	60	
5	GTGTATGAAA	ACACCTACAT	GAATATTACA	CTCCCTCCAC	CATTCCAGCA	TCCTGACCTC	120
	AGTCCATTGC	TTAGATATAG	TTTGAAACC	ATGGCTCCA	CTGGTTTGAG	TTCCATTGACC	180
	GTGAATAGTA	CAGCTGTGCC	CACAACACCA	GCAGCATTAA	AGAGCCTAAA	CTTGCCTCTT	240
10	CAGATCACCC	TTTCTGCTAT	AATGATATTTC	ATTCTGTTG	TGTCTTTCT	TGGGAACCTG	300
	GTTGTTTGCC	TCATGGTTTA	CCAAAAAGCT	GCCATGAGGT	CTGCAATTAA	CATCCTCCTT	360
10	GCCAGCCTAG	CTTTTGAGA	CATGTTGCTT	GCAGTGCTGA	ACATGCCCTT	TGCCCTGGTA	420
	ACTATTCTTA	CTACCCGATG	GATTTTGAGG	AAATTCTTCT	GTAGGGTATC	TGCTATGTTT	480
	TTCTGGTTAT	TTGTGATAGA	AGGAGTAGCC	ATCCTGCTCA	TCATTAGCAT	AGATAGGTTTC	540
15	CTTATTATAG	TCCAGAGGCA	GGATAAGCTA	AACCCATATA	GAGCTAAGGT	TCTGATTGCA	600
	GTTTCTTGAGG	CAACTTCCCTT	TTGTGATAGCT	TTTCCTTTAG	CCGTAGGAAA	CCCCGACCTG	660
15	CAGATACCTT	CCCGAGCTCC	CCAGTGTGTG	TTTGGGTACA	CAACCAATCC	AGGCTACCAG	720
	GCTTATGTGA	TTTGATTTC	TCTCATTTCT	TTCTTCATAC	CCTTCCTGGT	AATACTGTAC	780
	TCATTTATGG	GCATACTCAA	CACCCCTCGG	CACAATGCCT	TGAGGATCCA	TAGCTACCCT	840
20	GAAGGTATAT	GCCTCAGCCA	GGCCAGCAAA	CTGGGTCTCA	TGAGTCTGCA	GAGACCTTTC	900
	CAGATGAGCA	TTGACATGGG	CTTTAAAACA	CGTGCCTTCA	CCACTATTTT	GATTCTCTTT	960
20	GCTGTCTTCA	TTGTCTGCTG	GGCCCCATTC	ACCACTTACA	GCCTTGTGGC	AACATTCACT	1020
	AAGCACTTTT	ACTATCAGCA	CAACTTTTTT	GAGATTAGCA	CCTGGCTACT	GTGGCTCTGC	1080
25	TACCTCAAGT	CTGCATTGAA	TCCGCTGATC	TACTACTGGA	GGATTAAGAA	ATTCCATGAT	1140
	GCTTGCTCTGG	ACATGATGCC	TAAGTCCTTC	AAGTTTTGAG	CGCAGCTCCC	TGGTCACACA	1200
25	AAGCGACGGA	TACGTCTTAG	TGCTGTCTAT	GTGTGTTGGG	AACATCGGAC	GGTGGTGTGA	1260

(3) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

20

30

(iii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

	Met Val Phe Ser Ala Val Leu Thr Ala Phe His Thr Gly Thr Ser Asn			
1	5	10	15	
5	Thr Thr Phe Val Val Tyr Glu Asn Thr Tyr Met Asn Ile Thr Leu Pro			
	20	25	30	
	Pro Pro Phe Gln His Pro Asp Leu Ser Pro Leu Leu Arg Tyr Ser Phe			
	35	40	45	
10	Glu Thr Met Ala Pro Thr Gly Leu Ser Ser Leu Thr Val Asn Ser Thr			
	50	55	60	
	Ala Val Pro Thr Thr Pro Ala Ala Phe Lys Ser Leu Asn Leu Pro Leu			
	65	70	75	80
	Gln Ile Thr Leu Ser Ala Ile Met Ile Phe Ile Leu Phe Val Ser Phe			
	85	90	95	
15	Leu Gly Asn Leu Val Val Cys Leu Met Val Tyr Gln Lys Ala Ala Met			
	100	105	110	
	Arg Ser Ala Ile Asn Ile Leu Leu Ala Ser Leu Ala Phe Ala Asp Met			
	115	120	125	
20	Leu Leu Ala Val Leu Asn Met Pro Phe Ala Leu Val Thr Ile Leu Thr			
	130	135	140	
	Thr Arg Trp Ile Phe Gly Lys Phe Cys Arg Val Ser Ala Met Phe			
	145	150	155	160
	Phe Trp Leu Phe Val Ile Glu Gly Val Ala Ile Leu Leu Ile Ile Ser			
	165	170	175	
25	Ile Asp Arg Phe Leu Ile Ile Val Gln Arg Gln Asp Lys Leu Asn Pro			
	180	185	190	
	Tyr Arg Ala Lys Val Leu Ile Ala Val Ser Trp Ala Thr Ser Phe Cys			
	195	200	205	
30	Val Ala Phe Pro Leu Ala Val Gly Asn Pro Asp Leu Gln Ile Pro Ser			
	210	215	220	
	Arg Ala Pro Gln Cys Val Phe Gly Tyr Thr Thr Asn Pro Gly Tyr Gln			
	225	230	235	240
	Ala Tyr Val Ile Leu Ile Ser Leu Ile Ser Phe Phe Ile Pro Phe Leu			
	245	250	255	
35	Val Ile Leu Tyr Ser Phe Met Gly Ile Leu Asn Thr Leu Arg His Asn			
	260	265	270	

Ala Leu Arg Ile His Ser Tyr Pro Glu Gly Ile Cys Leu Ser Gln Ala  
 275 280 285  
 Ser Lys Leu Gly Leu Met Ser Leu Gln Arg Pro Phe Gln Met Ser Ile  
 290 295 300  
 5 Asp Met Gly Phe Lys Thr Arg Ala Phe Thr Thr Ile Leu Ile Leu Phe  
 305 310 315 320  
 Ala Val Phe Ile Val Cys Trp Ala Pro Phe Thr Thr Tyr Ser Leu Val  
 325 330 335  
 Ala Thr Phe Ser Lys His Phe Tyr Tyr Gln His Asn Phe Phe Glu Ile  
 10 340 345 350  
 Ser Thr Trp Leu Leu Trp Leu Cys Tyr Leu Lys Ser Ala Leu Asn Pro  
 355 360 365  
 Leu Ile Tyr Tyr Trp Arg Ile Lys Lys Phe His Asp Ala Cys Leu Asp  
 370 375 380  
 15 Met Met Pro Lys Ser Phe Lys Phe Leu Pro Gln Leu Pro Gly His Thr  
 385 390 395 400  
 Lys Arg Arg Ile Arg Pro Ser Ala Val Tyr Val Cys Gly Glu His Arg  
 405 410 415  
 Thr Val Val  
 20 420

## (4) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1119 base pairs
  - (B) TYPE: nucleic acid
  - 25 (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGTTAGCCA ACAGCTCCTC AACCAACAGT TCTGTTCTCC CGTGTCTGA CTACCGACCT	30 60
30 ACCCACCGCC TGCACTTGGT GGTCTACAGC TTGGTGCTGG CTGCCGGGCT CCCCCCTCAAC	120
GCGCTAGCCC TCTGGGTCTT CCTGCGCGCG CTGCGCGTGC ACTCGGTGGT GAGCGTGTAC	180
ATGTGTAACC TGGCGGCCAG CGACCTGCTC TTCACCCCTCT CGCTGCCCGT TCGTCTCTCC	240
TACTACGCAC TGCACCACTG GCCCTTCCCC GACCTCCTGT GCCAGACGAC GGGCGCCATC	300
TTCCAGATGA ACATGTACGG CAGCTGCATC TTCTGATGC TCATCAACGT GGACCGCTAC	360

GCCGCCATCG TGCACCCGCT GCGACTGCGC CACCTGCGGC GGCCCCGCGT GGCGCGGCTG 420  
 CTCTGCCTGG CGCTGTGGGC GCTCATCCTG GTGTTGCCG TGCCCGCCGC CCGCGTGCAC 480  
 AGGCCCTCGC GTTGGCGCTA CGGGGACCTC GAGGTGCGCC TATGCTTCGA GAGCTTCAGC 540  
 GACGAGCTGT GGAAAGGCAG GCTGCTGCC CTCGTGCTGC TGGCCGAGGC GCTGGGCTTC 600  
 5 CTGCTGCCCC TGGCGCGGT GGTCTACTCG TCGGGCCGAG TCTTCTGGAC GCTGGCGC 660  
 CCCGACGCCA CGCAGAGCCA GCGGCGGGCGG AAGACCGTGC GCCTCCTGCT GGCTAACCTC 720  
 GTCATCTTCC TGCTGTGCTT CGTGCCTAC AACAGCACGC TGGCGGTCTA CGGGCTGCTG 780 10  
 CGGAGCAAGC TGGTGGCGGC CAGCGTGCCT GCCCACGATC GCGTGCAGGG GGTGCTGATG 840  
 GTGATGGTGC TGCTGGCCGG CGCCAACACTGC GTGCTGGACC CGCTGGTGT 900  
 10 GCCGAGGGCT TCCGCAACAC CCTGCGCGGC CTGGGCACTC CGCACCGGGC CAGGACCTCG 960  
 GCCACCAACG GGACCGGGC GGCACGCGCG CAATCCGAAA GGTCCGCCGT CACCACCGAC 1020  
 GCCACCAGGC CGGATGCCGC CAGTCAGGGG CTGCTCCGAC CCTCCGACTC CCACTCTCTG 1080  
 TCTTCCTTCA CACAGTGTCC CCAGGATTCC GCCCTCTGA 1119

(5) INFORMATION FOR SEQ ID NO:4: 20

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 372 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

20 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Leu	Ala	Asn	Ser	Ser	Ser	Thr	Asn	Ser	Ser	Val	Leu	Pro	Cys	Pro
1															
															15
Asp	Tyr	Arg	Pro	Thr	His	Arg	Leu	His	Leu	Val	Val	Tyr	Ser	Leu	Val
25															
															30
Leu	Ala	Ala	Gly	Leu	Pro	Leu	Asn	Ala	Leu	Ala	Leu	Trp	Val	Phe	Leu
															35
Arg	Ala	Leu	Arg	Val	His	Ser	Val	Val	Ser	Val	Tyr	Met	Cys	Asn	Leu
															40
50															
Ala	Ala	Ser	Asp	Leu	Leu	Phe	Thr	Leu	Ser	Leu	Pro	Val	Arg	Leu	Ser
65															
															50
Tyr	Tyr	Ala	Leu	His	His	Trp	Pro	Phe	Pro	Asp	Leu	Leu	Cys	Gln	Thr

	85	90	95	
	Thr Gly Ala Ile Phe Gln Met Asn Met Tyr Gly Ser Cys Ile Phe Leu			
	100	105	110	
5	Met Leu Ile Asn Val Asp Arg Tyr Ala Ala Ile Val His Pro Leu Arg			
	115	120	125	
	Leu Arg His Leu Arg Arg Pro Arg Val Ala Arg Leu Leu Cys Leu Gly			
	130	135	140	
	Val Trp Ala Leu Ile Leu Val Phe Ala Val Pro Ala Ala Arg Val His			
	145	150	155	160
10	Arg Pro Ser Arg Cys Arg Tyr Arg Asp Leu Glu Val Arg Leu Cys Phe			
	165	170	175	
	Glu Ser Phe Ser Asp Glu Leu Trp Lys Gly Arg Leu Leu Pro Leu Val			
	180	185	190	
15	Leu Leu Ala Glu Ala Leu Gly Phe Leu Leu Pro Leu Ala Ala Val Val			
	195	200	205	
	Tyr Ser Ser Gly Arg Val Phe Trp Thr Leu Ala Arg Pro Asp Ala Thr			
	210	215	220	
	Gln Ser Gln Arg Arg Arg Lys Thr Val Arg Leu Leu Leu Ala Asn Leu			
	225	230	235	240
20	Val Ile Phe Leu Leu Cys Phe Val Pro Tyr Asn Ser Thr Leu Ala Val			
	245	250	255	
	Tyr Gly Leu Leu Arg Ser Lys Leu Val Ala Ala Ser Val Pro Ala Arg			
	260	265	270	
25	Asp Arg Val Arg Gly Val Leu Met Val Met Val Leu Leu Ala Gly Ala			
	275	280	285	
	Asn Cys Val Leu Asp Pro Leu Val Tyr Tyr Phe Ser Ala Glu Gly Phe			
	290	295	300	
	Arg Asn Thr Leu Arg Gly Leu Gly Thr Pro His Arg Ala Arg Thr Ser			
	305	310	315	320
30	Ala Thr Asn Gly Thr Arg Ala Ala Leu Ala Gln Ser Glu Arg Ser Ala			
	325	330	335	
	Val Thr Thr Asp Ala Thr Arg Pro Asp Ala Ala Ser Gln Gly Leu Leu			
	340	345	350	
	Arg Pro Ser Asp Ser His Ser Leu Ser Ser Phe Thr Gln Cys Pro Gln			
35	355	360	365	
	Asp Ser Ala Leu			
	370			

## (6) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1107 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATGGCCA	ACT CCACAGGGCT	GAACGCCTCA	GAAGTCGCAG	GCTCGTTGGG	GTTGATCCTG	60	10
10 GCAGCTGTCG	TGGAGGTGGG	GGCACTGCTG	GGCAACGGCG	CGCTGCTGGT	CGTGGTGCTG	120	
CGCACGCCGG	GA	CTGCGCGA	CGCGCTCTAC	CTGGCGCACC	TGTGCGTCGT	180	
GCGGCCG	CCT CCATCATGCC	GCTGGGCCTG	CTGGCGCAC	CGCCGCCCGG	GCTGGGCCGC	240	
GTGCGC	CTGGGCGCC	ATGCCCGGCC	GCTCGCTTCC	TCTCCGCCG	TCTGCTGCCG	300	
GCCTGCACGC	TCGGGGTGGC	CGCACTTGGC	CTGGCACGCT	ACCGCCTCAT	CGTGCACCCG	360	
15 CTGCGGCCAG	GCTCGCGGCC	GCCGCCTGTG	CTCGTGCTCA	CCGCCGTGTG	GGCCGCCGGCG	420	
GGACTGCTGG	GCGCGCTCTC	CCTGCTCGGC	CCGCCGCCCG	CACCGCCCCC	TGCTCCTGCT	480	20
CGCTGCTCGG	TCCTGGCTGG	GGGCCTCGGG	CCCTTCGGC	CGCTCTGGGC	CCTGCTGGCC	540	
TTCGCGCTGC	CCGCCCTCCT	GCTGCTCGGC	GCCTACGGCG	GCATCTTCGT	GGTGGCGCGT	600	
CGCGCTGCC	TGAGGCCCCC	ACGGCCGGCG	CGCGGGTCCC	GACTCCGCTC	GGACTCTCTG	660	
20 GATAGCCGCC	TTTCCATCTT	GCCGCCGCTC	CGGCCTCGCC	TGCCCCGGGG	CAAGGGCGGCC	720	
CTGGCCCCAG	CGCTGGCCGT	GGGCAATT	GCAGCCTGCT	GGCTGCCTTA	TGGCTGCGCG	780	
TGCCTGGCGC	CCGCAGGGCG	GGCCGCGGAA	GGCGAACGGG	CTGTCACCTG	GGTGGCCCTAC	840	
TCGGCCTTCG	CGGCTCACCC	CTTCCTGTAC	GGGCTGCTGC	AGCGCCCCGT	GCGCTTGGCA	900	30
CTGGGCCGCGC	TCTCTCGCG	TGCACTGCCT	GGACCTGTGC	GGGCCTGCAC	TCCGCAAGCC	960	
25 TGGCACCGC	GGGCACTCTT	GCAATGCCTC	CAGAGACCCC	CAGAGGGCCC	TGCCGTAGGC	1020	
CCTTCTGAGG	CTCCAGAAC	GACCCCCGAG	TTGGCAGGAG	GGCGGAGCCC	CGCATACCAAG	1080	
GGGCCACCTG	AGAGTTCTCT	CTCCTGA				1107	

## (7) INFORMATION FOR SEQ ID NO:6:

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 368 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

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

Gly Leu Ile Leu Ala Ala Val Val Glu Val Gly Ala Leu Leu Gly Asn  
20 25 30

10 Gly Ala Leu Leu Val Val Val Leu Arg Thr Pro Gly Leu Arg Asp Ala  
35 40 45

Leu Tyr Leu Ala His Leu Cys Val Val Asp Leu Leu Ala Ala Ala Ser  
50 55 60

15 Ile Met Pro Leu Gly Leu Leu Ala Ala Pro Pro Pro Gly Leu Gly Arg  
65 70 75 80

Val Arg Leu Gly Pro Ala Pro Cys Arg Ala Ala Arg Phe Leu Ser Ala  
85 90 95

Ala Leu Leu Pro Ala Cys Thr Leu Gly Val Ala Ala Leu Gly Leu Ala  
100 105 110

20 Arg Tyr Arg Leu Ile Val His Pro Leu Arg Pro Gly Ser Arg Pro Pro  
115 120 125

Pro Val Leu Val Leu Thr Ala Val Trp Ala Ala Ala Gly Leu Leu Gly  
130 135 140

25 Ala Leu Ser Leu Leu Gly Pro Pro Pro Ala Pro Pro Pro Ala Pro Ala  
145 150 155 160

Arg Cys Ser Val Leu Ala Gly Leu Gly Pro Phe Arg Pro Leu Trp  
165 170 175

Ala Leu Leu Ala Phe Ala Leu Pro Ala Leu Leu Leu Gly Ala Tyr  
180 185 190

30 Gly Gly Ile Phe Val Val Ala Arg Arg Ala Ala Leu Arg Pro Pro Arg  
195 200 205

Pro Ala Arg Gly Ser Arg Leu Arg Ser Asp Ser Leu Asp Ser Arg Leu  
210 215 220

35 Ser Ile Leu Pro Pro Leu Arg Pro Arg Leu Pro Gly Gly Lys Ala Ala  
225 230 235 240

Leu Ala Pro Ala Leu Ala Val Gly Gln Phe Ala Ala Cys Trp Leu Pro

10

20

30

40

	245	250	255	
	Tyr Gly Cys Ala Cys Leu Ala Pro Ala Ala Arg Ala Ala Glu Ala Glu			
	260	265	270	
5	Ala Ala Val Thr Trp Val Ala Tyr Ser Ala Phe Ala Ala His Pro Phe			
	275	280	285	
	Leu Tyr Gly Leu Leu Gln Arg Pro Val Arg Leu Ala Leu Gly Arg Leu			
	290	295	300	
	Ser Arg Arg Ala Leu Pro Gly Pro Val Arg Ala Cys Thr Pro Gln Ala			
	305	310	315	320
10	Trp His Pro Arg Ala Leu Leu Gln Cys Leu Gln Arg Pro Pro Glu Gly			
	325	330	335	
	Pro Ala Val Gly Pro Ser Glu Ala Pro Glu Gln Thr Pro Glu Leu Ala			
	340	345	350	
15	Gly Gly Arg Ser Pro Ala Tyr Gln Gly Pro Pro Glu Ser Ser Leu Ser			
	355	360	365	

## (8) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1008 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATGGAATCAT CTTTCTCATT TGGAGTGATC CTTGCTGTCC TGGCCTCCCT CATCATTGCT	60
25 ACTAACACAC TAGTGGCTGT GGCTGTGCTG CTGTTGATCC ACAAGAATGA TGGTGTCACT	120
CTCTGCTTCA CCTTGAATCT GGCTGTGGCT GACACCTTGA TTGGTGTGGC CATCTCTGGC	180
CTACTCACAG ACCAGCTCTC CAGCCCTTCT CGGCCACAC AGAAGACCCGT GTCAGCCTG	240
CGGATGGCAT TTGTCACTTC CTCCGCAGCT GCCTCTGTCC TCACGGTCAT GCTGATCACC	300
TTTGACAGGT ACCTTGCCAT CAAGCAGCCC TTCCGCTACT TGAAGATCAT GAGTGGGTTTC	360
30 GTGGCCGGGG CCTGCATTGC CGGGCTGTGG TTAGTGTCTT ACCTCATTGG CTTCCTCCCA	420
CTCGGAATCC CCATGTTCCA GCAGACTGCC TACAAAGGGC AGTGCAGCTT CTTTGCTGTA	480
TTTCACCCCTC ACTTCGTGCT GACCCCTCTCC TGGCTTGCGT TCTTCCCAGC CATGCTCCTC	540
TTTGTCTTCT TCTACTGCGA CATGCTCAAG ATTGCCTCCA TGCACAGCCA GCAGATTGCA	600

30

40

AAGATGGAAC ATGCAGGAGC CATGGCTGGA GGTTATCGAT CCCCCACGGAC TCCCAGCGAC	660
TTCAAAGCTC TCCGTACTGT GTCTGTTCTC ATTGGGAGCT TTGCTCTATC CTGGACCCCC	720
TTCCTTATCA CTGGCATTGT GCAGGTGGCC TGCCAGGAGT GTCACCTCTA CCTAGTGCTG	780
GAACGGTACC TGTGGCTGCT CGGCCTGGGC AACTCCCTGC TCAACCCACT CATCTATGCC	840
5 TATTGGCAGA AGGAGGTGCG ACTGCAGCTC TACCACATGG CCCTAGGAGT GAAGAAGGTG	900
CTCACCTCAT TCCTCCTCTT TCTCTCGGCC AGGAATTGTG GCCCAGAGAG GCCCAGGGAA	960
AGTTCCCTGTC ACATCGTCAC TATCTCCAGC TCAGAGTTG ATGGCTAA	1008

10

## (9) INFORMATION FOR SEQ ID NO:8:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 335 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

## 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Glu Ser Ser Phe Ser Phe Gly Val Ile Leu Ala Val Leu Ala Ser	20
1 5 10 15	
Leu Ile Ile Ala Thr Asn Thr Leu Val Ala Val Ala Val Leu Leu Leu	
20 25 30	
20 Ile His Lys Asn Asp Gly Val Ser Leu Cys Phe Thr Leu Asn Leu Ala	20
35 40 45	
Val Ala Asp Thr Leu Ile Gly Val Ala Ile Ser Gly Leu Leu Thr Asp	
50 55 60	
25 Gln Leu Ser Ser Pro Ser Arg Pro Thr Gln Lys Thr Leu Cys Ser Leu	25
65 70 75 80	
30 Arg Met Ala Phe Val Thr Ser Ser Ala Ala Ala Ser Val Leu Thr Val	30
85 90 95	
Met Leu Ile Thr Phe Asp Arg Tyr Leu Ala Ile Lys Gln Pro Phe Arg	
100 105 110	
30 Tyr Leu Lys Ile Met Ser Gly Phe Val Ala Gly Ala Cys Ile Ala Gly	30
115 120 125	
Leu Trp Leu Val Ser Tyr Leu Ile Gly Phe Leu Pro Leu Gly Ile Pro	
130 135 140	
Met Phe Gln Gln Thr Ala Tyr Lys Gly Gln Cys Ser Phe Phe Ala Val	

20

25

30

40

145	150	155	160		
Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro					
165	170	175			
Ala Met Leu Leu Phe Val Phe Phe Tyr Cys Asp Met Leu Lys Ile Ala					
5	180	185	190		
Ser Met His Ser Gln Gln Ile Arg Lys Met Glu His Ala Gly Ala Met					
195	200	205			
Ala Gly Gly Tyr Arg Ser Pro Arg Thr Pro Ser Asp Phe Lys Ala Leu					
210	215	220		10	
Arg Thr Val Ser Val Leu Ile Gly Ser Phe Ala Leu Ser Trp Thr Pro					
10	225	230	235	240	
Phe Leu Ile Thr Gly Ile Val Gln Val Ala Cys Gln Glu Cys His Leu					
15	245	250	255		
Tyr Leu Val Leu Glu Arg Tyr Leu Trp Leu Leu Gly Val Gly Asn Ser					
15	260	265	270		
Leu Leu Asn Pro Leu Ile Tyr Ala Tyr Trp Gln Lys Glu Val Arg Leu					
275	280	285			
Gln Leu Tyr His Met Ala Leu Gly Val Lys Lys Val Leu Thr Ser Phe					
290	295	300		20	
20	Leu Leu Phe Leu Ser Ala Arg Asn Cys Gly Pro Glu Arg Pro Arg Glu				
305	310	315	320		
Ser Ser Cys His Ile Val Thr Ile Ser Ser Ser Glu Phe Asp Gly					
325	330	335			
(10) INFORMATION FOR SEQ ID NO:9:					
25	(i) SEQUENCE CHARACTERISTICS:				
	(A) LENGTH: 1413 base pairs				
	(B) TYPE: nucleic acid				
	(C) STRANDEDNESS: single				
	(D) TOPOLOGY: linear				30
30	(ii) MOLECULE TYPE: DNA (genomic)				
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:					

ATGGACACTA CCATGGAAGC TGACCTGGGT GCCACTGGCC ACAGGGCCCCG CACAGAGCTT	60
GATGATGAGG ACTCCTACCC CCAAGGTGGC TGGGACACGG TCTTCCTGGT GGCCCTGCTG	120
CTCCTTGGGC TGCCAGCCAA TGGGTTGATG GCGTGGCTGG CCGGCTCCCA GGCCCGGCAT	180
35 GGAGCTGGCA CGCGTCTGGC GCTGCTCCTG CTCAGCCTGG CCCTCTCTGA CTTCTTGTTC	240

CTGGCAGCAG CGGCCTCCA GATCCTAGAG ATCCGGCATG GGGGACACTG GCCGCTGGGG	300
ACAGCTGCCT GCCGCTTCTA CTACTTCCTA TGGGGCGTGT CCTACTCCTC CGGCCTCTTC	360
CTGCTGGCCG CCCTCAGCCT CGACCGCTGC CTGCTGGCGC TGTGCCACCA CTGGTACCCCT	420
GGGCACCGCC CAGTCCGCCT GCCCCTCTGG GTCTGCGCCG GTGTCTGGGT GCTGCCACCA	480
5 CTCTTCAGCG TGCCCTGGCT GGTCTTCCCC GAGGCTGCCG TCTGGTGGTA CGACCTGGTC	540
ATCTGCCTGG ACTTCTGGGA CAGCGAGGAG CTGTCGCTGA GGATGCTGGA GGTCTGGGG	600
GGCTTCCCTGC CTTTCCCTCT GCTGCTCGTC TGCCACGTGC TCACCCAGGC CACAGCCTGT	660
CGCACCTGCC ACCGCCAACA GCAGCCCGCA GCCTGCCGG GCTTCGCCCG TGTGCCAGG	720
ACCATTCTGT CAGCCTATGT GGTCTGAGG CTGCCCTACC AGCTGCCCA GCTGCTCTAC	780
10 CTGGCCTTCC TGTGGGACGT CTACTCTGGC TACCTGCTCT GGGAGGCCCT GGTCTACTCC	840
GACTACCTGA TCCTACTCAA CAGCTGCCTC AGCCCCTTCC TCTGCCTCAT GGCCAGTGCC	900
GACCTCCGGGA CCCTGCTGCG CTCCGTGCTC TCGTCCTTCG CGGCAGCTCT CTGCGAGGAG	960
CGGCCGGGCA GCTTCACGCC CACTGAGCCA CAGACCCAGC TAGATTCTGA GGGTCCAAC	1020
CTGCCAGAGC CGATGGCAGA GGCCCAGTCA CAGATGGATC CTGTGCCCA GCCTCAGGTG	1080
15 AACCCCCACAC TCCAGCCACG ATCGGATCCC ACAGCTCAGC CACAGCTGAA CCCTACGGCC	1140
CAGCCACAGT CGGATCCCAC AGCCCAGCCA CAGCTGAACC TCATGCCCA GCCACAGTCA	1200
GATTCTGTGG CCCAGCCACA GGCAGACACT AACGTCCAGA CCCCTGCACC TGCTGCCAGT	1260
TCTGTGCCCA GTCCCTGTGA TGAAGCTTCC CCAACCCAT CCTCGCATCC TACCCCAGGG	1320
GCCCTTGAGG ACCCAGCCAC ACCTCCTGCC TCTGAAGGAG AAAGCCCCAG CAGCACCCCG	1380
20 CCAGAGGGCGG CCCCCGGCGC AGGCCCCACG TGA	1413

## (11) INFORMATION FOR SEQ ID NO:10:

- 30  
(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 468 amino acids  
 (B) TYPE: amino acid  
 25 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Asp Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly His Arg Pro			
1	5	10	15

Arg Thr Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly Gly Trp Asp  
 20 25 30

Thr Val Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Pro Ala Asn Gly  
 35 40 45

5 Leu Met Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr  
 50 55 60

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

10 Leu Ala Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His Gly His  
 85 90 95

Trp Pro Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe Leu Trp Gly  
 100 105 110

Val Ser Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu Ser Leu Asp  
 115 120 125

15 Arg Cys Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly His Arg Pro  
 130 135 140

Val Arg Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val Leu Ala Thr  
 145 150 155 160

20 Leu Phe Ser Val Pro Trp Leu Val Phe Pro Glu Ala Ala Val Trp Trp  
 165 170 175

Tyr Asp Leu Val Ile Cys Leu Asp Phe Trp Asp Ser Glu Glu Leu Ser  
 180 185 190

Leu Arg Met Leu Glu Val Leu Gly Gly Phe Leu Pro Phe Leu Leu Leu  
 195 200 205

25 Leu Val Cys His Val Leu Thr Gln Ala Thr Arg Thr Cys His Arg Gln  
 210 215 220

Gln Gln Pro Ala Ala Cys Arg Gly Phe Ala Arg Val Ala Arg Thr Ile  
 225 230 235 240

30 Leu Ser Ala Tyr Val Val Leu Arg Leu Pro Tyr Gln Leu Ala Gln Leu  
 245 250 255

Leu Tyr Leu Ala Phe Leu Trp Asp Val Tyr Ser Gly Tyr Leu Leu Trp  
 260 265 270

Glu Ala Leu Val Tyr Ser Asp Tyr Leu Ile Leu Leu Asn Ser Cys Leu  
 275 280 285

35 Ser Pro Phe Leu Cys Leu Met Ala Ser Ala Asp Leu Arg Thr Leu Leu  
 290 295 300

Arg Ser Val Leu Ser Ser Phe Ala Ala Leu Cys Glu Glu Arg Pro

305	310	315	320		
Gly Ser Phe Thr Pro Thr Glu Pro Gln Thr Gln Leu Asp Ser Glu Gly					
	325	330	335		
5	Pro Thr Leu Pro Glu Pro Met Ala Glu Ala Gln Ser Gln Met Asp Pro				
	340	345	350		
Val Ala Gln Pro Gln Val Asn Pro Thr Leu Gln Pro Arg Ser Asp Pro					
	355	360	365		
10	Thr Ala Gln Pro Gln Leu Asn Pro Thr Ala Gln Pro Gln Ser Asp Pro				
	370	375	380	10	
10	Thr Ala Gln Pro Gln Leu Asn Leu Met Ala Gln Pro Gln Ser Asp Ser				
	385	390	395	400	
Val Ala Gln Pro Gln Ala Asp Thr Asn Val Gln Thr Pro Ala Pro Ala					
	405	410	415		
15	Ala Ser Ser Val Pro Ser Pro Cys Asp Glu Ala Ser Pro Thr Pro Ser				
	420	425	430		
Ser His Pro Thr Pro Gly Ala Leu Glu Asp Pro Ala Thr Pro Pro Ala					
	435	440	445		
Ser Glu Gly Glu Ser Pro Ser Ser Thr Pro Pro Glu Ala Ala Pro Gly					
	450	455	460	20	
20	Ala Gly Pro Thr				
	465				
(12) INFORMATION FOR SEQ ID NO:11:					
(i) SEQUENCE CHARACTERISTICS:					
25	(A) LENGTH: 1248 base pairs				
	(B) TYPE: nucleic acid				
	(C) STRANDEDNESS: single				
	(D) TOPOLOGY: linear				
(ii) MOLECULE TYPE: DNA (genomic)					
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:					
30	ATGTCAGGG A TGGAAAAACT TCAGAATGCT TCCTGGATCT ACCAGCAGAA ACTAGAAGAT				60
	CCATTCCAGA AACACCTGAA CAGCACCGAG GAGTATCTGG CCTTCCTCTG CGGACCTCGG				120
	CGCAGCCACT TCTTCCTCCC CGTGTCTGTG GTGTATGTGC CAATTTTGTT GGTGGGGGTC				180
	ATTGGCAATG TCCTGGTGTG CCTGGTGATT CTGCAGCACC AGGCTATGAA GACGCCACC				240
	AACTACTACC TCTTCAGCCT GGCAGGTCTCT GACCTCCTGG TCCTGCTCCT TGGAATGCC				300

CTGGAGGTCT ATGAGATGTG GCGCAACTAC CCTTTCTTGT TCGGGCCCGT GGGCTGCTAC	360
TTCAAGACGG CCCTCTTGA GACCGTGTGC TTCGCCTCCA TCCTCAGCAT CACCACCGTC	420
AGCGTGGAGC GCTACGTGGC CATCCTACAC CCGTTCCGCG CCAAACGTGCA GAGCACCCGG	480
CGCCGGGCC CGAGGATCCT CGGCATCGTC TGGGGCTTCT CCGTGCTCTT CTCCCTGCC	540
5 AACACCAGCA TCCATGGCAT CAAGTTCCAC TACTTCCCCA ATGGGTCCCT GGTCCCAGGT	600
TCGGCCACCT GTACGGTCAT CAAGCCCATG TGGATCTACA ATTTCATCAT CCAGGTCACC	660
TCCTTCCTAT TCTACCTCCT CCCCATGACT GTCATCAGTG TCCTCTACTA CCTCATGGCA	720
CTCAGACTAA AGAAAGACAA ATCTCTTGAG GCAGATGAAG GGAATGCAA TATTCAAAGA	780
CCCTGCAGAA AATCAGTCAA CAAGATGCTG TTTGTCTTGG TCTTAGTGTG TGCTATCTGT	840
10 TGGGGCCCGT TCCACATTGA CCGACTCTTC TTCAGCTTTG TGGAGGAGTG GAGTGAATCC	900
CTGGCTGCTG TGTTCAACCT CGTCCATGTG GTGTCAGGTG TCTTCTTCTA CCTGAGCTCA	960
GCTGTCAACC CCATTATCTA TAAACCTACTG TCTCGCCGCT TCCAGGCAGC ATTCCAGAAT	1020
GTGATCTCTT CTTTCCACAA ACAGTGGCAC TCCCAGCATG ACCCACAGTT GCCACCTGCC	1080
CAGCGGAACA TCTTCCTGAC AGAATGCCAC TTTGTGGAGC TGACCGAAGA TATAGGTCCC	1140
15 CAATTCCCAT GTCAGTCATC CATGCACAAAC TCTCACCTCC CAACAGCCCT CTCTAGTGAA	1200
CAGATGTCAA GAACAAACTA TCAAAGCTTC CACTTTAACAA AACCTGA	1248

## (13) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 415 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

25 Met Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln	
1 5 10 15	
Lys Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr	
20 25 30	
Leu Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val	
35 40 45	
Ser Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val	

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

25 Met Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln	
1 5 10 15	
Lys Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr	
20 25 30	
Leu Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val	
35 40 45	
Ser Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val	

40

	50	55	60	
	Leu Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr			
65	65	70	75	80
5	Asn Tyr Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu			
	85	90	95	
	Leu Gly Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe			
	100	105	110	
	Leu Phe Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr			
	115	120	125	
10	Val Cys Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg			
	130	135	140	
	Tyr Val Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg			
	145	150	155	160
15	Arg Arg Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu			
	165	170	175	
	Phe Ser Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe			
	180	185	190	
	Pro Asn Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys			
	195	200	205	
20	Pro Met Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe			
	210	215	220	
	Tyr Leu Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala			
	225	230	235	240
25	Leu Arg Leu Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala			
	245	250	255	
	Asn Ile Gln Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Phe Val			
	260	265	270	
	Leu Val Leu Val Phe Ala Ile Cys Trp Ala Pro Phe His Ile Asp Arg			
	275	280	285	
30	Leu Phe Phe Ser Phe Val Glu Glu Trp Ser Glu Ser Leu Ala Ala Val			
	290	295	300	
	Phe Asn Leu Val His Val Val Ser Gly Val Phe Phe Tyr Leu Ser Ser			
	305	310	315	320
35	Ala Val Asn Pro Ile Ile Tyr Asn Leu Leu Ser Arg Arg Phe Gln Ala			
	325	330	335	
	Ala Phe Gln Asn Val Ile Ser Ser Phe His Lys Gln Trp His Ser Gln			
	340	345	350	

His	Asp	Pro	Gln	Leu	Pro	Pro	Ala	Gln	Arg	Asn	Ile	Phe	Leu	Thr	Glu	
355							360						365			
Cys	His	Phe	Val	Glu	Leu	Thr	Glu	Asp	Ile	Gly	Pro	Gln	Phe	Pro	Cys	
370							375					380				
5	Gln	Ser	Ser	Met	His	Asn	Ser	His	Leu	Pro	Thr	Ala	Leu	Ser	Ser	Glu
385							390				395		400			
	Gln	Met	Ser	Arg	Thr	Asn	Tyr	Gln	Ser	Phe	His	Phe	Asn	Lys	Thr	
							405				410		415			

## (14) INFORMATION FOR SEQ ID NO:13:

10

- 10 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1173 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- 15 (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATGCCAGATA	CTAATAGCAC	AATCAATTAA	TCACTAAGCA	CTCGTGTAC	TTTAGCATT	60		
TTTATGTCCT	TAGTAGCTTT	TGCTATAATG	CTAGGAAATG	CTTTGGTCAT	TTTAGCTTTT	120	20	
GTGGTGGACA	AAAACCTTAG	ACATCGAAGT	AGTTATTTT	TTCTTAACCT	GGCCATCTCT	180		
20 GACTTCTTGT	TGGGTGTGAT	CTCCATTCT	TTGTACATCC	CTCACACGCT	GTTGAAATGG	240		
GATTTTGAA	AGGAAATCTG	TGTATTTGG	CTCACTACTG	ACTATCTGTT	ATGTACAGCA	300		
TCTGTATATA	ACATTGTCCT	CATCAGCTAT	GATCGATAACC	TGTCAGTCTC	AAATGCTGTG	360		
TCTTATAGAA	CTCAACATAC	TGGGGTCTTG	AAGATTGTTA	CTCTGATGGT	GGCGTTTGG	420		
GTGCTGGCCT	TCTTAGTGAA	TGGGCCATG	ATTCTAGTTT	CAGAGTCTTG	GAAGGGATGAA	480		
25 GGTAGTGAAT	GTGAACCTGG	ATTTTTTCG	GAATGGTACA	TCCTTGCCT	CACATCATTC	540	30	
TTGGAATTG	TGATCCCAGT	CATCTTAGTC	GCTTATTCA	ACATGAATAT	TTATTGGAGC	600		
CTGTGGAAAGC	GTGATCATCT	CAGTAGGTGC	CAAAGCCATC	CTGGACTGAC	TGCTGTCTCT	660		
TCCAACATCT	GTGGACACTC	ATTCAAGAGT	AGACTATCTT	CAAGGAGATC	TCTTCTGCA	720		
TCGACAGAAG	TTCCTGCATC	CTTTCATTCA	GAGAGACAGA	GGAGAAAGAG	TAGTCTCATG	780		
30 TTTTCCTCAA	GAACCAAGAT	GAATAGCAAT	ACAATTGCTT	CCAAAATGGG	TTCCCTCTCC	840		
CAATCAGATT	CTGTAGCTCT	TCACCAAAGG	GAACATGTTG	AACTGCTTAG	AGCCAGGAGA	900		

TTAGCCAAGT CACTGGCCAT TCTCTTAGGG GTTTTGCTG TTTGCTGGC TCCATATTCT 960  
 CTGTTCACAA TTGTCCTTTC ATTTTATTCC TCAGCAACAG GTCCTAAATC AGTTTGGTAT 1020  
 AGAATTGCAT TTTGGCTTCA GTGGTTCAAT TCCTTGTCA ATCCTCTTT GTATCCATTG 1080  
 TGTCACAAGC GCTTTCAAAA GGCTTCTTG AAAATATTTT GTATAAAAAA GCAACCTCTA 1140  
 5 CCATCACAAAC ACAGTCGGTC AGTATCTTCT TAA 1173

## (15) INFORMATION FOR SEQ ID NO:14:

- 10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 390 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

15 Met Pro Asp Thr Asn Ser Thr Ile Asn Leu Ser Leu Ser Thr Arg Val  
 1 5 10 15  
 Thr Leu Ala Phe Phe Met Ser Leu Val Ala Phe Ala Ile Met Leu Gly  
 20 25 30  
 Asn Ala Leu Val Ile Leu Ala Phe Val Val Asp Lys Asn Leu Arg His  
 35 40 45  
 20 Arg Ser Ser Tyr Phe Phe Leu Asn Leu Ala Ile Ser Asp Phe Phe Val  
 50 55 60  
 Gly Val Ile Ser Ile Pro Leu Tyr Ile Pro His Thr Leu Phe Glu Trp  
 65 70 75 80  
 Asp Phe Gly Lys Glu Ile Cys Val Phe Trp Leu Thr Thr Asp Tyr Leu  
 25 85 90 95  
 Leu Cys Thr Ala Ser Val Tyr Asn Ile Val Leu Ile Ser Tyr Asp Arg  
 100 105 110  
 Tyr Leu Ser Val Ser Asn Ala Val Ser Tyr Arg Thr Gln His Thr Gly  
 115 120 125  
 30 Val Leu Lys Ile Val Thr Leu Met Val Ala Val Trp Val Leu Ala Phe  
 130 135 140  
 Leu Val Asn Gly Pro Met Ile Leu Val Ser Glu Ser Trp Lys Asp Glu  
 145 150 155 160  
 35 Gly Ser Glu Cys Glu Pro Gly Phe Phe Ser Glu Trp Tyr Ile Leu Ala  
 165 170 175

10

20

30

40

Ile Thr Ser Phe Leu Glu Phe Val Ile Pro Val Ile Leu Val Ala Tyr  
 180 185 190  
 Phe Asn Met Asn Ile Tyr Trp Ser Leu Trp Lys Arg Asp His Leu Ser  
 195 200 205  
 5 Arg Cys Gln Ser His Pro Gly Leu Thr Ala Val Ser Ser Asn Ile Cys  
 210 215 220  
 Gly His Ser Phe Arg Gly Arg Leu Ser Ser Arg Arg Ser Leu Ser Ala  
 225 230 235 240  
 Ser Thr Glu Val Pro Ala Ser Phe His Ser Glu Arg Gln Arg Arg Lys  
 10 245 250 255  
 Ser Ser Leu Met Phe Ser Ser Arg Thr Lys Met Asn Ser Asn Thr Ile  
 260 265 270  
 Ala Ser Lys Met Gly Ser Phe Ser Gln Ser Asp Ser Val Ala Leu His  
 275 280 285  
 15 Gln Arg Glu His Val Glu Leu Leu Arg Ala Arg Arg Leu Ala Lys Ser  
 290 295 300  
 Leu Ala Ile Leu Leu Gly Val Phe Ala Val Cys Trp Ala Pro Tyr Ser  
 305 310 315 320  
 Leu Phe Thr Ile Val Leu Ser Phe Tyr Ser Ser Ala Thr Gly Pro Lys  
 20 325 330 335  
 Ser Val Trp Tyr Arg Ile Ala Phe Trp Leu Gln Trp Phe Asn Ser Phe  
 340 345 350  
 Val Asn Pro Leu Leu Tyr Pro Leu Cys His Lys Arg Phe Gln Lys Ala  
 355 360 365  
 25 Phe Leu Lys Ile Phe Cys Ile Lys Lys Gln Pro Leu Pro Ser Gln His  
 370 375 380  
 Ser Arg Ser Val Ser Ser  
 385 390  
 (16) INFORMATION FOR SEQ ID NO:15: 30  
 30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 35 (ii) MOLECULE TYPE: DNA (genomic)  
 (iv) ANTI-SENSE: NO  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGAAAGCTTA ACGATCCCCA GGAGCAACAT

30

## (17) INFORMATION FOR SEQ ID NO:16:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iv) ANTI-SENSE: YES
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
- CTGGGATCCT ACGAGAGCAT TTTTCACACA G  
 31
- (18) INFORMATION FOR SEQ ID NO:17:
- 15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1128 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATGGCGAACG CGAGCGAGCC GGGTGGCAGC GGCGGGCG AGGCAGGCCGC CCTGGGCCTC	60
AAGCTGGCCA CGCTCAGCCT GCTGCTGTGC GTGAGCCTAG CGGGCAACGT GCTGTTCGCG	120
CTGCTGATCG TGCGGGAGCG CAGCCTGCAC CGCGCCCCGT ACTACCTGCT GCTCGACCTG	180
TGCCCTGGCG ACGGGCTGCG CGCGCTCGCC TGCCCTCCGG CGGTCAATGCT GGCGGCGCGG	240
25 CGTGCGGCCGG CGCGGGCGGG GGCAGGCCGC GGCGCGCTGG GCTGCAAGCT GCTCGCCTTC	300
CTGGCCCGCGC TCTTCTGCTT CCACGCCGCC TTCCCTGCTGC TGGGCGTGGG CGTCACCCGC	360
TACCTGGCCA TCGCGCACCA CCGCTTCTAT GCAGAGCGCC TGGCCGGCTG GCCGTGCGCC	420
GCCATGCTGG TGTGCGCCGC CTGGGCGCTG GCGCTGGCG CGGCCTTCCC GCCAGTGCTG	480
GACGGCGGTG GCGACGACGA GGACGCGCCG TGCGCCCTGG AGCAGCGGCC CGACGGCGCC	540
30 CCCGGCGCGC TGGGCTTCCT GCTGCTGCTG GCCGTGGTGG TGGCGCCAC GCACCTCGTC	600
TACCTCCGCC TGCTCTTCTT CATCCACGAC CGCCGCAAGA TGCGGCCCCGC CGCCTGGTG	660

CCCGCCGTCA GCCACGACTG GACCTTCCAC GGCCCCGGCG CCACCGGCCA GGCGGCCGCC 720  
 AACTGGACGG CGGGCTTCGG CGCGGGGCCA ACGCCGCCCG CGCTTGTGGG CATCCGGCCC 780  
 GCAGGGCCGG 840  
 GCAGGGCCGG 840  
 AGGCTGTGCA AGATGTTCTA CGCCGTCACG CTGCTCTTCC TGCTCCTCTG GGGGCCCTAC 900  
 5 GTCGTGGCCA GCTACCTGCG GGTCTGGTG CGGCCCGGCG CCGTCCCCCA GGCCTACCTG 960  
 ACGGCCTCCG TGTGGCTGAC CTTCGCGCAG GCCGGCATCA ACCCCGTCGT GTGCTTCCTC 1020  
 TTCAACAGGG AGCTGAGGG 1080 10  
 ACCACCCAGG CGACCCATCC CTGCGACCTG AAAGGCATTG GTTTATGA 1128

## (19) INFORMATION FOR SEQ ID NO:18:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 375 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

15 (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18: 20

Met Ala Asn Ala Ser Glu Pro Gly Gly Ser Gly Gly Gly Glu Ala Ala  
 1 5 10 15  
 Ala Leu Gly Leu Lys Leu Ala Thr Leu Ser Leu Leu Leu Cys Val Ser  
 20 25 30  
 Leu Ala Gly Asn Val Leu Phe Ala Leu Leu Ile Val Arg Glu Arg Ser  
 35 40 45  
 Leu His Arg Ala Pro Tyr Tyr Leu Leu Leu Asp Leu Cys Leu Ala Asp  
 50 55 60  
 25 Gly Leu Arg Ala Leu Ala Cys Leu Pro Ala Val Met Leu Ala Ala Arg  
 65 70 75 80 30  
 Arg Ala Ala Ala Ala Gly Ala Pro Pro Gly Ala Leu Gly Cys Lys  
 85 90 95  
 Leu Leu Ala Phe Leu Ala Ala Leu Phe Cys Phe His Ala Ala Phe Leu  
 100 105 110  
 Leu Leu Gly Val Gly Val Thr Arg Tyr Leu Ala Ile Ala His His Arg  
 115 120 125  
 Phe Tyr Ala Glu Arg Leu Ala Gly Trp Pro Cys Ala Ala Met Leu Val  
 130 135 140

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

## (20) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1002 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ATGAAACACCA	CA GTGATGCA	AGGCTTCAAC	AGATCTGAGC	GGTCCCCAG	AGACACTCGG	60	
ATAGTACAGC	TGGTATTCCC	AGCCCTCTAC	ACAGTGGTT	TCTTGACCGG	CATCCTGCTG	120	
AATACTTTGG	CTCTGTGGGT	GTTTGTTCAC	ATCCCCAGCT	CCTCACCTT	CATCATCTAC	180	
5	CTCAAAAACA	CTTGGTGGC	CGACTTGATA	ATGACACTCA	TGCTTCCTTT	CAAATCCTC	240
	TCTGACTCAC	ACCTGGCACC	CTGGCAGCTC	AGAGCTTTG	TGTGTCGTTT	TTCTTCGGTG	300
	ATATTTATG	AGACCATGTA	TGTGGGCATC	GTGCTGTTAG	GGCTCATAGC	CTTGACAGA	360
	TTCCTCAAGA	TCATCAGACC	TTTGAGAAAT	ATTTTCTAA	AAAAACCTGT	TTTGCAAAA	420
	ACGGTCTCAA	TCTTCATCTG	GTTCTTTTG	TTCTTCATCT	CCCTGCCAAA	TACGATCTTG	480
10	AGCAACAAAGG	AAGCAACACC	ATCGTCTGTG	AAAAAGTGTG	CTTCCTTAAA	GGGGCCTCTG	540
	GGGCTGAAAT	GGCATCAAAT	GGTAAATAAC	ATATGCCAGT	TTATTTCTG	GACTGTTTT	600
	ATCCTAATGC	TTGTGTTTA	TGTGGTTATT	GCAAAAAAAG	TATATGATTC	TTATAGAAAG	660
	TCCAAAAGTA	AGGACAGAAA	AAACAAACAAA	AAGCTGGAAG	GCAAAGTATT	TGTTGTCGTG	720
	GCTGTCTTCT	TTGTGTTTT	TGCTCCATT	CATTTGCCA	GAGTCCATA	TACTCACAGT	780
15	CAAACCAACA	ATAAGACTGA	CTGTAGACTG	CAAAATCAAC	TGTTTATTGC	TAAAGAAACA	840
	ACTCTCTTTT	TGGCAGCAAC	TAACATTTGT	ATGGATCCCT	TAATATACAT	ATTCTTATGT	900
	AAAAAATTCA	CAGAAAAGCT	ACCATGTATG	CAAGGGAGAA	AGACCACAGC	ATCAAGCCAA	960
	GAAAATCATA	GCAGTCAGAC	AGACAACATA	ACCTTAGGCT	GA		1002

## (21) INFORMATION FOR SEQ ID NO:20:

20	(i) SEQUENCE CHARACTERISTICS:						
	(A) LENGTH: 333 amino acids						
	(B) TYPE: amino acid						
	(C) STRANDEDNESS:						
	(D) TOPOLOGY: not relevant						
25	(ii) MOLECULE TYPE: protein						30

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met	Asn	Thr	Thr	Val	Met	Gln	Gly	Phe	Asn	Arg	Ser	Glu	Arg	Cys	Pro	
1					5				10				15			
30	Arg	Asp	Thr	Arg	Ile	Val	Gln	Leu	Val	Phe	Pro	Ala	Leu	Tyr	Thr	Val
					20				25				30			

Val Phe Leu Thr Gly Ile Leu Leu Asn Thr Leu Ala Leu Trp Val Phe  
 35 40 45

Val His Ile Pro Ser Ser Ser Thr Phe Ile Ile Tyr Leu Lys Asn Thr  
 50 55 60

5 Leu Val Ala Asp Leu Ile Met Thr Leu Met Leu Pro Phe Lys Ile Leu  
 65 70 75 80

Ser Asp Ser His Leu Ala Pro Trp Gln Leu Arg Ala Phe Val Cys Arg  
 85 90 95

Phe Ser Ser Val Ile Phe Tyr Glu Thr Met Tyr Val Gly Ile Val Leu  
 10 100 105 110

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

Arg Asn Ile Phe Leu Lys Lys Pro Val Phe Ala Lys Thr Val Ser Ile  
 130 135 140

15 Phe Ile Trp Phe Phe Leu Phe Phe Ile Ser Leu Pro Asn Thr Ile Leu  
 145 150 155 160

Ser Asn Lys Glu Ala Thr Pro Ser Ser Val Lys Lys Cys Ala Ser Leu  
 165 170 175

Lys Gly Pro Leu Gly Leu Lys Trp His Gln Met Val Asn Asn Ile Cys  
 20 180 185 190

Gln Phe Ile Phe Trp Thr Val Phe Ile Leu Met Leu Val Phe Tyr Val  
 195 200 205

Val Ile Ala Lys Lys Val Tyr Asp Ser Tyr Arg Lys Ser Lys Ser Lys  
 210 215 220

25 Asp Arg Lys Asn Asn Lys Lys Leu Glu Gly Lys Val Phe Val Val Val  
 225 230 235 240

Ala Val Phe Phe Val Cys Phe Ala Pro Phe His Phe Ala Arg Val Pro  
 245 250 255

Tyr Thr His Ser Gln Thr Asn Asn Lys Thr Asp Cys Arg Leu Gln Asn  
 30 260 265 270

Gln Leu Phe Ile Ala Lys Glu Thr Thr Leu Phe Leu Ala Ala Thr Asn  
 275 280 285

Ile Cys Met Asp Pro Leu Ile Tyr Ile Phe Leu Cys Lys Lys Phe Thr  
 290 295 300

35 Glu Lys Leu Pro Cys Met Gln Gly Arg Lys Thr Thr Ala Ser Ser Gln  
 305 310 315 320

Glu Asn His Ser Ser Gln Thr Asp Asn Ile Thr Leu Gly

325

330

## (22) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1122 base pairs  
 5 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

10	ATGGCCAACA CTACCGGAGA GCCTGAGGAG GTGAGCGGCG CTCTGTCCCC ACCGTCCGCA	60
	TCAGCTTATG TGAAGCTGGT ACTGCTGGGA CTGATTATGT GCGTGAGCCT GGCGGGTAAC	120
	GCCATCTTGT CCCTGCTGGT GCTCAAGGAG CGTGCCCTGC ACAAGGCTCC TTACTACTTC	180
	CTGCTGGACC TGTGCCTGGC CGATGGCATA CGCTCTGCCG TCTGCTTCCC CTTTGTGCTG	240
	GCTTCTGTGC GCCACGGCTC TTCATGGACC TTCAGTGCAC TCAGCTGCAA GATTGTGGCC	300
15	TTTATGGCCG TGCTCTTTG CTTCCATGCG GCCTTCATGC TGTTCTGCAT CAGCGTCACC	360
	CGCTACATGG CCATCGCCCA CCACCGCTTC TACGCCAAGC GCATGACACT CTGGACATGC	420
	GCGGCTGTCA TCTGCATGGC CTGGACCCCTG TCTGTGGCCA TGGCCTTCCC ACCTGTCTTT	480
	GACGTGGGCA CCTACAAAGTT TATTGGGAG GAGGACCAAGT GCATCTTGA GCATCGCTAC	540
	TTCAAGGCCA ATGACACGCT GGGCTTCATG CTTATGTTGG CTGTGCTCAT GGCAGCTACC	600
20	CATGCTGTCT ACGGCAAGCT GCTCCTCTTC GAGTATCGTC ACCGCAAGAT GAAGCCAGTG	660
	CAGATGGTGC CAGCCATCAG CCAGAACTGG ACATCCATG GTCCCCGGGC CACCGGCCAG	720
	GCTGCTGCCA ACTGGATCGC CGGCTTTGGC CGTGGGCCA TGCCACCAAC CCTGCTGGGT	780
	ATCCGGCAGA ATGGGCATGC AGCCAGCCGG CGGCTACTGG GCATGGACGA GGTCAAGGGT	840
	GAAAAGCAGC TGGGCCGCAT GTTCTACGCG ATCACACTGC TCTTCTGCT CCTCTGGTCA	900
25	CCCTACATCG TGGCTGCTA CTGGCGAGTG TTTGTAAAG CCTGTCGTGT GCCCCACCGC	960
	TACCTGGCCA CTGCTGTTG GATGAGCTTC GCCCAGGCTG CCGTCAACCC AATTGTCTGC	1020
	TTCTGCTCA ACAAGGACCT CAAGAAGTGC CTGACCACTC ACGCCCCCTG CTGGGGCACA	1080
	GGAGGTGCCG CGGCTCCCAG AGAACCCCTAC TGTGTCATGT GA	1122

## (23) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 373 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

5

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

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

10

Pro Pro Ser Ala Ser Ala Tyr Val Lys Leu Val Leu Leu Gly Leu Ile  
 20 25 30

15

Met Cys Val Ser Leu Ala Gly Asn Ala Ile Leu Ser Leu Leu Val Leu  
 35 40 45

15

Lys Glu Arg Ala Leu His Lys Ala Pro Tyr Tyr Phe Leu Leu Asp Leu  
 50 55 60

Cys Leu Ala Asp Gly Ile Arg Ser Ala Val Cys Phe Pro Phe Val Leu  
 65 70 75 80

Ala Ser Val Arg His Gly Ser Ser Trp Thr Phe Ser Ala Leu Ser Cys  
 85 90 95

20

Lys Ile Val Ala Phe Met Ala Val Leu Phe Cys Phe His Ala Ala Phe  
 100 105 110

Met Leu Phe Cys Ile Ser Val Thr Arg Tyr Met Ala Ile Ala His His  
 115 120 125

25

Arg Phe Tyr Ala Lys Arg Met Thr Leu Trp Thr Cys Ala Ala Val Ile  
 130 135 140

Cys Met Ala Trp Thr Leu Ser Val Ala Met Ala Phe Pro Pro Val Phe  
 145 150 155 160

Asp Val Gly Thr Tyr Lys Phe Ile Arg Glu Glu Asp Gln Cys Ile Phe  
 165 170 175

30

Glu His Arg Tyr Phe Lys Ala Asn Asp Thr Leu Gly Phe Met Leu Met  
 180 185 190

Leu Ala Val Leu Met Ala Ala Thr His Ala Val Tyr Gly Lys Leu Leu  
 195 200 205

35

Leu Phe Glu Tyr Arg His Arg Lys Met Lys Pro Val Gln Met Val Pro  
 210 215 220

Ala Ile Ser Gln Asn Trp Thr Phe His Gly Pro Gly Ala Thr Gly Gln  
 225 230 235 240

10

20

30

40

Ala Ala Ala Asn Trp Ile Ala Gly Phe Gly Arg Gly Pro Met Pro Pro		
245	250	255
Thr Leu Leu Gly Ile Arg Gln Asn Gly His Ala Ala Ser Arg Arg Leu		
260	265	270
5 Leu Gly Met Asp Glu Val Lys Gly Glu Lys Gln Leu Gly Arg Met Phe		
275	280	285
Tyr Ala Ile Thr Leu Leu Phe Leu Leu Trp Ser Pro Tyr Ile Val		
290	295	300
10 Ala Cys Tyr Trp Arg Val Phe Val Lys Ala Cys Ala Val Pro His Arg		
305	310	315
Tyr Leu Ala Thr Ala Val Trp Met Ser Phe Ala Gln Ala Ala Val Asn		
325	330	335
Pro Ile Val Cys Phe Leu Leu Asn Lys Asp Leu Lys Lys Cys Leu Thr		
340	345	350
15 Thr His Ala Pro Cys Trp Gly Thr Gly Gly Ala Pro Ala Pro Arg Glu		
355	360	365
Pro Tyr Cys Val Met		
370		
(24) INFORMATION FOR SEQ ID NO:23:		
20	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1053 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
25	(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:		
ATGGCTTTGG AACAGAACCA GTCAACAGAT TATTATTATG AGGAAAATGA AATGAATGGC	60	
ACTTATGACT ACAGTCAATA TGAATTGATC TGTATCAAAG AAGATGTCAG AGAATTGCA	120	30
AAAGTTTTC CTCCTGTATT CCTCACAATA GCTTTCGTCA TTGGACTTGC AGGCAATTCC	180	
30 ATGGTAGTGG CAATTTATGC CTATTACAAG AAACAGAGAA CCAAAACAGA TGTGTACATC	240	
CTGAATTGAG CTGTAGCAGA TTTACTCCTT CTATTCACTC TGCCTTTTG GGCTGTTAAT	300	
GCAGTTCATG GGTGGGTTTT AGGGAAAATA ATGTGAAAA TAACTTCAGC CTTGTACACA	360	
CTAAACTTTG TCTCTGGAAT GCAGTTCTG GCTTGCATCA GCATAGACAG ATATGTGGCA	420	
GTAACATAATG TCCCCAGCCA ATCAGGAGTG GGAAAACCAT GCTGGATCAT CTGTTTCTGT	480	

GTCTGGATGG CTGCCATCTT GCTGAGCATA CCCCAGCTGG TTTTTTATAC AGTAAATGAC 540  
 AATGCTAGGT GCATTCCAT TTTCCCCGC TACCTAGGAA CATCAATGAA AGCATTGATT 600  
 CAAATGCTAG AGATCTGCAT TGGATTGTA GTACCCTTC TTATTATGGG GGTGTGCTAC 660  
 TTTATCACCG CAAGGACACT CATGAAGATG CCAAACATTA AAATATCTCG ACCCCTAAAA 720  
 5 GTTCTGCTCA CAGTCGTTAT AGTTTCATT GTCACTAAC TGCCCTATAA CATTGTCAAG 780  
 TTCTGCCGAG CCATAGACAT CATCTACTCC CTGATCACCA GCTGCAACAT GAGCAAACGC 840  
 ATGGACATCG CCATCCAAGT CACAGAAAGC ATTGCACTCT TTCACAGCTG CCTCAACCCA 900 10  
 ATCCTTTATG TTTTTATGGG AGCATCTTC AAAAACTACG TTATGAAAGT GGCCAAGAAA 960  
 TATGGGTCCCT GGAGAAGACA GAGACAAAGT GTGGAGGAGT TTCCTTTGA TTCTGAGGGT 1020  
 10 CCTACAGAGC CAACCAGTAC TTTAGCATT TAA 1053

## (25) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 350 amino acids  
 (B) TYPE: amino acid  
 15 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant 20

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Ala Leu Glu Gln Asn Gln Ser Thr Asp Tyr Tyr Tyr Glu Glu Asn  
 20 1 5 10 15

Glu Met Asn Gly Thr Tyr Asp Tyr Ser Gln Tyr Glu Leu Ile Cys Ile  
 20 25 30

Lys Glu Asp Val Arg Glu Phe Ala Lys Val Phe Leu Pro Val Phe Leu  
 35 40 45

25 Thr Ile Ala Phe Val Ile Gly Leu Ala Gly Asn Ser Met Val Val Ala  
 50 55 60

Ile Tyr Ala Tyr Tyr Lys Lys Gln Arg Thr Lys Thr Asp Val Tyr Ile  
 65 70 75 80

Leu Asn Leu Ala Val Ala Asp Leu Leu Leu Phe Thr Leu Pro Phe  
 30 85 90 95

Trp Ala Val Asn Ala Val His Gly Trp Val Leu Gly Lys Ile Met Cys  
 100 105 110

Lys Ile Thr Ser Ala Leu Tyr Thr Leu Asn Phe Val Ser Gly Met Gln

20

30

40

	115	120	125	
	Phe Leu Ala Cys Ile Ser Ile Asp Arg Tyr Val Ala Val Thr Asn Val			
	130	135	140	
5	Pro Ser Gln Ser Gly Val Gly Lys Pro Cys Trp Ile Ile Cys Phe Cys			
	145	150	155	160
	Val Trp Met Ala Ala Ile Leu Leu Ser Ile Pro Gln Leu Val Phe Tyr			
	165	170	175	
	Thr Val Asn Asp Asn Ala Arg Cys Ile Pro Ile Phe Pro Arg Tyr Leu			
	180	185	190	
10	Gly Thr Ser Met Lys Ala Leu Ile Gln Met Leu Glu Ile Cys Ile Gly			
	195	200	205	
	Phe Val Val Pro Phe Leu Ile Met Gly Val Cys Tyr Phe Ile Thr Ala			
	210	215	220	
15	Arg Thr Leu Met Lys Met Pro Asn Ile Lys Ile Ser Arg Pro Leu Lys			
	225	230	235	240
	Val Leu Leu Thr Val Val Ile Val Phe Ile Val Thr Gln Leu Pro Tyr			
	245	250	255	
	Asn Ile Val Lys Phe Cys Arg Ala Ile Asp Ile Ile Tyr Ser Leu Ile			
	260	265	270	
20	Thr Ser Cys Asn Met Ser Lys Arg Met Asp Ile Ala Ile Gln Val Thr			
	275	280	285	
	Glu Ser Ile Ala Leu Phe His Ser Cys Leu Asn Pro Ile Leu Tyr Val			
	290	295	300	
25	Phe Met Gly Ala Ser Phe Lys Asn Tyr Val Met Lys Val Ala Lys Lys			
	305	310	315	320
	Tyr Gly Ser Trp Arg Arg Gln Arg Gln Ser Val Glu Glu Phe Pro Phe			
	325	330	335	
	Asp Ser Glu Gly Pro Thr Glu Pro Thr Ser Thr Phe Ser Ile			
	340	345	350	
30	(26) INFORMATION FOR SEQ ID NO:25:			
	(i) SEQUENCE CHARACTERISTICS:			
	(A) LENGTH: 1116 base pairs			
	(B) TYPE: nucleic acid			
	(C) STRANDEDNESS: single			
35	(D) TOPOLOGY: linear			
	(ii) MOLECULE TYPE: DNA (genomic)			
	30			

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATGCCAGGAA	ACGCCACCCC	AGTGACCACC	ACTGCCCCGT	GGGCCTCCCT	GGGCCTCTCC	60	
GCCAAGACCT	GCAACAACGT	GTCCTTCGAA	GAGAGCAGGA	TAGTCCTGGT	CGTGGTGTAC	120	
AGCGCGGTGT	GCACGCTGGG	GGTGCCGGCC	AACTGCCTGA	CTGCGTGGCT	GGCGCTGCTG	180	
5	CAGGTACTGC	AGGGCAACGT	GCTGGCCGTC	TACCTGCTCT	GCCTGGCACT	CTGCGAACTG	240
CTGTACACAG	GCACGCTGCC	ACTCTGGGTC	ATCTATATCC	GCAACCAGCA	CCGCTGGACC	300	
CTAGGCCTGC	TGGCCTCGAA	GGTGACCGCC	TACATCTTCT	TCTGCAACAT	CTACGTCAGC	360	
ATCCTCTTCC	TGTGCTGCAT	CTCCTGGCAC	CGCTTCGTGG	CCGTGGTGT	CCGGCTGGAG	420	
AGTCGGGGCC	GCCGCCGCCG	GAGGACCGCC	ATCCTCATCT	CCGCCTGCAT	CTTCATCCTC	480	
10	GTCGGGATCG	TTCACTACCC	GGTGTTCAG	ACGGAAGACA	AGGAGACCTG	CTTGACATG	540
CTGCAGATGG	ACAGCAGGAT	TGCCGGGTAC	TACTACGCCA	GGTCACCGT	TGGCTTGCC	600	
ATCCTCTCT	CCATCATCGC	CTTCACCAAC	CACCGGATT	TCAGGAGCAT	CAAGCAGAGC	660	
ATGGGCTTAA	GCGCTGCCA	GAAGGCCAAG	GTGAAGCACT	CGGCCATCGC	GGTGGTTGTC	720	
ATCTTCCTAG	TCTGCTTCGC	CCCGTACAC	CTGGTTCTCC	TCGTCAAAGC	CGCTGCCTTT	780	
15	TCCTACTACA	GAGGAGACAG	GAACGCCATG	TGCGGCTTGG	AGGAAAGGCT	GTACACAGCC	840
TCTGTGGTGT	TTCTGTGCCT	GTCCACGGTG	AACGGCGTGG	CTGACCCCAT	TATCTACGTG	900	
CTGGCCACGG	ACCATTCCCG	CCAAGAAAGTG	TCCAGAAATCC	ATAAGGGGTG	GAAAGAGTGG	960	
TCCATGAAGA	CAGACGTCAC	CAGGCTCACC	CACAGCAGGG	ACACCGAGGA	GCTGCAGTCG	1020	
CCCGTGGCCC	TTGCAGACCA	CTACACCTTC	TCCAGGCCCG	TGCACCCACC	AGGGTCACCA	1080	
20	TGCCCTGCAA	AGAGGCTGAT	TGAGGAGTCC	TGCTGA		1116	

## (28) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:	30
(A) LENGTH: 371 amino acids	
(B) TYPE: amino acid	
25	(C) STRANDEDNESS:
	(D) TOPOLOGY: not relevant
(ii) MOLECULE TYPE: protein	

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met	Pro	Gly	Asn	Ala	Thr	Pro	Val	Thr	Thr	Thr	Ala	Pro	Trp	Ala	Ser
30	1				5			10				15			

Leu Gly Leu Ser Ala Lys Thr Cys Asn Asn Val Ser Phe Glu Glu Ser  
 20 25 30

Arg Ile Val Leu Val Val Val Tyr Ser Ala Val Cys Thr Leu Gly Val  
 35 40 45

5 Pro Ala Asn Cys Leu Thr Ala Trp Leu Ala Leu Leu Gln Val Leu Gln  
 50 55 60

Gly Asn Val Leu Ala Val Tyr Leu Leu Cys Leu Ala Leu Cys Glu Leu  
 65 70 75 80

10 Leu Tyr Thr Gly Thr Leu Pro Leu Trp Val Ile Tyr Ile Arg Asn Gln  
 10 85 90 95

His Arg Trp Thr Leu Gly Leu Leu Ala Ser Lys Val Thr Ala Tyr Ile  
 100 105 110

Phe Phe Cys Asn Ile Tyr Val Ser Ile Leu Phe Leu Cys Cys Ile Ser  
 115 120 125

15 Cys Asp Arg Phe Val Ala Val Val Tyr Ala Leu Glu Ser Arg Gly Arg  
 130 135 140

Arg Arg Arg Arg Thr Ala Ile Leu Ile Ser Ala Cys Ile Phe Ile Leu  
 145 150 155 160

20 Val Gly Ile Val His Tyr Pro Val Phe Gln Thr Glu Asp Lys Glu Thr  
 165 170 175

Cys Phe Asp Met Leu Gln Met Asp Ser Arg Ile Ala Gly Tyr Tyr Tyr  
 180 185 190

Ala Arg Phe Thr Val Gly Phe Ala Ile Pro Leu Ser Ile Ile Ala Phe  
 195 200 205

25 Thr Asn His Arg Ile Phe Arg Ser Ile Lys Gln Ser Met Gly Leu Ser  
 210 215 220

Ala Ala Gln Lys Ala Lys Val Lys His Ser Ala Ile Ala Val Val Val  
 225 230 235 240

30 Ile Phe Leu Val Cys Phe Ala Pro Tyr His Leu Val Leu Val Lys  
 245 250 255

Ala Ala Ala Phe Ser Tyr Tyr Arg Gly Asp Arg Asn Ala Met Cys Gly  
 260 265 270

Leu Glu Glu Arg Leu Tyr Thr Ala Ser Val Val Phe Leu Cys Leu Ser  
 275 280 285

35 Thr Val Asn Gly Val Ala Asp Pro Ile Ile Tyr Val Leu Ala Thr Asp  
 290 295 300

	His Ser Arg Gln Glu Val Ser Arg Ile His Lys Gly Trp Lys Glu Trp		
305	310	315	320
	Ser Met Lys Thr Asp Val Thr Arg Leu Thr His Ser Arg Asp Thr Glu		
325	330	335	
5	Glu Leu Gln Ser Pro Val Ala Leu Ala Asp His Tyr Thr Phe Ser Arg		
	340	345	350
	Pro Val His Pro Pro Gly Ser Pro Cys Pro Ala Lys Arg Leu Ile Glu		
	355	360	365
10	Glu Ser Cys		
	370		

## (26) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1113 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATGGCGAACT ATAGCCATGC AGCTGACAAC ATTTTGCAAA ATCTCTGCC TCTAACAGCC	60	20
20 TTTCTGAAAC TGACTTCCTT GGGTTTCATA ATAGGAGTCA GCGTGGTGGG CAACCTCCTG	120	
ATCTCCATTT TGCTAGTGAA AGATAAGACC TTGCATAGAG CACCTTACTA CTTCTGTTG	180	
GATCTTGCT GTTCAGATAT CCTCAGATCT GCAATTGTT TCCCATTGTT GTTCAACTCT	240	
GTCAAAAATG GCTCTACCTG GACTTATGGG ACTCTGACTT GCAAAGTGAT TGCCTTCTG	300	
GGGGTTTGT CCTGTTCCA CACTGCTTTC ATGCTCTCT GCATCAGTGT CACCAGATAC	360	
25 TTAGCTATCG CCCATCACCG CTTCTATACA AAGAGGCTGA CCTTTGGAC GTGTCTGGCT	420	
GTGATCTGTA TGGTGTGGAC TCTGCTGTG GCCATGGCAT TTCCCCGGT TTTAGACGTG	480	30
GGCACTTACT CATTCTATTAG GGAGGAAGAT CAATGCACCT TCCAACACCG CTCCCTCAGG	540	
GCTAATGATT CCTTAGGATT TATGCTGCTT CTTGCTCTCA TCCTCCTAGC CACACAGCTT	600	
GTCTACCTCA AGCTGATATT TTTCGTCCAC GATCGAAGAA AAATGAAGCC AGTCCAGTTT	660	
30 GTAGCAGCAG TCAGCCAGAA CTGGACTTTT CATGGTCTG GAGCCAGTGG CCAGGCAGCT	720	
GCCAATTGGC TAGCAGGATT TGGAAAGGGGT CCCACACAC CCACCTTGCT GGGCAGTCAGG	780	
CAAAATGCAA ACACCACAGG CAGAAGAAGG CTATTGGTCT TAGACGAGTT CAAAATGGAG	840	

AAAAGAATCA	GCAGAATGTT	CTATATAATG	ACTTTTCTGT	TTCTAACCTT	GTTGGGGCCCC	900
TACCTGGTGG	CCTGTATTG	GAGAGTTTT	GCAAGAGGGC	CTGTAGTACC	AGGGGGATTT	960
CTAACAGCTG	CTGTCTGGAT	GAGTTTGCC	CAAGCAGGAA	TCAATCCTT	TGTCCTGCATT	1020
TTCTCAAACA	GGGAGCTGAG	GCGCTTTTC	AGCACAAACCC	TTCTTTACTG	CAGAAAATCC	1080
5	AGGTTACCAA	GGGAACCTTA	CTGTGTTATA	TGA		1113

## (29) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 370 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: not relevant

10

- (ii) MOLECULE TYPE: protein

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met	Ala	Asn	Tyr	Ser	His	Ala	Ala	Asp	Asn	Ile	Leu	Gln	Asn	Leu	Ser
1															
															15

Pro	Leu	Thr	Ala	Phe	Leu	Lys	Leu	Thr	Ser	Leu	Gly	Phe	Ile	Ile	Gly
															20
															30

Val	Ser	Val	Val	Gly	Asn	Leu	Leu	Ile	Ser	Ile	Leu	Leu	Val	Lys	Asp
															35
															40
															45

20	Lys	Thr	Leu	His	Arg	Ala	Pro	Tyr	Tyr	Phe	Leu	Leu	Asp	Leu	Cys	Cys
															50	
															55	
															60	

25	Ser	Asp	Ile	Leu	Arg	Ser	Ala	Ile	Cys	Phe	Pro	Phe	Val	Phe	Asn	Ser
															65	
															70	
															75	
															80	

25	Val	Lys	Asn	Gly	Ser	Thr	Trp	Thr	Tyr	Gly	Thr	Leu	Thr	Cys	Lys	Val
															85	
															90	
															95	

30	Ile	Ala	Phe	Leu	Gly	Val	Leu	Ser	Cys	Phe	His	Thr	Ala	Phe	Met	Leu
															100	
															105	
															110	

30	Phe	Cys	Ile	Ser	Val	Thr	Arg	Tyr	Leu	Ala	Ile	Ala	His	His	Arg	Phe
															115	
															120	
															125	

30	Tyr	Thr	Lys	Arg	Leu	Thr	Phe	Trp	Thr	Cys	Leu	Ala	Val	Ile	Cys	Met
															130	
															135	
															140	

30	Val	Trp	Thr	Leu	Ser	Val	Ala	Met	Ala	Phe	Pro	Pro	Val	Leu	Asp	Val
															145	
															150	
															155	
															160	

Gly Thr Tyr Ser Phe Ile Arg Glu Glu Asp Gln Cys Thr Phe Gln His

	165	170	175	
	Arg Ser Phe Arg Ala Asn Asp Ser Leu Gly Phe Met Leu Leu Ala			
	180	185	190	
5	Leu Ile Leu Ala Thr Gln Leu Val Tyr Leu Lys Leu Ile Phe Phe			
	195	200	205	
	Val His Asp Arg Arg Lys Met Lys Pro Val Gln Phe Val Ala Ala Val			
	210	215	220	
	Ser Gln Asn Trp Thr Phe His Gly Pro Gly Ala Ser Gly Gln Ala Ala			
	225	230	235	240
10	Ala Asn Trp Leu Ala Gly Phe Gly Arg Gly Pro Thr Pro Pro Thr Leu			
	245	250	255	
	Leu Gly Ile Arg Gln Asn Ala Asn Thr Thr Gly Arg Arg Arg Leu Leu			
	260	265	270	
15	Val Leu Asp Glu Phe Lys Met Glu Lys Arg Ile Ser Arg Met Phe Tyr			
	275	280	285	
	Ile Met Thr Phe Leu Phe Leu Thr Leu Trp Gly Pro Tyr Leu Val Ala			
	290	295	300	
	Cys Tyr Trp Arg Val Phe Ala Arg Gly Pro Val Val Pro Gly Gly Phe			
	305	310	315	320
20	Leu Thr Ala Ala Val Trp Met Ser Phe Ala Gln Ala Gly Ile Asn Pro			
	325	330	335	
	Phe Val Cys Ile Phe Ser Asn Arg Glu Leu Arg Arg Cys Phe Ser Thr			
	340	345	350	
25	Thr Leu Leu Tyr Cys Arg Lys Ser Arg Leu Pro Arg Glu Pro Tyr Cys			
	355	360	365	
	Val Ile			
	370			
	(30) INFORMATION FOR SEQ ID NO:29:			
	30			
	(i) SEQUENCE CHARACTERISTICS:			
30	(A) LENGTH: 1080 base pairs			
	(B) TYPE: nucleic acid			
	(C) STRANDEDNESS: single			
	(D) TOPOLOGY: linear			
	(ii) MOLECULE TYPE: DNA (genomic)			
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:			
	ATGCAGGTCC CGAACAGCAC CGGCCGGAC AACGCGACGC TGCAGATGCT GCGGAACCCG 60			
	40			

120	GCGATCGCGG TGGCCCTGCC CGTGGTGTAC TCGCTGGTGG CGGCGGTCAG CATCCCAGGC
180	AACCTCTTCT CTCTGTGGGT GCTGTGCCGG CGCATGGGGC CCAGATCCCC GTCGGTCATC
240	TTCATGATCA ACCTGAGCGT CACGGACCTG ATGCTGGCCA GCGTGGTGCCTTCCAAATC
300	TACTACCATT GCAACCGCCA CCACTGGTA TTCGGGGTGC TGCTTGCAA CGTGGTGACC
360	5 GTGGCCTTT ACGCAAACAT GTATTCCAGC ATCCTCACCA TGACCTGTAT CAGCGTGGAG
420	CGCTTCTGG GGGTCTGTAA CCCGCTCAGC TCCAAGCGCT GGCGCCGCCG TCGTTACGCG
480	GTGGCCGGT GTGCAGGGAC CTGGCTGCTG CTCCGTACCG CCCTGTGCCG GCTGGCGCGC
540	ACCGATCTCA CCTACCCGGT GCACGCCCTG GGCATCATCA CCTGCTTCGA CGTCCTCAAG
600	TGGACGATGC TCCCCAGCGT GGCCATGTGG GCCGTGTTCC TCTTCACCAT CTTCATCCTG
660	10 CTGTTCTCA TCCCCTTCGT GATCACCGTG GCTTGTACA CGGCCACCAT CCTCAAGCTG
720	TTGCGCACGG AGGAGGGCGCA CGGCCGGGAG CAGCGGAGGC GCGCGGTGGG CCTGGCGCG
780	GTGGTCTTGC TGGCTTTGT CACCTGCTTC GCCCCCAACA ACTTCGTGCT CCTGGCGCAC
840	ATCGTGAGCC GCCTGTTCTA CGGCAAGAGC TACTACCACG TGTACAAGCT CACGCTGTGT
900	CTCAGCTGCC TCAACAACTG TCTGGACCCG TTTGTTTATT ACTTTGCGTC CGGGGAATT
960	15 CAGCTGCGCC TGCAGGAATA TTTGGGCTGC CGCCGGGTGC CCAGAGACAC CCTGGACACG
1020	CGCCGCGAGA GCCTCTTCTC CGCCAGGACC ACGTCCGTGC GCTCCGAGGC CGGTGCGCAC
1080	CCTGAAGGGA TGGAGGGAGC CACCAGGCCG GGCCTCCAGA GGCAGGAGAG TGTGTTCTGA

## (31) INFORMATION FOR SEQ ID NO:30:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 359 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

30 (ii) MOLECULE TYPE: protein

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Gln Val Pro Asn Ser Thr Gly Pro Asp Asn Ala Thr Leu Gln Met  
 1 5 10 15

Leu Arg Asn Pro Ala Ile Ala Val Ala Leu Pro Val Val Tyr Ser Leu  
 20 25 30

30 Val Ala Ala Val Ser Ile Pro Gly Asn Leu Phe Ser Leu Trp Val Leu

35

40

45

5	Cys Arg Arg Met Gly Pro Arg Ser Pro Ser Val Ile Phe Met Ile Asn 50 55 60
	Leu Ser Val Thr Asp Leu Met Leu Ala Ser Val Leu Pro Phe Gln Ile 65 70 75 80
	Tyr Tyr His Cys Asn Arg His His Trp Val Phe Gly Val Leu Leu Cys 85 90 95
10	Asn Val Val Thr Val Ala Phe Tyr Ala Asn Met Tyr Ser Ser Ile Leu 100 105 110
	Thr Met Thr Cys Ile Ser Val Glu Arg Phe Leu Gly Val Leu Tyr Pro 115 120 125
	Leu Ser Ser Lys Arg Trp Arg Arg Arg Tyr Ala Val Ala Ala Cys 130 135 140
15	Ala Gly Thr Trp Leu Leu Leu Leu Thr Ala Leu Cys Pro Leu Ala Arg 145 150 155 160
	Thr Asp Leu Thr Tyr Pro Val His Ala Leu Gly Ile Ile Thr Cys Phe 165 170 175
	Asp Val Leu Lys Trp Thr Met Leu Pro Ser Val Ala Met Trp Ala Val 180 185 190
20	Phe Leu Phe Thr Ile Phe Ile Leu Leu Phe Leu Ile Pro Phe Val Ile 195 200 205
	Thr Val Ala Cys Tyr Thr Ala Thr Ile Leu Lys Leu Leu Arg Thr Glu 210 215 220
25	Glu Ala His Gly Arg Glu Gln Arg Arg Arg Ala Val Gly Leu Ala Ala 225 230 235 240
	Val Val Leu Leu Ala Phe Val Thr Cys Phe Ala Pro Asn Asn Phe Val 245 250 255
	Leu Leu Ala His Ile Val Ser Arg Leu Phe Tyr Gly Lys Ser Tyr Tyr 260 265 270
30	His Val Tyr Lys Leu Thr Leu Cys Leu Ser Cys Leu Asn Asn Cys Leu 275 280 285
	Asp Pro Phe Val Tyr Tyr Phe Ala Ser Arg Glu Phe Gln Leu Arg Leu 290 295 300
35	Arg Glu Tyr Leu Gly Cys Arg Arg Val Pro Arg Asp Thr Leu Asp Thr 305 310 315 320
	Arg Arg Glu Ser Leu Phe Ser Ala Arg Thr Thr Ser Val Arg Ser Glu 325 330 335

Ala Gly Ala His Pro Glu Gly Met Glu Gly Ala Thr Arg Pro Gly Leu  
 340 345 350  
 Gln Arg Gln Glu Ser Val Phe  
 355

## 5 (32) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1503 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ATGGAGCGTC CCTGGGAGGA CAGCCCAGGC CCGGAGGGGG CAGCTGAGGG CTCGCCTGTG	60
CCAGTCGCCG CCGGGGCGCG CTCCGGTGCC GCGGCGAGTG GCACAGGCTG GCAGCCATGG	120
15 GCTGAGTGCC CGGGACCCAA GGGGAGGGGG CAACTGCTGG CGACCGCCGG CCCTTTGCGT	180
CGCTGGCCCG CCCCCTCGCC TGCCAGCTCC AGCCCCGCC CGGGAGCGGC GTCCGCTCAC	240
TCGGTTCAAG GCAGCGCGAC TGCGGGTGCG GCACGACCAG GGCGCAGACC TTGGGGCGCG	300
20 CGGCCCCATGG AGTCGGGGCT GCTGCAGCGC GCGCCGGTGA GCGAGGTCAT CGTCCTGCAT	360
TACAACATACA CGCGCAAGCT CGCGGGTGCG AGCTACCAAGC CGGGTGCCGG CCTGCGCGCC	420
25 GACGCCGTGG TGTGCCTGGC GGTGTGCGCC TTCATCGTGC TAGAGAATCT AGCCGTGTTG	480
TTGGTGCTCG GACGCCACCC GCGCTTCCAC GCTCCCATGT TCCTGCTCCT GGGCAGCCTC	540
ACGTTGTCGG ATCTGCTGGC AGGCGCCGCC TACGCCCCA ACATCCTACT GTGGGGCGCG	600
CTCACGCTGA AACTGTCCCC CGCGCTCTGG TTGCACTGGGG AGGGAGGCGT CTTCGTGGCA	660
30 CTCACGTGAG CGTGCAGCAGC CCTCCTGGCC ATCGCGCTGG AGCGCAGCCT CACCATGGCG	720
CGCAGGGGGC CGCGCCCGT CTCCAGTCGG GGGCGCACGC TGGCGATGGC AGCCGCGGCC	780
TGGGGCGTGT CGCTGCTCCT CGGGCTCCTG CCAGCGCTGG GCTGGAATTG CCTGGGTGCG	840
CTGGACGCTT GCTCCACTGT CTTGCCGCTC TACGCCAAGG CCTACGTGCT CTTCTGCGTG	900
CTCGCCTTCG TGGGCATCCT GGCGCGATC TGTGCACTCT ACGCGCGCAT CTACTGCCAG	960
GTACCGGCCA ACGCGCGGCC CCTGCCGGCA CGGCCCGGGGA CTGCGGGGAC CACCTCGACC	1020
35 CGGGCGCGTC GCAAGCCGCG CTCTCTGGCC TTGCTGCGCA CGCTCAGCGT GGTGCTCCTG	1080

GCCTTTGTGG	CATGTTGGGG	CCCCCTCTTC	CTGCTGCTGT	TGCTCGACGT	GGCGTGCCCG	1140	
GCGCGCACCT	GTCCTGTACT	CCTGCAGGCC	GATCCCTTCC	TGGGACTGGC	CATGGCCAAC	1200	
TCACTTCTGA	ACCCCATCAT	CTACACGCTC	ACCAACCGCG	ACCTGCGCCA	CGCGCTCCTG	1260	
CGCCTGGTCT	GCTGCGGACG	CCACTCCTGC	GGCAGAGACC	CGAGTGGCTC	CCAGCAGTCG	1320	
5	GCGAGCGCGG	CTGAGGCTTC	CGGGGGCCTG	CGCCGCTGCC	TGCCCCGGG	CCTTGATGGG	1380
AGCTTCAGCG	GCTCGGAGCG	CTCATCGCCC	CAGCGCGACG	GGCTGGACAC	CAGCGGCTCC	1440	
ACAGGCAGCC	CCGGTGCACC	CACAGCCGCC	CGGACTCTGG	TATCAGAACCC	GGCTGCAGAC	1500	
TGA						1503	

## (33) INFORMATION FOR SEQ ID NO:32:

10 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 500 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

15 (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met	Glu	Arg	Pro	Trp	Glu	Asp	Ser	Pro	Gly	Pro	Glu	Gly	Ala	Ala	Glu	
1				5					10				15			
Gly	Ser	Pro	Val	Pro	Val	Ala	Ala	Gly	Ala	Arg	Ser	Gly	Ala	Ala	Ala	
20			20			25				30						
Ser	Gly	Thr	Gly	Trp	Gln	Pro	Trp	Ala	Glu	Cys	Pro	Gly	Pro	Lys	Gly	
	35				40					45						
Arg	Gly	Gln	Leu	Leu	Ala	Thr	Ala	Gly	Pro	Leu	Arg	Arg	Trp	Pro	Ala	
	50				55					60						
25	Pro	Ser	Pro	Ala	Ser	Ser	Ser	Pro	Ala	Pro	Gly	Ala	Ala	Ser	Ala	His
	65			70				75		80						30
Ser	Val	Gln	Gly	Ser	Ala	Thr	Ala	Gly	Gly	Ala	Arg	Pro	Gly	Arg	Arg	
	85				90					95						
30	Pro	Trp	Gly	Ala	Arg	Pro	Met	Glu	Ser	Gly	Leu	Leu	Arg	Pro	Ala	Pro
		100				105					110					
Val	Ser	Glu	Val	Ile	Val	Leu	His	Tyr	Asn	Tyr	Thr	Gly	Lys	Leu	Arg	
	115				120					125						
Gly	Ala	Ser	Tyr	Gln	Pro	Gly	Ala	Gly	Leu	Arg	Ala	Asp	Ala	Val	Val	
	130				135					140						

Cys Leu Ala Val Cys Ala Phe Ile Val Leu Glu Asn Leu Ala Val Leu  
 145 150 155 160  
 Leu Val Leu Gly Arg His Pro Arg Phe His Ala Pro Met Phe Leu Leu  
 165 170 175  
 5 Leu Gly Ser Leu Thr Leu Ser Asp Leu Leu Ala Gly Ala Ala Tyr Ala  
 180 185 190  
 Ala Asn Ile Leu Leu Ser Gly Pro Leu Thr Leu Lys Leu Ser Pro Ala  
 195 200 205  
 Leu Trp Phe Ala Arg Glu Gly Val Phe Val Ala Leu Thr Ala Ser  
 10 210 215 220  
 Val Leu Ser Leu Leu Ala Ile Ala Leu Glu Arg Ser Leu Thr Met Ala  
 225 230 235 240  
 Arg Arg Gly Pro Ala Pro Val Ser Ser Arg Gly Arg Thr Leu Ala Met  
 245 250 255  
 15 Ala Ala Ala Ala Trp Gly Val Ser Leu Leu Leu Gly Leu Leu Pro Ala  
 260 265 270  
 Leu Gly Trp Asn Cys Leu Gly Arg Leu Asp Ala Cys Ser Thr Val Leu  
 275 280 285  
 Pro Leu Tyr Ala Lys Ala Tyr Val Leu Phe Cys Val Leu Ala Phe Val  
 20 290 295 300  
 Gly Ile Leu Ala Ala Ile Cys Ala Leu Tyr Ala Arg Ile Tyr Cys Gln  
 305 310 315 320  
 Val Arg Ala Asn Ala Arg Arg Leu Pro Ala Arg Pro Gly Thr Ala Gly  
 325 330 335  
 25 Thr Thr Ser Thr Arg Ala Arg Arg Lys Pro Arg Ser Leu Ala Leu Leu  
 340 345 350  
 Arg Thr Leu Ser Val Val Leu Leu Ala Phe Val Ala Cys Trp Gly Pro  
 355 360 365  
 30 Leu Phe Leu Leu Leu Leu Asp Val Ala Cys Pro Ala Arg Thr Cys  
 370 375 380  
 Pro Val Leu Leu Gln Ala Asp Pro Phe Leu Gly Leu Ala Met Ala Asn  
 385 390 395 400  
 Ser Leu Leu Asn Pro Ile Ile Tyr Thr Leu Thr Asn Arg Asp Leu Arg  
 405 410 415  
 35 His Ala Leu Leu Arg Leu Val Cys Cys Gly Arg His Ser Cys Gly Arg  
 420 425 430  
 Asp Pro Ser Gly Ser Gln Gln Ser Ala Ser Ala Glu Ala Ser Gly

	435	440	445
	Gly Leu Arg Arg Cys Leu Pro Pro Gly Leu Asp Gly Ser Phe Ser Gly		
	450	455	460
5	Ser Glu Arg Ser Ser Pro Gln Arg Asp Gly Leu Asp Thr Ser Gly Ser		
	465	470	475
	Thr Gly Ser Pro Gly Ala Pro Thr Ala Ala Arg Thr Leu Val Ser Glu		
	485	490	495
	Pro Ala Ala Asp		
	500		

## 10 (34) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1029 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ATGCAAGCCG TCGACAATCT CACCTCTGCG CCTGGGAACA CCAGTCTGTG CACCAGAGAC	60	20
TACAAAATCA CCCAGGTCTT CTTCCCCTG CTCTACACTG TCCTGTTTT TGTTGGACTT	120	
20 ATCACAAATG GCCTGGCGAT GAGGATTTTC TTTCAAATCC GGAGTAAATC AAACTTTATT	180	
ATTTTCTTA AGAACACAGT CATTCTGAT CTTCTCATGA TTCTGACTTT TCCATTCAAA	240	
ATTCTTAGTG ATGCCAAACT GGGAACAGGA CCACTGAGAA CTTTTGTGTG TCAAGTTACC	300	
TCCGTCATAT TTTATTCAC AATGTATATC AGTATTCAT TCCTGGGACT GATAACTATC	360	
GATCGCTACC AGAACGACCAC CAGGCCATT AAAACATCCA ACCCAAAAAA TCTCTGGGG	420	
25 GCTAAGATTC TCTCTGTTGT CATCTGGCA TTCATGTTCT TACTCTCTTT GCCTAACATG	480	30
ATTCTGACCA ACAGGCAGCC GAGAGACAAG AATGTGAAGA AATGCTCTTT CCTTAAATCA	540	
GAGTTGGTC TAGTCTGGCA TGAAATAGTA AATTACATCT GTCAAGTCAT TTTCTGGATT	600	
AATTTCTTAA TTGTTATTGT ATGTTATACA CTCATTACAA AAGAACTGTA CCGGTCATAC	660	
GTAAGAACGA GGGGTGAGG TAAAGTCCCC AGGAAAAGG TGAACGTCAA AGTTTCATT	720	
30 ATCATTGCTG TATTCTTAT TTGTTTGTT CCTTTCCATT TTGCCGAAT TCCTTACACC	780	
CTGAGCCAAA CCCGGGATGT CTTTGACTGC ACTGCTGAAA ATACTCTGTT CTATGTGAAA	840	

GAGAGCACTC TGTGGTTAAC TTCCCTTAAAT GCATGCCTGG ATCCGTTCAT CTATTTTTTC 900  
 CTTTGCAAGT CCTTCAGAAA TTCCCTGATA AGTATGCTGA AGTGCCTCAA TTCTGCAACA 960  
 TCTCTGTCCC AGGACAATAG GAAAAAAAGAA CAGGATGGTG GTGACCCAAA TGAAGAGACT 1020  
 CCAATGTAA 1029

## 5 (35) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 342 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Gln Ala Val Asp Asn Leu Thr Ser Ala Pro Gly Asn Thr Ser Leu  
 1 5 10 15

15 Cys Thr Arg Asp Tyr Lys Ile Thr Gln Val Leu Phe Pro Leu Leu Tyr  
 20 25 30

Thr Val Leu Phe Phe Val Gly Leu Ile Thr Asn Gly Leu Ala Met Arg  
 35 40 45

20 Ile Phe Phe Gln Ile Arg Ser Lys Ser Asn Phe Ile Ile Phe Leu Lys  
 50 55 60

Asn Thr Val Ile Ser Asp Leu Leu Met Ile Leu Thr Phe Pro Phe Lys  
 65 70 75 80

Ile Leu Ser Asp Ala Lys Leu Gly Thr Gly Pro Leu Arg Thr Phe Val  
 85 90 95

25 Cys Gln Val Thr Ser Val Ile Phe Tyr Phe Thr Met Tyr Ile Ser Ile  
 100 105 110

Ser Phe Leu Gly Leu Ile Thr Ile Asp Arg Tyr Gln Lys Thr Thr Arg  
 115 120 125

30 Pro Phe Lys Thr Ser Asn Pro Lys Asn Leu Leu Gly Ala Lys Ile Leu  
 130 135 140

Ser Val Val Ile Trp Ala Phe Met Phe Leu Leu Ser Leu Pro Asn Met  
 145 150 155 160

Ile Leu Thr Asn Arg Gln Pro Arg Asp Lys Asn Val Lys Lys Cys Ser  
 165 170 175

35 Phe Leu Lys Ser Glu Phe Gly Leu Val Trp His Glu Ile Val Asn Tyr

	180	185	190		
	Ile Cys Gln Val Ile Phe Trp Ile Asn Phe Leu Ile Val Ile Val Cys				
	195	200	205		
5	Tyr Thr Leu Ile Thr Lys Glu Leu Tyr Arg Ser Tyr Val Arg Thr Arg				
	210	215	220		
	Gly Val Gly Lys Val Pro Arg Lys Lys Val Asn Val Lys Val Phe Ile				
	225	230	235	240	
	Ile Ile Ala Val Phe Phe Ile Cys Phe Val Pro Phe His Phe Ala Arg				
	245	250	255		
10	Ile Pro Tyr Thr Leu Ser Gln Thr Arg Asp Val Phe Asp Cys Thr Ala				
	260	265	270		
	Glu Asn Thr Leu Phe Tyr Val Lys Glu Ser Thr Leu Trp Leu Thr Ser				
	275	280	285		
15	Leu Asn Ala Cys Leu Asp Pro Phe Ile Tyr Phe Phe Leu Cys Lys Ser				
	290	295	300		
	Phe Arg Asn Ser Leu Ile Ser Met Leu Lys Cys Pro Asn Ser Ala Thr				
	305	310	315	320	
	Ser Leu Ser Gln Asp Asn Arg Lys Lys Glu Gln Asp Gly Gly Asp Pro				
	325	330	335		
20	Asn Glu Glu Thr Pro Met				
	340				
(36) INFORMATION FOR SEQ ID NO:35:					
	(i) SEQUENCE CHARACTERISTICS:				
	(A) LENGTH: 1077 base pairs				
25	(B) TYPE: nucleic acid				
	(C) STRANDEDNESS: single				
	(D) TOPOLOGY: linear				
	(ii) MOLECULE TYPE: DNA (genomic)				
					30
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:				
30	ATGTCGGTCT GCTACCGTCC CCCAGGGAAC GAGACACTGC TGAGCTGGAA GACTTCGGCG				60
	GCCACAGGCA CAGCCTTCCT GCTGCTGGCG GCGCTGCTGG GGCTGCCTGG CAACGGCTTC				120
	GTGGTGTGGA GCTTGGCGGG CTGGCGGCCT GCACGGGGC GACCGCTGGC GGCCACGCTT				180
	GTGCTGCACC TGGCGCTGGC CGACGGCGCG GTGCTGCTGC TCACGCCGCT CTTTGTGGCC				240
	TTCCTGACCC GGCAGGCCTG GCCGCTGGGC CAGGCAGGGCT GCAAGGCGGT GTACTACGTG				300

TCGCGCGCTCA	GCATGTACGC	CAGCGTGCTG	CTCACCGGCC	TGCTCAGCCT	GCAGCGCTGC	360	
CTCGCAGTCA	CCCGCCCCTT	CCTGGCGCCT	CGGCTGCGCA	GCCCCGGCCCT	GGCCCGCCGC	420	
CTGCTGCTGG	CGGTCTGGCT	GGCCGCCCTG	TTGCTCGCCG	TCCCAGGCCG	CGTCTACCGC	480	
CACCTGTGGA	GGGACCGCGT	ATGCCAGCTG	TGCCACCCGT	CGCCGGTCCA	CGCCGCCGCC	540	
5	CACCTGAGCC	TGGAGACTCT	GACCGCTTTC	GTGCTTCCTT	TCGGGCTGAT	GCTCGGCTGC	600
	TACAGCGTGA	CGCTGGCAGC	GCTGGGGGC	GCCCCTGGGG	GCTCCGGGCG	GCACGGGGCG	660
	CGGGTGGGCC	GGCTGGTGAG	CGCCATCGTG	CTTGCCTTCG	GCTTGCTCTG	GGCCCCCTAC	720
	CACGCAGTCA	ACCTTCTGCA	GGCGGTGCGA	GCGCTGGCTC	CACCGGAAGG	GGCCTTGGCG	780
	AAGCTGGGCC	GAGCCGGCCA	GGCGGCCGGA	GCGGGAACTA	CGGCCTTGGC	CTTCTTCAGT	840
10	TCTAGCGTCA	ACCCGGTGCT	CTACGTCTTC	ACCGCTGGAG	ATCTGCTGCC	CGGGCAGGGT	900
	CCCCGTTCC	TCACCGGGCT	CTTCGAAGGC	TCTGGGGAGG	CCCGAGGGGG	CGGCCGCTCT	960
	AGGGAAAGGGA	CCATGGAGCT	CCGAACCTACC	CCTCAGCTGA	AAGTGGTGGG	GCAGGGCCGC	1020
	GGCAATGGAG	ACCCGGGGGG	TGGGATGGAG	AAGGACGGTC	CGGAATGGGA	CCTTTGA	1077

(37) INFORMATION FOR SEQ ID NO:36:

20

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 358 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: not relevant

20 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Met	Ser	Val	Cys	Tyr	Arg	Pro	Pro	Gly	Asn	Glu	Thr	Leu	Leu	Ser	Trp
1										10				15	

Lys Thr Ser Arg Ala Thr Gly Thr Ala Phe Leu Leu Leu Ala Ala Leu  
20 25 30

Leu Gly Leu Pro Gly Asn Gly Phe Val Val Trp Ser Leu Ala Gly Trp  
35 40 45

Ala Leu Ala Asp Gly Ala Val Leu Leu Leu Thr Pro Leu Phe Val Ala  
65 70 75 80

Phe Leu Thr Arg Gln Ala Trp Pro Leu Gly Gln Ala Gly Cys Lys Ala

	85	90	95	
	Val Tyr Tyr Val Cys Ala Leu Ser Met Tyr Ala Ser Val Leu Leu Thr			
	100	105	110	
5	Gly Leu Leu Ser Leu Gln Arg Cys Leu Ala Val Thr Arg Pro Phe Leu			
	115	120	125	
	Ala Pro Arg Leu Arg Ser Pro Ala Leu Ala Arg Arg Leu Leu Leu Ala			
	130	135	140	
	Val Trp Leu Ala Ala Leu Leu Leu Ala Val Pro Ala Ala Val Tyr Arg			
	145	150	155	160
10	His Leu Trp Arg Asp Arg Val Cys Gln Leu Cys His Pro Ser Pro Val			
	165	170	175	
	His Ala Ala Ala His Leu Ser Leu Glu Thr Leu Thr Ala Phe Val Leu			
	180	185	190	
15	Pro Phe Gly Leu Met Leu Gly Cys Tyr Ser Val Thr Leu Ala Arg Leu			
	195	200	205	
	Arg Gly Ala Arg Trp Gly Ser Gly Arg His Gly Ala Arg Val Gly Arg			
	210	215	220	
	Leu Val Ser Ala Ile Val Leu Ala Phe Gly Leu Leu Trp Ala Pro Tyr			
	225	230	235	240
20	His Ala Val Asn Leu Leu Gln Ala Val Ala Ala Leu Ala Pro Pro Glu			
	245	250	255	
	Gly Ala Leu Ala Lys Leu Gly Gly Ala Gly Gln Ala Ala Arg Ala Gly			
	260	265	270	
25	Thr Thr Ala Leu Ala Phe Phe Ser Ser Ser Val Asn Pro Val Leu Tyr			
	275	280	285	
	Val Phe Thr Ala Gly Asp Leu Leu Pro Arg Ala Gly Pro Arg Phe Leu			
	290	295	300	
	Thr Arg Leu Phe Glu Gly Ser Gly Glu Ala Arg Gly Gly Arg Ser			
	305	310	315	320
30	Arg Glu Gly Thr Met Glu Leu Arg Thr Thr Pro Gln Leu Lys Val Val			
	325	330	335	
	Gly Gln Gly Arg Gly Asn Gly Asp Pro Gly Gly Met Glu Lys Asp			
	340	345	350	
35	Gly Pro Glu Trp Asp Leu			
	355			

(38) INFORMATION FOR SEQ ID NO:37:

- 5 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1005 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

ATGCTGGGGA TCATGGCATG GAATGCAACT TGCAAAACT GGCTGGCAGC AGAGGCTGCC	60	10
CTGGAAAAGT ACTACCTTTC CATTTTTAT GGGATTGAGT TCGTTGTGGG AGTCCTTGGA	120	
10 AATACCATTG TTGTTTACGG CTACATCTTC TCTCTGAAGA ACTGGAACAG CAGTAATATT	180	
TATCTCTTTA ACCTCTCTGT CTCTGACTTA GCTTTCTGT GCACCCCTCCC CATGCTGATA	240	
AGGAGTTATG CCAATGGAAA CTGGATATAT GGAGACGTGC TCTGCATAAG CAACCGATAT	300	
GTGCTTCATG CCAACCTCTA TACCAGCATT CTCTTCTCA CTTTTATCAG CATAAGATCGA	360	
TACTTGATAA TTAAGTATCC TTTCCGAGAA CACCTTCTGC AAAAGAAAAGA GTTTGCTATT	420	
15 TTAATCTCCT TGGCCATTG GGTTTAGTA ACCTTAGAGT TACTACCCAT ACTTCCCCTT	480	
ATAAATCCTG TTATAACTGA CAATGGCACC ACCTGTAATG ATTTTGCAAG TTCTGGAGAC	540	20
CCCAACTACA ACCTCATTG CAGCATGTGT CTAACACTGT TGGGGTTCTT TATTCCCTTT	600	
TTTGTGATGT GTTTCTTTA TTACAAGATT GCTCTCTCC TAAAGCAGAG GAATAGGCAG	660	
GTTGCTACTG CTCTGCCCT TGAAAAGCCT CTCAACTTGG TCATCATGGC AGTGGTAATC	720	
20 TTCTCTGTGC TTTTACACC CTATCACGTC ATGCCGAATG TGAGGATCGC TTACGCCCTG	780	
GGGAGTTGGA AGCAGTATCA GTGCACTCAG GTCGTCACTCA ACTCCTTTA CATTGTGACA	840	
CGGCCTTGG CCTTTCTGAA CAGTGTCACTC AACCCTGTCT TCTATTTCT TTTGGGAGAT	900	
CACTTCAGGG ACATGCTGAT GAATCAACTG AGACACAACT TCAAATCCCT TACATCCTT	960	30
AGCAGATGGG CTCATGAACT CCTACTTTCA TTCAGAGAAA AGTGA	1005	

25 (39) INFORMATION FOR SEQ ID NO:38:

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 334 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

	Met	Leu	Gly	Ile	Met	Ala	Trp	Asn	Ala	Thr	Cys	Lys	Asn	Trp	Leu	Ala	
1					5				10						15		
5		Ala	Glu	Ala	Ala	Leu	Glu	Lys	Tyr	Tyr	Leu	Ser	Ile	Phe	Tyr	Gly	Ile
			20				25							30			
	Glu	Phe	Val	Val	Gly	Val	Leu	Gly	Asn	Thr	Ile	Val	Val	Tyr	Gly	Tyr	
		35				40							45				
	Ile	Phe	Ser	Leu	Lys	Asn	Trp	Asn	Ser	Ser	Asn	Ile	Tyr	Leu	Phe	Asn	
		50			55						60					10	
10		Leu	Ser	Val	Ser	Asp	Leu	Ala	Phe	Leu	Cys	Thr	Leu	Pro	Met	Leu	Ile
		65				70					75				80		
	Arg	Ser	Tyr	Ala	Asn	Gly	Asn	Trp	Ile	Tyr	Gly	Asp	Val	Leu	Cys	Ile	
		85					90							95			
15		Ser	Asn	Arg	Tyr	Val	Leu	His	Ala	Asn	Leu	Tyr	Thr	Ser	Ile	Leu	Phe
		100				105								110			
	Leu	Thr	Phe	Ile	Ser	Ile	Asp	Arg	Tyr	Leu	Ile	Ile	Lys	Tyr	Pro	Phe	
		115				120							125				
	Arg	Glu	His	Leu	Leu	Gln	Lys	Lys	Glu	Phe	Ala	Ile	Leu	Ile	Ser	Leu	
		130				135							140				20
20		Ala	Ile	Trp	Val	Leu	Val	Thr	Leu	Glu	Leu	Leu	Pro	Ile	Leu	Pro	Leu
		145				150						155			160		
	Ile	Asn	Pro	Val	Ile	Thr	Asp	Asn	Gly	Thr	Thr	Cys	Asn	Asp	Phe	Ala	
		165					170						175				
25		Ser	Ser	Gly	Asp	Pro	Asn	Tyr	Asn	Leu	Ile	Tyr	Ser	Met	Cys	Leu	Thr
		180				185							190				
	Leu	Leu	Gly	Phe	Leu	Ile	Pro	Leu	Phe	Val	Met	Cys	Phe	Phe	Tyr	Tyr	
		195				200							205				
	Lys	Ile	Ala	Leu	Phe	Leu	Lys	Gln	Arg	Asn	Arg	Gln	Val	Ala	Thr	Ala	
		210				215						220					30
30		Leu	Pro	Leu	Glu	Lys	Pro	Leu	Asn	Leu	Val	Ile	Met	Ala	Val	Val	Ile
		225			230						235			240			
	Phe	Ser	Val	Leu	Phe	Thr	Pro	Tyr	His	Val	Met	Arg	Asn	Val	Arg	Ile	
		245					250						255				
35		Ala	Ser	Arg	Leu	Gly	Ser	Trp	Lys	Gln	Tyr	Gln	Cys	Thr	Gln	Val	Val
			260			265							270				
	Ile	Asn	Ser	Phe	Tyr	Ile	Val	Thr	Arg	Pro	Leu	Ala	Phe	Leu	Asn	Ser	

275 280 285

Val Ile Asn Pro Val Phe Tyr Phe Leu Leu Gly Asp His Phe Arg Asp  
 290 295 300

5 Met Leu Met Asn Gln Leu Arg His Asn Phe Lys Ser Leu Thr Ser Phe  
 305 310 315 320

Ser Arg Trp Ala His Glu Leu Leu Ser Phe Arg Glu Lys  
 325 330

## (40) INFORMATION FOR SEQ ID NO:39:

- 10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1296 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

ATGCAGGCAG	TTAACATTAC	CCCGGAGCAG	TTCTCTCGC	TGCTGCGGA	CCACAAACCTG	60
ACGCAGGAGC	AGTTCATCGC	TCTGTACCGG	CTGCGACCGC	TCGTCTACAC	CCCAGAGCTG	120
CCGGGACGCG	CCAAGCTGGC	CCTCGTGCTC	ACCGGGCGTGC	TCATCTTCGC	CCTGGCGCTC	180
TTTGGCAATG	CTCTGGTGT	CTACGTGGTG	ACCCGCAGCA	AGGCCATGCG	CACCGTCACC	240
20 AACATCTTA	TCTGCTCCTT	GGCGCTCAGT	GACCTGCTCA	TCACCTTCTT	CTGCATTCCC	300
GTCACCATGC	TCCAGAACAT	TTCCGACAAAC	TGGCTGGGG	GTGCTTTCAT	TTGCAAGATG	360
GTGCCATTG	TCCAGTCTAC	CGCTGTTGTG	ACAGAAATGC	TCACTATGAC	CTGCATTGCT	420
GTGGAAAGGC	ACCAGGGACT	TGTGCATCCT	TTTAAAATGA	AGTGGCAATA	CACCAACCGA	480
AGGGCTTCA	CAATGCTAGG	TGTGGTCTGG	CTGGTGGCAG	TCATCGTAGG	ATCACCCATG	540
25 TGGCACGTGC	AACAACCTGA	GATCAAATAT	GACTTCCTAT	ATGAAAAGGA	ACACATCTGC	600
TGCTTAGAAG	AGTGGACCAAG	CCCTGTGCAC	CAGAAGATCT	ACACCACCTT	CATCCTTGTC	660
ATCCTCTTCC	TCCTGCCTCT	TATGGTGATG	CTTATTCTGT	ACAGTAAAAT	TGGTTATGAA	720
CTTTGGATAA	AGAAAAGAGT	TGGGGATGGT	TCAGTGCTTC	GAACATTCA	TGGAAAAGAA	780
ATGTCCAAAA	TAGCCAGGAA	GAAGAAACGA	GCTGTCAATTA	TGATGGTGAC	AGTGGTGGCT	840
30 CTCTTTGCTG	TGTGCTGGGC	ACCATTCCAT	GTTGTCCATA	TGATGATTGA	ATACAGTAAT	900
TTTGAAAAGG	AATATGATGA	TGTCACAATC	AAGATGATT	TTGCTATCGT	GCAGATTATT	960

GGATTTCCA ACTCCATCTG TAATCCCATT GTCTATGCAT TTATGAATGA AAACTTCAA 1020  
 AAAATGTT TGTCTGCAGT TTGTTATTGC ATAGTAAATA AAACCTTCTC TCCAGCACAA 1080  
 AGGCATGGAA ATTCAAGGAAT TACAATGATG CGGAAGAAAG CAAAGTTTC CCTCAGAGAG 1140  
 AATCCAGTGG AGGAAACCAA AGGAGAAGCA TTCAGTGATG GCAACATTGA AGTCAAATTG 1200  
 5 TGTGAACAGA CAGAGGAGAA GAAAAAGCTC AAACGACATC TTGCTCTCT TAGGTCTGAA 1260  
 CTGGCTGAGA ATTCTCCTTT AGACAGTGGG CATTAA 1296

## (41) INFORMATION FOR SEQ ID NO:40:

10

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 431 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

15 Met Gln Ala Leu Asn Ile Thr Pro Glu Gln Phe Ser Arg Leu Leu Arg  
 1 5 10 15  
 Asp His Asn Leu Thr Arg Glu Gln Phe Ile Ala Leu Tyr Arg Leu Arg  
 20 25 30  
 20 Pro Leu Val Tyr Thr Pro Glu Leu Pro Gly Arg Ala Lys Leu Ala Leu  
 35 40 45  
 Val Leu Thr Gly Val Leu Ile Phe Ala Leu Ala Leu Phe Gly Asn Ala  
 50 55 60  
 Leu Val Phe Tyr Val Val Thr Arg Ser Lys Ala Met Arg Thr Val Thr  
 65 70 75 80  
 25 Asn Ile Phe Ile Cys Ser Leu Ala Leu Ser Asp Leu Leu Ile Thr Phe  
 85 90 95  
 Phe Cys Ile Pro Val Thr Met Leu Gln Asn Ile Ser Asp Asn Trp Leu  
 100 105 110  
 Gly Gly Ala Phe Ile Cys Lys Met Val Pro Phe Val Gln Ser Thr Ala  
 30 115 120 125  
 Val Val Thr Glu Met Leu Thr Met Thr Cys Ile Ala Val Glu Arg His  
 130 135 140  
 Gln Gly Leu Val His Pro Phe Lys Met Lys Trp Gln Tyr Thr Asn Arg  
 145 150 155 160

20

30

Arg Ala Phe Thr Met Leu Gly Val Val Trp Leu Val Ala Val Ile Val			
165	170	175	
Gly Ser Pro Met Trp His Val Gln Gln Leu Glu Ile Lys Tyr Asp Phe			
180	185	190	
5 Leu Tyr Glu Lys Glu His Ile Cys Cys Leu Glu Glu Trp Thr Ser Pro			
195	200	205	
Val His Gln Lys Ile Tyr Thr Thr Phe Ile Leu Val Ile Leu Phe Leu			
210	215	220	
10 Leu Pro Leu Met Val Met Leu Ile Leu Tyr Ser Lys Ile Gly Tyr Glu			
225	230	235	240
Leu Trp Ile Lys Lys Arg Val Gly Asp Gly Ser Val Leu Arg Thr Ile			
245	250	255	
His Gly Lys Glu Met Ser Lys Ile Ala Arg Lys Lys Arg Ala Val			
260	265	270	
15 Ile Met Met Val Thr Val Val Ala Leu Phe Ala Val Cys Trp Ala Pro			
275	280	285	
Phe His Val Val His Met Met Ile Glu Tyr Ser Asn Phe Glu Lys Glu			
290	295	300	
20 Tyr Asp Asp Val Thr Ile Lys Met Ile Phe Ala Ile Val Gln Ile Ile			
305	310	315	320
Gly Phe Ser Asn Ser Ile Cys Asn Pro Ile Val Tyr Ala Phe Met Asn			
325	330	335	
Glu Asn Phe Lys Lys Asn Val Leu Ser Ala Val Cys Tyr Cys Ile Val			
340	345	350	
25 Asn Lys Thr Phe Ser Pro Ala Gln Arg His Gly Asn Ser Gly Ile Thr			
355	360	365	
Met Met Arg Lys Lys Ala Lys Phe Ser Leu Arg Glu Asn Pro Val Glu			
370	375	380	
30 Glu Thr Lys Gly Glu Ala Phe Ser Asp Gly Asn Ile Glu Val Lys Leu			
385	390	395	400
Cys Glu Gln Thr Glu Glu Lys Lys Leu Lys Arg His Leu Ala Leu			
405	410	415	
Phe Arg Ser Glu Leu Ala Glu Asn Ser Pro Leu Asp Ser Gly His			
420	425	430	
35 (42) INFORMATION FOR SEQ ID NO:41:			
(i) SEQUENCE CHARACTERISTICS:			
(A) LENGTH: 24 base pairs			

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CTGTGTACAG CAGTCGCAG AGTG

24

(43) INFORMATION FOR SEQ ID NO:42:

- 10 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

15 GAGTGCCAGG CAGAGCAGGT AGAC

24

(44) INFORMATION FOR SEQ ID NO:43:

20

- 20 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 31 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

25 CCCGAATTCC TGCTTGCTCC CAGCTTGGCC C

31

30

(45) INFORMATION FOR SEQ ID NO:44:

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 32 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TGTGGATCCT GCTGTCAAAG GTCCCATTCC GG

32

(46) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

10 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TCACAATGCT AGGTGTGGTC

20

(47) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

20 (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TGCATAGACA ATGGGATTAC AG

22

(48) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 511 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TCACAATGCT AGGTGTGGTC TGGCTGGTGG CAGTCATCGT AGGATCACCC ATGTGGCAGG

60

TGCAAACAACT TGAGATCAAATATGACTTCC TATATGAAAAA GGAACACATC TGCTGCTTAG

120

AAGAGTGGAC CAGCCCTGTG CACCAGAAGA TCTACACCAC CTTCATCCTT GTCATCCTCT	180
TCCTCCTGCC TCTTATGGTG ATGCTTATTG TGTACGTAAA ATTGGTTATG AACTTTGGAT	240
AAAGAAAAAGA GTTGGGGATG GTTCAGTGCT TCGAACTATT CATGGAAAAG AAATGTCCAA	300
AATAGCCAGG AAGAAGAAC GAGCTGTCAT TATGATGGTG ACAGTGGTGG CTCTCTTGC	360
5 TGTGTGCTGG GCACCATTCC ATGTTGTCCA TATGATGATT GAATACAGTA ATTTTGAAAAA	420
GGAATATGAT GATGTCACAA TCAAGATGAT TTTTGCTATC GTGCAAATTAA TTGGATTTTC	480
CAACTCCATC TGTAATCCCA TTGTCTATGC A	511

10

(49) INFORMATION FOR SEQ ID NO:48:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

15 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

20

CTGCTTAGAA GAGTGGACCA G

21

(50) INFORMATION FOR SEQ ID NO:49:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (iv) ANTI-SENSE: NO

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CTGTGCACCA GAAGATCTAC AC

22

(51) INFORMATION FOR SEQ ID NO:50:

30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

CAAGGATGAA GGTGGTGTAG A

21

5 (52) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GTGTAGATCT TCTGGTGCAC AGG

23

15 (53) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GCAATGCAGG TCATAGTGAG C

21

(54) INFORMATION FOR SEQ ID NO:53:

30

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TGGAGCATGG TGACGGGAAT GCAGAAG

27

(55) INFORMATION FOR SEQ ID NO:54:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

10 (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GTGATGAGCA GGTCACTGAG CGCCAAG

27

(56) INFORMATION FOR SEQ ID NO:55:

- 15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

20 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

GCAATGCAGG CGCTTAACAT TAC

23

(57) INFORMATION FOR SEQ ID NO:56:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: DNA (genomic)

30 (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

TTGGGTTACA ATCTGAAGGG CA

22

(58) INFORMATION FOR SEQ ID NO:57:

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

10 ACTCCGTGTC CAGCAGGACT CTG

23

(58) INFORMATION FOR SEQ ID NO:58:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

20 TGC GTGTTCC TGG ACCCTCA CGTG

24

(58) INFORMATION FOR SEQ ID NO:59:

25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

30 CAGGCCTTGG ATTTAATGT CAGGGATGG

29

(61) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

GGAGAGTCAG CTCTGAAAGA ATTCAAGG

27

(62) INFORMATION FOR SEQ ID NO:61:

10

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

15 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

TGATGTGATG CCAGATACTA ATAGCAC

27

20

(63) INFORMATION FOR SEQ ID NO:62:

- 20 (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (iv) ANTI-SENSE: YES

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CCTGATTCA TTAGGTGAGA TTGAGAC

27

(64) INFORMATION FOR SEQ ID NO:63:

- 30 (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

CCCAAGCTTC CCCAGGTGTA TTTGAT

26

(3) INFORMATION FOR SEQ ID NO:63:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

10

10 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GTTGGATCCA CATAATGCAT TTTCTC

26

(66) INFORMATION FOR SEQ ID NO:65:

- 15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1080 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

ATGATTCTCA ACTCTTCTAC TGAAGATGGT ATTAAAAAGAA TCCAAGATGA TTGTCCAAA 60

GCTGGAAGGC ATAATTACAT ATTTGTCTG ATTCTACTT TATACAGTAT CATCTTGTG 120

GTGGGAATAT TTGGAAACAG CTTGGTGGTG ATAGTCATTT ACTTTTATAT GAAGCTGAAG 180

ACTGTGGCCA GTGTTTTCT TTTGAATTAA GCACTGGCTG ACTTATGCTT TTTACTGACT 240

30

25 TTGCCACTAT GGGCTGTCTA CACAGCTATG GAATACCGCT GGCCCTTTGG CAATTACCTA 300

TGTAAGATTG CTTCAGCCAG CGTCAGTTTC AACCTGTACG CTAGTGTGTT TCTACTCAGC 360

TGTCTCAGCA TTGATCGATA CCTGGCTATT GTTCACCCAA TGAAGTCCCG CCTTCGACGC 420

ACAATGCTTG TAGCCAAAGT CACCTGCATC ATCATTGGC TGCTGGCAGG CTTGGCCAGT 480

TTGCCAGCTA TAATCCATCG AAATGTATTT TTCATTGAGA ACACCAATAT TACAGTTGT 540

30 GCTTCCATT ATGAGTCCCA AAATTCAACC CTTCCGATAG GGCTGGCCT GACCAAAAT 600

ATACTGGGTT TCCTGTTCC TTTCTGATC ATTCTTACAA GTTATACTCT TATTTGGAAG 660  
 GCCCTAAAGA AGGCTTATGA AATTCAAGAAG AACAAACCAA GAAATGATGA TATTTTAAG 720  
 ATAATTATGG CAATTGTGCT TTTCTTTTC TTTCTCTGGA TTCCCCACCA AATATTCACT 780  
 TTTCTGGATG TATTGATTCA ACTAGGCATC ATACGTGACT GTAGAATTGC AGATATTGTG 840  
 5 GACACGGCCA TGCCTATCAC CATTGTATA GCTTATTTA ACAATTGCCT GAATCCTCTT 900  
 TTTTATGGCT TTCTGGGAA AAAATTAAA AGATATTTC TCCAGCTCT AAAATATATT 960  
 CCCCCAAAAG CCAAATCCC CTCAAAACCTT TCAACAAAAA TGAGCACGCT TTCCTACCGC 1020  
 CCCTCAGATA ATGTAAGCTC ATCCACCAAG AAGCCTGCAC CATGTTTGA GGTTGAGTGA 1080

## (67) INFORMATION FOR SEQ ID NO:66:

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 359 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

15 (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

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

20 Asp Cys Pro Lys Ala Gly Arg His Asn Tyr Ile Phe Val Met Ile Pro  
 20 25 30

Thr Leu Tyr Ser Ile Ile Phe Val Val Gly Ile Phe Gly Asn Ser Leu  
 35 40 45

Val Val Ile Val Ile Tyr Phe Tyr Met Lys Leu Lys Thr Val Ala Ser  
 50 55 60

25 Val Phe Leu Leu Asn Leu Ala Leu Ala Asp Leu Cys Phe Leu Leu Thr  
 65 70 75 80

Leu Pro Leu Trp Ala Val Tyr Thr Ala Met Glu Tyr Arg Trp Pro Phe  
 85 90 95

30 Gly Asn Tyr Leu Cys Lys Ile Ala Ser Ala Ser Val Ser Phe Asn Leu  
 100 105 110

Tyr Ala Ser Val Phe Leu Leu Thr Cys Leu Ser Ile Asp Arg Tyr Leu  
 115 120 125

Ala Ile Val His Pro Met Lys Ser Arg Leu Arg Arg Thr Met Leu Val

10

20

30

40

	130	135	140	
	Ala Lys Val Thr Cys Ile Ile Trp Leu Leu Ala Gly Leu Ala Ser			
145	145	150	155	160
5	Leu Pro Ala Ile Ile His Arg Asn Val Phe Phe Ile Glu Asn Thr Asn			
	165	170	175	
	Ile Thr Val Cys Ala Phe His Tyr Glu Ser Gln Asn Ser Thr Leu Pro			
	180	185	190	
	Ile Gly Leu Gly Leu Thr Lys Asn Ile Leu Gly Phe Leu Phe Pro Phe			
	195	200	205	
10	Leu Ile Ile Leu Thr Ser Tyr Thr Leu Ile Trp Lys Ala Leu Lys Lys			
	210	215	220	
	Ala Tyr Glu Ile Gln Lys Asn Lys Pro Arg Asn Asp Asp Ile Phe Lys			
	225	230	235	240
15	Ile Ile Met Ala Ile Val Leu Phe Phe Phe Ser Trp Ile Pro His			
	245	250	255	
	Gln Ile Phe Thr Phe Leu Asp Val Leu Ile Gln Leu Gly Ile Ile Arg			
	260	265	270	
	Asp Cys Arg Ile Ala Asp Ile Val Asp Thr Ala Met Pro Ile Thr Ile			
	275	280	285	
20	Cys Ile Ala Tyr Phe Asn Asn Cys Leu Asn Pro Leu Phe Tyr Gly Phe			
	290	295	300	
	Leu Gly Lys Lys Phe Lys Arg Tyr Phe Leu Gln Leu Leu Lys Tyr Ile			
	305	310	315	320
25	Pro Pro Lys Ala Lys Ser His Ser Asn Leu Ser Thr Lys Met Ser Thr			
	325	330	335	
	Leu Ser Tyr Arg Pro Ser Asp Asn Val Ser Ser Ser Thr Lys Lys Pro			
	340	345	350	
	Ala Pro Cys Phe Glu Val Glu			
	355			30
30	(68) INFORMATION FOR SEQ ID NO:67:			
	(i) SEQUENCE CHARACTERISTICS:			
	(A) LENGTH: 27 base pairs			
	(B) TYPE: nucleic acid			
	(C) STRANDEDNESS: single			
35	(D) TOPOLOGY: linear			
	(ii) MOLECULE TYPE: DNA (genomic)			

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

ACCATGGGCA GCCCCTGGAA CGGCAGC

27

(69) INFORMATION FOR SEQ ID NO:68:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

AGAACCCACCA CCAGCAGGAC GCGGACGGTC TGCCGGTGG

39

(70) INFORMATION FOR SEQ ID NO:69:

- 15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

20 GTCCCGCGTCC TGCTGGTGGT GGTTCTGGCA TTTATAATT

39

(71) INFORMATION FOR SEQ ID NO:70:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: DNA (genomic)

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CCTGGATCCT TATCCCATCG TCTTCACGTT AGC

33

30 (72) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

40

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

5 CTGGAATTCT CCTGCCAGCA TGGTGA  
26

(73) INFORMATION FOR SEQ ID NO:72:

10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

GCAGGATCCT ATATTGCGTG CTCTGTCCCC  
30

20

(74) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 999 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

ATGGTGAAC	CCACCCACCG	TGGGATGCAC	ACTTCTCTGC	ACCTCTGGAA	CCGCAGCAGT	60	30
TACAGACTGC	ACAGCAATGC	CAGTGAGTCC	CTTGGAAAAG	GCTACTCTGA	TGGAGGGTGC	120	
TACGAGCAAC	TTTTGTCTC	TCCTGAGGTG	TTTGTGACTC	TGGGTGTCA	CAGCTTGTG	180	
GAGAATATCT	TAGTGATTGT	GGCAATAGCC	AAGAACAAAGA	ATCTGCATTC	ACCCATGTAC	240	
30 TTTTTCATCT	GCAGCTTGGC	TGTGGCTGAT	ATGCTGGTGA	GCGTTTCAA	TGGATCAGAA	300	
ACCATTATCA	TCACCCATT	AAACAGTACA	GATACGGATG	CACAGAGTTT	CACAGTGAAT	360	
ATTGATAATG	TCATTGACTC	GGTGATCTGT	AGCTCCTTGC	TTGCATCCAT	TTGCAGCCTG	420	

40

CTTTCAATTG CAGTGGACAG GTACTTTACT ATCTTCTATG CTCTCCAGTA CCATAACATT	480
ATGACAGTTA AGCGGGTTGG GATCAGCATA AGTTGTATCT GGGCAGCTTG CACGGTTCA	540
GGCATTGGT TCATCATTAA CTCAGATAGT AGTGCTGTCA TCATCTGCCT CATCACCATG	600
TTCTTCACCA TGCTGGCTCT CATGGCTTCT CTCTATGTCC ACATGTTCCCT GATGCCAGG	660
5 CTTCACATTA AGAGGATTGC TGCCTCCCC GGCACTGGTG CCATCCGCCA AGGTGCCAAT	720
ATGAAGGGAG CGATTACCTT GACCATCCTG ATTGGCGTCT TTGTTGTCTG CTGGGCCCA	780
TTCTTCCTCC ACTTAATATT CTACATCTCT TGTCTCAGA ATCCATATTG TGTGTGCTTC	840
ATGTCTCACT TTAACTTGTA TCTCATACTG ATCATGTGTA ATTCAATCAT CGATCCTCTG	900
ATTATGCAC TCCGGAGTCA AGAACTGAGG AAAACCTCA AAGAGATCAT CTGTTGCTAT	960
10 CCCCTGGGAG GCCTTTGTGA CTTGTCTAGC AGATATTAA	999

## (75) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 332 amino acids	
(B) TYPE: amino acid	
15 (C) STRANDEDNESS:	
(D) TOPOLOGY: not relevant	20
(ii) MOLECULE TYPE: protein	

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

Met Val Asn Ser Thr His Arg Gly Met His Thr Ser Leu His Leu Trp	
20 1 5 10 15	
Asn Arg Ser Ser Tyr Arg Leu His Ser Asn Ala Ser Glu Ser Leu Gly	
25 20 25 30	
Lys Gly Tyr Ser Asp Gly Gly Cys Tyr Glu Gln Leu Phe Val Ser Pro	
30 35 40 45	
25 Glu Val Phe Val Thr Leu Gly Val Ile Ser Leu Leu Glu Asn Ile Leu	30
50 55 60	
Val Ile Val Ala Ile Ala Lys Asn Lys Asn Leu His Ser Pro Met Tyr	
65 70 75 80	
30 Phe Phe Ile Cys Ser Leu Ala Val Ala Asp Met Leu Val Ser Val Ser	35
85 90 95	
Asn Gly Ser Glu Thr Ile Ile Thr Leu Leu Asn Ser Thr Asp Thr	
100 105 110	
Asp Ala Gln Ser Phe Thr Val Asn Ile Asp Asn Val Ile Asp Ser Val	

	115	120	125	
	Ile Cys Ser Ser Leu Leu Ala Ser Ile Cys Ser Leu Leu Ser Ile Ala			
	130	135	140	
5	Val Asp Arg Tyr Phe Thr Ile Phe Tyr Ala Leu Gln Tyr His Asn Ile			
	145	150	155	160
	Met Thr Val Lys Arg Val Gly Ile Ser Ile Ser Cys Ile Trp Ala Ala			
	165	170	175	
10	Cys Thr Val Ser Gly Ile Leu Phe Ile Ile Tyr Ser Asp Ser Ser Ala			
	180	185	190	
	Val Ile Ile Cys Leu Ile Thr Met Phe Phe Thr Met Leu Ala Leu Met			
	195	200	205	
	Ala Ser Leu Tyr Val His Met Phe Leu Met Ala Arg Leu His Ile Lys			
	210	215	220	
15	Arg Ile Ala Val Leu Pro Gly Thr Gly Ala Ile Arg Gln Gly Ala Asn			
	225	230	235	240
	Met Lys Gly Ala Ile Thr Leu Thr Ile Leu Ile Gly Val Phe Val Val			
	245	250	255	
	Cys Trp Ala Pro Phe Phe Leu His Leu Ile Phe Tyr Ile Ser Cys Pro			
	260	265	270	
20	Gln Asn Pro Tyr Cys Val Cys Phe Met Ser His Phe Asn Leu Tyr Leu			
	275	280	285	
	Ile Leu Ile Met Cys Asn Ser Ile Ile Asp Pro Leu Ile Tyr Ala Leu			
	290	295	300	
25	Arg Ser Gln Glu Leu Arg Lys Thr Phe Lys Glu Ile Ile Cys Cys Tyr			
	305	310	315	320
	Pro Leu Gly Gly Leu Cys Asp Leu Ser Ser Arg Tyr			
	325	330		

## (76) INFORMATION FOR SEQ ID NO:75:

	30
	(i) SEQUENCE CHARACTERISTICS:
30	(A) LENGTH: 32 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: single
	(D) TOPOLOGY: linear
	(ii) MOLECULE TYPE: DNA (genomic)
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

CCGAAGCTTC GAGCTGAGTA AGGCAGGCGGG CT

32

40

(77) INFORMATION FOR SEQ ID NO:76:

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

GTGGAATTCA TTTGCCCTGC CTCAACCCCC A

31

10

10 (78) INFORMATION FOR SEQ ID NO:77:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1344 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

20 ATGGAGCTGC TAAAGCTGAA CGGGAGCGTG CAGGGAACCG GACCCGGGCC GGGGGCTTCC 60  
 CTGTGCCGCC CGGGGGCGCC TCTCCTCAAC AGCAGCAGTG TGGGCAACCT CAGCTGCGAG 120  
 20 CCCCTCGCA TTCGCGGAGC CGGGACACGA GAATTGGAGC TGGCCATTAG AATCACTCTT 180  
 TACGCAGTGA TCTTCCTGAT GAGCGTTGGA GGAAATATGC TCATCATCGT GGTCTGGGA 240  
 CTGAGCCGCC GCCTGAGGAC TGTCACCAAT GCCTTCCTCC TCTCACTGGC AGTCAGCGAC 300  
 CTCCTGCTGG CTGTGGCTTG CATGCCCTTC ACCCTCCTGC CCAATCTCAT GGGCACATTC 360  
 ATCTTGCA CCGTCATCTG CAAGGCGGTT TCCTACCTCA TGGGGGTGTC TGTGAGTGTG 420  
 25 TCCACGCTAA GCCTCGTGGC CATCGCACTG GAGCGATATA GCGCCATCTG CCGACCACTG 480  
 CAGGCCACGAG TGTGGCAGAC GCGCTCCCAC GCGGCTCGCG TGATTGTAGC CACGTGGCTG 540  
 CTGTCCGGAC TACTCATGGT GCCCTACCCC GTGTACACTG TCGTGCAACC AGTGGGGCCT 600  
 CGTGTGCTGC AGTGCCTGCA TCGCTGGCCC AGTGCCTGGG TCCGCCAGAC CTGGTCCGTA 660  
 CTGCTGCTTC TGCTCTTGTGTT CTTCATCCCA GGTGTGGTTA TGGCCGTGGC CTACGGGTT 720  
 30 ATCTCTCGCG AGCTCTACTT AGGGCTTCGC TTTGACGGCG ACAGTGACAG CGACAGCAA 780  
 AGCAGGGTCC GAAACCAAGG CGGGCTGCCA GGGGCTGTTC ACCAGAACGG GCGTTGCCGG 840

20

30

40

CCTGAGACTG	GCGCGGTTGG	CAAAGACAGC	GATGGCTGCT	ACGTGCAACT	TCCACGTTCC	900	
CGGCCTGCC	TGGAGCTGAC	GGCGCTGACG	GCTCCTGGC	CGGGATCCGG	CTCCCGGCC	960	
ACCCAGGCCA	AGCTGCTGGC	TAAGAAGCGC	GTGGTGCGAA	TGTTGCTGGT	GATCGTTGTG	1020	
CTTTTTTTC	TGTGTTGGTT	GCCAGTTTAT	AGTGCCAACA	CGTGGCGCGC	CTTTGATGGC	1080	
5	CCGGGTGCAC	ACCGAGCACT	CTCGGGTGCT	CCTATCTCCT	TCATTCACTT	GCTGAGCTAC	1140
	GCCTCGGCCT	GTGTCAACCC	CCTGGTCTAC	TGCTTCATGC	ACCGTCGCTT	TCGCCAGGCC	1200
	TGCCTGGAAA	CTTGCCTCG	CTGCTGCC	CGGCCTCCAC	GAGCTCGCCC	CAGGGCTCTT	1260
10	CCCGATGAGG	ACCCCTCCCAC	TCCCTCCATT	GCTTCGCTGT	CCAGGCTTAG	CTACACCACC	1320
	ATCAGCACAC	TGGGCCCTGG	CTGA				1344

## 10 (79) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 447 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: not relevant

15 (ii) MOLECULE TYPE: protein

## 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Met	Glu	Leu	Leu	Lys	Leu	Asn	Arg	Ser	Val	Gln	Gly	Thr	Gly	Pro	Gly
1															15
Pro	Gly	Ala	Ser	Leu	Cys	Arg	Pro	Gly	Ala	Pro	Leu	Leu	Asn	Ser	Ser
20															30
Ser	Val	Gly	Asn	Leu	Ser	Cys	Glu	Pro	Pro	Arg	Ile	Arg	Gly	Ala	Gly
35															45
Thr	Arg	Glu	Leu	Glu	Leu	Ala	Ile	Arg	Ile	Thr	Leu	Tyr	Ala	Val	Ile
25															55
Phe	Leu	Met	Ser	Val	Gly	Gly	Asn	Met	Leu	Ile	Ile	Val	Val	Leu	Gly
65															75
Leu	Ser	Arg	Arg	Leu	Arg	Thr	Val	Thr	Asn	Ala	Phe	Leu	Leu	Ser	Leu
85															95
Ala	Val	Ser	Asp	Leu	Leu	Ala	Val	Ala	Cys	Met	Pro	Phe	Thr	Leu	
100															110
Leu	Pro	Asn	Leu	Met	Gly	Thr	Phe	Ile	Phe	Gly	Thr	Val	Ile	Cys	Lys
115															125

10

20

30

Ala Val Ser Tyr Leu Met Gly Val Ser Val Ser Val Ser Thr Leu Ser  
 130 135 140  
 Leu Val Ala Ile Ala Leu Glu Arg Tyr Ser Ala Ile Cys Arg Pro Leu  
 145 150 155 160  
 5 Gln Ala Arg Val Trp Gln Thr Arg Ser His Ala Ala Arg Val Ile Val  
 165 170 175  
 Ala Thr Trp Leu Leu Ser Gly Leu Leu Met Val Pro Tyr Pro Val Tyr  
 180 185 190  
 10 Thr Val Val Gln Pro Val Gly Pro Arg Val Leu Gln Cys Val His Arg  
 195 200 205  
 Trp Pro Ser Ala Arg Val Arg Gln Thr Trp Ser Val Leu Leu Leu  
 210 215 220  
 Leu Leu Phe Phe Ile Pro Gly Val Val Met Ala Val Ala Tyr Gly Leu  
 225 230 235 240  
 15 Ile Ser Arg Glu Leu Tyr Leu Gly Leu Arg Phe Asp Gly Asp Ser Asp  
 245 250 255  
 Ser Asp Ser Gln Ser Arg Val Arg Asn Gln Gly Gly Leu Pro Gly Ala  
 260 265 270  
 20 Val His Gln Asn Gly Arg Cys Arg Pro Glu Thr Gly Ala Val Gly Lys  
 275 280 285  
 Asp Ser Asp Gly Cys Tyr Val Gln Leu Pro Arg Ser Arg Pro Ala Leu  
 290 295 300  
 Glu Leu Thr Ala Leu Thr Ala Pro Gly Pro Gly Ser Gly Ser Arg Pro  
 305 310 315 320  
 25 Thr Gln Ala Lys Leu Leu Ala Lys Lys Arg Val Val Arg Met Leu Leu  
 325 330 335  
 Val Ile Val Val Leu Phe Phe Leu Cys Trp Leu Pro Val Tyr Ser Ala  
 340 345 350  
 Asn Thr Trp Arg Ala Phe Asp Gly Pro Gly Ala His Arg Ala Leu Ser  
 355 360 365  
 30 Val Ala Pro Ile Ser Phe Ile His Leu Leu Ser Tyr Ala Ser Ala Cys  
 370 375 380  
 Val Asn Pro Leu Val Tyr Cys Phe Met His Arg Arg Phe Arg Gln Ala  
 385 390 395 400  
 35 Cys Leu Glu Thr Cys Ala Arg Cys Cys Pro Arg Pro Pro Arg Ala Arg  
 405 410 415  
 Pro Arg Ala Leu Pro Asp Glu Asp Pro Pro Thr Pro Ser Ile Ala Ser

420

425

430

Leu Ser Arg Leu Ser Tyr Thr Thr Ile Ser Thr Leu Gly Pro Gly  
 435 440 445

## (80) INFORMATION FOR SEQ ID NO:79:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

TGCAAGCTTA AAAAGGAAAA AATGAACAGC

30

## (81) INFORMATION FOR SEQ ID NO:80:

- 15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

TAAGGATCCC TTCCCTTCAA AACATCCTTG

30

## (82) INFORMATION FOR SEQ ID NO:81:

- 25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1014 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

30 ATGAACAGCA CATGTATTGA AGAACAGCAT GACCTGGATC ACTATTTGTT TCCCATTGTT 60  
 TACATCTTTG TGATTATAGT CAGCATTCCA GCCAATATTG GATCTCTGTG TGTGTCTTTC 120  
 CTGCAACCCA AGAAGGAAAG TGAACCTAGGA ATTTACCTCT TCAGTTTGTC ACTATCAGAT 180  
 TTACTCTATG CATTAACTCT CCCTTTATGG ATTGATTATA CTTGGAATAA AGACAACCTGG 240

ACTTTCTCTC CTGCCTTGTG CAAAGGGAGT GCTTTCTCA TGTACATGAA GTTTTACAGC 300  
 AGCACAGCAT TCCTCACCTG CATTGCCGTT GATCGGTATT TGGCTGTTGT CTACCCTTG 360  
 AAGTTTTTT TCCTAAGGAC AAGAAGAATT GCACTCATGG TCAGCCTGTC CATCTGGATA 420  
 TTGGAAACCA TCTTCAATGC TGTCAATGTTG TGGGAAGATG AAACAGTTGT TGAATATTGC 480  
 5 GATGCCGAAA AGTCTAATTT TACTTTATGC TATGACAAAT ACCCTTTAGA GAAATGGCAA 540  
 ATCAACCTCA ACTTGTTCAAG GACGTGTACA GGCTATGCAA TACCTTTGGT CACCACCTG 600  
 ATCTGTAACC GGAAAGTCTA CCAAGCTGTG CGGCACAATA AAGCCACGGA AAACAAGGAA 660  
 10 AAGAAGAGAA TCATAAAACT ACTTGTCAGC ATCACAGTTA CTTTTGTCCTT ATGCTTTACT  
 CCCTTTCATG TGATGTTGCT GATTCGCTGC ATTTTAGAGC ATGCTGTGAA CTTCGAAGAC 780  
 15 CACAGCAATT CTGGGAAGCG AACTTACACA ATGTATAGAA TCACGGTTGC ATTAACAAGT 840  
 TTAAATTGTG TTGCTGATCC AATTCTGTAC TGTGTTGTTA CCGAAACAGG AAGATATGAT 900  
 ATGTGGAATA TATTAAAATT CTGCACTGGG AGGTGTAATA CATCACAAAG ACAAAAGAAAA 960  
 CGCATACTTT CTGTGTCTAC AAAAGATACT ATGGAATTAG AGGTCCCTGA GTAG 1014

## (83) INFORMATION FOR SEQ ID NO:82:

20 15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 337 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

20 (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

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

	85	90	95	
	Lys Phe Tyr Ser Ser Thr Ala Phe Leu Thr Cys Ile Ala Val Asp Arg			
	100	105	110	
5	Tyr Leu Ala Val Val Tyr Pro Leu Lys Phe Phe Leu Arg Thr Arg			
	115	120	125	
	Arg Ile Ala Leu Met Val Ser Leu Ser Ile Trp Ile Leu Glu Thr Ile			
	130	135	140	
	Phe Asn Ala Val Met Leu Trp Glu Asp Glu Thr Val Val Glu Tyr Cys			
	145	150	155	160
10	Asp Ala Glu Lys Ser Asn Phe Thr Leu Cys Tyr Asp Lys Tyr Pro Leu			
	165	170	175	
	Glu Lys Trp Gln Ile Asn Leu Asn Leu Phe Arg Thr Cys Thr Gly Tyr			
	180	185	190	
15	Ala Ile Pro Leu Val Thr Ile Leu Ile Cys Asn Arg Lys Val Tyr Gln			
	195	200	205	
	Ala Val Arg His Asn Lys Ala Thr Glu Asn Lys Glu Lys Lys Arg Ile			
	210	215	220	
	Ile Lys Leu Leu Val Ser Ile Thr Val Thr Phe Val Leu Cys Phe Thr			
	225	230	235	240
20	Pro Phe His Val Met Leu Leu Ile Arg Cys Ile Leu Glu His Ala Val			
	245	250	255	
	Asn Phe Glu Asp His Ser Asn Ser Gly Lys Arg Thr Tyr Thr Met Tyr			
	260	265	270	
	Arg Ile Thr Val Ala Leu Thr Ser Leu Asn Cys Val Ala Asp Pro Ile			
25	275	280	285	
	Leu Tyr Cys Phe Val Thr Glu Thr Gly Arg Tyr Asp Met Trp Asn Ile			
	290	295	300	
	Leu Lys Phe Cys Thr Gly Arg Cys Asn Thr Ser Gln Arg Gln Arg Lys			
	305	310	315	320
30	Arg Ile Leu Ser Val Ser Thr Lys Asp Thr Met Glu Leu Glu Val Leu			
	325	330	335	
	Glu			

## (84) INFORMATION FOR SEQ ID NO:83:

- 35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 40 base pairs  
 (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

5 CAGGAAGAAG AAACGAGCTG TCATTATGAT GGTGACAGTG  
40

(85) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 40 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

15 CACTGTCACC ATCATAATGA CAGCTCGTTT CTTCTTCCTG  
40

(86) INFORMATION FOR SEQ ID NO:85:

20

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

25 GGCCACCGGC AGACCAAACG CGTCCTGCTG  
30

30

(87) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 31 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

CTCCTTCGGT CCTCCTATCG TTGTCAGAAG T  
31

(88) INFORMATION FOR SEQ ID NO:87:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 37 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

GGAAAAGAAG AGAATCAAAA AACTACTTGT CAGCATC

37

(89) INFORMATION FOR SEQ ID NO:88:

- 15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

20 CTCCTTCGGT CCTCCTATCG TTGTCAGAAG T

31

(90) INFORMATION FOR SEQ ID NO:89:

- 25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1080 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

ATGATTCTCA ACTCTTCTAC TGAAGATGGT ATTAAAAGAA TCCAAGATGA TTGTCCAAA 60

30 GCTGGAAGGC ATAATTACAT ATTTGTCTG ATTCCTACTT TATACAGTAT CATCTTG 120

GTGGGAATAT TTGGAAACAG CTTGGTGGTG ATAGTCATTT ACTTTTATAT GAAGCTGAAG 180

ACTGTGGCCA GTGTTTTCT TTTGAATTAA GCACTGGCTG ACTTATGCTT TTTACTGACT 240

TTGCCACTAT GGGCTGTCTA CACAGCTATG GAATACCGCT GCCCCTTGG CAATTACCTA 300

TGTAAGATTG CTTCAGCCAG CGTCAGTTTC AACCTGTACG CTAGTGTGTT TCTACTCAGC 360  
 TGTCTCAGCA TTGATCGATA CCTGGCTATT GTTCACCCAA TGAAGTCCCG CCTTCGACGC 420  
 ACAATGCTTG TAGCCAAAGT CACCTGCATC ATCATTGGC TGCTGGCAGG CTTGGCCAGT 480  
 TTGCCAGCTA TAATCCATCG AAATGTATTT TTCATTGAGA ACACCAATAT TACAGTTTGT 540  
 5 GCTTTCCATT ATGAGTCCA AAATTCAACC CTTCCGATAG GGCTGGGCCT GACCAAAAT 600  
 ATACTGGGTT TCCTGTTCC TTTCTGATC ATTCTTACAA GTTATACTCT TATTGGAAAG 660  
 GCCCTAAAGA AGGCTTATGA AATTCAAGAACAA GAAATGATGA TATTAAGAAAG 720  
 10 ATAATTATGG CAAATTGTGCT TTTCTTTTC TTTTCCTGGA TTCCCCACCA AATATTCACT 780  
 TTTCTGGATG TATTGATTCA ACTAGGCATC ATACGTGACT GTAGAATTGC AGATATTGTG 840  
 GACACGGCCA TGCCTATCAC CATTGTATA GCTTATTTA ACAATTGCCT GAATCCTCTT 900  
 TTTTATGGCT TTCTGGGAA AAAATTAAA AGATATTTTC TCCAGCTTCT AAAATATATT 960  
 CCCCCAAAAG CCAAATCCCA CTCAAACCTT TCAACAAAAA TGAGCACGCT TTCCTACCGC 1020  
 CCCTCAGATA ATGTAAGCTC ATCCACCAAG AAGCCTGCAC CATGTTTGA GGTTGAGTGA 1080

(91) INFORMATION FOR SEQ ID NO:90:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 359 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

20 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

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

10

20

30

	85	90	95	
	Gly Asn Tyr Leu Cys Lys Ile Ala Ser Ala Ser Val Ser Phe Asn Leu			
	100	105	110	
5	Tyr Ala Ser Val Phe Leu Leu Thr Cys Leu Ser Ile Asp Arg Tyr Leu			
	115	120	125	
	Ala Ile Val His Pro Met Lys Ser Arg Leu Arg Arg Thr Met Leu Val			
	130	135	140	
	Ala Lys Val Thr Cys Ile Ile Trp Leu Leu Ala Gly Leu Ala Ser			
	145	150	155	160
10	Leu Pro Ala Ile Ile His Arg Asn Val Phe Phe Ile Glu Asn Thr Asn			
	165	170	175	
	Ile Thr Val Cys Ala Phe His Tyr Glu Ser Gln Asn Ser Thr Leu Pro			
	180	185	190	
15	Ile Gly Leu Gly Leu Thr Lys Asn Ile Leu Gly Phe Leu Phe Pro Phe			
	195	200	205	
	Leu Ile Ile Leu Thr Ser Tyr Thr Leu Ile Trp Lys Ala Leu Lys Lys			
	210	215	220	
	Ala Tyr Glu Ile Gln Lys Asn Lys Pro Arg Asn Asp Asp Ile Lys Lys			
	225	230	235	240
20	Ile Ile Met Ala Ile Val Leu Phe Phe Phe Ser Trp Ile Pro His			
	245	250	255	
	Gln Ile Phe Thr Phe Leu Asp Val Leu Ile Gln Leu Gly Ile Ile Arg			
	260	265	270	
25	Asp Cys Arg Ile Ala Asp Ile Val Asp Thr Ala Met Pro Ile Thr Ile			
	275	280	285	
	Cys Ile Ala Tyr Phe Asn Asn Cys Leu Asn Pro Leu Phe Tyr Gly Phe			
	290	295	300	
	Leu Gly Lys Lys Phe Lys Arg Tyr Phe Leu Gln Leu Leu Lys Tyr Ile			
	305	310	315	320
30	Pro Pro Lys Ala Lys Ser His Ser Asn Leu Ser Thr Lys Met Ser Thr			
	325	330	335	
	Leu Ser Tyr Arg Pro Ser Asp Asn Val Ser Ser Ser Thr Lys Lys Pro			
	340	345	350	
35	Ala Pro Cys Phe Glu Val Glu			
	355			

(92) INFORMATION FOR SEQ ID NO:91:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 35 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

CCAAGAAATG ATGATATTAA AAAGATAATT ATGGC

35

(93) INFORMATION FOR SEQ ID NO:92:

- 10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

CTCCTTCGGT CCTCCTATCG TTGTCAGAAG T

31

(94) INFORMATION FOR SEQ ID NO:93:

- 20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1080 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

ATGATTCTCA ACTCTTCTAC TGAAGATGGT ATTAAAAGAA TCCAAGATGA TTGTCCAAA 60

30

GCTGGAAGGC ATAATTACAT ATTTGTCTAG ATTCCTACTT TATACTGAT CATCTTTGTG 120

GTGGGAATAT TTGGAAACAG CTTGGTGGTG ATAGTCATTT ACTTTTATAT GAAGCTGAAG 180

ACTGTGGCCA GTGTTTTCT TTTGAATTAA GCACTGGCTG ACTTATGCTT TTTACTGACT 240

30 TTGCCACTAT GGGCTGTCTA CACAGCTATG GAATACCGCT GGCCCTTTGG CAATTACCTA 300

TGTAAGATTG CTTCAGCCAG CGTCAGTTTC GCCCTGTACG CTAGTGTGTT TCTACTCAGC 360

TGTCTCAGCA TTGATCGATA CCTGGCTATT GTTCACCCAA TGAAGTCCCG CCTTCGACGC 420

ACAATGCTTG TAGCCAAAGT CACCTGCATC ATCATTGGC TGCTGGCAGG CTTGGCCAGT 480  
 TTGCCAGCTA TAATCCATCG AAATGTATTT TTCATTGAGA ACACCAATAT TACAGTTTGT 540  
 GCTTTCCATT ATGAGTCCC AATTCAACC CTTCCGATAG GGCTGGGCCT GACCAAAAT 600  
 ATACTGGGTT TCCTGTTCC TTTCTGATC ATTCTTACAA GTTATACTCT TATTTGGAAG 660  
 5 GCCCTAAAGA AGGCTTATGA AATTCAAGG AACAAACCA GAAATGATGA TATTTTAAG 720  
 ATAATTATGG CAATTGTGCT TTTCTTTTC TTTTCTGGA TTCCCCACCA AATATTCACT 780  
 TTTCTGGATG TATTGATTCA ACTAGGCATC ATACGTGACT GTAGAATTGC AGATATTGTG 840  
 10 GACACGGCCA TGCCTATCAC CATTGTATA GCTTATTTA ACAATTGCCT GAATCCTCTT 900  
 TTTTATGGCT TTCTGGGAA AAAATTTAAA AGATATTTTC TCCAGCTTCT AAAATATATT 960  
 15 CCCCCAAAG CCAAATCCC CTCAAACCTT TCAACAAAAA TGAGCACGCT TTCTACCGC 1020  
 CCCTCAGATA ATGTAAGCTC ATCCACCAAG AAGCCTGCAC CATGTTTGA GGTTGAGTGA 1080

## (95) INFORMATION FOR SEQ ID NO:94:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 359 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

20 Met Ile Leu Asn Ser Ser Thr Glu Asp Gly Ile Lys Arg Ile Gln Asp  
 1 5 10 15  
 Asp Cys Pro Lys Ala Gly Arg His Asn Tyr Ile Phe Val Met Ile Pro  
 20 25 30  
 25 Thr Leu Tyr Ser Ile Ile Phe Val Val Gly Ile Phe Gly Asn Ser Leu  
 35 40 45  
 Val Val Ile Val Ile Tyr Phe Tyr Met Lys Leu Lys Thr Val Ala Ser  
 50 55 60  
 Val Phe Leu Leu Asn Leu Ala Leu Ala Asp Leu Cys Phe Leu Leu Thr  
 65 70 75 80  
 30 Leu Pro Leu Trp Ala Val Tyr Thr Ala Met Glu Tyr Arg Trp Pro Phe  
 85 90 95  
 Gly Asn Tyr Leu Cys Lys Ile Ala Ser Ala Ser Val Ser Phe Ala Leu

	100	105	110	
	Tyr Ala Ser Val Phe Leu Leu Thr Cys Leu Ser Ile Asp Arg Tyr Leu			
	115	120	125	
5	Ala Ile Val His Pro Met Lys Ser Arg Leu Arg Arg Thr Met Leu Val			
	130	135	140	
	Ala Lys Val Thr Cys Ile Ile Trp Leu Leu Ala Gly Leu Ala Ser			
	145	150	155	160
10	Leu Pro Ala Ile Ile His Arg Asn Val Phe Phe Glu Asn Thr Asn			
	165	170	175	
10	Ile Thr Val Cys Ala Phe His Tyr Glu Ser Gln Asn Ser Thr Leu Pro			
	180	185	190	
	Ile Gly Leu Gly Leu Thr Lys Asn Ile Leu Gly Phe Leu Phe Pro Phe			
	195	200	205	
15	Leu Ile Ile Leu Thr Ser Tyr Thr Leu Ile Trp Lys Ala Leu Lys Lys			
	210	215	220	
	Ala Tyr Glu Ile Gln Lys Asn Lys Pro Arg Asn Asp Asp Ile Phe Lys			
	225	230	235	240
	Ile Ile Met Ala Ile Val Leu Phe Phe Phe Ser Trp Ile Pro His			
	245	250	255	
20	Gln Ile Phe Thr Phe Leu Asp Val Leu Ile Gln Leu Gly Ile Ile Arg			
	260	265	270	
	Asp Cys Arg Ile Ala Asp Ile Val Asp Thr Ala Met Pro Ile Thr Ile			
	275	280	285	
25	Cys Ile Ala Tyr Phe Asn Asn Cys Leu Asn Pro Leu Phe Tyr Gly Phe			
	290	295	300	
	Leu Gly Lys Phe Lys Arg Tyr Phe Leu Gln Leu Leu Lys Tyr Ile			
	305	310	315	320
	Pro Pro Lys Ala Lys Ser His Ser Asn Leu Ser Thr Lys Met Ser Thr			
	325	330	335	
30	Leu Ser Tyr Arg Pro Ser Asp Asn Val Ser Ser Ser Thr Lys Lys Pro			
	340	345	350	
	Ala Pro Cys Phe Glu Val Glu			
	355			

## (97) INFORMATION FOR SEQ ID NO:95:

- 35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 base pairs  
 (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

CCCAAGCTTC CCCAGGTGTA TTTGAT

26

(97) INFORMATION FOR SEQ ID NO:96:

10

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

CCTGCAGGCG AAACTGACTC TGGCTGAAG

29

20

(98) INFORMATION FOR SEQ ID NO:97:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CTGTACGCTA GTGTGTTCT ACTCACGTGT CTCAGCATTG AT

42

30

(99) INFORMATION FOR SEQ ID NO:98:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

GTTGGATCCA CATAATGCAT TTTCTC

26

(100) INFORMATION FOR SEQ ID NO:99:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1080 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

10

10 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

ATGATTCTCA ACTCTTCTAC TGAAGATGGT ATTAAAAGAA TCCAAGATGA TTGTCCAAA	60
GCTGGAAGGC ATAATTACAT ATTTGTCAATG ATTCCTACTT TATACTGTAT CATCTTTGTG	120
GTGGAAATAT TTGGAAACAG CTTGGTGGTG ATAGTCATTT ACTTTTATAT GAAGCTGAAG	180
15 ACTGTGGCCA GTGTTTTCT TTTGAATTAA GCACTGGCTG ACTTATGCTT TTTACTGACT	240
TTGCCACTAT GGGCTGTCTA CACAGCTATG GAATACCGCT GGCCCTTTGG CAATTACCTA	300
TGTAAGATTG CTTCAGCCAG CGTCAGTTTC AACCTGTACG CTAGTGTGTT TCTACTCACG	360
TGTCTCAGCA TTGATCGATA CCTGGCTATT GTTCACCCAA TGAAGTCCCG CCTTCGACGC	420
ACAATGCTTG TAGCCAAAGT CACCTGCATC ATCATTGGC TGCTGGCAGG CTTGGCCAGT	480
20 TTGCCAGCTA TAATCCATCG AAATGTATTT TTCATTGAGA ACACCAATAT TACAGTTGT	540
GCTTTCCATT ATGAGTCCCA AAATTCAACC CTTCCGATAG GGCTGGGCCT GACCAAAAAT	600
ATACTGGTT TCCTGTTCC TTTCTGATC ATTCTTACAA GTTATTTGG AATTGAAAAA	660
CACTTACTGA AGACGAATAG CTATGGGAAG AACAGGATAA CCCGTGACCA AGTTAAGAAG	720
ATAATTATGG CAATTGTGCT TTTCTTTTC TTTTCCTGGA TTCCCCACCA AATATTCACT	780
25 TTTCTGGATG TATTGATTCA ACTAGGCATC ATACGTGACT GTAGAATTGC AGATATTGTG	840
GACACGGCCA TGCTTATCAC CATTGTATA GCTTATTTA ACAATTGCCT GAATCCTCTT	900
TTTTATGGCT TTCTGGGAA AAAATTAAA AGATATTTC TCCAGCTTCT AAAATATATT	960
CCCCAAAG CCAAATCCCA CTCAAACCTT TCAACAAAAA TGAGCACGCT TTCCTACCGC	1020
CCCTCAGATA ATGTAAGCTC ATCCACCAAG AAGCCTGCAC CATGTTTGA GGTTGAGTGA	1080

20

30

## (101) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 359 amino acids  
 (B) TYPE: amino acid  
 5 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

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

Asp Cys Pro Lys Ala Gly Arg His Asn Tyr Ile Phe Val Met Ile Pro  
 10 20 25 30

Thr Leu Tyr Ser Ile Ile Phe Val Val Gly Ile Phe Gly Asn Ser Leu  
 15 35 40 45

Val Val Ile Val Ile Tyr Phe Tyr Met Lys Leu Lys Thr Val Ala Ser  
 15 50 55 60

Val Phe Leu Leu Asn Leu Ala Ala Asp Leu Cys Phe Leu Leu Thr  
 10 65 70 75 80

Leu Pro Leu Trp Ala Val Tyr Thr Ala Met Glu Tyr Arg Trp Pro Phe  
 20 85 90 95

Gly Asn Tyr Leu Cys Lys Ile Ala Ser Ala Ser Val Ser Phe Asn Leu  
 100 105 110

Tyr Ala Ser Val Phe Leu Leu Thr Cys Leu Ser Ile Asp Arg Tyr Leu  
 115 120 125

25 Ala Ile Val His Pro Met Lys Ser Arg Leu Arg Arg Thr Met Leu Val  
 130 135 140

Ala Lys Val Thr Cys Ile Ile Trp Leu Leu Ala Gly Leu Ala Ser  
 145 150 155 160

Leu Pro Ala Ile Ile His Arg Asn Val Phe Phe Ile Glu Asn Thr Asn  
 30 165 170 175

Ile Thr Val Cys Ala Phe His Tyr Glu Ser Gln Asn Ser Thr Leu Pro  
 180 185 190

Ile Gly Leu Gly Leu Thr Lys Asn Ile Leu Gly Phe Leu Phe Pro Phe  
 195 200 205

35 Leu Ile Ile Leu Thr Ser Tyr Phe Gly Ile Arg Lys His Leu Leu Lys  
 210 215 220

10

20

30

Thr Asn Ser Tyr Gly Lys Asn Arg Ile Thr Arg Asp Gln Val Lys Lys  
 225 230 235 240  
 Ile Ile Met Ala Ile Val Leu Phe Phe Phe Ser Trp Ile Pro His  
 245 250 255  
 5 Gln Ile Phe Thr Phe Leu Asp Val Leu Ile Gln Leu Gly Ile Ile Arg  
 260 265 270  
 Asp Cys Arg Ile Ala Asp Ile Val Asp Thr Ala Met Pro Ile Thr Ile  
 275 280 285  
 10 Cys Ile Ala Tyr Phe Asn Asn Cys Leu Asn Pro Leu Phe Tyr Gly Phe  
 290 295 300  
 Leu Gly Lys Phe Lys Arg Tyr Phe Leu Gln Leu Leu Lys Tyr Ile  
 305 310 315 320  
 Pro Pro Lys Ala Lys Ser His Ser Asn Leu Ser Thr Lys Met Ser Thr  
 325 330 335  
 15 Leu Ser Tyr Arg Pro Ser Asp Asn Val Ser Ser Ser Thr Lys Lys Pro  
 340 345 350  
 Ala Pro Cys Phe Glu Val Glu  
 355

(102) INFORMATION FOR SEQ ID NO:101: 20

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 37 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: DNA (genomic)  
 (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:  
 TCCGAATTCC AAAATAACTT GTAAGAATGA TCAGAAA 37 30

(103) INFORMATION FOR SEQ ID NO:102:  
 30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 33 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: DNA (genomic)  
 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

AGATCTTAAG AAGATAATTA TGGCAATTGT GCT 33

(104) INFORMATION FOR SEQ ID NO:103:

- 5 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 62 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic) 10

10 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

AATTCGAAAA CACTTACTGA AGACGAATAG CTATGGGAAG AACAGGATAA CCCGTGACCA 60

AG 62

(105) INFORMATION FOR SEQ ID NO:104:

- 15 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 62 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear 20

20 (ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

TTAACTTGGT CACGGGTTAT CCTGTTCTTC CCATAGCTAT TCGTCTTCAG TAAGTGT 60

CG 62

25 (106) INFORMATION FOR SEQ ID NO:105:

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1083 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear 30

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

ATGATTCTCA ACTCTTCTAC TGAAGATGGT ATTAAGAGAA TCCAAGATGA TTGTCCAAA 60  
 GCTGGAAGGC ATAATTACAT ATTTGTCATG ATTCCTACTT TATACAGTAT CATCTTGTT 120  
 GTGGAAATAT TTGGAAACAG CTTGGTGGTG ATAGTCATTT ACTTTATAT GAAGCTGAAG 180  
 ACTGTGGCCA GTGTTTTCT TTTGAATTAA GCACTGGCTG ACTTATGCTT TTTACTGACT 240  
 5 TTGCCACTAT GGGCTGTCTA CACAGCTATG GAATACCGCT GGCCCTTGG CAATTACCTA 300  
 TGTAAGATTG CTTCAGCCAG CGTCAGTTTC AACCTGTACG CTAGTGTGTT TCTACTCACG 360  
 TGTCTCAGCA TTGATCGATA CCTGGCTATT GTTCACCCAA TGAAGTCCCG CCTTCGACGC 420 10  
 ACAATGCTTG TAGCCAAAGT CACCTGCATC ATCATTGTC TGCTGGCAGG CTTGGCCAGT 480  
 TTGCCAGCTA TAATCCATCG AAATGTATTT TTCATTGAGA ACACCAATAT TACAGTTGT 540  
 10 GCTTCCATT ATGAGTCCCA AAATTCAACC CTTCCGATAG GGCTGGGCCT GACCAAAAT 600  
 ATACTGGGTT TCCTGTTCC TTTCTGATC ATTCTTACAA GTTATACTCT TATTTGGAAG 660  
 GCCCTAAAGA AGGCTTATGA AATTCAAGG AACAAACCAA GAAATGATGA TATTTTTAAG 720  
 ATAATTATGG CAGCAATTGT GCTTTCTTT TTCTTTCTT GGATTCCCA CCAAATATTC 780  
 ACTTTCTGG ATGTATTGAT TCAACTAGGC ATCATACTGT ACTGTAGAAT TGCAGATATT 840  
 15 GTGGACACGG CCATGCCTAT CACCATTGT ATAGCTTATT TTAACAATTG CCTGAATCCT 900  
 CTTTTTATG GCTTCTGGG GAAAAAATT AAAAGATATT TTCTCCAGCT TCTAAAATAT 960  
 ATTCCCCCAA AAGCCAAATC CCACTCAAAC CTTCAACAA AAATGAGCAC GCTTTCTAC 1020  
 CGCCCCCTCAG ATAATGTAAG CTCATCCACC AAGAAGCCTG CACCATGTTT TGAGGTTGAG 1080  
 TGA 1083

20 (107) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 360 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 25 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

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

30 Asp Cys Pro Lys Ala Gly Arg His Asn Tyr Ile Phe Val Met Ile Pro

	20	25	30	
35	Thr Leu Tyr Ser Ile Ile Phe Val Val Gly Ile Phe Gly Asn Ser Leu	40	45	
5	50 Val Val Ile Val Ile Tyr Phe Tyr Met Lys Leu Lys Thr Val Ala Ser	55	60	
65	70 Val Phe Leu Leu Asn Leu Ala Leu Ala Asp Leu Cys Phe Leu Leu Thr	75	80	
85	80 Leu Pro Leu Trp Ala Val Tyr Thr Ala Met Glu Tyr Arg Trp Pro Phe	90	95	10
10	100 Gly Asn Tyr Leu Cys Lys Ile Ala Ser Ala Ser Val Ser Phe Asn Leu	105	110	
115	120 Tyr Ala Ser Val Phe Leu Leu Thr Cys Leu Ser Ile Asp Arg Tyr Leu	125		
15	130 Ala Ile Val His Pro Met Lys Ser Arg Leu Arg Arg Thr Met Leu Val	135	140	
145	150 Ala Lys Val Thr Cys Ile Ile Trp Leu Leu Ala Gly Leu Ala Ser	155	160	
165	170 Leu Pro Ala Ile Ile His Arg Asn Val Phe Phe Ile Glu Asn Thr Asn	175		20
20	180 Ile Thr Val Cys Ala Phe His Tyr Glu Ser Gln Asn Ser Thr Leu Pro	185	190	
195	200 Ile Gly Leu Gly Leu Thr Lys Asn Ile Leu Gly Phe Leu Phe Pro Phe	205		
25	210 Leu Ile Ile Leu Thr Ser Tyr Thr Leu Ile Trp Lys Ala Leu Lys Lys	215	220	
225	230 Ala Tyr Glu Ile Gln Lys Asn Lys Pro Arg Asn Asp Asp Ile Phe Lys	235	240	
245	250 Ile Ile Met Ala Ala Ile Val Leu Phe Phe Phe Ser Trp Ile Pro	255		30
30	260 His Gln Ile Phe Thr Phe Leu Asp Val Leu Ile Gln Leu Gly Ile Ile	265	270	
275	280 Arg Asp Cys Arg Ile Ala Asp Ile Val Asp Thr Ala Met Pro Ile Thr	285		
290	295 Ile Cys Ile Ala Tyr Phe Asn Asn Cys Leu Asn Pro Leu Phe Tyr Gly	300		
35	305 Phe Leu Gly Lys Lys Phe Lys Arg Tyr Phe Leu Gln Leu Leu Lys Tyr	310	315	320

Ile Pro Pro Lys Ala Lys Ser His Ser Asn Leu Ser Thr Lys Met Ser  
325 330 335

Thr Leu Ser Tyr Arg Pro Ser Asp Asn Val Ser Ser Ser Thr Lys Lys  
340 345 350

5 Pro Ala Pro Cys Phe Glu Val Glu  
355 360

(108) INFORMATION FOR SEQ ID NO:107:

- (i) SEQUENCE CHARACTERISTICS:  
10 (A) LENGTH: 26 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

CCCAAGCTTC CCCAGGTGTA TTTGAT

26

(109) INFORMATION FOR SEQ ID NO:108:

- 20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 38 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

AAGCACAAATT GCTGCATAAT TATCTTAAAA ATATCATC

38

30

(110) INFORMATION FOR SEQ ID NO:109:

- 30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 39 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

AAGATAATTA TGGCAGCAAT TGTGCTTTTC TTTTTCTTT

39

(111) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 26 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

10 (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

GTTGGATCCA CATAATGCAT TTTCTC

26

(112) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 1344 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: DNA (genomic)

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

ATGGAGCTGC TAAAGCTGAA CCGGAGCGTG CAGGGAACCG GACCCGGGCC GGGGGCTTCC

60

CTGTGCCGCC CGGGGGCGCC TCTCCTCAAC AGCAGCAGTG TGGGCAACCT CAGCTGCGAG

120

CCCCCTCGCA TTCGGGGAGC CGGGACACGA GAATTGGAGC TGGCCATTAG AATCACTCTT

180

TACGCAGTGA TCTTCCTGAT GAGCGTTGGA GGAAATATGC TCATCATCGT GGTCCCTGGGA

240

25 CTGAGCCGCC GCCTGAGGAC TGTCACCAAT GCCTTCCTCC TCTCACTGGC AGTCAGCGAC

30

CTCCTGCTGG CTGTGGCTTG CATGCCCTTC ACCCTCCTGC CCAATCTCAT GGGCACATTG

360

ATCTTTGGCA CCGTCATCTG CAAGGCGGTT TCCTACCTCA TGGGGGTGTC TGTGAGTGTG

420

TCCACGCTAA GCCTCGTGGC CATGGCACTG GAGCGATATA GCGCCATCTG CCGACCAC TG

480

CAGGCACGAG TGTGGCAGAC GCGCTCCCAC GCGGCTCGCG TGATTGTAGC CACGTGGCTG

540

30 CTGTCCGGAC TACTCATGGT GCCCTACCCC GTGTACACTG TCGTGCACCC AGTGGGGCCT

600

CGTGTGCTGC AGTGCCTGCA TCGCTGGCCC AGTGCCTGGG TCCGCCAGAC CTGGTCCGTA

660

CTGCTGCTTC TGCTCTTGTGTT CTTCATCCCCA GGTGTGGTTA TGGCCGTGGC CTACGGGCTT 720  
 ATCTCTCGCG AGCTCTACTT AGGGCTTCGC TTTGACGGCG ACAGTGACAG CGACAGCCAA 780  
 AGCAGGGTCC GAAACCAAGG CGGGCTGCCA GGGGCTGTTG ACCAGAACGG GCGTTGCCGG 840  
 CCTGAGACTG GCGCGGTTGG CAAAGACAGC GATGGCTGCT ACGTGCAACT TCCACGTTCC 900  
 5 CGGCCTGCCCG TGGAGCTGAC GGCCTGACG GCTCCTGGC CGGGATCCGG CTCCCGGGCC 960  
 ACCCAGGCCA AGCTGCTGGC TAAGAACGCG GTGAAACGAA TGTTGCTGGT GATCGTTGTG 1020  
 CTTTTTTTTC TGTGTTGGTT GCCAGTTTAT AGTGCCAACA CGTGGCGCGC CTTTGATGGC 1080 10  
 CGGGGTGCAC ACCGAGCACT CTCGGGTGCT CCTATCTCCT TCATTCACTT GCTGAGCTAC 1140  
 GCCTCGGCCT GTGTCAACCC CCTGGTCTAC TGCTTCATGC ACCGTCGCTT TCGCCAGGCC 1200  
 10 TGCCTGGAAA CTTGCGCTCG CTGCTGCCCG CGGCCTCCAC GAGCTGCCCG CAGGGCTCTT 1260  
 CCCGATGAGG ACCCTCCAC TCCCTCCATT GCTTCGCTGT CCAGGCTTAG CTACACCACC 1320  
 ATCAGCACAC TGGGCCCTGG CTGA 1344

## (113) INFORMATION FOR SEQ ID NO:112:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 447 amino acids 20  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

## 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Met Glu Leu Leu Lys Leu Asn Arg Ser Val Gln Gly Thr Gly Pro Gly  
 1 5 10 15

Pro Gly Ala Ser Leu Cys Arg Pro Gly Ala Pro Leu Leu Asn Ser Ser  
 20 25 30

25 Ser Val Gly Asn Leu Ser Cys Glu Pro Pro Arg Ile Arg Gly Ala Gly  
 35 40 45

Thr Arg Glu Leu Glu Leu Ala Ile Arg Ile Thr Leu Tyr Ala Val Ile  
 50 55 60

30 Phe Leu Met Ser Val Gly Gly Asn Met Leu Ile Ile Val Val Leu Gly  
 65 70 75 80

Leu Ser Arg Arg Leu Arg Thr Val Thr Asn Ala Phe Leu Leu Ser Leu  
 85 90 95

Ala Val Ser Asp Leu Leu Leu Ala Val Ala Cys Met Pro Phe Thr Leu  
 100 105 110  
 Leu Pro Asn Leu Met Gly Thr Phe Ile Phe Gly Thr Val Ile Cys Lys  
 115 120 125  
 5 Ala Val Ser Tyr Leu Met Gly Val Ser Val Ser Val Ser Thr Leu Ser  
 130 135 140  
 Leu Val Ala Ile Ala Leu Glu Arg Tyr Ser Ala Ile Cys Arg Pro Leu  
 145 150 155 160  
 10 Gln Ala Arg Val Trp Gln Thr Arg Ser His Ala Ala Arg Val Ile Val  
 165 170 175  
 Ala Thr Trp Leu Leu Ser Gly Leu Leu Met Val Pro Tyr Pro Val Tyr  
 180 185 190  
 Thr Val Val Gln Pro Val Gly Pro Arg Val Leu Gln Cys Val His Arg  
 195 200 205  
 15 Trp Pro Ser Ala Arg Val Arg Gln Thr Trp Ser Val Leu Leu Leu  
 210 215 220  
 Leu Leu Phe Phe Ile Pro Gly Val Val Met Ala Val Ala Tyr Gly Leu  
 225 230 235 240  
 Ile Ser Arg Glu Leu Tyr Leu Gly Leu Arg Phe Asp Gly Asp Ser Asp  
 20 245 250 255  
 Ser Asp Ser Gln Ser Arg Val Arg Asn Gln Gly Gly Leu Pro Gly Ala  
 260 265 270  
 Val His Gln Asn Gly Arg Cys Arg Pro Glu Thr Gly Ala Val Gly Lys  
 275 280 285  
 25 Asp Ser Asp Gly Cys Tyr Val Gln Leu Pro Arg Ser Arg Pro Ala Leu  
 290 295 300  
 Glu Leu Thr Ala Leu Thr Ala Pro Gly Pro Gly Ser Gly Ser Arg Pro  
 305 310 315 320  
 Thr Gln Ala Lys Leu Leu Ala Lys Lys Arg Val Lys Arg Met Leu Leu  
 320 325 330 335  
 30 Val Ile Val Val Leu Phe Phe Leu Cys Trp Leu Pro Val Tyr Ser Ala  
 340 345 350  
 Asn Thr Trp Arg Ala Phe Asp Gly Pro Gly Ala His Arg Ala Leu Ser  
 355 360 365  
 35 Val Ala Pro Ile Ser Phe Ile His Leu Leu Ser Tyr Ala Ser Ala Cys  
 370 375 380  
 Val Asn Pro Leu Val Tyr Cys Phe Met His Arg Arg Phe Arg Gln Ala

385 390 395 400  
Cys Leu Glu Thr Cys Ala Arg Cys Cys Pro Arg Pro Pro Arg Ala Arg  
405 410 415  
Pro Arg Ala Leu Pro Asp Glu Asp Pro Pro Thr Pro Ser Ile Ala Ser  
5 420 425 430  
Leu Ser Arg Leu Ser Tyr Thr Thr Ile Ser Thr Leu Gly Pro Gly  
435 440 445

## (114) INFORMATION FOR SEQ ID NO:113:

- 10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

## 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

CAGCAGCATG CGCTTCACGC GCTTCTTAGC CCAG 34

## (115) INFORMATION FOR SEQ ID NO:114:

- 20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

25 AGAAGCGCGT GAAGCGCATG CTGCTGGTGA TCGTT 35

## (116) INFORMATION FOR SEQ ID NO:115:

- 30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

ATGGAGAAAA GAATCAAAAG AATGTTCTAT ATA

33

(117) INFORMATION FOR SEQ ID NO:116:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

TATATAGAAC ATTCTTTGA TTCTTTCTC CAT

33

(118) INFORMATION FOR SEQ ID NO:117:

- 15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 30 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

CGCTCTCTGG CCTTGAAGCG CACGCTCAGC

30

(119) INFORMATION FOR SEQ ID NO:118:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 30 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

GCTGAGCGTG CGCTTCAAGG CCAGAGAGCG

30

(120) INFORMATION FOR SEQ ID NO:119:

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 30 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

10 CCCAGGAAAA AGGTGAAAGT CAAAGTTTTC 30

10 (121) INFORMATION FOR SEQ ID NO:120:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 30 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

20 GAAAACTTTG ACTTTCACCT TTTTCCTGGG 30

20 (122) INFORMATION FOR SEQ ID NO:121:

25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

30 GGGGCGCGGG TGAAACGGCT GGTGAGC 27

30 (123) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

5 GCTCACCAAGC CGTTTCACCC GCGCCCC 27

(124) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

10 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

15 CCCCTTGAAA AGCCTAAGAA CTTGGTCATC 30 20

(125) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

20 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

25 GATGACCAAG TTCTTAGGCT TTTCAAGGGG 30

(126) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

30 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

GATCTCTAGA ATGAAACAGCA CATGTATTGA AG 32

(127) INFORMATION FOR SEQ ID NO:126:

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 35 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

CTAGGGTACC CGCTCAAGGA CCTCTAATTC CATA 35

(128) INFORMATION FOR SEQ ID NO:127:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1296 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

ATGCAGGCGC TTAACATTAC CCCGGAGCAG TTCTCTCGGC TGCTGCGGGA CCACAAACCTG 60

ACGCAGGAGC AGTTCATCGC TCTGTACCGG CTGCGACCGC TCGTCTACAC CCCAGAGCTG 120

CCGGGACGCG CCAAGCTGGC CCTCGTGCTC ACCGGCGTGC TCATTTCGC CCTGGCGCTC 180

25 TTTGGCAATG CTCTGGTGTG CTACGTGGTG ACCCGCAGCA AGGCCATGCG CACCGTCACC 240

AACATCTTA TCTGCTCCTT GGCAGTCAGT GACCTGCTCA TCACCTTCTT CTGCATTCCC 300

GTCACCATGC TCCAGAACAT TTCCGACAAAC TGGCTGGGG GTGCTTCAT TTGCAAGATG 360

GTGCCATTG TCCAGTCTAC CGCTGTTGTG ACAGAAATGC TCACTATGAC CTGCATTGCT 420

GTGGAAAGGC ACCAGGGACT TGTGCATCCT TTTAAAATGA AGTGGCAATA CACCAACCGA 480

10

20

30

AGGGCTTC	CAATGCTAGG	TGTGGCTGG	CTGGTGGCAG	TCATCGTAGG	ATCACCCATG	540	
TGGCACGTG	AAACAACCTG	GATCAAATAT	GACTTCCTAT	ATGAAAAGGA	ACACATCTGC	600	
TGCTTAGAAG	AGTGGACCA	CCCTGTGCAC	CAGAAGATCT	ACACCACCTT	CATCCTTGTC	660	
ATCCTCTTCC	TCCTGCCTCT	TATGGTGATG	CTTATTCTGT	ACAGTAAAAT	TGGTTATGAA	720	
5	CTTTGGATAA	AGAAAAGAGT	TGGGGATGGT	TCAGTGCTTC	GAACATTCA	TGGAAAAGAA	780
ATGTCCAAAA	TAGCCAGGAA	GAAGAAACGA	GCTAAGATTA	TGATGGTGAC	AGTGGTGGCT	840	
CTCTTTGCTG	TGTGCTGGGC	ACCATTCCAT	GTTGTCCATA	TGATGATTGA	ATACAGTAAT	900	
10	TTTGAAAAGG	AATATGATGA	TGTCACAATC	AAGATGATTT	TTGCTATCGT	GCAAATTATT	960
GGATTTCCA	ACTCCATCTG	TAATCCCATT	GTCTATGCAT	TTATGAATGA	AAACTTCAAA	1020	
10	AAAAATGTTT	TGTCTGCAGT	TTGTTATTGC	ATAGTAAATA	AAACCTTCTC	TCCAGCACAA	1080
AGGCATGGAA	ATTCAAGGAAT	TACAATGATG	CGGAAGAAAG	CAAAGTTTTC	CCTCAGAGAG	1140	
AATCCAGTGG	AGGAAACCAA	AGGAGAACGA	TTCAGTGATG	GCAACATTGA	AGTCAAATTG	1200	
TGTGAACAGA	CAGAGGGAGAA	GAAAAAGCTC	AAACGACATC	TTGCTCTCTT	TAGGTCTGAA	1260	
CTGGCTGAGA	ATTCTCCTTT	AGACAGTGGG	CATTAAC			1296	
15	(129) INFORMATION FOR SEO ID NO:128:					20	

15 (129) INFORMATION FOR SEQ ID NO:128:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 431 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

20

25 Asp His Asn Leu Thr Arg Glu Gln Phe Ile Ala Leu Tyr Arg Leu Arg

Pro Leu Val Tyr Thr Pro Glu Leu Pro Gly Arg Ala Lys Leu Ala Leu

Val Leu Thr Gly Val Leu Ile Phe Ala Leu Ala Leu Phe Gly Asn Ala  
36 50 54 58 62

Leu Val Phe Tyr Val Val Thr Arg Ser Lys Ala Met Arg Thr Val Thr  
65 70 75 80

Asn Ile Phe Ile Cys Ser Leu Ala Leu Ser Asp Leu Leu Ile Thr Phe  
 85 90 95  
 Phe Cys Ile Pro Val Thr Met Leu Gln Asn Ile Ser Asp Asn Trp Leu  
 100 105 110  
 5 Gly Gly Ala Phe Ile Cys Lys Met Val Pro Phe Val Gln Ser Thr Ala  
 115 120 125  
 Val Val Thr Glu Met Leu Thr Met Thr Cys Ile Ala Val Glu Arg His  
 130 135 140  
 Gln Gly Leu Val His Pro Phe Lys Met Lys Trp Gln Tyr Thr Asn Arg  
 10 145 150 155 160  
 Arg Ala Phe Thr Met Leu Gly Val Val Trp Leu Val Ala Val Ile Val  
 165 170 175  
 Gly Ser Pro Met Trp His Val Gln Gln Leu Glu Ile Lys Tyr Asp Phe  
 180 185 190  
 15 Leu Tyr Glu Lys Glu His Ile Cys Cys Leu Glu Glu Trp Thr Ser Pro  
 195 200 205  
 Val His Gln Lys Ile Tyr Thr Thr Phe Ile Leu Val Ile Leu Phe Leu  
 210 215 220  
 Leu Pro Leu Met Val Met Leu Ile Leu Tyr Ser Lys Ile Gly Tyr Glu  
 20 225 230 235 240  
 Leu Trp Ile Lys Lys Arg Val Gly Asp Gly Ser Val Leu Arg Thr Ile  
 245 250 255  
 His Gly Lys Glu Met Ser Lys Ile Ala Arg Lys Lys Lys Arg Ala Lys  
 260 265 270  
 25 Ile Met Met Val Thr Val Val Ala Leu Phe Ala Val Cys Trp Ala Pro  
 275 280 285  
 Phe His Val Val His Met Met Ile Glu Tyr Ser Asn Phe Glu Lys Glu  
 290 295 300  
 Tyr Asp Asp Val Thr Ile Lys Met Ile Phe Ala Ile Val Gln Ile Ile  
 30 305 310 315 320  
 Gly Phe Ser Asn Ser Ile Cys Asn Pro Ile Val Tyr Ala Phe Met Asn  
 325 330 335  
 Glu Asn Phe Lys Lys Asn Val Leu Ser Ala Val Cys Tyr Cys Ile Val  
 340 345 350  
 35 Asn Lys Thr Phe Ser Pro Ala Gln Arg His Gly Asn Ser Gly Ile Thr  
 355 360 365  
 Met Met Arg Lys Lys Ala Lys Phe Ser Leu Arg Glu Asn Pro Val Glu

370 375 380

Glu Thr Lys Gly Glu Ala Phe Ser Asp Gly Asn Ile Glu Val Lys Leu  
385 390 395 400

Cys Glu Gln Thr Glu Glu Lys Lys Leu Lys Arg His Leu Ala Leu  
5 405 410 415

Phe Arg Ser Glu Leu Ala Glu Asn Ser Pro Leu Asp Ser Gly His  
420 425 430

(130) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 2040 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

ATGGGCAGCC CCTGGAACGG CAGCGACGGC CCCGAGGGGG CGCGGGAGCC GCCGTGGCCC  
60

20 GCGCTGCCGC CTTGCGACGA CGCGCGCTGC TCGCCCTTTC CCCTGGGGGC GCTGGTGCGC  
120

25 GTGACCGCTG TGTGCCTGTG CCTGTTCGTC GTCGGGGTGA GCGGCAACGT GGTGACCGTG  
180

ATGCTGATCG GGCGCTACCG GGACATGCGG ACCACCACCA ACTTGTACCT GGGCAGCATG  
240

25 GCGGTGTCG ACCTACTCAT CCTGCTCGGG CTGCCGTTCG ACCTGTACCG CCTCTGGCGC  
300

TCGCGGCCCT GGGTGTTCGG GCCGCTGCTC TGCCGCCTGT CCCTCTACGT GGGCGAGGGC  
360

30 TGCACCTACG CCACGCTGCT GCACATGACC GCGCTCAGCG TCGAGCGCTA CCTGGCCATC  
420

TGCCGCCCGC TCCGCGCCCG CGTCTTGGTC ACCCGGCGCC GCGTCCGCGC GCTCATCGCT  
480

35 GTGCTCTGGG CCGTGGCGCT GCTCTCTGCC GGTCCCTTCT TGTTCTGGT GGGCGTCGAG  
540

CAGGACCCCG GCATCTCCGT AGTCCCGGGC CTCAAATGGCA CGCGCGGGAT CGCCTCCTCG  
600

40 CCTCTCGCCT CGTCGCCGCC TCTCTGGCTC TCGCGGGCGC CACCGCCGTC CCCGCCGTCG

660  
GGGCCGGAGA CGCGGGAGGC CGCGGCGCTG TTCAGCCGCG AATGCCGGCC GAGCCCCGCG  
720  
5 CAGCTGGCG CGCTGCGTGT CATGCTGTGG GTCACCACCG CCTACTTCTT CCTGCCCTTT  
780  
CTGTGCCTCA GCATCCTCTA CGGGCTCATC GGGCGGGAGC TGTGGAGCAG CGGGCGGCCG  
10 840  
CTGCGAGGCC CGGCCGCCCTC GGGGCGGGAG AGAGGCCACC GGCAGACCAA ACGCGTCCTG  
900  
15 CGTAAGTGA GCGCCCGTGG TTCCAAAGAC GCCTGCCCTGC AGTCCGCCCC GCGGGGGACC  
960  
GCGCAAACGC TGGGTCCCCT TCCCCTGCTC GCCCAGCTCT GGGGCCGCT TCCAGCTCCC  
1020  
20 TTTCCCTATTT CGATTCCAGC CTCCACCCGC CGGTACTTCC CATCCCCCGA GAAAACCATG  
1080  
TCCTGTCCCC CAGGAGCTCT GGGGGACCCC AGGGCGCTT GAGGGTGGGA TCCCCGGATC  
25 1140  
CGATTCAAGTA ACCAGCAGTG CTTTCCAGA GCCTCTGAGA CCAGAAAGGA GAGTTGGTAA  
1200  
30 TTCTTAATCC AACACACGT TAGATGCCAC AAATGAGGAG TCCTCACAGT GCTCTTGAGA  
1260  
AGACGAGGGA GATTCAATTA AGCTAAAATT TTTTATTAA TGTTAAGTGA TGCTGAAGGC  
1320  
35 TAAAGTAAAC CTTGCTCGTA TCAAAAAGTA AAGATTGTGC AGACCTGTTG TAGAATTCTT  
1380  
TTCAACAGAG AACAGAAAAC TTGTCTCCGA AGTGGGTTG TGGAAGGAAG CCTGCCAAGG  
40 1440  
CGGCTTGTTC AGAGAAATTG CTCCTCTGG TTTATGTCCA GCCTTGATAA CACATATGGG  
1500  
45 AGCCTACTAT GCAGTTTAA AGCAAGTATC CATGCAGCCT GCAGCCTGGT CATTTCCT  
1560  
GGGTGAGGA TCTGCCTAGG TAGAAGTTTT CTCTAATTAA TTTTGCTGTT ACTTGTTATT  
1620  
50 GCAGATGGTT CCTTGTCGGG GTGGGGGGTT TATTTGCTTC CCAATGCTTT TGTTAATCCC  
1680  
GGTGCTGTGT CTTATGTTGC AGTGGTGGTG GTTCTGGCAT TTATAATTG CTGGTTGCC  
55 1740

TTCCACGTTG GCAGAACAT TTACATAAAC ACGGAAGATT CGCGGATGAT GTACTTCTCT  
1800

5 CAGTACTTTA ACATCGTCGC TCTGCAACTT TTCTATCTGA GCGCATCTAT CAACCCAATC  
1860

CTCTACAAACC TCATTCAAA GAAGTACAGA GCGGCGGCCT TTAAACTGCT GCTCGCAAGG  
1920

10 AAGTCCAGGC CGAGAGGCTT CCACAGAACG AGGGACACTG CGGGGGAAGT TGCAGGGAC  
1980

15 ACTGGAGGAG ACACGGTGGG CTACACCGAG ACAAGCGCTA ACGTGAAGAC GATGGGATAA  
2040

10

(131) INFORMATION FOR SEQ ID NO:130:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 412 amino acids
  - (B) TYPE: amino acid
  - 20 (C) STRANDEDNESS:
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

20

25 Met Gly Ser Pro Trp Asn Gly Ser Asp Gly Pro Glu Gly Ala Arg Glu  
1 5 10 15

Pro Pro Trp Pro Ala Leu Pro Pro Cys Asp Glu Arg Arg Cys Ser Pro  
20 25 30

Phe Pro Leu Gly Ala Leu Val Pro Val Thr Ala Val Cys Leu Cys Leu  
35 40 45

30 Phe Val Val Gly Val Ser Gly Asn Val Val Thr Val Met Leu Ile Gly  
50 55 60

Arg Tyr Arg Asp Met Arg Thr Thr Asn Leu Tyr Leu Gly Ser Met  
65 70 75 80

30

35 Ala Val Ser Asp Leu Leu Ile Leu Gly Leu Pro Phe Asp Leu Tyr  
85 90 95

Arg Leu Trp Arg Ser Arg Pro Trp Val Phe Gly Pro Leu Leu Cys Arg  
100 105 110

Leu Ser Leu Tyr Val Gly Glu Gly Cys Thr Tyr Ala Thr Leu Leu His  
115 120 125

40 Met Thr Ala Leu Ser Val Glu Arg Tyr Leu Ala Ile Cys Arg Pro Leu  
130 135 140

Arg Ala Arg Val Leu Val Thr Arg Arg Arg Val Arg Ala Leu Ile Ala			
145	150	155	160
Val Leu Trp Ala Val Ala Leu Leu Ser Ala Gly Pro Phe Leu Phe Leu			
165	170	175	
5 Val Gly Val Glu Gln Asp Pro Gly Ile Ser Val Val Pro Gly Leu Asn			
180	185	190	
Gly Thr Ala Arg Ile Ala Ser Ser Pro Leu Ala Ser Ser Pro Pro Leu			
195	200	205	
10 Trp Leu Ser Arg Ala Pro Pro Pro Ser Pro Pro Ser Gly Pro Glu Thr			
210	215	220	
Ala Glu Ala Ala Ala Leu Phe Ser Arg Glu Cys Arg Pro Ser Pro Ala			
225	230	235	240
Gln Leu Gly Ala Leu Arg Val Met Leu Trp Val Thr Thr Ala Tyr Phe			
245	250	255	
15 Phe Leu Pro Phe Leu Cys Leu Ser Ile Leu Tyr Gly Leu Ile Gly Arg			
260	265	270	
Glu Leu Trp Ser Ser Arg Arg Pro Leu Arg Gly Pro Ala Ala Ser Gly			
275	280	285	
Arg Glu Arg Gly His Arg Gln Thr Lys Arg Val Leu Leu Val Val Val			
290	295	300	
20 305 Leu Ala Phe Ile Ile Cys Trp Leu Pro Phe His Val Gly Arg Ile Ile			
310	315	320	
Tyr Ile Asn Thr Glu Asp Ser Arg Met Met Tyr Phe Ser Gln Tyr Phe			
325	330	335	
25 Asn Ile Val Ala Leu Gln Leu Phe Tyr Leu Ser Ala Ser Ile Asn Pro			
340	345	350	
Ile Leu Tyr Asn Leu Ile Ser Lys Lys Tyr Arg Ala Ala Ala Phe Lys			
355	360	365	
30 30 Leu Leu Leu Ala Arg Lys Ser Arg Pro Arg Gly Phe His Arg Ser Arg			
370	375	380	
Asp Thr Ala Gly Glu Val Ala Gly Asp Thr Gly Gly Asp Thr Val Gly			
385	390	395	400
Tyr Thr Glu Thr Ser Ala Asn Val Lys Thr Met Gly			
405	410		
35 (132) INFORMATION FOR SEQ ID NO:131:			
(i) SEQUENCE CHARACTERISTICS:			
(A) LENGTH: 1344 base pairs			

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

ATGGAGCTGC TAAAGCTGAA CGGGAGCGTG CAGGGAACCG GACCCGGGCC GGGGGCTTCC  
60

CTGTGCCGCC CGGGGGCGCC TCTCCTCAAC AGCAGCAGTG TGGGCAACCT CAGCTGCGAG  
120 10

10 CCCCCCTCGCA TTCGCGGAGC CGGGACACGA GAATTGGAGC TGGCCATTAG AATCACTCTT  
180

TACGCAGTGA TCTTCCTGAT GAGCGTTGGA GGAAATATGC TCATCATCGT GGTCTGGGA  
240

15 CTGAGCCGCC GCCTGAGGAC TGTCACCAAT GCCTTCCTCC TCTCACTGGC AGTCAGCGAC  
300

CTCCTGCTGG CTGTGGCTTG CATGCCCTTC ACCCTCCTGC CCAATCTCAT GGGCACATT  
360

ATCTTTGGCA CCGTCATCTG CAAGGCGGTT TCCTACCTCA TGGGGGTGTC TGTGAGTGTG  
420 20

20 TCCACGCTAA GCCTCGTGGC CATCGCACTG GAGCGATATA GCGCCATCTG CCGACCACTG  
480

CAGGCACGAG TGTGGCAGAC GCGCTCCCAC GCGGCTCGCG TGATTGTAGC CACGTGGCTG  
540

25 CTGTCCGGAC TACTCATGGT GCCCTACCCC GTGTACACTG TCGTGCAACC AGTGGGGCCT  
600

CGTGTGCTGC AGTGCCTGCA TCGCTGGCCC AGTGCCTGGGG TCCGCCAGAC CTGGTCCGTA  
660

CTGCTGCTTC TGCTCTTGTGTT CTTCATCCCA GGTGTGGTTA TGGCCGTGGC CTACGGGCTT  
720 30

30 ATCTCTCGCG AGCTCTACTT AGGGCTTCGC TTTGACGGCG ACAGTGACAG CGACAGCCAA  
780

AGCAGGGTCC GAAACCAAGG CGGGCTGCCA GGGGCTGTTG ACCAGAACGG GCGTTGCCGG  
840

35 CCTGAGACTG GCGCGGTTGG CAAAGACAGC GATGGCTGCT ACGTGCAACT TCCACGTTCC  
900

CGGCCTGCCCG TGGAGCTGAC GGCGCTGACG GCTCCTGGC CGGGATCCGG CTCCCGGCC

960

ACCCAGGCCA AGCTGCTGGC TAAGAAGCGC GTGAAACGAA TGTTGCTGGT GATCGTTGTG  
1020

CTTTTTTTC TGTGTTGGTT GCCAGTTAT AGTGCCAACA CGTGGCGCGC CTTTGATGGC  
5 1080

CCGGGTGCAC ACCGAGCACT CTCGGGTGCT CCTATCTCCT TCATTCACCT GCTGAGCTAC  
1140

GCCTCGGCCT GTGTCAACCC CCTGGTCTAC TGCTTCATGC ACCGTCGCTT TCGCCAGGCC  
1200

10 TGCCTGGAAA CTTGGCGCTCG CTGCTGCCAC CGGCCTCCAC GAGCTCGCCC CAGGGCTCTT  
1260

CCCGATGAGG ACCCTCCCAC TCCCTCCATT GCTTCGCTGT CCAGGCTTAG CTACACCACC  
1320

15 ATCAGCACAC TGGGCCCTGG CTGA  
1344

(133) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 447 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

25 Met Glu Leu Leu Lys Leu Asn Arg Ser Val Gln Gly Thr Gly Pro Gly  
1 5 10 15

Pro Gly Ala Ser Leu Cys Arg Pro Gly Ala Pro Leu Leu Asn Ser Ser  
20 25 30

30 Ser Val Gly Asn Leu Ser Cys Glu Pro Pro Arg Ile Arg Gly Ala Gly  
35 40 45

35 Thr Arg Glu Leu Glu Leu Ala Ile Arg Ile Thr Leu Tyr Ala Val Ile  
50 55 60

65 Phe Leu Met Ser Val Gly Gly Asn Met Leu Ile Ile Val Val Leu Gly  
65 70 75 80

Leu Ser Arg Arg Leu Arg Thr Val Thr Asn Ala Phe Leu Leu Ser Leu  
35 85 90 95

Ala Val Ser Asp Leu Leu Ala Val Ala Cys Met Pro Phe Thr Leu

10

20

30

40

	100	105	110	
	Leu Pro Asn Leu Met Gly Thr Phe Ile Phe Gly Thr Val Ile Cys Lys			
	115	120	125	
5	Ala Val Ser Tyr Leu Met Gly Val Ser Val Ser Val Ser Thr Leu Ser			
	130	135	140	
	Leu Val Ala Ile Ala Leu Glu Arg Tyr Ser Ala Ile Cys Arg Pro Leu			
	145	150	155	160
	Gln Ala Arg Val Trp Gln Thr Arg Ser His Ala Ala Arg Val Ile Val			
	165	170	175	
10	Ala Thr Trp Leu Leu Ser Gly Leu Leu Met Val Pro Tyr Pro Val Tyr			
	180	185	190	
	Thr Val Val Gln Pro Val Gly Pro Arg Val Leu Gln Cys Val His Arg			
	195	200	205	
15	Trp Pro Ser Ala Arg Val Arg Gln Thr Trp Ser Val Leu Leu Leu Leu			
	210	215	220	
	Leu Leu Phe Phe Ile Pro Gly Val Val Met Ala Val Ala Tyr Gly Leu			
	225	230	235	240
	Ile Ser Arg Glu Leu Tyr Leu Gly Leu Arg Phe Asp Gly Asp Ser Asp			
	245	250	255	
20	Ser Asp Ser Gln Ser Arg Val Arg Asn Gln Gly Gly Leu Pro Gly Ala			
	260	265	270	
	Val His Gln Asn Gly Arg Cys Arg Pro Glu Thr Gly Ala Val Gly Lys			
	275	280	285	
25	Asp Ser Asp Gly Cys Tyr Val Gln Leu Pro Arg Ser Arg Pro Ala Leu			
	290	295	300	
	Glu Leu Thr Ala Leu Thr Ala Pro Gly Pro Gly Ser Gly Ser Arg Pro			
	305	310	315	320
	Thr Gln Ala Lys Leu Leu Ala Lys Lys Arg Val Lys Arg Met Leu Leu			
	325	330	335	
30	Val Ile Val Val Leu Phe Phe Leu Cys Trp Leu Pro Val Tyr Ser Ala			
	340	345	350	
	Asn Thr Trp Arg Ala Phe Asp Gly Pro Gly Ala His Arg Ala Leu Ser			
	355	360	365	
35	Val Ala Pro Ile Ser Phe Ile His Leu Leu Ser Tyr Ala Ser Ala Cys			
	370	375	380	
	Val Asn Pro Leu Val Tyr Cys Phe Met His Arg Arg Phe Arg Gln Ala			
	385	390	395	400

Cys Leu Glu Thr Cys Ala Arg Cys Cys Pro Arg Pro Pro Arg Ala Arg  
 405 410 415  
 Pro Arg Ala Leu Pro Asp Glu Asp Pro Pro Thr Pro Ser Ile Ala Ser  
 420 425 430  
 5 Leu Ser Arg Leu Ser Tyr Thr Ile Ser Thr Leu Gly Pro Gly  
 435 440 445

## (134) INFORMATION FOR SEQ ID NO:133:

- 10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1014 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

15	ATGAAACAGCA CATGTATTGA AGAACAGCAT GACCTGGATC ACTATTTGTT TCCCATTGTT	60
	TACATCTTG TGATTATAGT CAGCATTCCA GCCAATATTG GATCTCTGTC TGTGTCTTTC	120
	CTGCAAGCAA AGAAGGAAAG TGAACCTAGGA ATTTACCTCT TCAGTTTGTGTC ACTATCAGAT	180
	TTACTCTATG CATTAACCTCT CCCTTTATGG ATTGATTATA CTTGGAATAA AGACAACTGG	240
	ACTTTCTCTC CTGCCTTGTG CAAAGGGAGT GCTTTCTCA TGTACATGAA TTTTTACAGC	300
20	AGCACAGCAT TCCTCACCTG CATTGCCGTT GATCGGTATT TGGCTGTTGT CTACCCCTTG	360
	AAGTTTTTT TCCTAAGGAC AAGAAGATT GCACTCATGG TCAGCCTGTC CATCTGGATA	420
	TTGGAAACCA TCTTCAATGC TGTCAATGTT GGGGAAGATG AAACAGTTGT TGAATATTGC	480
	GATGCCGAAA AGTCTAATTT TACTTTATGC TATGACAAAT ACCCTTTAGA GAAATGGCAA	540
	ATCAACCTCA ACTTGTTCAG GACGTGTACA GGCTATGCAA TACCTTTGGT CACCATCCTG	600
25	ATCTGTAACC GGAAAGTCTA CCAAGCTGTG CGGCACAATA AAGCCACGGAA AAACAAGGAA	660
	AAGAAGAGAA TCAAAAAACT ACTTGTCAAGC ATCACAGTTA CTTTTGTCCTT ATGCTTTACT	720
	CCCTTTCATG TGATGTTGCT GATTCGCTGC ATTTAGAGC ATGCTGTGAA CTTCGAAGAC	780
	CACAGCAATT CTGGGAAGCG AACTTACACA ATGTATAGAA TCACGGTTGC ATTAACAAGT	840
	TTAAATTGTG TTGCTGATCC AATTCTGTAC TGTTTGTGAA CCGAAACAGG AAGATATGAT	900
30	ATGTGGAATA TATTAAAATT CTGCACTGGG AGGTGTAATA CATCACAAAG ACAAAAGAAAA	960
	CGCATACTTT CTGTGTCTAC AAAAGATACT ATGGAATTAG AGGTCCCTGAA GTAG	1014

10

20

30

(135) INFORMATION FOR SEQ ID NO:134:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 337 amino acids  
 (B) TYPE: amino acid  
 5 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

10 Met Asn Ser Thr Cys Ile Glu Glu Gln His Asp Leu Asp His Tyr Leu  
 1 5 10 15

Phe Pro Ile Val Tyr Ile Phe Val Ile Ile Val Ser Ile Pro Ala Asn  
 20 25 30

Ile Gly Ser Leu Cys Val Ser Phe Leu Gln Ala Lys Lys Glu Ser Glu  
 35 40 45

15 Leu Gly Ile Tyr Leu Phe Ser Leu Ser Leu Ser Asp Leu Leu Tyr Ala  
 50 55 60

Leu Thr Leu Pro Leu Trp Ile Asp Tyr Thr Trp Asn Lys Asp Asn Trp  
 65 70 75 80

20 Thr Phe Ser Pro Ala Leu Cys Lys Gly Ser Ala Phe Leu Met Tyr Met  
 85 90 95

Asn Phe Tyr Ser Ser Thr Ala Phe Leu Thr Cys Ile Ala Val Asp Arg  
 100 105 110

Tyr Leu Ala Val Val Tyr Pro Leu Lys Phe Phe Leu Arg Thr Arg  
 115 120 125

25 Arg Phe Ala Leu Met Val Ser Leu Ser Ile Trp Ile Leu Glu Thr Ile  
 130 135 140

Phe Asn Ala Val Met Leu Trp Glu Asp Glu Thr Val Val Glu Tyr Cys  
 145 150 155 160

30 Asp Ala Glu Lys Ser Asn Phe Thr Leu Cys Tyr Asp Lys Tyr Pro Leu  
 165 170 175

Glu Lys Trp Gln Ile Asn Leu Asn Leu Phe Arg Thr Cys Thr Gly Tyr  
 180 185 190

Ala Ile Pro Leu Val Thr Ile Leu Ile Cys Asn Arg Lys Val Tyr Gln  
 195 200 205

35 Ala Val Arg His Asn Lys Ala Thr Glu Asn Lys Glu Lys Lys Arg Ile  
 210 215 220

10

20

30

Lys Lys Leu Leu Val Ser Ile Thr Val Thr Phe Val Leu Cys Phe Thr  
 225 230 235 240  
 Pro Phe His Val Met Leu Leu Ile Arg Cys Ile Leu Glu His Ala Val  
 245 250 255  
 5 Asn Phe Glu Asp His Ser Asn Ser Gly Lys Arg Thr Tyr Thr Met Tyr  
 260 265 270  
 Arg Ile Thr Val Ala Leu Thr Ser Leu Asn Cys Val Ala Asp Pro Ile  
 275 280 285  
 10 Leu Tyr Cys Phe Val Thr Glu Thr Gly Arg Tyr Asp Met Trp Asn Ile  
 290 295 300  
 Leu Lys Phe Cys Thr Gly Arg Cys Asn Thr Ser Gln Arg Gln Arg Lys  
 305 310 315 320  
 Arg Ile Leu Ser Val Ser Thr Lys Asp Thr Met Glu Leu Glu Val Leu  
 325 330 335  
 15 Glu

## (136) INFORMATION FOR SEQ ID NO:135:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 999 base pairs  
 20 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: DNA (genomic)  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:  
 25 ATGGTGAACT CCACCCACCG TGGGATGCAC ACTTCTCTGC ACCTCTGGAA CCGCAGCAGT  
 60  
 TACAGACTGC ACAGCAATGC CAGTGAGTCC CTTGGAAAAG GCTACTCTGA TGGAGGGTGC  
 120  
 30 TACGAGCAAC TTTTTGTCTC TCCTGAGGTG TTTGTGACTC TGGGTGTCAT CAGCTTGTG  
 180  
 GAGAATATCT TAGTGATTGT GGCAATAGCC AAGAACAAAGA ATCTGCATTC ACCCATGTAC  
 240  
 TTTTTCATCT GCAGCTTGGC TGTGGCTGAT ATGCTGGTGA GCGTTCAAA TGGATCAGAA  
 300  
 35 ACCATTATCA TCACCCATTAAACAGTACA GATACGGATG CACAGAGTTT CACAGTGAAT  
 360

ATTGATAATG TCATTGACTC GGTGATCTGT AGCTCCTTGC TTGCATCCAT TTGCAGCCTG  
 420

CTTTCAATTG CAGTGGACAG GTACTTTACT ATCTTCTATG CTCTCCAGTA CCATAACATT  
 480

5 ATGACAGTTA AGCGGGTTGG GATCAGCATA AGTTGTATCT GGGCAGCTTG CACGGTTCA  
 540

GGCATTTGT TCATCATTAA CTCAGATAGT AGTGCTGTCA TCATCTGCCT CATCACCAGT  
 600

TTCTTCACCA TGCTGGCTCT CATGGCTTCT CTCTATGTCC ACATGTTCT GATGGCCAGG  
 10 660

CTTCACATTA AGAGGATTGC TGTCCCTCCC GGCACGGTG CCATCCGCCA AGGTGCCAAT  
 720

ATGAAGGGAA AAATTACCTT GACCATCCTG ATTGGCGTCT TTGTTGCTG CTGGGGCCCA  
 780

15 TTCTTCCTCC ACTTAATATT CTACATCTCT TGTCCCTCAGA ATCCATATTG TGTGTGCTTC  
 840

ATGTCTCACT TTAACTTGTA TCTCATACTG ATCATGTGTA ATTCAATCAT CGATCCTCTG  
 900

20 ATTTATGCAC TCCGGAGTCA AGAACTGAGG AAAACCTTCA AAGAGATCAT CTGTTGCTAT  
 960

CCCCTGGGAG GCCTTTGTGA CTTGTCTAGC AGATATTAA  
 999

(137) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:  
 25 (A) LENGTH: 332 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein  
 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

Met Val Asn Ser Thr His Arg Gly Met His Thr Ser Leu His Leu Trp  
 1 5 10 15

Asn Arg Ser Ser Tyr Arg Leu His Ser Asn Ala Ser Glu Ser Leu Gly  
 20 25 30

35 Lys Gly Tyr Ser Asp Gly Gly Cys Tyr Glu Gln Leu Phe Val Ser Pro  
 35 40 45

	Glu Val Phe Val Thr Leu Gly Val Ile Ser Leu Leu Glu Asn Ile Leu		
50	55	60	
	Val Ile Val Ala Ile Ala Lys Asn Lys Asn Leu His Ser Pro Met Tyr		
65	70	75	80
5	Phe Phe Ile Cys Ser Leu Ala Val Ala Asp Met Leu Val Ser Val Ser		
	85	90	95
	Asn Gly Ser Glu Thr Ile Ile Thr Leu Leu Asn Ser Thr Asp Thr		
	100	105	110
10	Asp Ala Gln Ser Phe Thr Val Asn Ile Asp Asn Val Ile Asp Ser Val		
	115	120	125
	Ile Cys Ser Ser Leu Leu Ala Ser Ile Cys Ser Leu Leu Ser Ile Ala		
	130	135	140
	Val Asp Arg Tyr Phe Thr Ile Phe Tyr Ala Leu Gln Tyr His Asn Ile		
	145	150	155
15	Met Thr Val Lys Arg Val Gly Ile Ser Ile Ser Cys Ile Trp Ala Ala		
	165	170	175
	Cys Thr Val Ser Gly Ile Leu Phe Ile Ile Tyr Ser Asp Ser Ser Ala		
	180	185	190
20	Val Ile Ile Cys Leu Ile Thr Met Phe Phe Thr Met Leu Ala Leu Met		
	195	200	205
	Ala Ser Leu Tyr Val His Met Phe Leu Met Ala Arg Leu His Ile Lys		
	210	215	220
	Arg Ile Ala Val Leu Pro Gly Thr Gly Ala Ile Arg Gln Gly Ala Asn		
	225	230	235
25	Met Lys Gly Lys Ile Thr Leu Thr Ile Leu Ile Gly Val Phe Val Val		
	245	250	255
	Cys Trp Ala Pro Phe Phe Leu His Leu Ile Phe Tyr Ile Ser Cys Pro		
	260	265	270
30	Gln Asn Pro Tyr Cys Val Cys Phe Met Ser His Phe Asn Leu Tyr Leu		
	275	280	285
	Ile Leu Ile Met Cys Asn Ser Ile Ile Asp Pro Leu Ile Tyr Ala Leu		
	290	295	300
	Arg Ser Gln Glu Leu Arg Lys Thr Phe Lys Glu Ile Ile Cys Cys Tyr		
	305	310	315
35	Pro Leu Gly Gly Leu Cys Asp Leu Ser Ser Arg Tyr		
	325	330	

(138) INFORMATION FOR SEQ ID NO:137:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 33 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

GCCTAATATGA AGGGAAAAAT TACCTTGACC ATC  
 33

10 (137) INFORMATION FOR SEQ ID NO:138:

- 15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

CTCCTTCGGT CCTCCTATCG TTGTCAGAAG T  
 31

20 (140) INFORMATION FOR SEQ ID NO:139:

- 25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1842 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

ATGGGGCCCA CCCTAGCGGT TCCCACCCCC TATGGCTGTA TTGGCTGTAA GCTACCCCAG	60
CCAGAATACC CACCGGCTCT AATCATCTT ATGTTCTGCG CGATGGTTAT CACCATCGTT	120
30 GTAGACCTAA TCGGCAACTC CATGGTCATT TTGGCTGTGA CGAAGAACAA GAAGCTCCGG	180
AATTCTGGCA ACATCTTCGT GGTCAGTCTC TCTGTGGCCG ATATGCTGGT GGCCATCTAC	240
CCATACCCTT TGATGCTGCA TGCCATGTCC ATTGGGGGCT GGGATCTGAG CCAGTTACAG	300
TGCCAGATGG TCGGGTTCAT CACAGGGCTG AGTGTGGTCG GCTCCATCTT CAACATCGTG	360

GCAATCGCTA	TCAACCGTTA	CTGCTACATC	TGCCACAGCC	TCCAGTACGA	ACGGATCTTC	420		
AGTGTGCGCA	ATACCTGCAT	CTACCTGGTC	ATCACCTGGA	TCATGACCGT	CCTGGCTGTC	480		
CTGCCAACAA	TGTACATTGG	CACCATCGAG	TACGATCCTC	GCACCTACAC	CTGCATCTTC	540		
AACTATCTGA	ACAACCCTGT	CTTCACTGTT	ACCATCGTCT	GCATCCACTT	CGTCCTCCCT	600		
5	CTCCTCATCG	TGGGTTTCTG	CTACGTGAGG	ATCTGGACCA	AAGTGCTGGC	GGCCCGTGAC	660	
	CCTGCAGGGC	AGAACCTCTGA	CAACCAACTT	GCTGAGGTTTC	GCAATTTCCT	AACCATGTTT	720	
	GTGATCTTCC	TCCTCTTGC	AGTGTGCTGG	TGCCCTATCA	ACGTGCTCAC	TGTCTTGGTG	780	10
	GCTGTCAGTC	CGAAGGAGAT	GGCAGGCAAG	ATCCCCAACT	GGCTTTATCT	TGCAGCCTAC	840	
	TTCATAGCCT	ACTTCAACAG	CTGCCTCAAC	GCTGTGATCT	ACGGGCTCCT	CAATGAGAAT	900	
10	TTCCGAAGAG	AATACTGGAC	CATCTTCCAT	GCTATGCGGC	ACCTATCAT	ATTCTTCCCT	960	
	GGCCTCATCA	GTGATATTG	TGAGATGCAG	GAGGCCCGTA	CCCTGGCCCG	CGCCCGTGCC	1020	
	CATGCTCGCG	ACCAAGCTCG	TGAACAAGAC	CGTGCCCATG	CCTGTCTGC	TGTGGAGGAA	1080	
	ACCCCGATGA	ATGTCCGGAA	TGTTCCATTA	CCTGGTGATG	CTGCAGCTGG	CCACCCCGAC	1140	
	CGTGCTCTG	GCCACCCCTAA	GCCCCATTCC	AGATCCTCCT	CTGCCTATCG	CAAATCTGCC	1200	
15	TCTACCCACC	ACAAGTCTGT	CTTTAGCCAC	TCCAAGGCTG	CCTCTGGTCA	CCTCAAGCCT	1260	20
	GTCTCTGGCC	ACTCCAAGCC	TGCCTCTGGT	CACCCCAAGT	CTGCCACTGT	CTACCCCTAAG	1320	
	CCTGCCTCTG	TCCATTTCAA	GGGTGACTCT	GTCCATTTCA	AGGGTGACTC	TGTCCATTTC	1380	
	AAGCCTGACT	CTGTTCAATT	CAAGCCTGCT	TCCAGCAACC	CCAAGCCCAT	CACTGGCCAC	1440	
	CATGTCTCTG	CTGGCAGCCA	CTCCAAGTCT	GCCTTCAGTG	CTGCCACCAG	CCACCCCTAAA	1500	
20	CCCATCAAGC	CAGCTACCAAG	CCATGCTGAG	CCCACCCTG	CTGACTATCC	CAAGCCTGCC	1560	
	ACTACCAGCC	ACCCTAAGCC	CGCTGCTGCT	GACAACCCTG	AGCTCTCTGC	CTCCCATTGC	1620	
	CCCGAGATCC	CTGCCATTGC	CCACCCCTGTG	TCTGACGACA	GTGACCTCCC	TGAGTCGGCC	1680	30
	TCTAGCCCTG	CCGCTGGGCC	CACCAAGCCT	GCTGCCAGCC	AGCTGGAGTC	TGACACCATC	1740	
	GCTGACCTTC	CTGACCCCTAC	TGTAGTCACT	ACCAGTACCA	ATGATTACCA	TGATGTCGTG	1800	
25	GTTGTTGATG	TTGAAGATGA	TCCTGATGAA	ATGGCTGTGT	GA		1842	

(141) INFORMATION FOR SEQ ID NO:140:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 613 amino acids
  - (B) TYPE: amino acid

(C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

5	Met	Gly	Pro	Thr	Leu	Ala	Val	Pro	Thr	Pro	Tyr	Gly	Cys	Ile	Gly	Cys	
	1				5			10					15				
	Lys	Leu	Pro	Gln	Pro	Glu	Tyr	Pro	Pro	Ala	Leu	Ile	Ile	Phe	Met	Phe	10
		20				25								30			
10	Cys	Ala	Met	Val	Ile	Thr	Ile	Val	Val	Asp	Leu	Ile	Gly	Asn	Ser	Met	
		35				40							45				
	Val	Ile	Leu	Ala	Val	Thr	Lys	Asn	Lys	Lys	Leu	Arg	Asn	Ser	Gly	Asn	
		50				55						60					
	Ile	Phe	Val	Val	Ser	Leu	Ser	Val	Ala	Asp	Met	Leu	Val	Ala	Ile	Tyr	
		65			70			75					80				
15	Pro	Tyr	Pro	Leu	Met	Leu	His	Ala	Met	Ser	Ile	Gly	Gly	Trp	Asp	Leu	
			85			90			95								
	Ser	Gln	Leu	Gln	Cys	Gln	Met	Val	Gly	Phe	Ile	Thr	Gly	Leu	Ser	Val	20
		100				105						110					
20	Val	Gly	Ser	Ile	Phe	Asn	Ile	Val	Ala	Ile	Ala	Ile	Asn	Arg	Tyr	Cys	
		115			120			125									
	Tyr	Ile	Cys	His	Ser	Leu	Gln	Tyr	Glu	Arg	Ile	Phe	Ser	Val	Arg	Asn	
		130			135			140									
	Thr	Cys	Ile	Tyr	Leu	Val	Ile	Thr	Trp	Ile	Met	Thr	Val	Leu	Ala	Val	
		145			150			155					160				
25	Leu	Pro	Asn	Met	Tyr	Ile	Gly	Thr	Ile	Glu	Tyr	Asp	Pro	Arg	Thr	Tyr	
			165			170			175								
	Thr	Cys	Ile	Phe	Asn	Tyr	Leu	Asn	Asn	Pro	Val	Phe	Thr	Val	Thr	Ile	30
		180			185			190									
30	Val	Cys	Ile	His	Phe	Val	Leu	Pro	Leu	Leu	Ile	Val	Gly	Phe	Cys	Tyr	
		195			200			205									
	Val	Arg	Ile	Trp	Thr	Lys	Val	Leu	Ala	Ala	Arg	Asp	Pro	Ala	Gly	Gln	
		210			215			220									
	Asn	Pro	Asp	Asn	Gln	Leu	Ala	Glu	Val	Arg	Asn	Phe	Leu	Thr	Met	Phe	
		225			230			235					240				
35	Val	Ile	Phe	Leu	Leu	Phe	Ala	Val	Cys	Trp	Cys	Pro	Ile	Asn	Val	Leu	
		245			250			255									

	Thr Val Leu Val Ala Val Ser Pro Lys Glu Met Ala Gly Lys Ile Pro			
	260	265	270	
	Asn Trp Leu Tyr Leu Ala Ala Tyr Phe Ile Ala Tyr Phe Asn Ser Cys			
	275	280	285	
5	Leu Asn Ala Val Ile Tyr Gly Leu Leu Asn Glu Asn Phe Arg Arg Glu			
	290	295	300	
	Tyr Trp Thr Ile Phe His Ala Met Arg His Pro Ile Ile Phe Phe Pro			
	305	310	315	320
10	Gly Leu Ile Ser Asp Ile Arg Glu Met Gln Glu Ala Arg Thr Leu Ala			10
	325	330	335	
	Arg Ala Arg Ala His Ala Arg Asp Gln Ala Arg Glu Gln Asp Arg Ala			
	340	345	350	
	His Ala Cys Pro Ala Val Glu Glu Thr Pro Met Asn Val Arg Asn Val			
	355	360	365	
15	Pro Leu Pro Gly Asp Ala Ala Gly His Pro Asp Arg Ala Ser Gly			
	370	375	380	
	His Pro Lys Pro His Ser Arg Ser Ser Ser Ala Tyr Arg Lys Ser Ala			
	385	390	395	400
20	Ser Thr His His Lys Ser Val Phe Ser His Ser Lys Ala Ala Ser Gly			20
	405	410	415	
	His Leu Lys Pro Val Ser Gly His Ser Lys Pro Ala Ser Gly His Pro			
	420	425	430	
	Lys Ser Ala Thr Val Tyr Pro Lys Pro Ala Ser Val His Phe Lys Gly			
	435	440	445	
25	Asp Ser Val His Phe Lys Gly Asp Ser Val His Phe Lys Pro Asp Ser			
	450	455	460	
	Val His Phe Lys Pro Ala Ser Ser Asn Pro Lys Pro Ile Thr Gly His			
	465	470	475	480
30	His Val Ser Ala Gly Ser His Ser Lys Ser Ala Phe Ser Ala Ala Thr			30
	485	490	495	
	Ser His Pro Lys Pro Ile Lys Pro Ala Thr Ser His Ala Glu Pro Thr			
	500	505	510	
	Thr Ala Asp Tyr Pro Lys Pro Ala Thr Thr Ser His Pro Lys Pro Ala			
	515	520	525	
35	Ala Ala Asp Asn Pro Glu Leu Ser Ala Ser His Cys Pro Glu Ile Pro			
	530	535	540	
	Ala Ile Ala His Pro Val Ser Asp Asp Ser Asp Leu Pro Glu Ser Ala			

545	550	555	560
Ser Ser Pro Ala Ala Gly Pro Thr Lys Pro Ala Ala Ser Gln Leu Glu			
565	570	575	
Ser Asp Thr Ile Ala Asp Leu Pro Asp Pro Thr Val Val Thr Thr Ser			
5	580	585	590
Thr Asn Asp Tyr His Asp Val Val Val Val Asp Val Glu Asp Asp Pro			
595	600	605	
Asp Glu Met Ala Val			
610			

10

## 10 (142) INFORMATION FOR SEQ ID NO:141:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1842 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

20

ATGGGGCCCA CCCTAGCGGT TCCCACCCCC TATGGCTGTA TTGGCTGTAA GCTACCCCCAG	60
CCAGAATACC CACCGGCTCT AATCATCTTT ATGTTCTGCG CGATGGTTAT CACCATCGTT	
20 GTAGACCTAA TCGGCAACTC CATGGTCATT TTGGCTGTGA CGAAGAACAA GAAGCTCCGG	
AATTCTGGCA ACATCTTCGT GGTCAGTCTC TCTGTGGCCG ATATGCTGGT GGCCATCTAC	
CCATACCCCTT TGATGCTGCA TGCCATGTCC ATTGGGGCT GGGATCTGAG CCAGTTACAG	
TGCCAGATGG TCGGGTTCAT CACAGGGCTG AGTGTGGTCG GCTCCATCTT CAACATCGTG	
GCAATCGCTA TCAACCGTTA CTGCTACATC TGCCACAGCC TCCAGTACGA ACGGATCTTC	
25 AGTGTGCGCA ATACCTGCAT CTACCTGGTC ATCACCTGGA TCATGACCGT CCTGGCTGTC	
CTGCCCAACA TGTACATTGG CACCATCGAG TACGATCCTC GCACCTACAC CTGCATCTTC	
AACTATCTGA ACAACCCTGT CTTCACTGTT ACCATCGTCT GCATCCACTT CGTCCTCCCT	
CTCCTCATCG TGGGTTCTG CTACGTGAGG ATCTGGACCA AAGTGTGGC GGCCCGTGAC	
CCTGCAGGGC AGAACCTGA CAACCAACTT GCTGAGGTT TC GCAATAAACT AACCATGTTT	
30 GTGATCTTCC TCCTCTTGC AGTGTGCTGG TGCCCTATCA ACGTGCTCAC TGTCTTGGTG	
GCTGTCAGTC CGAAGGAGAT GGCAGGCAAG ATCCCCAACT GGCTTTATCT TGCAGCCTAC	
840	

30

40

TTCATAGCCT ACTTCAACAG CTGCCTCAAC GCTGTGATCT ACGGGCTCCT CAATGAGAAT 900  
 TTCCGAAGAG AATACTGGAC CATCTTCCAT GCTATGCGGC ACCCTATCAT ATTCTTCTCT 960  
 GGCCTCATCA GTGATATTG TGAGATGCAG GAGGCCCGTA CCCTGGCCCG CGCCCGTGCC 1020  
 CATGCTCGCG ACCAAGCTCG TGAACAAGAC CGTGCCTCATG CCTGTCCCTGC TGTGGAGGAA 1080  
 5 ACCCCGATGA ATGTCGGAA TGTTCCATTA CCTGGTGATG CTGCAGCTGG CCACCCCGAC 1140  
 CGTGCTCTG GCCACCCCTAA GCCCCATTCC AGATCCTCCT CTGCCTATCG CAAATCTGCC 1200  
 TCTACCCACC ACAAGTCTGT CTTTAGCCAC TCCAAGGCTG CCTCTGGTCA CCTCAAGCCT 1260 10  
 GTCTCTGGCC ACTCCAAGCC TGCCTCTGGT CACCCCAAGT CTGCCACTGT CTACCCCTAAG 1320  
 CCTGCCTCTG TCCATTCAA GGCTGACTCT GTCCATTCA AGGGTGACTC TGTCCATTTC 1380  
 10 AAGCCTGACT CTGTTCATTT CAAGCCTGCT TCCAGCAACC CCAAGCCCAT CACTGGCCAC 1440  
 CATGTCTCTG CTGGCAGCCA CTCCAAGTCT GCCTTCAATG CTGCCACCAG CCACCCCTAAA 1500  
 CCCATCAAGC CAGCTACCAAG CCATGCTGAG CCCACCACTG CTGACTATCC CAAGCCTGCC 1560  
 ACTACCAGCC ACCCTAAGCC CGCTGCTGCT GACAACCCTG AGCTCTCTGC CTCCCATTGC 1620  
 CCCGAGATCC CTGCCATTGC CCACCCGTG TCTGACGACA GTGACCTCCC TGAGTCGGCC 1680  
 15 TCTAGCCCTG CCGCTGGCC CACCAAGCCT GCTGCCAGCC AGCTGGAGTC TGACACCATC 1740 20  
 GCTGACCTTC CTGACCCCTAC TGTAGTCACT ACCAGTACCA ATGATTACCA TGATGTCGTG 1800  
 GTTGTGATG TTGAAGATGA TCCTGATGAA ATGGCTGTGT GA 1842

(143) INFORMATION FOR SEQ ID NO:142:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 613 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein 30

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

Met Gly Pro Thr Leu Ala Val Pro Thr Pro Tyr Gly Cys Ile Gly Cys  
 1 5 10 15

Lys Leu Pro Gln Pro Glu Tyr Pro Pro Ala Leu Ile Ile Phe Met Phe  
 20 25 30

30 Cys Ala Met Val Ile Thr Ile Val Val Asp Leu Ile Gly Asn Ser Met  
 35 40 45

Val Ile Leu Ala Val Thr Lys Asn Lys Lys Leu Arg Asn Ser Gly Asn  
 50 55 60

Ile Phe Val Val Ser Leu Ser Val Ala Asp Met Leu Val Ala Ile Tyr  
 65 70 75 80

5 Pro Tyr Pro Leu Met Leu His Ala Met Ser Ile Gly Gly Trp Asp Leu  
 85 90 95

Ser Gln Leu Gln Cys Gln Met Val Gly Phe Ile Thr Gly Leu Ser Val  
 100 105 110

Val Gly Ser Ile Phe Asn Ile Val Ala Ile Ala Ile Asn Arg Tyr Cys  
 10 115 120 125

Tyr Ile Cys His Ser Leu Gln Tyr Glu Arg Ile Phe Ser Val Arg Asn  
 130 135 140

Thr Cys Ile Tyr Leu Val Ile Thr Trp Ile Met Thr Val Leu Ala Val  
 145 150 155 160

15 Leu Pro Asn Met Tyr Ile Gly Thr Ile Glu Tyr Asp Pro Arg Thr Tyr  
 165 170 175

Thr Cys Ile Phe Asn Tyr Leu Asn Asn Pro Val Phe Thr Val Thr Ile  
 180 185 190

Val Cys Ile His Phe Val Leu Pro Leu Leu Ile Val Gly Phe Cys Tyr  
 20 195 200 205

Val Arg Ile Trp Thr Lys Val Leu Ala Ala Arg Asp Pro Ala Gly Gln  
 210 215 220

Asn Pro Asp Asn Gln Leu Ala Glu Val Arg Asn Lys Leu Thr Met Phe  
 225 230 235 240

25 Val Ile Phe Leu Leu Phe Ala Val Cys Trp Cys Pro Ile Asn Val Leu  
 245 250 255

Thr Val Leu Val Ala Val Ser Pro Lys Glu Met Ala Gly Lys Ile Pro  
 260 265 270

Asn Trp Leu Tyr Leu Ala Ala Tyr Phe Ile Ala Tyr Phe Asn Ser Cys  
 30 275 280 285

Leu Asn Ala Val Ile Tyr Gly Leu Leu Asn Glu Asn Phe Arg Arg Glu  
 290 295 300

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

35 Gly Leu Ile Ser Asp Ile Arg Glu Met Gln Glu Ala Arg Thr Leu Ala  
 325 330 335

Arg Ala Arg Ala His Ala Arg Asp Gln Ala Arg Glu Gln Asp Arg Ala

	340	345	350	
	His Ala Cys Pro Ala Val Glu Glu Thr Pro Met Asn Val Arg Asn Val			
	355	360	365	
5	Pro Leu Pro Gly Asp Ala Ala Ala Gly His Pro Asp Arg Ala Ser Gly			
	370	375	380	
	His Pro Lys Pro His Ser Arg Ser Ser Ala Tyr Arg Lys Ser Ala			
	385	390	395	400
	Ser Thr His His Lys Ser Val Phe Ser His Ser Lys Ala Ala Ser Gly			
	405	410	415	
10	His Leu Lys Pro Val Ser Gly His Ser Lys Pro Ala Ser Gly His Pro			
	420	425	430	
	Lys Ser Ala Thr Val Tyr Pro Lys Pro Ala Ser Val His Phe Lys Ala			
	435	440	445	
15	Asp Ser Val His Phe Lys Gly Asp Ser Val His Phe Lys Pro Asp Ser			
	450	455	460	
	Val His Phe Lys Pro Ala Ser Ser Asn Pro Lys Pro Ile Thr Gly His			
	465	470	475	480
	His Val Ser Ala Gly Ser His Ser Lys Ser Ala Phe Asn Ala Ala Thr			
	485	490	495	
20	Ser His Pro Lys Pro Ile Lys Pro Ala Thr Ser His Ala Glu Pro Thr			
	500	505	510	
	Thr Ala Asp Tyr Pro Lys Pro Ala Thr Thr Ser His Pro Lys Pro Ala			
	515	520	525	
25	Ala Ala Asp Asn Pro Glu Leu Ser Ala Ser His Cys Pro Glu Ile Pro			
	530	535	540	
	Ala Ile Ala His Pro Val Ser Asp Asp Ser Asp Leu Pro Glu Ser Ala			
	545	550	555	560
	Ser Ser Pro Ala Ala Gly Pro Thr Lys Pro Ala Ala Ser Gln Leu Glu			
	565	570	575	
30	Ser Asp Thr Ile Ala Asp Leu Pro Asp Pro Thr Val Val Thr Thr Ser			
	580	585	590	
	Thr Asn Asp Tyr His Asp Val Val Val Asp Val Glu Asp Asp Pro			
	595	600	605	
35	Asp Glu Met Ala Val			
	610			

(144) INFORMATION FOR SEQ ID NO:143:

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 33 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

GCTGAGGTTC GCAATAACT AACCATGTTT GTG

33

(145) INFORMATION FOR SEQ ID NO:144:

10

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 30 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

CTCCTTCGGT CCTCCTATCG TTGTCAGAAG T

31

20

(146) INFORMATION FOR SEQ ID NO:145:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

TTAGATATCG GGGCCCACCC TAGCGGT

33

30

(147) INFORMATION FOR SEQ ID NO:146:

30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 29 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

GGTACCCCCA CAGCCATTTC ATCAGGATC

33

【図 1】

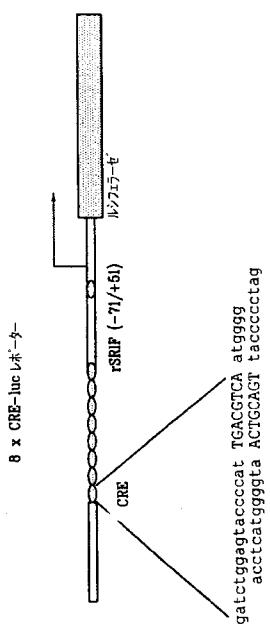


FIG. 1

【図 2】

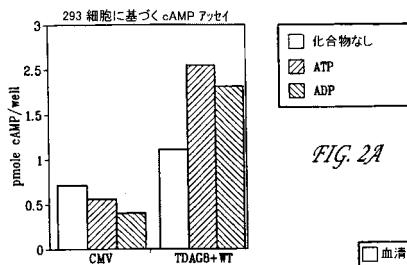


FIG. 2A

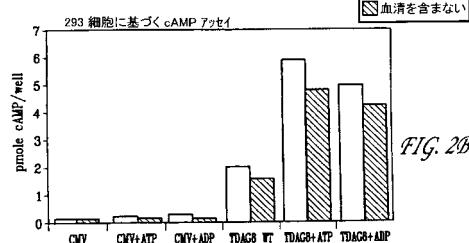


FIG. 2B

【図 3】

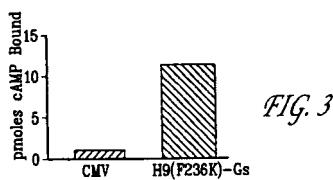


FIG. 3

## フロントページの続き

(51) Int.CI.	F I	テーマコード(参考)
C 1 2 N	5/10	(2006.01)
	C 1 2 N	5/00
		1 0 1

(31) 優先権主張番号 60/120,416  
 (32) 優先日 平成11年2月16日(1999.2.16)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/121,852  
 (32) 優先日 平成11年2月26日(1999.2.26)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/123,944  
 (32) 優先日 平成11年3月12日(1999.3.12)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/123,945  
 (32) 優先日 平成11年3月12日(1999.3.12)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/123,948  
 (32) 優先日 平成11年3月12日(1999.3.12)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/123,949  
 (32) 優先日 平成11年3月12日(1999.3.12)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/123,951  
 (32) 優先日 平成11年3月12日(1999.3.12)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/136,436  
 (32) 優先日 平成11年5月28日(1999.5.28)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/136,437  
 (32) 優先日 平成11年5月28日(1999.5.28)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/136,439  
 (32) 優先日 平成11年5月28日(1999.5.28)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/137,567  
 (32) 優先日 平成11年5月28日(1999.5.28)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/137,127  
 (32) 優先日 平成11年5月28日(1999.5.28)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/137,131  
 (32) 優先日 平成11年5月28日(1999.5.28)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/141,448  
 (32) 優先日 平成11年6月30日(1999.6.30)  
 (33) 優先権主張国 米国(US)

(31) 優先権主張番号 60/151,114  
 (32) 優先日 平成11年8月27日(1999.8.27)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/152,524  
 (32) 優先日 平成11年9月3日(1999.9.3)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/156,653  
 (32) 優先日 平成11年9月9日(1999.9.9)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/156,633  
 (32) 優先日 平成11年9月29日(1999.9.29)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/156,555  
 (32) 優先日 平成11年9月29日(1999.9.29)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/156,634  
 (32) 優先日 平成11年9月29日(1999.9.29)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/157,280  
 (32) 優先日 平成11年10月1日(1999.10.1)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/157,294  
 (32) 優先日 平成11年10月1日(1999.10.1)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/157,281  
 (32) 優先日 平成11年10月1日(1999.10.1)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/157,293  
 (32) 優先日 平成11年10月1日(1999.10.1)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 60/157,282  
 (32) 優先日 平成11年10月1日(1999.10.1)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 09/417,044  
 (32) 優先日 平成11年10月12日(1999.10.12)  
 (33) 優先権主張国 米国(US)  
 (31) 優先権主張番号 09/416,760  
 (32) 優先日 平成11年10月12日(1999.10.12)  
 (33) 優先権主張国 米国(US)

(72) 発明者 ドミニク・ピー・ベハン  
 アメリカ合衆国カリフォルニア州 92131 サンディエゴ・ロックスピロコート 11472  
 (72) 発明者 カリン・レーマン - ブルインスマ  
 アメリカ合衆国カリフォルニア州 92129 サンディエゴ・パトライン 12565  
 (72) 発明者 デレク・ティ・シヤルマース  
 アメリカ合衆国カリフォルニア州 92150 ソラナビーチ・ロングデンレイン 347  
 (72) 発明者 チエン・ルオピング  
 アメリカ合衆国カリフォルニア州 92130 サンディエゴ・テインバーブランチウエイ 5296  
 (72) 発明者 フォング・ティ・ダング  
 アメリカ合衆国カリフォルニア州 92105 サンディエゴ・オークパークドライブ 5352

(72)発明者 マーテイン ゴア  
アメリカ合衆国カリフォルニア州 92120 サンディエゴ・エストレラアベニュー 6868

(72)発明者 チエン・ダブリュー リオウ  
アメリカ合衆国カリフォルニア州 92129 サンディエゴ・サリツクスプレイス 7668

(72)発明者 アイ・リン リン  
アメリカ合衆国カリフォルニア州 92126 サンディエゴ・ゴールドコーストドライブ 8291-7

(72)発明者 ケビン ローウィツツ  
アメリカ合衆国カリフォルニア州 92108 サンディエゴ・カミニトデピツツア 8031・アパートメントシー

(72)発明者 キヤロル ホワイト  
アメリカ合衆国カリフォルニア州 92103 サンディエゴ・クリーブランドアベニュー 4260

F ターム(参考) 4B024 AA01 AA11 BA63 CA04 CA09 CA11 CA20 DA12 EA04 GA11  
HA01 HA11  
4B065 AA01X AA57X AA87X AA90X AA90Y AB01 AC14 BA01 CA24 CA44  
CA46  
4H045 AA10 AA30 BA10 CA40 DA50 EA20 EA50 FA71 FA74