

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-500013

(P2004-500013A)

(43) 公表日 平成16年1月8日(2004. 1. 8)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A
A O 1 K 67/027	A O 1 K 67/027	2 G O 4 5
A 6 1 K 31/711	A 6 1 K 31/711	4 B O 2 4
A 6 1 K 38/00	A 6 1 P 7/00	4 B O 6 3
A 6 1 P 7/00	A 6 1 P 7/06	4 B O 6 4
		4 B O 6 5
	審査請求 未請求 予備審査請求 有	(全 371 頁) 最終頁に続く

(21) 出願番号	特願2000-577297 (P2000-577297)	(71) 出願人	500287639
(86) (22) 出願日	平成11年10月18日 (1999. 10. 18)		ミレニアム・ファーマシューティカルズ・
(85) 翻訳文提出日	平成13年4月16日 (2001. 4. 16)		インコーポレイテッド
(86) 国際出願番号	PCT/US1999/024368		アメリカ合衆国 O 2 1 3 9 マサチューセッ
(87) 国際公開番号	W02000/023588		ツ州ケンブリッジ、シドニー・ストリート
(87) 国際公開日	平成12年4月27日 (2000. 4. 27)		7 5 番
(31) 優先権主張番号	09/173. 869	(74) 代理人	100099623
(32) 優先日	平成10年10月16日 (1998. 10. 16)		弁理士 奥山 尚一
(33) 優先権主張国	米国 (US)	(74) 代理人	100096769
(31) 優先権主張番号	09/420. 187		弁理士 有原 幸一
(32) 優先日	平成11年10月18日 (1999. 10. 18)	(74) 代理人	100107319
(33) 優先権主張国	米国 (US)		弁理士 松島 鉄男

最終頁に続く

(54) 【発明の名称】 新規Gタンパク質共役受容体

(57) 【要約】

本発明は、Gタンパク質共役受容体のスーパーファミリーに属する、新規に同定された受容体に関する。本発明はまた、受容体をコードするポリヌクレオチドにも関する。本発明はさらに、受容体の媒介する疾患の診断および治療の標的として、受容体ポリペプチドおよびポリヌクレオチドを使用する方法に関する。本発明はさらに、診断および治療用のアゴニストおよびアンタゴニストを同定するための、受容体ポリペプチドおよびポリヌクレオチドを使用した薬物スクリーニング法に関する。本発明はさらに、受容体ポリペプチドおよびポリヌクレオチドをベースとしたアゴニストおよびアンタゴニストを包含する。本発明はさらに、受容体ポリペプチドおよびポリヌクレオチドを製造する方法に関する。

【特許請求の範囲】

【請求項 1】

(a) 配列番号 1 に示したアミノ酸配列、配列番号 3 に示したアミノ酸配列、および配列番号 5 に示したアミノ酸配列と、

(b) 配列番号 1 に示したアミノ酸配列の対立遺伝子変異体のアミノ酸配列、配列番号 3 に示したアミノ酸配列の対立遺伝子変異体のアミノ酸配列、および配列番号 5 に示したアミノ酸配列の対立遺伝子変異体のアミノ酸配列と、

(c) 配列番号 1 に示したアミノ酸配列の配列変異体であって、ストリンジェントな条件下で、配列番号 2 に示した核酸分子にハイブリッドする核酸分子によりコードされる配列変異体のアミノ酸配列、配列番号 3 に示したアミノ酸配列の配列変異体であって、ストリンジェントな条件下で、配列番号 4 に示した核酸分子にハイブリッドする核酸分子によりコードされる配列変異体のアミノ酸配列、および配列番号 5 に示したアミノ酸配列の配列変異体であって、ストリンジェントな条件下で配列番号 6 に示した核酸分子にハイブリッドする核酸分子によりコードされる配列変異体のアミノ酸配列と、

(d) アミノ酸 1～264 のうち少なくとも 8 つの連続したアミノ酸を含む配列番号 1 に示したアミノ酸配列の断片、アミノ酸 1～244 のうち少なくとも 9 つの連続したアミノ酸を含む配列番号 3 に示したアミノ酸配列の断片、および少なくとも 8 つ連続したアミノ酸を含む配列番号 5 に示したアミノ酸配列の断片と、

(e) 配列番号 1 に示したアミノ酸約 6～アミノ酸約 319 番目までの成熟受容体ポリペプチドのアミノ酸配列、配列番号 3 に示したアミノ酸約 6～アミノ酸約 337 番目までの成熟受容体ポリペプチドのアミノ酸配列、および配列番号 5 に示したアミノ酸約 6～アミノ酸約 372 番目までの成熟受容体ポリペプチドのアミノ酸配列と、

(f) アミノ酸約 1～アミノ酸約 24 番目までの配列番号 1 に示したポリペプチドのアミノ酸配列、アミノ酸約 1～アミノ酸約 28 番目までの配列番号 3 に示したポリペプチドのアミノ酸配列、およびアミノ酸約 1～アミノ酸約 25 番目までの配列番号 5 に示したポリペプチドのアミノ酸配列と、

(g) (a)～(f) のポリペプチドのいずれか 1 つのエピトープ含有領域のアミノ酸配列と

からなる群から選択されたアミノ酸配列を有する単離ポリペプチド。

【請求項 2】

請求項 1 のポリペプチド (a)～(g) に選択的に結合する単離抗体。

【請求項 3】

(a) 配列番号 2 に示したヌクレオチド配列、配列番号 4 に示したヌクレオチド配列、および配列番号 6 に示したヌクレオチド配列と、

(b) 配列番号 1 に示したアミノ酸配列をコードするヌクレオチド配列、配列番号 3 に示したアミノ酸配列をコードするヌクレオチド配列、および配列番号 5 に示したアミノ酸配列をコードするヌクレオチド配列と、

(c) (a) または (b) のヌクレオチド配列のいずれかに相補的なヌクレオチド配列とからなる群から選択されたヌクレオチド配列を有する単離された核酸分子。

【請求項 4】

(a) 配列番号 2 に示したヌクレオチド配列にストリンジェントな条件下でハイブリッド形成する配列番号 1 に示したアミノ酸配列の配列変異体のアミノ酸配列をコードするヌクレオチド配列、配列番号 4 に示したヌクレオチド配列にストリンジェントな条件下でハイブリッド形成する配列番号 3 に示したアミノ酸配列の配列変異体のアミノ酸配列をコードするヌクレオチド配列、および、配列番号 6 に示したヌクレオチド配列にストリンジェントな条件下でハイブリッド形成する配列番号 5 に示したアミノ酸配列の配列変異体のアミノ酸配列をコードするヌクレオチド配列と、

(b) (a) のヌクレオチド配列に相補的なヌクレオチド配列と

からなる群から選択されたヌクレオチド配列を有する単離された核酸分子。

【請求項 5】

(a) アミノ酸 1～264 のうち少なくとも 8 つの連続したアミノ酸を含む配列番号 1 に示したアミノ酸配列の断片をコードするヌクレオチド配列、アミノ酸 1～244 のうち少なくとも 9 つの連続したアミノ酸を含む配列番号 3 に示したアミノ酸配列の断片をコードするヌクレオチド配列、および、少なくとも 8 つの連続したアミノ酸を含む配列番号 5 に示したアミノ酸配列の断片をコードするヌクレオチド配列と、

(b) (a) のヌクレオチド配列に相補的なヌクレオチド配列とからなる群から選択されたヌクレオチド配列を有する単離された核酸分子。

【請求項 6】

請求項 3～5 のいずれか一の核酸配列を含む核酸ベクター。

【請求項 7】

請求項 6 のベクターを含む宿主細胞。

【請求項 8】

(a)～(f) のポリペプチド配列のいずれかをコードするヌクレオチド配列を宿主細胞に導入するステップと、タンパク質が核酸から発現される条件下で宿主細胞を培養するステップとを含む請求項 1 のポリペプチドのいずれかを製造する方法。

【請求項 9】

サンプル中における請求項 1 のポリペプチドのいずれかの存在を検出する方法であって、サンプル中のポリペプチドの存在を特異的に検出できる物質とサンプルとを接触させるステップと、ポリペプチドの存在を検出するステップとを含む方法。

【請求項 10】

前記物質は、前記ポリペプチドとの選択的な物理的会合が可能である請求項 9 の方法。

【請求項 11】

前記物質は前記ポリペプチドに結合する請求項 10 の方法。

【請求項 12】

前記物質は抗体である請求項 11 の方法。

【請求項 13】

前記物質はリガンドである請求項 11 の方法。

【請求項 14】

請求項 9 の方法に使用するための試薬を含むキットであって、試薬が前記ポリペプチドに特異的に結合する物質を含むことを特徴とするキット。

【請求項 15】

サンプル中における、請求項 3～5 のいずれか一の核酸配列のいずれかの存在を検出する方法であって、核酸配列とストリンジェントな条件下でハイブリッド形成するオリゴヌクレオチドとサンプルとを接触させるステップと、オリゴヌクレオチドがサンプル中の核酸配列に結合するかを決定するステップとを含む方法。

【請求項 16】

存在が検出される核酸は mRNA である請求項 15 の方法。

【請求項 17】

請求項 15 の方法に使用するための試薬を含むキットであって、試薬がストリンジェントな条件下で核酸分子のいずれかにハイブリッドする化合物を含むことを特徴とするキット。

【請求項 18】

請求項 1 のポリペプチドのいずれかと相互作用する物質を同定する方法であって、前記ポリペプチドが前記物質と相互作用できるように、前記ポリペプチドと前記物質の間の相互作用を可能とする細胞と前記物質とを接触させるステップと、その相互作用を測定するステップとを含む方法。

【請求項 19】

請求項 1 のポリペプチドのいずれかと相互作用する物質を同定するために細胞をスクリーニングする方法であって、前記ポリペプチドが前記物質と相互作用できるように、前記ポリペプチドと前記物質の間の相互作用を可能とする細胞と前記物質とを接触させるステッ

10

20

30

40

50

ブと、その相互作用を測定するステップとを含む方法。

【請求項 20】

細胞において請求項 1 のポリペプチドのいずれかのレベルまたは活性を調節する物質を同定する方法であって、前記ポリペプチドのレベルまたは活性を前記物質により前記細胞で調節できるように、前記ポリペプチドを発現できる細胞と前記物質とを接触させるステップと、前記ポリペプチドのレベルまたは活性を測定するステップとを含む方法。

【請求項 21】

細胞において請求項 1 のポリペプチドのいずれかのレベルまたは活性を調節する物質を同定するために細胞をスクリーニングする方法であって、前記ポリペプチドのレベルまたは活性を前記物質により前記細胞において調節できるように、前記物質を前記ポリペプチドを発現できる細胞と接触させるステップと、前記ポリペプチドのレベルまたは活性を測定するステップとを含む方法。

10

【請求項 22】

前記相互作用は結合である請求項 19 または 20 のいずれかの方法。

【請求項 23】

ポリペプチドの断片を接触させる請求項 22 の方法。

【請求項 24】

請求項 1 におけるポリペプチドのいずれかの活性を調節する方法であって、物質がポリペプチドの活性を調節できる条件下で、請求項 1 のポリペプチドのいずれかを前記物質と接触させるステップを含む方法。

20

【請求項 25】

請求項 3～5 の核酸分子のいずれかと相互作用する物質を同定する方法であって、前記核酸分子が前記物質と相互作用できるように、前記核酸分子と前記物質の間の相互作用を可能とする細胞と前記物質とを接触させるステップと、その相互作用を測定するステップとを含む方法。

【請求項 26】

請求項 3～5 の核酸分子のいずれかと相互作用する物質を同定するために細胞をスクリーニングする方法であって、前記核酸分子が前記物質と相互作用できるように、前記核酸分子と前記物質の間の相互作用を可能とする細胞と前記物質とを接触させるステップと、その相互作用を測定するステップとを含む方法。

30

【請求項 27】

細胞において請求項 3～5 の核酸分子のいずれかのレベルまたは活性を調節する物質を同定する方法であって、前記核酸分子のレベルまたは活性を前記物質により前記細胞において調節できるように、前記物質を前記核酸分子を発現できる細胞と接触させるステップと、前記核酸分子のレベルまたは活性を測定するステップとを含む方法。

【請求項 28】

細胞において請求項 3～5 の核酸分子のいずれかのレベルまたは活性を調節する物質を同定するために細胞をスクリーニングする方法であって、前記核酸分子のレベルまたは活性を前記物質により前記細胞において調節できるように、前記核酸分子を発現できる細胞と前記物質とを接触させるステップと、前記核酸分子のレベルまたは活性を測定するステップとを含む方法。

40

【請求項 29】

前記相互作用が結合である請求項 25 または 26 のいずれかの方法。

【請求項 30】

請求項 3～5 の核酸分子のいずれかの活性を調節する方法であって、物質が核酸分子の活性を調節できる条件下で、請求項 1 の核酸分子のいずれかを前記物質と接触させるステップを含む方法。

【請求項 31】

医薬的に許容される担体に請求項 1 のポリペプチドのいずれかを含んでなる医薬組成物。

【請求項 32】

50

医薬的に許容される担体に請求項 3～5 の核酸分子のいずれかを含んでなる医薬組成物。

【請求項 33】

非ヒトトランスジェニック動物であって、1つ以上の細胞が請求項 3～5 の核酸配列のいずれかを含むトランスジェニック動物。

【請求項 34】

非ヒトトランスジェニック動物であって、1つ以上の細胞が、請求項 3～5 の核酸配列のいずれかを含み、前記細胞は、請求項 1 のポリペプチドのいずれかを発現することの特徴とするトランスジェニック動物。

【請求項 35】

請求項 32 および 33 に記載のトランスジェニック動物の作成方法であって、請求項 1 の核酸配列のいずれかを細胞に導入するステップを含み、前記細胞は、前記動物に存在するか、または前記動物に生じることを特徴とする作成方法。 10

【請求項 36】

請求項 19～22 または 25～28 の方法のいずれかを使用して同定した物質。

【請求項 37】

請求項 31 または 32 のいずれかの医薬組成物を投与することを含む被検者の疾患の治療方法。

【請求項 38】

被検者における請求項 1 のポリペプチドのいずれかのレベルまたは発現を調節することによる被検者の疾患の治療方法。 20

【請求項 39】

被検者における請求項 3～5 の核酸分子のいずれかのレベルまたは活性を調節することによる前記被検者の治療方法。

【請求項 40】

前記検出は、前記細胞に關与する疾患を有する被検者から得られた細胞で行われる請求項 9 または 15 のいずれかの方法。

【請求項 41】

前記調節は、前記細胞に關与する疾患を有する被検者で行われる請求項 24 または 30 のいずれかの方法。

【請求項 42】

前記疾患は、貧血、好中球減少症または血小板減少症である請求項 39 の方法。 30

【請求項 43】

前記疾患は、貧血、好中球減少症または血小板減少症である請求項 41 の方法。

【発明の詳細な説明】

【0001】

(発明の分野)

本発明は、Gタンパク質共役受容体のスーパーファミリーに属する、新規に同定した受容体に関する。本発明はまた、受容体をコードするポリヌクレオチドにも関する。本発明はさらに、受容体媒介または関連疾患の診断および治療の標的として、受容体ポリペプチドおよびポリヌクレオチドを使用する方法に関する。本発明はさらに、診断および治療用のアゴニストおよびアンタゴニストを同定するための、受容体ポリペプチドおよびポリヌクレオチドを使用した薬物スクリーニング法に関する。本発明はさらに、受容体ポリペプチドおよびポリヌクレオチドをベースとしたアゴニストおよびアンタゴニストを包含する。本発明はさらに、受容体ポリペプチドおよびポリヌクレオチドを製造する手順に関する。 40

【0002】

(発明の背景)

Gタンパク質共役受容体

Gタンパク質共役受容体 (G-protein coupled receptor; GPCR) は、細胞内へのシグナルの伝達に關与する主要なクラスのタンパク質を構成する。GPCRは、3つの構造ドメイン、即ち、アミノ末端細胞外ドメインと、7回膜貫通セ 50

グメント、3つの細胞外ループおよび3つの細胞内ループを含む膜貫通ドメインと、カルボキシ末端細胞内ドメインとを有する。リガンドがGPCRの細胞外部分に結合すると、シグナルが細胞内に伝達され、細胞の生物学的または生理学的特性が変化する。GPCRは、Gタンパク質およびエフェクター（Gタンパク質により調節される細胞内酵素およびチャネル）と共に、細胞内二次メッセンジャーの状態を細胞外入力に連結するモジュラーシグナル伝達系の成分である。

【0003】

GPCR遺伝子および遺伝子産物は、疾病の病因の可能性がある（Spiegelら（1993）J. Clin. Invest. 92:1119-1125；McKusickら（1993）J. Med. Genet. 30:1-26）。

10

ロドプシン遺伝子およびV2バソプレッシン受容体遺伝子の特異的欠陥により、様々な形の網膜色素変性症（Nathansら（1992）Annu. Rev. Genet. 26:403-424）および腎性尿崩症（Holtzmanら（1993）Hum. Mol. Genet. 2:1201-1204）が引き起こされることが示された。これらの受容体は、中枢神経系および末梢生理学的プロセスの両方に非常に重要である。進化学的解析により、これらの受容体の祖先は、当初、複雑な体制プランおよび神経系と協奏して発達したことが示唆される。

【0004】

GPCRタンパク質スーパーファミリーは、次の5つのファミリーに分類できる。ファミリーI、ロドプシンおよび β 2-アドレナリン受容体を典型とし、現在200以上の独特なメンバーにより示される受容体（Dohlmanら（1991）Annu. Rev. Biochem. 60:653-688）。ファミリーII、副甲状腺ホルモン／カルシトニン／セレクトイン受容体ファミリー（Juppnerら（1991）Science 254:1024-1026；Linら（1991）Science 254:1022-1024）。ファミリーIII、代謝指向型グルタミン酸受容体ファミリー（Nakanishi（1992）Science 258 597:603）。ファミリーIV、D. discoideumの走化性および発達に重要なcAMP受容体ファミリー（Kleinら（1988）Science 241:1467-1472）。ファミリーV、STE2などの真菌接合フェロモン受容体（Kurjan（1992）Annu. Rev. Biochem. 61:1097-1129）。

20

30

【0005】

7つの推定疎水性部を示し、GPCRには無関係のようである他のタンパク質も少数存在する。それらはGタンパク質に結合することが示されていない。ショウジョウバエは、十分に研究されたがGPCRである証拠を示さない、光受容体特異的タンパク質である、bride of sevenless (boss)の7回膜貫通セグメントタンパク質を発現する（Hartら（1993）Proc. Natl. Acad. Sci. USA 90:5047-5051）。ショウジョウバエの遺伝子fizzled (fz)も、7回膜貫通セグメントを有するタンパク質であると考えられる。bossのように、fzは、Gタンパク質に結合することが示されていない（Vinsonら（1989）Nature 338:263-264）。

40

【0006】

Gタンパク質は、グアニンヌクレオチドに結合している、 α 、 β および γ サブユニットからなるヘテロ三量体タンパク質のファミリーを示す。これらのタンパク質は、通常、7回膜貫通セグメントを含む受容体などの細胞表面受容体に連結している。リガンドがGPCRに結合した後に、立体構造の変化がGタンパク質に伝達され、これにより α サブユニットは、結合しているGDP分子をGTPと交換し、 $\beta\gamma$ サブユニットから解離する。GTP結合形の α サブユニットは、典型的には、エフェクター調節部分として機能し、cAMP（例えば、アデニル酸シクラーゼの活性化により）、ジアシルグリセロールまたはイノシトールリン酸などの二次メッセンジャーを発生する。ヒトにおいて20以上の異なる型の α サブユニットが知られている。これらのサブユニットは、プールのより少ない β およ

50

び γ サブユニットと会合する。哺乳動物 G タンパク質の例は、G i、G o、G q、G s および G t を含む。G タンパク質は、L o d i s h ら、M o l e c u l a r C e l l B i o l o g y (S c i e n t i f i c A m e r i c a n B o o k s I n c.、ニューヨーク、N. Y.、1995) に十分に記載され、その内容は参照することによって本明細書に組み込まれる。G P C R、G タンパク質および G タンパク質連結エフェクターおよび二次メッセンジャー系は、G p r o t e i n L i n k e d R e c e p t o r F a c t B o o k、W a t s o n ら編、A c a d e m i c P r e s s (1994) に論評されている。

【0007】

プリン受容体

プリン、および特にアデノシンおよびアデニンヌクレオチドは、細胞表面受容体を介して媒介される広範な薬理効果を有する。一般的な論評については、G p r o t e i n L i n k e d R e c e p t o r F a c t B o o k におけるアデノシンおよびアデニンヌクレオチド、W a t s o n ら (編)、A c a d e m i c P r e s s 1994、p. 19-31 参照。

【0008】

A T P の作用には、平滑筋活性の調節、腸平滑筋弛緩および膀胱収縮の刺激、血管内皮から遊離された場合の A D P による血小板活性化の刺激、および中枢神経系の興奮効果を含む。アデノシンの作用には、血管拡張、気管支収縮、免疫抑制、血小板凝集阻害、心抑制、侵害受容求心神経刺激、神経伝達物質遊離阻害、前および後シナプス抑制作用、運動活性の減少、呼吸抑制、睡眠誘導、不安寛解、およびホルモンなどの因子の遊離阻害を含む。

【0009】

アデノシンおよびアデニンヌクレオチドに別個の受容体が存在する。メチルキサンチン類 (例えばテオフィリンおよびカフェイン) などのそのような類似体の臨床作用は、アデノシン受容体と拮抗することによりその作用を達成すると考えられている。アデノシンは、アデニンヌクレオチド受容体に親和性が低く、アデニンヌクレオチドは、アデノシン受容体に親和性が低い。

【0010】

A₁、A_{2A}、A_{2B}、および A₃ と称される、4つの承認されているアデノシン受容体のサブタイプが存在する。さらに、A₄ 受容体は、2-フェニルアミノアデノシンによる標識に基づいて提唱されている (C o r n f i e l d ら (1992) M o l . P h a r m a c o l . 42: 552-561)。

【0011】

P_{2X} 受容体は、A T P 作動性カチオンチャネルである (N e u r o p h a r m a c o l o g y 36 (1997) 参照)。P_{2X} 受容体の提唱されている形態は、2回膜貫通領域、大きな細胞外ループ、および細胞内 N 末端および C 末端である。

【0012】

P_{2Y} と称される多くのクローニングされた受容体が、G タンパク質共役ファミリーのメンバーであることが提唱されている。U D P、U T P、A D P および A T P がアゴニストとして同定された。現在までに、P_{2Y1-7} は特徴づけられているが、P_{2Y7} はロイコトリエン B₄ 受容体であり得ると提唱されている (Y o k o m i z o ら (1997) N a t u r e 387: 620-624)。しかし、P_{2Y1, 2, 4} および ₆ は、P_{2Y} 受容体の G タンパク質共役ファミリーのメンバーであることが広く受け入れられている。

【0013】

造血細胞系 H E L から少なくとも3つの P₂ プリン受容体が、細胞内カルシウム動員により、および光親和性標識により同定されている (A k b a r ら (1996) J . B i o c h e m . 271: 18363-18567)。

【0014】

A₁ アデノシン受容体は、アデニルシクラーゼの阻害能に鑑みて名付けられた。この受容

10

20

30

40

50

体は、心臓、脂肪、腎臓、胃および膵臓などの多くの末梢組織に分布している。それらはまた、例えば、腸および輸精管などの末梢神経にも見られる。それらは、大脳皮質、海馬、小脳、視床および線条体を含む中枢神経系、並びに、いくつかの細胞系に高いレベルで存在する。アゴニストおよびアンタゴニストは、G protein Linked Receptor Facts Bookの22ページに見られ、これは参照することによって本明細書に組み込まれるものとする。これらの受容体は、アデニルシクラーゼおよび電位依存性カルシウムチャネルを阻害し、 G_i / G_o クラスであると示唆されている百日咳毒素感受性Gタンパク質を介してカリウムチャネルを活性化すると報告されている。 A_1 受容体はまた、ホスホリパーゼCの活性化を誘導し、この経路を他の受容体が活性化する能力を増強することが報告されている。

10

【0015】

A_{2A} アデノシン受容体は、線条体、嗅結節および側坐核などの脳に見られる。末梢では、 A_2 受容体は、血管拡張、免疫抑制、血小板凝集阻害、およびグルコース生成を媒介する。アゴニストおよびアンタゴニストは、本明細書に参考として取り込んだ、上記に引用したG PROTEIN LINKED RECEPTOR FACTS BOOKの25ページに見られる。この受容体は G_s を介してアデニルシクラーゼの活性化を媒介する。

【0016】

A_{2B} 受容体は、ヒト脳およびラット腸および膀胱に存在することが示された。アゴニストおよびアンタゴニストは、参照することにより本明細書に組み込んだ、上記に引用したG protein Linked Receptor Facts Bookの27ページに議論される。この受容体は G_s を介してcAMPの刺激を媒介する。

20

【0017】

A_3 アデノシン受容体は、精巣、肺、腎臓、心臓、中枢神経系（大脳皮質、線条体および嗅球を含む）に発現される。アゴニストおよびアンタゴニストの議論は、本明細書に参考として取り込んだ、上記に引用したG PROTEIN LINKED RECEPTOR FACTS BOOKの28ページに見られる。この受容体は、 G_i / G_o クラスであると示唆されている百日咳毒素感受性Gタンパク質を介してアデニルシクラーゼの阻害を媒介する。

【0018】

P_{2Y} プリン受容体は、ATPおよびADPに類似の親和性を示すが、AMPには親和性が低い。この受容体は、taeniacaeなどの平滑筋および血管組織に見られ、ここで内皮依存性一酸化窒素の遊離により、血管拡張を誘導する。それはトリ赤血球にも示された。アゴニストおよびアンタゴニストは、本明細書に参照することにより組み込まれた、上記に引用したG PROTEIN LINKED RECEPTOR FACTS BOOKの30ページに議論される。この受容体は、 G_i / G_o クラスであると示唆されている百日咳毒素非感受性Gタンパク質を介したホスホイノシチド代謝の活性化を介して機能する。

30

【0019】

従って、GPCR、および特にプリン受容体が、薬物作用および開発の主要な標的である。従って、以前には不明であったGPCRを同定し特徴づけることは、医薬開発分野にとって価値がある。本発明は、プリン受容体に相同性を有する、以前には同定されていなかったヒトGPCRを提供することにより従来技術を前進させる。

40

【0020】

（発明の概要）

本発明の目的は、新規GPCRを同定することである。

【0021】

本発明のさらなる目的は、GPCR媒介疾患の治療および診断に適用可能な受容体アッセイにおける試薬または標的として有用である、新規GPCRポリペプチドを提供することである。

50

【0022】

本発明のさらなる目的は、GPCR媒介疾患の治療および診断に適用可能な受容体アッセイにおける標的および試薬として有用であり、組換え法による新規受容体ポリペプチドの製造に有用な新規GPCR受容体ポリペプチドに対応するポリヌクレオチドを提供することである。

【0023】

本発明の具体的な目的は、新規受容体のアゴニストおよびアンタゴニストとして作用し、新規受容体の発現を調節する化合物を同定することである。

【0024】

本発明のさらに具体的な目的は、GPCR関連疾患の治療および診断のために、受容体の発現を調節する化合物を提供することである。 10

【0025】

従って、本発明は、2838、14618および15334受容体と称される、新規GPCRの同定に基づく。

【0026】

本発明は、配列番号1に示したアミノ酸配列を有するポリペプチドを含む単離2838受容体ポリペプチドを提供する。

本発明は、配列番号3に示したアミノ酸配列を有するポリペプチドを含む単離14618受容体ポリペプチドを提供する。

本発明は、配列番号5に示したアミノ酸配列を有するポリペプチドを含む単離15334受容体ポリペプチドを提供する。 20

【0027】

本発明はまた、配列番号2に示した配列を有する単離2838受容体核酸分子を提供する。

本発明はまた、配列番号4に示した配列を有する単離14618受容体核酸分子を提供する。

本発明はまた、配列番号6に示した配列を有する15334受容体核酸分子を提供する。

【0028】

本発明はまた、配列番号1に示したアミノ酸配列に実質的に相同であるアミノ酸配列を有する変異体ポリペプチドを提供する。 30

本発明はまた、配列番号3に示したアミノ酸配列に実質的に相同であるアミノ酸配列を有する変異体ポリペプチドを提供する。

本発明はまた、配列番号5に示したアミノ酸配列に実質的に相同であるアミノ酸配列を有する変異体ポリペプチドを提供する。

【0029】

本発明はまた、配列番号2に示したヌクレオチド配列に実質的に相同である変異体核酸配列を提供する。

本発明はまた、配列番号4に示したヌクレオチド配列に実質的に相同である変異体核酸配列を提供する。

本発明はまた、配列番号6に示したヌクレオチド配列に実質的に相同である変異体核酸配列を提供する。 40

【0030】

本発明は、配列番号1に示したポリペプチド断片および配列番号2に示したヌクレオチド配列、並びに、実質的に相同なポリペプチドまたは核酸の断片を提供する。

本発明はまた、配列番号3に示したポリペプチド断片および配列番号4に示したヌクレオチド配列、並びに、実質的に相同なポリペプチドまたは核酸の断片を提供する。

本発明はまた、配列番号5に示したポリペプチド断片および配列番号6に示したヌクレオチド配列、並びに、実質的に相同なポリペプチドまたは核酸の断片を提供する。

【0031】

本発明はさらに、上記の核酸分子を含む核酸作成物を提供する。好ましい実施形態におい 50

て、本発明の核酸分子は、調節配列に作動可能に連鎖している。

【0032】

本発明はまた、受容体核酸分子およびポリペプチドを発現するためのベクターおよび宿主細胞、および特に組換えベクターおよび宿主細胞を提供する。

【0033】

本発明はまた、ベクターおよび宿主細胞の製造方法、およびそれを使用して、受容体核酸分子およびポリペプチドを調製する方法を提供する。

【0034】

本発明はまた、受容体ポリペプチドおよび断片に選択的に結合する、抗体またはその抗原結合断片を提供する。

【0035】

本発明はまた、受容体ポリペプチドまたは核酸（RNAまたはDNA）の発現または活性を調節する化合物をスクリーニングする方法を提供する。

【0036】

本発明はまた、特にスクリーニングした化合物を使用して、受容体ポリペプチドまたは核酸の発現または活性を調節する方法を提供する。受容体ポリペプチドまたは核酸の異常な活性または発現に関連した状態を治療するために調節を使用し得る。

【0037】

本発明はまた、疾病診断を含む、生物学的サンプル中での受容体ポリペプチドまたは核酸分子の有無およびレベルを決定するためのアッセイを提供する。

【0038】

本発明はまた、疾病診断を含む、受容体ポリペプチドまたは核酸分子における変異の存在を決定するためのアッセイを提供する。

【0039】

またさらなる実施形態において、本発明は、本発明の核酸およびポリペプチドのヌクレオチドおよび／またはアミノ酸配列をそれぞれ含む、コンピューターで解読可能な手段を提供する。

【0040】

（発明の詳細な説明）

受容体機能／シグナル伝達経路

2838、14618および15334受容体タンパク質は、シグナル伝達経路に関与するGPCRである。本明細書に使用した「シグナル伝達経路」は、リガンドがGPCR（2838、14618または15334タンパク質）に結合した時の、細胞機能／活性の調節（例えば刺激または阻害）を意味する。このような機能の例は、例えば、ホスファチジルイノシトール4，5-二リン酸（PIP₂）、イノシトール1，4，5-三リン酸（IP₃）およびアデニル酸シクラーゼなどの、シグナル伝達経路に関与する細胞内分子の動員；形質膜の分極；分子の生成または分泌；細胞成分の構造変化；細胞増殖、例えばDNAの合成；細胞移動；細胞分化；および細胞生存を含む。2838受容体タンパク質は、胸腺、リンパ節、脾臓、精巣、大腸および末梢血リンパ球（活性化Tヘルパー細胞-1、活性化Tヘルパー細胞-2、CD3（CD4およびCD8の両方）、活性化B細胞、および顆粒球を含むがこれに限定されない）に発現されるので、2838受容体タンパク質シグナル伝達経路に関与する細胞は、これらの組織および細胞を含むがこれに限定されない。14618受容体タンパク質は、乳房、骨格筋、甲状腺、リンパ節、脾臓および末梢血リンパ球（CD34⁺細胞、休止B細胞、および巨核球を含むがこれに限定されない）に発現されるので、14618受容体タンパク質シグナル伝達経路に関与する細胞は、これらの組織および細胞に由来する細胞を含むがこれに限定されない。15334受容体タンパク質は、大腸、脾臓、扁桃腺、リンパ節、脾臓、胸腺、副腎、心臓および末梢血細胞（図17に示したもの、巨核球、および赤血球を含むがこれに限定されない）に発現されるので、15334受容体タンパク質シグナル伝達経路に関与する細胞は、これらの組織および細胞に由来する細胞を含むがこれに限定されない。

10

20

30

40

50

【0041】

受容体タンパク質により媒介される応答は、細胞の種類により変化する。例えば、ある細胞では、リガンドの受容体タンパク質への結合は、ホスファチジルイノシトールまたはサイクリックAMP代謝および代謝回転を介して、化合物の遊離、チャネルの開閉、細胞接着、移動、分化等などの活性を刺激し得るが、別の細胞では、リガンドの結合は、異なる結果を引き起こす。受容体タンパク質により調節される細胞活性／応答に関わらず、タンパク質がGPCRであり、および、細胞において、Gタンパク質と相互作用して、様々な細胞内シグナル伝達経路で、例えば、ホスファチジルイノシトールまたはサイクリックAMP代謝および代謝回転を介して、1つ以上の二次シグナルを発生することは不変である。

10

【0042】

本明細書に使用した「ホスファチジルイノシトール代謝回転および代謝」は、ホスファチジルイノシトール4, 5-ニリン酸 (PIP_2) の代謝回転および代謝に参与する分子、並びに、これらの分子の活性を意味する。 PIP_2 は、形質膜の細胞質ゾルリーフレットに見られるリン脂質である。いくつかの細胞では、リガンドの受容体への結合により、形質膜酵素ホスホリパーゼCが活性化され、これは次いで、 PIP_2 を加水分解して、1, 2-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_3) を生成できる。一旦形成された IP_3 が、小胞体表面に拡散し得ると、そこで、 IP_3 受容体、例えば IP_3 結合部位を含むカルシウムチャネルタンパク質に結合できる。 IP_3 結合は、チャネルの開口を誘導し、カルシウムイオンを細胞質に遊離できる。 IP_3 はまた、特異的キナーゼによりリン酸化されて、イノシトール1, 3, 4, 5-四リン酸 (IP_4) を形成でき、この分子は、細胞外媒体から細胞質へのカルシウム流入を引き起こすことができる。 IP_3 および IP_4 は、その後、非常に迅速に、不活性産物であるイノシトール1, 4-ニリン酸 (IP_2) およびイノシトール1, 3, 4-三リン酸にそれぞれ加水分解され得る。これらの不活性産物は細胞によりリサイクルされて PIP_2 を合成できる。 PIP_2 の加水分解により生じる他の二次メッセンジャー、すなわち1, 2-ジアシルグリセロール (DAG) は、細胞膜に留まり、そこで作用して酵素プロテインキナーゼCを活性化できる。プロテインキナーゼCは、通常、細胞の細胞質に可溶性で見られるが、細胞内カルシウム濃度が増加すると、この酵素は、形質膜に移動し、そこで DAG により活性化され得る。異なる細胞でのプロテインキナーゼCの活性化は、グリコーゲンシンターゼのリン酸化、または様々な転写因子 (例えば $\text{NF-}\kappa\text{B}$) のリン酸化といった様々な細胞応答をもたらす。本明細書に使用した「ホスファチジルイノシトール活性」なる語は、 PIP_2 の活性またはその代謝物の活性を意味する。

20

30

【0043】

受容体が関与し得る別のシグナル伝達経路は、cAMP代謝回転経路である。本明細書に使用した「サイクリックAMP代謝回転および代謝」は、サイクリックAMP (cAMP) の代謝回転および代謝に参与する分子、並びに、これらの分子の活性を意味する。サイクリックAMPは、特定のGタンパク質共役受容体のリガンド誘導刺激に応答して産生される二次メッセンジャーである。 cAMP シグナル伝達経路では、リガンドのGPCRへの結合により、 cAMP の合成を触媒する、酵素アデニルシクラーゼが活性化され得る。新規に合成された cAMP は、次に、 cAMP 依存性プロテインキナーゼを活性化できる。この活性化キナーゼは、電位作動性カリウムチャネルタンパク質または会合タンパク質をリン酸化でき、よってカリウムチャネルは活動電位中に開口できなくなる。カリウムチャネルが開口できないことにより、カリウムの外向き流は減少し、これにより通常ニューロンの膜は再分極し、長期の膜の脱分極が起こる。

40

【0044】

ポリペプチド

本発明は、新規G共役タンパク質受容体の発見に基づく。特に、発現配列タグ ($\text{expressed sequence tag}$; EST) を、Gタンパク質共役受容体配列に対する相同性に基づいて選択した。この EST を、それが含む配列を基にしたプライマーを

50

設計するために使用し、B細胞cDNAライブラリーから2838cDNA、肝臓および脾臓cDNAライブラリーから14618cDNA、および脾臓cDNAライブラリーから15334cDNAを同定するために使用した。陽性クローンをシークエンスにかけ、重複断片を集合させた。集合した配列の解析により、3つのクローニングcDNA分子がGタンパク質共役受容体をコードすることが判明した。

【0045】

従って、本発明は、配列番号1に示した推定アミノ酸配列を有する新規GPCRに関する。

本発明はまた、配列番号3に示した推定アミノ酸配列を有する新規GPCRに関する。

本発明はまた、配列番号5に示した推定アミノ酸配列を有する新規GPCRに関する。

10

【0046】

配列表は、2838ヌクレオチド配列（配列番号2）および推定2838アミノ酸配列（配列番号1）を示す。アミノ酸1～約24番目が、アミノ末端細胞外ドメインを構成し、アミノ酸約25～292番目が、膜貫通ドメインにおよぶ領域を構成し、アミノ酸約293～319番目が、カルボキシ末端細胞内ドメインを構成すると予測される。膜貫通ドメインは、7回膜貫通セグメント、3つの細胞外ループおよび3つの細胞内ループを含む。膜貫通部は、アミノ酸約25～アミノ酸約49番目、アミノ酸約56～アミノ酸約79番目、アミノ酸約100～アミノ酸約117番目、アミノ酸約138～アミノ酸約159番目、アミノ酸約187～アミノ酸約210番目、アミノ酸約224～アミノ酸約248番目、およびアミノ酸約268～アミノ酸約292番目に見られる。全膜貫通ドメインにおよぶ領域内には、3つの細胞内ループおよび3つの細胞外ループが存在する。3つの細胞内ループが、アミノ酸約50～アミノ酸約55番目、アミノ酸約118～アミノ酸約137番目、およびアミノ酸約211～アミノ酸約223番目に見られる。3つの細胞外ループが、アミノ酸約80～アミノ酸約99番目、アミノ酸約160～アミノ酸約186番目、およびアミノ酸約249～アミノ酸約267番目に見られる。

20

【0047】

配列表は、14618ヌクレオチド配列（配列番号4）および推定14618アミノ酸配列（配列番号3）を示す。アミノ酸1～約28番目が、アミノ末端細胞外ドメインを構成し、アミノ酸約29～297番目が、膜貫通ドメインにおよぶ領域を構成し、アミノ酸約298～337番目が、カルボキシ末端細胞内ドメインを構成すると予測される。膜貫通ドメインは、7回膜貫通セグメント、3つの細胞外ループおよび3つの細胞内ループを含む。膜貫通部は、アミノ酸約29～アミノ酸約49番目、アミノ酸約60～アミノ酸約84番目、アミノ酸約103～アミノ酸約127番目、アミノ酸約142～アミノ酸約161番目、アミノ酸約194～アミノ酸約217番目、アミノ酸約231～アミノ酸約247番目、およびアミノ酸約276～アミノ酸約297番目に見られる。全膜貫通ドメインにおよぶ領域内には、3つの細胞内ループおよび3つの細胞外ループが存在する。3つの細胞内ループが、アミノ酸約50～アミノ酸約59番目、アミノ酸約128～アミノ酸約141、およびアミノ酸約218～アミノ酸約230に見られる。3つの細胞外ループが、アミノ酸約85～アミノ酸約102番目、アミノ酸約162～アミノ酸約193番目、およびアミノ酸約248～アミノ酸約275番目に見られる。

30

40

【0048】

配列表は、15334ヌクレオチド配列（配列番号6）および推定15334アミノ酸配列（配列番号5）を示す。アミノ酸1～約25番目が、アミノ末端細胞外ドメインを構成し、アミノ酸約26～299番目が、膜貫通ドメインにおよぶ領域を構成し、アミノ酸約300～372番目が、カルボキシ末端細胞内ドメインを構成すると予測される。膜貫通ドメインは、7回膜貫通セグメント、3つの細胞外ループおよび3つの細胞内ループを含む。膜貫通部は、アミノ酸約26～アミノ酸約48番目、アミノ酸約56～アミノ酸約77番目、アミノ酸約99～アミノ酸約115番目、アミノ酸約140～アミノ酸約157番目、アミノ酸約188～アミノ酸約209番目、アミノ酸約235～アミノ酸約259番目、アミノ酸約277～アミノ酸約299番目に見られる。全膜貫通ドメインにおよぶ

50

領域内には、3つの細胞内ループおよび3つの細胞外ループが存在する。3つの細胞内ループが、アミノ酸約49～アミノ酸約55番目、アミノ酸約116～アミノ酸約139番目、およびアミノ酸約210～アミノ酸約234番目に見られる。3つの細胞外ループが、アミノ酸約78～アミノ酸約98番目、アミノ酸約158～アミノ酸約187番目、およびアミノ酸約260～アミノ酸約276番目に見られる。

【0049】

「2838受容体ポリペプチド」または「2838受容体タンパク質」は、配列番号1のポリペプチドを意味する。「14618受容体ポリペプチド」または「14618受容体タンパク質」は、配列番号3のポリペプチドを意味する。「15334受容体ポリペプチド」または「15334受容体タンパク質」は、配列番号5のポリペプチドを意味する。しかし、「受容体タンパク質」または「受容体ポリペプチド」なる語は、本明細書に記載した2838、14618または15334ポリペプチドの多くの変異体、並びに、全長2838、14618または15334ポリペプチドおよび変異体から得られた断片をさらに含む。

10

【0050】

従って、本発明は、単離または精製2838、14618および15334受容体ポリペプチドおよびその変異体および断片を提供する。

【0051】

2838ポリペプチドは、3つの主要な構造ドメインを示す319残基のタンパク質である。アミノ酸末端細胞外ドメインは、配列番号1の残基1～約24番目に存在すると同定された。膜貫通ドメインは、配列番号1の約25～約292番目の残基内に存在すると同定された。カルボキシ末端細胞内ドメインは、配列番号1の約293～319番目の残基内に存在すると同定された。膜貫通ドメインは、膜におよぶ7つの部分を含む。膜貫通部は、アミノ酸約25～アミノ酸約49、アミノ酸約56～アミノ酸約79番目、アミノ酸約100～アミノ酸約117番目、アミノ酸約138～アミノ酸約159番目、アミノ酸約187～アミノ酸約210番目、アミノ酸約224～アミノ酸約248番目、およびアミノ酸約268～アミノ酸約292番目に見られる。全膜貫通ドメインにおよぶ領域内に、3つの細胞内ループおよび3つの細胞外ループが存在する。3つの細胞内ループが、アミノ酸50～アミノ酸約55番目、アミノ酸約118～アミノ酸約137番目、およびアミノ酸約211～アミノ酸約223番目に見られる。3つの細胞外ループが、アミノ酸約80～アミノ酸約99番目、アミノ酸約160～アミノ酸約186番目、およびアミノ酸約249～アミノ酸約267番目に見られる。

20

30

【0052】

膜貫通ドメインは、GPCRシグナル伝達サインであるDRFを残基118～120番目に含む。配列は、GPCRにおける不変アミノ酸である、アルギニンを残基119に含む。

【0053】

BLAST探索に基づき、最も高い相同性がプリン受容体に示された。

【0054】

14618ポリペプチドは、3つの主要な構造ドメインを示す、337残基のタンパク質である。アミノ末端細胞外ドメインが、配列番号3の残基1～約28番目内に存在すると同定される。膜貫通ドメインが、配列番号3の約29～約297番目の残基内に存在すると同定される。カルボキシ末端細胞内ドメインが、配列番号3の約298～337番目の残基内に存在すると同定される。膜貫通ドメインは、膜におよぶ7つの部を含む。膜貫通部は、アミノ酸約29～アミノ酸約49番目、アミノ酸約84～アミノ酸約60番目、アミノ酸約103～アミノ酸約127番目、アミノ酸約142～アミノ酸約161番目、アミノ酸約194～アミノ酸約217番目、アミノ酸約231～アミノ酸約247番目、およびアミノ酸約276～アミノ酸約297番目に見られる。全膜貫通ドメインにおよぶ領域内に、3つの細胞内ループおよび3つの細胞外ループが存在する。3つの細胞内ループが、アミノ酸50～アミノ酸約59番目、アミノ酸約128～アミノ酸約141番目、およ

40

50

びアミノ酸約218～アミノ酸約230番目に見られる。3つの細胞外ループが、アミノ酸約85～アミノ酸約102番目、アミノ酸約162～アミノ酸約193番目、およびアミノ酸約248～アミノ酸約275番目に見られる。

【0055】

膜貫通ドメインは、GPCRシグナル伝達サインのFRCを残基121～123番目に含む。配列は、GPCRにおける不変アミノ酸であるアルギニンを残基122に含む。

【0056】

BLAST探索に基づき、最も高い相同性がプリン受容体に示された。

【0057】

15334ポリペプチドは、3つの主な構造ドメインを示す、372残基のタンパク質である。アミノ末端細胞外ドメインが、配列番号5の残基1～約25番目内に存在すると同定される。膜貫通ドメインが、配列番号5の約26～約299番目の残基内に存在すると同定される。カルボキシ末端細胞内ドメインが、配列番号5の約300～372番目の残基内に存在すると同定される。膜貫通ドメインは、膜におよぶ7つの部を含む。膜貫通部は、アミノ酸約26～アミノ酸約48番目、アミノ酸約56～アミノ酸約77番目、アミノ酸約99～アミノ酸約115番目、アミノ酸約140～アミノ酸約157番目、アミノ酸約188～アミノ酸約209番目、アミノ酸約235～アミノ酸約259番目、およびアミノ酸約277～アミノ酸約299番目に見られる。全膜貫通ドメインにおよぶ領域内に、3つの細胞内ループおよび3つの細胞外ループが存在する。3つの細胞内ループが、アミノ酸49～アミノ酸約55番目、アミノ酸約116～アミノ酸約139番目、およびアミノ酸約210～アミノ酸約234番目に見られる。3つの細胞外ループが、アミノ酸約78～アミノ酸約98番目、アミノ酸約158～アミノ酸約187番目、およびアミノ酸約260～アミノ酸約276番目に見られる。

10

20

【0058】

膜貫通ドメインは、GPCRシグナル伝達サインのDRYを残基118～120番目に含む。配列は、GPCRにおける不変アミノ酸であるアルギニンを残基119に含む。

【0059】

BLAST探索に基づき、最も高い相同性がプリン受容体に示された。

【0060】

本明細書に使用したように、ポリペプチドは、組換えおよび非組換え細胞から単離した場合に細胞物質を実質的に含まない場合に、または化学合成した場合に化学物質前駆体または他の化学物質を実質的に含まない場合に、「単離」または「精製」されたという。しかし、ポリペプチドを、細胞中で通常会合していない別のポリペプチドと結合でき、依然として「単離」または「精製」と考えることができる。

30

【0061】

受容体ポリペプチドは精製して均一にすることができる。しかし、ポリペプチドを精製して均一にしていない調製物も有用であり、単離形のポリペプチドを含むと考えられると理解される。重要な特徴は、調製物には、かなりの量の他の成分の存在下でさえ、ポリペプチドの望む機能が可能であることである。従って、本発明は、様々な純度を包含する。

【0062】

1つの実施形態において、「細胞物質を実質的に含まない」なる語は、約30%以下（乾燥重量で）の他のタンパク質（すなわち、汚染タンパク質）、約20%以下の他のタンパク質、約10%以下の他のタンパク質、または約5%以下の他のタンパク質を有する受容体ポリペプチドの調製物を含む。受容体ポリペプチドを組換え生産した場合、それは実質的に培養培地も含まず、すなわち、培養培地は、約20%以下、約10%以下、または約5%以下の容量のタンパク質調製物を示す。

40

【0063】

受容体ポリペプチドは、それが膜調製物の一部であるか、または精製し、次いで膜小胞またはリポソームで復元される場合に単離されたと考える。

【0064】

50

「実質的に化学物質前駆体または他の化学物質を含まない」なる語は、その合成に關与する化学物質前駆体または他の化学物質から分離されている、受容体ポリペプチドの調製物を含む。1つの実施形態において、「実質的に化学物質前駆体または他の化学物質を含まない」なる語は、約30%以下（乾燥重量で）の化学物質前駆体または他の化学物質、約20%以下の化学物質前駆体または他の化学物質、約10%以下の化学物質前駆体または他の化学物質、または約5%以下の化学物質前駆体または他の化学物質の調製物を含む。

【0065】

1つの実施形態において、受容体ポリペプチドは、配列番号1に示したアミノ酸配列を含む。しかし、本発明は、配列変異体も含む。変異体は、生物の同じ遺伝子座、すなわち対立遺伝子変異体によりコードされた実質的に相同なタンパク質を含む。受容体は、WI-7921に近位の、第2染色体に位置する。変異体はまた、生物の他の遺伝子座から得られるが、配列番号1の2838受容体タンパク質に実質的に相同である、タンパク質も包含する。変異体はまた、2838受容体タンパク質に実質的に相同であるが、別の生物から得られた、すなわち相同分子種であるタンパク質も含む。変異体はまた、化学合成により製造された2838受容体タンパク質に実質的に相同なタンパク質も含む。変異体はまた、組換え法により製造された2838受容体タンパク質に実質的に相同なタンパク質も含む。しかし、変異体は、本発明より前に開示された全てのアミノ酸配列を排除すると理解する。

10

【0066】

別の実施形態において、受容体ポリペプチドは、配列番号3に示したアミノ酸配列を含む。しかし、本発明はまた、配列変異体も包含する。変異体は対立遺伝子変異体を含む。変異体はまた、生物の他の遺伝子座から得られるが、配列番号3の14618受容体タンパク質に実質的に相同なタンパク質も包含する。変異体はまた、実質的に相同な相同分子種も含む。変異体はまた、化学合成により製造された14618受容体タンパク質に実質的に相同なタンパク質も含む。変異体はまた、組換え法により製造された14618受容体タンパク質に実質的に相同なタンパク質も含む。しかし、変異体は、本発明より前に開示された全てのアミノ酸配列を排除すると理解する。

20

【0067】

別の実施形態において、受容体ポリペプチドは、配列番号5に示したアミノ酸配列を含む。しかし、本発明はまた、配列変異体も包含する。変異体は対立遺伝子変異体を含む。受容体は、SHGC-30262に近位の第12染色体に位置する。変異体はまた、生物の他の遺伝子座から得られるが、配列番号5の15334受容体タンパク質に実質的に相同なタンパク質も包含する。変異体はまた、実質的に相同な相同分子種も含む。変異体はまた、化学合成により製造された15334受容体タンパク質に実質的に相同なタンパク質も含む。変異体はまた、組換え法により製造された15334受容体タンパク質に実質的に相同なタンパク質も含む。しかし、変異体は、本発明より前に開示された全てのアミノ酸配列を排除すると理解する。

30

【0068】

本明細書に使用したように、2つのタンパク質（またはタンパク質の領域）は、アミノ酸配列が、少なくとも40~45%、45~50%、50~55%、55~60%、典型的には少なくとも70~75%、より典型的には少なくとも80~85%、最も典型的には少なくとも90~95%またはそれ以上相同性である場合に、2838タンパク質に実質的に相同である。本発明による、実質的に相同なアミノ酸配列は、以下により完全に記載したようにストリンジントな条件下で、配列番号2に示した配列の核酸配列またはその一部にハイブリッド形成する核酸配列によりコードされる。

40

【0069】

本明細書に使用したように、2つのタンパク質（またはタンパク質の領域）は、アミノ酸配列が、少なくとも35~40%、40~45%、45~50%、55~60%、典型的には少なくとも70~75%、より典型的には少なくとも80~85%、最も典型的には少なくとも90~95%またはそれ以上相同性である場合に、14618タンパク質に実

50

質的に相同である。本発明による、実質的に相同なアミノ酸配列は、以下により完全に記載したようにストリンジェントな条件下で、配列番号4に示した配列の核酸配列またはその一部にハイブリッド形成する核酸配列によりコードされる。

【0070】

本明細書に使用したように、2つのタンパク質（またはタンパク質の領域）は、アミノ酸配列が、少なくとも40～45%、45～50%、50～55%、55～60%、典型的には少なくとも70～75%、より典型的には少なくとも80～85%、最も典型的には少なくとも90～95%またはそれ以上相同性である場合に、15334タンパク質に実質的に相同である。本発明による、実質的に相同なアミノ酸配列は、以下により完全に記載したようにストリンジェントな条件下で、配列番号6に示した配列の核酸配列またはその一部にハイブリッド形成する核酸配列によりコードされる。

10

【0071】

2つのアミノ酸配列の、または2つの核酸の相同性%を決定するために、配列を、最適に比較するためにアラインメントを作成する（例えば、他のタンパク質または核酸と最適にアラインメントを作成するために、そのタンパク質または核酸の配列に、ギャップを導入できる）。次いで、対応するアミノ酸位置またはヌクレオチド位置のアミノ酸残基またはヌクレオチドを比較する。ある配列の位置が、他の配列の対応する位置において同じアミノ酸残基またはヌクレオチドにより占められる場合、分子はその位置で相同である。本明細書に使用したように、アミノ酸または核酸「相同性」は、アミノ酸または核酸「同一性」に等価である。2つの配列間の相同性%は、配列により共有される同一位置の数の関数である（すなわち、相同性%は、同一位置の数／全位置数×100に等価である）。

20

【0072】

本発明はまた、同一度は低い、十分な類似性を有し、よって2838、14618または15334ポリペプチドにより実施される同機能の1つ以上を実施する、ポリペプチドを包含する。類似性は、保存されたアミノ酸置換により決定される。かかる置換は、ポリペプチド中のあるアミノ酸を、類似の特徴をもつ別のアミノ酸に置換するものである。保存的置換は、表現型的にサイレントのようである。保存的置換として典型的なものは、脂肪族アミノ酸Ala、Val、Leu、およびIle間の互いの置換であり；ヒドロキシル残基SerおよびThrの交換、酸性残基AspおよびGluの交換、アミド残基AsnおよびGln間の置換、塩基性残基LysおよびArgの交換、並びに、芳香族残基Phe、Tyr間の置換である。どのアミノ酸変化が、表現型的にサイレントであるかに関する指針はBowieら（1990）Science 247:1306-1310に見られる。

30

【0073】

【表1】

表1 保存的アミノ酸置換

芳香族	フェニルアラニン トリプトファン チロシン
疎水性	ロイシン イソロイシン バリン
極性	グルタミン アスパラギン
塩基性	アルギニン リジン ヒスチジン
酸性	アスパラギン酸 グルタミン酸
小さい	アラニン セリン トレオニン メチオニン グリシン

10

20

30

40

50

【0074】

同一性および類似性の両方が、容易に計算できる（コンピューター分子生物学、Lesk, A. M. 編、オックスフォード大学出版、ニューヨーク、1988；バイオコンピューティング；情報科学およびゲノムプロジェクト、Smith, D. W. 編、Academic Press、ニューヨーク、1993；配列データのコンピューター解析、第1部、Griffin, A. M. およびGriffin, H. G. 編、Humana Press、ニュージャージー、1994；分子生物学の配列解析、von Heinje, G.、Academic Press、1987；および配列解析プライマー、Grishkov, M. およびDevereux, J. 編、M Stockton Press、ニューヨーク、1991）。

【0075】

かかる数学的アルゴリズムの好ましい非制限的な例は、Karlinら（1993）Proc. Natl. Acad. Sci. USA 90:5873-5877に記載されている。かかるアルゴリズムは、Altschul（1997）Nucleic Acids Res. 25:3389-3402に記載のように、NBLASTおよびXBLAST（バージョン2.0）に取り込まれている。BLASTおよびギャップを入れたBLASTプログラムを使用する場合、それぞれのプログラム（例えばNBLAST）のデフォルトパラメータを使用できる。<http://www.ncbi.nlm.nih.gov> 参照。1つの実施形態において、配列比較のパラメータは、スコア=100、語長=12で設定できるか、または不変である（例えば、W=5またはW=20）。

【0076】

好ましい実施形態において、2つのアミノ酸配列間の同一性%は、B L O S U M 6 2 マトリックスまたはP A M 2 5 0 マトリックス、およびギャップ負荷16、14、12、10、8、6または4および長さ負荷1、2、3、4、5または6を使用して、G C G ソフトウェアパッケージ (<http://www.gcg.com>で入手可能) のG A P プログラムに取り込んだN e e d l e m a n ら (1970) (J. Mol. Biol. 48: 444-453) のアルゴリズムを使用して決定する。さらに別の好ましい実施形態において、2つのヌクレオチド配列間の同一性%は、N W S g a p d n a . C M P マトリックスおよびギャップ負荷40、50、60、70または80および長さ負荷1、2、3、4、5または6を使用して、G C G ソフトウェアパッケージ (D e v e r e u x ら (1984) N u c l e i c A c i d s R e s . 12 (1): 387) (<http://www.gcg.com>で入手可能) のG A P アルゴリズムを使用して決定する。 10

【0077】

配列の比較に使用する数学的アルゴリズムの別の好ましい非制限的な例は、M y e r s およびM i l l e r、C A B I O S (1989) のアルゴリズムである。かかるアルゴリズムは、G C C 配列アラインメントソフトウェアパッケージの一部であるA L I G N (バージョン2.0) に取り込まれている。A L I G N プログラムをアミノ酸配列の比較に使用する場合、P A M 1 2 0 負荷残基テーブル、ギャップ長さペナルティー12、およびギャップペナルティー4を使用できる。配列解析のさらなるアルゴリズムが当分野で公知であり、T o r e l l i s ら (1994) C o m p u t . A p p l . B i o s c i . 10: 3-5 に記載のA D V A N C E およびA D A M ; およびP e a r s o n ら (1988) P N A S 85: 2444-8 に記載のF A S T A を含む。 20

【0078】

変異体ポリペプチドは、アミノ酸配列において、1つ以上の置換、欠失、挿入、逆位、融合、および切断短縮、或いはこれらの任意の組合せが異なり得る。

【0079】

変異体ポリペプチドは、完全に機能し得るか、または1つ以上の活性が欠失し得る。従って、今回の場合では、変異は、例えば、リガンド結合、膜結合、Gタンパク質結合およびシグナル伝達に対応する1つ上の領域の機能に影響を及ぼし得る。

【0080】

完全に機能する変異体は、典型的には、保存的変異、すなわち重要でない残基または重要でない領域の変異のみを含む。機能的変異体はまた、機能を全く変化させないかまたは取るに足りない変化をもたらす、類似アミノ酸の置換を含み得る。別に、かかる置換は、ある程度機能に正または負に影響を及ぼし得る。 30

【0081】

非機能的変異体は、典型的には、1つ以上の非保存的アミノ酸置換、欠失、挿入、逆位、または切断短縮、或いは重要な残基または重量な領域の置換、挿入、逆位、または欠失を含む。

【0082】

示したように、変異体は、天然型であっても、受容体ポリペプチドに有用で新規な特徴を付与するために、組換え手段または化学合成により製造してもよい。これは、タンパク質凝集を防ぐことにより、医薬製剤の免疫原性を防ぐことを含む。 40

【0083】

有用な変異はさらに、リガンド結合特徴の変化を含む。例えば、1つの実施形態は、リガンドの結合をもたらすが、遊離または遅延放出はもたらさない、結合部位での変化を含む。同じ部位でのさらに有用な変異は、リガンドへのより高い親和性をもたらし得る。有用な変異はまた、別のリガンドへのより高い親和性を与える変化を含む。別の有用な変異は、リガンドへの結合は可能であるが、リガンドによる活性化は防ぐものである。別の有用な変異は、適切なGタンパク質による結合の減少または増加、または、受容体が通常会合しているGタンパク質とは異なるGタンパク質による結合を与える、膜貫通Gタンパク質結合/シグナル伝達ドメインの変異を含む。別の有用な変異は、1つ以上のドメインまた 50

は小領域が、作動可能に、別の G タンパク質共役受容体の 1 つ以上のドメインまたは小領域に融合している、融合タンパク質を提供する。

【0084】

機能に必須なアミノ酸は、部位特異的変異誘発またはアラニン走査型変異誘発などの、当分野で公知の方法により同定できる (Cunninghamら (1989) Science 244:1081-1085)。後者の手順は、分子のどの残基にも 1 つのアラニン変異を導入する。次いで、得られた変異分子を、受容体結合または *in vitro* または *in vivo* 増殖活性などの生物活性について試験する。リガンド-受容体結合に重要な部位はまた、結晶化、核磁気共鳴または光親和性標識などの構造解析により決定できる (Smithら (1992) J. Mol. Biol. 224:899-904; de Vosら (1992) Science 255:306-312)。

10

【0085】

実質的な相同性は、全核酸またはアミノ酸配列またはこれらの配列の断片にあり得る。

【0086】

従って、本発明は、2838 受容体タンパク質のポリペプチド断片を含む。断片は、配列番号 1 に示したアミノ酸配列から得られ得る。しかし、本発明はまた、本明細書に記載した 2838 受容体タンパク質の変異体の断片も包含する。

【0087】

従って、本発明はまた、14618 受容体タンパク質のポリペプチド断片も含む。断片は、配列番号 3 に示したアミノ酸配列から得られ得る。しかし、本発明はまた、本明細書に記載した 14618 受容体タンパク質の変異体の断片も包含する。

20

【0088】

従って、本発明はまた、15334 受容体タンパク質のポリペプチド断片も含む。断片は、配列番号 5 に示したアミノ酸配列から得られ得る。しかし、本発明はまた、本明細書に記載した 15334 受容体タンパク質の変異体の断片も包含する。

【0089】

しかし、本発明が関する断片は、本発明より前に開示され得る断片を含むとは捉えない。

【0090】

断片は、例えば、G タンパク質またはリガンドへの結合能などのタンパク質の 1 つ以上の生物活性、並びに、受容体抗体を作成するために免疫原として使用できる断片も保持し得る。

30

【0091】

2838 受容体 (例えば、8、12、15、20、30、35、36、37、38、39、40、50、100 またはそれ以上のアミノ酸長であるペプチド) の生物学的に活性な断片は、例えば、細胞外または細胞内ドメインまたはループ、1 つ以上の膜貫通部、またはその一部、G タンパク質結合部位、または GPCR サイン、グリコシル化部位、プロテインキナーゼ C およびカゼインキナーゼ II リン酸化部位、N-ミリスチル化およびアミド化部位などのドメインまたはモチーフを含む。かかるドメインまたはモチーフは、慣用的なコンピューター化相同探索手順により同定できる。

【0092】

可能な断片は、1) 配列番号 1 のアミノ酸 1 ~ アミノ酸約 24 番目の全アミノ末端細胞外ドメインを含む可溶性断片、またはその一部、2) 配列番号 1 のアミノ酸約 293 ~ アミノ酸 319 番目の全カルボキシ末端細胞内ドメインを含むペプチド、またはその一部、3) アミノ酸約 25 ~ アミノ酸約 292 番目の全膜貫通ドメインにおよぶ領域を含むペプチド、またはその一部、4) アミノ酸約 25 ~ アミノ酸約 49 番目、アミノ酸約 56 ~ アミノ酸約 79 番目、アミノ酸約 100 ~ アミノ酸約 117 番目、アミノ酸約 138 ~ アミノ酸約 159 番目、アミノ酸約 187 ~ アミノ酸約 210 番目、アミノ酸約 224 ~ アミノ酸約 248 番目、およびアミノ酸約 268 ~ アミノ酸約 292 番目の特定の膜貫通部のいずれか、またはその一部、5) アミノ酸約 50 ~ アミノ酸約 55 番目、アミノ酸約 118 ~ アミノ酸約 137 番目、アミノ酸約 211 ~ アミノ酸約 223 番目、アミノ酸約 80 ~

40

50

アミノ酸約 99 番目、アミノ酸約 160～アミノ酸約 186 番目、およびアミノ酸約 249～アミノ酸約 267 番目の 3 つの細胞内または 3 つの細胞外ループのいずれか、またはその一部を含むがこれに限定されない。断片はさらに、1 つ以上の膜貫通部および添付細胞外または細胞内ループまたは 1 つ以上の膜貫通部、および添付細胞内または細胞外ループとカルボキシ末端ドメインと組合せたアミノ末端ドメインなどの、上記断片の組合せを含む。従って、任意の上記断片を組合せることができる。他の断片は、アミノ酸約 6～319 番目の成熟タンパク質を含む。他の断片は、リン酸化部位、グリコシル化部位、およびミリスチル化部位および GPCR サイン配列を含む配列などの、本明細書に記載の様々な機能的部位を含む。断片は、例えば、機能的部位から一方または両方向に伸長して、5、10、15、20、30、40、50 または 100 までのアミノ酸を包含できる。さらに、断片は、上記の特定のドメインの小断片を含み得、この小断片は、それが由来するドメインの機能を保持する。断片はまた、アミノ酸 1～アミノ酸約 264 までの 7 つ以上のアミノ酸のアミノ酸配列を含む。断片はまた、抗原性断片および特に図 2 の高い抗原性係数を有することが示された断片を含む。

10

【0093】

従って、可能な断片は、リガンド結合部位を規定する断片、グリコシル化部位を規定する断片、膜会合を規定する断片、リン酸化部位を規定する断片、G タンパク質との相互作用およびシグナル伝達を規定する断片、およびミリスチル化部位を規定する断片を含む。これにより、関連した機能を提供するまたは関連した機能を同定できる別個の断片が意図される。好ましい実施形態において、断片はリガンド結合部位を含む。

20

【0094】

14618 受容体の生物学的に活性な断片（例えば、9、12、15、20、30、35、36、37、38、39、40、50、100 またはそれ以上のアミノ酸長であるペプチド）は、例えば、細胞外または細胞内ドメインまたはループ、1 つ以上の膜貫通部、またはその一部、G タンパク質結合部位、または GPCR サイン、グリコシル化部位、および cAMP および cGMP 依存性プロテインキナーゼ C、およびカゼインキナーゼ II リン酸化部位などのドメインまたはモチーフを含み得る。

【0095】

可能な断片は、1) 配列番号 3 のアミノ酸約 1～アミノ酸約 28 番目の全アミノ末端細胞外ドメインを含む可溶性ペプチド、またはその一部、2) 配列番号 3 のアミノ酸約 298～アミノ酸 337 番目の全カルボキシ末端細胞内ドメインを含むペプチド、またはその一部、3) アミノ酸約 29～アミノ酸約 297 番目の全膜貫通ドメインにおよぶ領域を含むペプチド、またはその一部、4) アミノ酸約 29～アミノ酸約 49 番目、アミノ酸約 60～アミノ酸約 84 番目、アミノ酸約 103～アミノ酸約 127 番目、アミノ酸約 142～アミノ酸約 161 番目、アミノ酸約 194～アミノ酸約 217 番目、アミノ酸約 231～アミノ酸約 247 番目、およびアミノ酸約 276～アミノ酸約 297 番目の特定の膜貫通部のいずれか、またはその一部、5) アミノ酸約 50～アミノ酸約 59 番目、アミノ酸約 128～アミノ酸約 141 番目、アミノ酸約 218～アミノ酸約 230 番目、アミノ酸約 85～アミノ酸約 102 番目、アミノ酸約 162～アミノ酸約 193 番目、およびアミノ酸約 248～アミノ酸約 275 番目の 3 つの細胞内または 3 つの細胞外ループのいずれか、またはその一部を含むがこれに限定されない。断片はさらに、1 つ以上の膜貫通部および添付細胞外または細胞内ループまたは 1 つ以上の膜貫通部、および添付細胞内または細胞外ループとカルボキシ末端ドメインと組合せたアミノ末端ドメインなどの、上記断片の組合せを含む。従って、任意の上記断片を組合せることができる。他の断片は、アミノ酸約 6～337 番目の成熟タンパク質を含む。他の断片は、リン酸化部位、グリコシル化部位、および GPCR サイン配列を含む配列などの、本明細書に記載の様々な機能的部位を含む。断片は、例えば、機能的部位から一方または両方向に伸長して、5、10、15、20、30、40、50 または 100 個までのアミノ酸を包含できる。さらに、断片は、上記の特定のドメインの小断片を含み得、この小断片は、それが由来するドメインの機能を保持する。断片はまた、アミノ酸 1～アミノ酸約 244 番目のうち 8 つ以上のアミノ酸

30

40

50

のアミノ酸配列を含む。

【0096】

断片はまた、抗原性断片および特に図2の高い抗原性係数を有することが示された断片を含む。

【0097】

従って、可能な断片は、リガンド結合部位を規定する断片、グリコシル化部位を規定する断片、膜会合を規定する断片、リン酸化部位を規定する断片、Gタンパク質との相互作用およびシグナル伝達を規定する断片、およびミリスチル化部位を規定する断片を含む。これにより、関連した機能を提供するまたは関連した機能を同定できる別個の断片が意図される。好ましい実施形態において、断片はリガンド結合部位を含む。

10

【0098】

生物学的に活性な断片（例えば、8、12、15、20、30、35、36、37、38、39、40、50、100個またはそれ以上のアミノ酸長であるペプチド）は、例えば、細胞外または細胞内ドメインまたはループ、1つ以上の膜貫通部、またはその一部、Gタンパク質結合部位、またはGPCRサイン、グリコシル化部位、cAMP、cGMP、プロテインキナーゼC、およびカゼインキナーゼIIリン酸化部位、およびN-ミリスチル化部位などのドメインまたはモチーフを含み得る。

【0099】

可能な断片は、1) 配列番号5のアミノ酸約1～アミノ酸約23番目の全アミノ末端細胞外ドメインを含む可溶性ペプチド、またはその一部、2) 配列番号5のアミノ酸約300～アミノ酸372番目の全カルボキシ末端細胞内ドメインを含むペプチド、またはその一部、3) アミノ酸約26～アミノ酸約299番目の全膜貫通ドメインにおよぶ領域を含むペプチド、またはその一部、4) アミノ酸約26～アミノ酸約48番目、アミノ酸約56～アミノ酸約77番目、アミノ酸約99～アミノ酸約115番目、アミノ酸約140～アミノ酸約157番目、アミノ酸約188～アミノ酸約209番目、アミノ酸約235～アミノ酸約259番目、およびアミノ酸約277～アミノ酸約299番目の特定の膜貫通部のいずれか、またはその一部、5) アミノ酸約49～アミノ酸約55番目、アミノ酸約78～アミノ酸約98番目、アミノ酸約116～アミノ酸約139番目、アミノ酸約158～アミノ酸約187番目、アミノ酸約210～アミノ酸約234番目、およびアミノ酸約260～アミノ酸約276番目の3つの細胞内または3つの細胞外ループのいずれか、またはその一部を含むがこれに限定されない。断片はさらに、1つ以上の膜貫通部および添付細胞外または細胞内ループまたは1つ以上の膜貫通部、および添付細胞内または細胞外ループとカルボキシ末端ドメインと組合せたアミノ末端ドメインなどの、上記断片の組合せを含む。従って、任意の上記断片を組合せることができる。他の断片は、アミノ酸約6～372番目の成熟タンパク質を含む。他の断片は、リン酸化部位、グリコシル化部位、ミリスチル化部位およびGPCRサイン配列を含む配列などの、本明細書に記載の様々な機能的部位を含む。断片は、例えば、機能的部位から一方または両方向に伸長して、5、10、15、20、30、40、50または100個までのアミノ酸を包含できる。さらに、断片は、上記の特定のドメインの小断片を含み得、この小断片は、それが由来するドメインの機能を保持する。断片はまた、7つ以上のアミノ酸のアミノ酸配列を含む。

20

30

40

【0100】

断片はまた、抗原性断片および特に図2の高い抗原性係数を有することが示された断片を含む。

【0101】

これらの領域は、コンピューター化相同性解析を含む公知の方法により同定できる。

【0102】

従って、可能な断片は、リガンド結合部位を規定する断片、グリコシル化部位を規定する断片、膜会合を規定する断片、リン酸化部位を規定する断片、Gタンパク質との相互作用およびシグナル伝達を規定する断片、およびミリスチル化部位を規定する断片を含む。これにより、関連した機能を提供するまたは関連した機能を同定できる別個の断片が意図

50

される。好ましい実施形態において、断片はリガンド結合部位を含む。

【0103】

本発明はまた、免疫原性特性を有する2838受容体断片を提供する。これらは、2838受容体タンパク質および変異体のエピトープ含有部分を含む。これらのエピトープ含有ペプチドは、受容体ポリペプチドまたは領域または断片に特異的に結合する抗体を生じるのに有用である。これらのペプチドは、少なくとも8、12、少なくとも14、または少なくとも15～約30個のアミノ酸を含み得る。

【0104】

本発明はまた、免疫原性特性を有する14618受容体断片を提供する。これらは、14618受容体タンパク質および変異体のエピトープ含有部分を含む。これらのエピトープ含有ペプチドは、受容体ポリペプチドまたは領域または断片に特異的に結合する抗体を生じるのに有用である。これらのペプチドは、少なくとも9、12、少なくとも14、または少なくとも約15～約30個のアミノ酸を含み得る。

10

【0105】

本発明はまた、免疫原性特性を有する15334受容体断片を提供する。これらは、15334受容体タンパク質および変異体のエピトープ含有部分を含む。これらのエピトープ含有ペプチドは、受容体ポリペプチドまたは領域または断片に特異的に結合する抗体を生じるのに有用である。これらのペプチドは、少なくとも8、12、少なくとも14、または少なくとも約15～約30個のアミノ酸を含み得る。

【0106】

抗体の産生に使用できる抗原性ポリペプチドの非制限的な例は、アミノ末端細胞外ドメインまたは任意の細胞外ループから得られるペプチドを含む。高い抗原性係数を有する領域を、図2、6および10に示す。

20

【0107】

エピトープ含有受容体およびポリペプチドは、任意の慣用的な手段により製造し得る (Houghten (1985) Proc. Natl. Acad. Sci. USA 82: 5131-5135)。同時多ペプチド合成が米国特許第4,631,211号に記載されている。

【0108】

断片は別個 (他のアミノ酸またはポリペプチドに融合されていない) であっても、大きなポリペプチド内にあってもよい。さらに、数個の断片が、1つのより大きなポリペプチド内に含まれ得る。1つの実施形態において、宿主での発現に設計した断片は、受容体断片のアミノ末端に融合した異種プレポリペプチドおよびプロポリペプチド領域および断片のカルボキシ末端に融合した追加の領域を有し得る。

30

【0109】

従って、本発明は、キメラまたは融合タンパク質を提供する。これらは、受容体タンパク質に実質的に相同ではないアミノ酸配列を有する異種タンパク質に作動可能に連鎖した受容体タンパク質を含む。「作動可能に連鎖した」は、受容体タンパク質および異種タンパク質がインフレーション融合することを示す。異種タンパク質を、受容体タンパク質のN末端またはC末端に融合できる。

40

【0110】

1つの実施形態において、融合タンパク質は、それ自体受容体の機能に影響を及ぼさない。例えば、融合タンパク質は、受容体配列がGST配列のC末端に融合している、GST融合タンパク質であり得る。他の種類の融合タンパク質は、酵素的融合タンパク質、例えばβ-ガラクトシダーゼ融合物、酵母二重ハイブリッド融合物、ポリHis融合物およびIg融合物を含むがこれに限定されない。かかる融合タンパク質、特にポリHis融合物は、組換え受容体タンパク質の精製を容易にし得る。ある宿主細胞 (例えば哺乳動物宿主細胞) で、タンパク質の発現および/または分泌は、異種シグナル配列の使用により増加し得る。それ故、別の実施形態において、融合タンパク質は、異種シグナル配列をそのN末端に含む。

50

【0111】

EP-A-0464 533は、様々な割合の免疫グロブリン定常領域を含む融合タンパク質を開示する。Fcは、治療および診断に有用であり、従って、例えば、薬物動態特性は向上する（EP-A-0232 262）。薬物発見において、例えば、ヒトタンパク質を、アンタゴニストを同定するためのハイスループットスクリーニングアッセイの目的で、Fcタンパク質と融合させる。Bennettら（1995）J. Mol. Recogn. 8:52-58およびJohanssonら（1995）J. Biol. Chem. 270:16:9459-9471。従って、本発明はまた、受容体ポリペプチド、並びに、様々な割合の様々なサブクラス（IgG、IgM、IgA、IgE）の免疫グロブリンの重鎖または軽鎖の定常領域を含む、可溶性融合タンパク質を包含する。免疫グロブリンとして好ましいのは、ヒトIgG、特にIgG1の重鎖の定常部であり、ここでは融合はヒンジ領域で起こる。ある用途では、例えば、融合タンパク質を免疫化の抗原として使用する場合などでは、融合タンパク質を意図する目的に使用した後にFcを除去することが望ましい。特定の実施形態において、Fc部は、切断配列（これも取り込まれている）により単純に除去でき、およびXa因子で切断できる。

10

【0112】

キメラまたは融合タンパク質は、標準的な組換えDNA技法により製造できる。例えば、異なるタンパク質配列をコードするDNA断片を、慣用的な技法に従って、共にインフレイムで連結する。別の実施形態において、融合遺伝子は、自動DNA合成器を含む慣用的な技法により合成できる。別に、PCRによる遺伝子断片の増幅は、アンカープライマーを使用して実施でき、これは2つの連続的な遺伝子断片間の相補的なオーバーハングをもたらし、これは続いてアニーリングし、再増幅してキメラ遺伝子配列を作成できる（Ausubelら（1992）分子生物学の現在のプロトコル参照）。さらに、すでに融合部分（例えばGSTタンパク質）をコードする、多くの発現ベクターが、市販されている。受容体タンパク質をコードする核酸を、発現ベクターにクローニングして、融合部分を受容体タンパク質にインフレイムで連結できる。

20

【0113】

別の形の融合タンパク質は、受容体機能に直接影響を及ぼすものである。従って、受容体ポリペプチドは、本発明により包含され、ここで1つ以上の受容体ドメイン（またはその一部）が、別のGタンパク質共役受容体または別の種類の受容体由来の相同的ドメイン（またはその一部）により置換されている。従って、様々な順列が可能である。アミノ末端細胞外ドメイン、またはその小領域（例えばリガンド結合）は、別のリガンド結合受容体タンパク質由来のドメインまたは小領域と置換できる。別に、全膜貫通ドメイン、または7つの部またはループのいずれか、またはその一部、例えば、Gタンパク質結合/シグナル伝達を置換できる。最後に、カルボキシ末端細胞内ドメインまたは小領域を置換できる。従って、1つ以上の天然ドメインまたは小領域が置換されたキメラ受容体を形成できる。

30

【0114】

単離された2838受容体タンパク質は、リンパ節、胸腺、脾臓、精巣、大腸、および末梢血リンパ球由来などの、それを天然に発現する細胞から、並びに、活性化Tヘルパー細胞（1および2）、低酸素Hep3B細胞、CD3細胞（CD4およびCD8の両方）、活性化B細胞、ジャーカット細胞などの、図15に示したように有意に発現されている細胞から精製でき、とりわけ、それを発現するように変化させた（組換え）細胞から精製できるか、または既知のタンパク質合成法を使用して合成できる。

40

【0115】

単離された14618受容体タンパク質は、乳房、骨格筋、リンパ節、脾臓および血液末梢リンパ球などの、それを天然に発現する細胞、並びに、それを発現するように変化させた（組換え）細胞から精製した細胞CD34⁺細胞および巨核球を含むがこれに限定されない、遺伝子が図13および14に示したように有意に発現されている組織から精製できるか、または既知のタンパク質合成法を使用して合成できる。

50

【0116】

単離された15334受容体タンパク質は、大腸、胎盤、膵臓、扁桃腺、リンパ節、脾臓、末梢血細胞、胸腺、副腎および心臓などの、それを天然に発現する細胞、並びに、それを発現するように変化させた（組換え）細胞から精製した、K562細胞、赤芽球、および巨核球を含む、図16～19に示した組織から精製できる、または既知のタンパク質合成法を使用して合成できる。

【0117】

1つの実施形態において、タンパク質は、組換えDNA技法により製造される。例えば、受容体ポリペプチドをコードする核酸分子を、発現ベクターにクローニングし、発現ベクターを、宿主細胞に導入し、タンパク質を宿主細胞で発現させる。次いで、タンパク質を細胞から、適切な精製スキームにより、標準的なタンパク質精製技法を使用して単離できる。

10

【0118】

ポリペプチドはしばしば、20の天然アミノ酸と一般的に称される20のアミノ酸以外のアミノ酸も含む。さらに、多くのアミノ酸（末端アミノ酸を含む）は、天然プロセス（例えばプロセッシングおよび他の翻訳後修飾）または当分野で公知の化学修飾技法により修飾され得る。ポリペプチドで天然で起こる一般的な修飾は、基本的な教科書、詳細な研究書および研究文献に記載され、それらは当業者に公知である。

【0119】

従って、ポリペプチドはまた、置換アミノ酸残基が、遺伝子コードによりコードされていないものである、置換基が含まれない、成熟ポリペプチドが別の化合物（例えばポリペプチドの半減期を増加させる化合物（例えばポリエチレングリコール））と融合していない、または、追加のアミノ酸が成熟ポリペプチド（リーダーまたは分泌配列または成熟ポリペプチドまたはプロタンパク質配列の精製用の配列）に融合している、誘導体または類似体も包含する。

20

【0120】

既知の修飾は、アセチル化、アシル化、ADP-リボシル化、アミド化、フラビンの共有結合的付着、ヘム部分の共有結合的付着、ヌクレオチドまたはヌクレオチド誘導体の共有結合的付着、脂質または脂質誘導体の共有結合的付着、ホスファチジルイノシトールの共有結合的付着、架橋、環化、ジスルフィド結合形成、ジメチル化、共有結合的架橋の形成、シスチンの形成、ピログルタメートの形成、ホルミル化、 γ カルボキシル化、グリコシル化、GPIアンカー形成、ヒドロキシル化、ヨウ素化、メチル化、ミリストイル化、酸化、タンパク質分解的プロセッシング、リン酸化、プレニル化、ラセミ化、セレノイル化、硫化、トランスファーRNAの媒介するアミノ酸のタンパク質への付加、例えばアルギニル化、およびユビキチン結合を含むがこれに限定されない。

30

【0121】

かかる修飾は、当業者には公知であり、科学文献に非常に詳細に記載されている。数個の特に一般的な修飾、グリコシル化、脂質付着、硫化、グルタミン酸残基の γ -カルボキシル化、ヒドロキシル化およびADP-リボシル化は、例えば、最も基本的な教科書、例えば、タンパク質—構造的および分子的特性、第2版、T. E. Creighton、W. H. Freeman and Company、ニューヨーク（1993）に記載されている。多くの詳細な論評が、この主題に関して、例えば、World, F、翻訳後共有結合的タンパク質修飾、B. C. Johnson編、Academic Press、ニューヨーク1-12（1983）；Seifterら（1990）Meth. Enzymol. 182: 626-646）およびRattanら（1992）Ann. N. Y. Acad. Sci. 663: 48-62により入手できる。

40

【0122】

これも公知であるように、ポリペプチドは必ずしも全体的に直鎖状ではない。例えば、ポリペプチドは、ユビキチン結合の結果として分枝し得、および、それらは一般的に、天然プロセッシング事象および天然に起こらないヒトによる操作によりもたらされる事象を含む

50

、翻訳後事象の結果として、分枝を伴いまたは伴わずに、環状であり得る。環状、分枝および分枝環状ポリペプチドは、非翻訳天然プロセスにより、および合成法により合成され得る。

【0123】

修飾は、ペプチド骨格、アミノ酸側鎖およびアミノまたはカルボキシ末端を含む、ポリペプチド中のいずれにおいても起こり得る。ポリペプチド中のアミノまたはカルボキシ基、または両方の、共有結合的修飾による遮断は、天然および合成ポリペプチドにおいて一般的である。例えば、タンパク質分解プロセッシングの前の、E. コリ中で作成されたポリペプチドのアミノ末端残基は、ほとんど不変的にN-ホルミルメチオニンである。

【0124】

修飾は、タンパク質の調製法の関数であり得る。組換えポリペプチドでは、例えば、修飾は、宿主細胞の翻訳後修飾能およびポリペプチドのアミノ酸配列の修飾シグナルにより決定される。従って、グリコシル化を望む場合、ポリペプチドは、グリコシル化宿主、一般には真核細胞で発現すべきである。昆虫細胞はしばしば、哺乳動物細胞と同じ翻訳後グリコシル化を行うので、この理由から、昆虫細胞発現系は、天然パターンのグリコシル化を有する哺乳動物タンパク質を効率的に発現するように開発されている。類似の考慮が他の修飾にも適用される。

10

【0125】

同じ種類の修飾が、同程度または異なる程度で、あるポリペプチドの数個の部位に存在し得る。また、あるポリペプチドは、1種類以上の修飾を含み得る。

20

【0126】

ポリペプチドの使用

受容体ポリペプチドは、2838、14618および15334受容体タンパク質、領域または断片に特異的な抗体を製造するのに有用である。高い抗原性係数スコアを有する領域を図2、6および10に示す。

【0127】

受容体ポリペプチド、変異体および断片（本発明より前に開示され得るものを含む）は、GPCRに関連した生物学的アッセイに有用である。かかるアッセイは、GPCR関連状態の診断および治療に有用な任意の既知のGPCR機能または活性または特性を含む。

【0128】

受容体ポリペプチドはまた、薬物スクリーニングアッセイ、細胞を基にしたまたは細胞非含有系に有用である。細胞を基にした系は、生検または細胞培養液中に増殖した、天然細胞であり得、すなわち、受容体タンパク質を通常発現する細胞であり得る。しかし、1つの実施形態において、細胞を基にしたアッセイは、受容体タンパク質を発現する組換え宿主細胞を含む。

30

【0129】

化合物のポリペプチドとの相互作用能の決定はまた、リガンドまたは生物学的に活性なその一部の、ポリペプチドへの結合能と比較した、試験化合物のポリペプチドへの優先的結合能を決定することを含み得る。

【0130】

ポリペプチドは、受容体活性を調節する化合物の同定に使用できる。かかる化合物は、例えば、既知のリガンドへの親和性または結合率を増加または減少させ得るか、リガンドと受容体への結合について競合し得るか、または受容体に結合したリガンドを置換し得る。2838、14618および15334タンパク質および適切な変異体および断片をハイスループットスクリーンに使用して、候補化合物を受容体への結合能についてアッセイできる。これらの化合物はさらに、受容体活性に対する化合物の効果を決定するために、機能的受容体に対してスクリーニングできる。所望の程度まで受容体を活性化（アゴニスト）または不活性化（アンタゴニスト）する化合物を同定できる。調節法は、*in vitro*で（例えば物質と共に細胞を培養することにより）または別に、*in vivo*で（例えば、物質を被検者に投与することにより）実施できる。例は、上記で議論したような

40

50

プリン類似体を含む。

【0131】

受容体ポリペプチドを使用して、受容体タンパク質と、受容体タンパク質と通常相互作用する標的分子の間の相互作用を刺激または阻害する能力について化合物をスクリーニングできる。標的は、受容体タンパク質が通常相互作用する、リガンドまたはシグナル経路の成分であり得る（例えば、cAMPまたはホスファチジルイノシトール代謝回転および／またはアデニル酸シクラーゼ、またはホスホリパーゼC活性化に関与する、Gタンパク質または他の相互作用物質）。アッセイは、受容体タンパク質または断片が標的分子と相互作用でき、タンパク質と標的の間の複合体形成を検出できるか、またはGタンパク質リン酸化、サイクリックAMPまたはホスファチジルイノシトール代謝回転、およびアデニル酸シクラーゼまたはホスホリパーゼC活性化などのシグナル伝達の任意の会合効果といった、受容体タンパク質と標的の間の相互作用の生化学的結果を検出できる条件下で、受容体タンパク質と候補化合物を合わせる段階を含む。

10

20

30

【0132】

タンパク質の標的分子への結合能の決定はまた、リアルタイム二分子相互作用解析（BIA）などの技術を使用して達成できる。Sjolander, SおよびUrbaniczky, C (1991) Anal. Chem. 63:2338-2345およびSzaboら (1995) Curr. Opin. Struct. Biol. 5:699-705。本明細書に使用した「BIA」は、相互作用物質を全く標識することなく、リアルタイムで生物特異的相互作用を研究する技術である（BIACore（登録商標））。光学現象表面プラズモン共鳴（SPR）の変化を、生物学的分子間のリアルタイム反応の指標として使用できる。

【0133】

本発明の試験化合物は、生物学的ライブラリー；空間的に扱うことのできる平行固相または液相ライブラリー；脱回旋の必要な合成ライブラリー法；「1ビーズ1化合物」ライブラリー法；およびアフィニティークロマトグラフィー選択を使用した合成ライブラリー法を含む、当分野で公知の組合せライブラリーの多くの任意のアプローチを使用して得ることができる。生物学的ライブラリーアプローチは、ポリペプチドライブラリーに限定され、他の4つのアプローチは、ポリペプチド、非ペプチドオリゴマー、または小分子化合物ライブラリーに適用できる（Lam, K. S. (1997) Anticancer Drug Des. 12:145）。

【0134】

分子ライブラリーの合成法の例は、当分野に、例えば、DeWittら (1993) Proc. Natl. Acad. Sci. USA 90:6909；Erbら (1994) Proc. Natl. Acad. Sci. USA 91:11422；Zuckermanら (1994) J. Med. Chem. 37:2678；Choら (1993) Science 261:1303；Carellら (1994) Angew. Chem. Int. Ed. Engl. 33:2059；Carellら (1994) Angew. Chem. Int. Ed. Engl. 33:2061；およびGalloppら (1994) J. Med. Chem. 37:1233に見られる。化合物ライブラリーは、溶液中（例えば、Houghten (1992) Biotechniques 13:412-421）、またはビーズ（Lam (1991) Nature 354:82-84）、チップ（Fodor (1993) Nature 364:555-556）、細菌（Lander USP 5, 223, 409）、胞子（Ladner USP '409）、プラスミド（Cullら (1992) Proc. Natl. Acad. Sci. USA 89:1865-1869）またはファージ（ScottおよびSmith (1990) Science 249:386-390）；（Devlin (1990) Science 249:404-406）；Cwirllaら (1990) Proc. Natl. Acad. Sci. 97:6378-6382）；（Felicci (1991) J. Mol. Biol. 222:301-310）；（Ladner上記）に提示され得る。

40

50

【0135】

候補化合物は、例えば、1) プリン類似体、2) Ig尾融合ペプチドを含む可溶性ペプチドなどのペプチドおよびランダムペプチドライブラリーのメンバー（例えば、Lamら（1991）Nature 354: 82-84; Houghtenら（1991）Nature 354: 84-86 参照）およびDおよびL立体配置アミノ酸から作成したコンビナトリアルケミストリー由来分子ライブラリー；3) ホスホペプチド（例えば、ランダムおよび部分変性指向ホスホペプチドライブラリーのメンバー、例えば、Songyanら（1993）Cell 72: 767-778 参照）；4) 抗体（例えば、ポリクローナル、モノクローナル、ヒト化、抗イディオタイプ、キメラ、および単鎖抗体、並びに、Fab、F(ab')₂、Fab発現ライブラリー断片、および抗体のエピトープ結合断片）；および5) 小有機および無機分子（例えばコンビナトリアルおよび天然産物ライブラリーから得られた分子）を含む。

10

【0136】

1つの候補化合物は、リガンド結合に競合する、可溶性全長受容体または断片である。他の候補化合物は、変異受容体、または、受容体機能に影響を及ぼし従ってリガンドと競合する変異を含む適切な断片を含む。従って、リガンドと例えば高い親和性で競合する断片、またはリガンドと結合するが遊離はしない断片は、本発明に包含される。

【0137】

本発明は、受容体活性を調節（刺激または阻害）する化合物を同定する他の末端点を提供する。アッセイは、典型的には、受容体活性を示す、シグナル伝達経路の事象のアッセイを含む。従って、受容体タンパク質依存性シグナルカスケードに応答してアップまたはダウンレギュレートされる遺伝子の発現を、アッセイできる。1つの実施形態において、かかる遺伝子の調節領域は、ルシフェラーゼなどの容易に検出可能なマーカーに作動可能に連鎖し得る。別に、受容体タンパク質のリン酸化、または受容体タンパク質標的も測定できる。

20

【0138】

化合物の結合および／または活性化はまた、アミノ末端細胞外ドメイン、またはその一部、全膜貫通ドメインまたは小領域、例えば、7回膜貫通セグメントのいずれかまたは細胞外または細胞内ループのいずれかおよびカルボキシ末端細胞内ドメイン、またはその一部が、異種ドメインまたは小領域により置換され得る、キメラ受容体タンパク質の使用によりスクリーニングできる。例えば、天然受容体により認識され、異なるGタンパク質と相互作用する、Gタンパク質結合領域を使用できる。従って、異なるセットのシグナル伝達成分が、活性化の末端点アッセイとして利用できる。別に、全膜貫通部分または小領域（例えば、膜貫通部または細胞内または細胞外ループ）を、アミノ末端細胞外ドメインおよび／またはGタンパク質結合領域の由来する宿主細胞とは異なる、宿主細胞に特異的な全膜貫通部分または小領域と置換できる。これにより、受容体が由来する特異的宿主細胞以外でアッセイを実施できる。別に、アミノ末端細胞外ドメイン（および／または他のリガンド結合領域）は、異なるリガンドに結合しているドメイン（および／または他の結合領域）により置換でき、従って、異種アミノ末端細胞外ドメイン（または領域）に相互作用するが、依然としてシグナル伝達は引き起こす試験化合物についてのアッセイが提供される。最後に、活性化は、天然シグナル伝達経路2の一部である転写調節配列に作動可能に連鎖した容易に検出可能なコード領域を含むレポーター遺伝子により検出できる。

30

40

【0139】

受容体ポリペプチドはまた、受容体と相互作用する化合物の発見のために設計された方法における競合的結合アッセイに有用である。従って、化合物を、化合物がポリペプチドと結合するかまたは別様に相互作用できる条件下で、受容体ポリペプチドに曝露させる。可溶性受容体ポリペプチドも混合物に加える。試験化合物が可溶性受容体ポリペプチドと相互作用すれば、形成される複合体の量または受容体標的の活性を減少させる。この種類のアッセイは、特に、受容体の特異的領域に相互作用する化合物を探索する場合に有用である。従って、標的受容体領域と競合する可溶性ポリペプチドを、目的の領域に対応するペ

50

プチド配列を含むように設計する。

【0140】

細胞非含有薬物スクリーニングアッセイを実施するために、受容体タンパク質、または断片、またはその標的分子を固定して、一方または両方のタンパク質の複合体非形成形からの複合体の分離を容易にし、並びに、アッセイの自動化に適応させることが望ましい。

【0141】

タンパク質をマトリックス上に固定する技法は、薬物スクリーニングアッセイに使用できる。1つの実施形態において、タンパク質がマトリックスに結合できるドメインの加わった融合タンパク質を提供できる。例えば、グルタチオン-S-トランスフェラーゼ/2838、14168および15334融合タンパク質を、グルタチオンセファロースビーズ（シグマケミカル、セントルイス、ミズーリ州）またはグルタチオン由来マイクロタイタープレートに吸着でき、次いで細胞溶解液（例えば³⁵S標識）および候補化合物と合わせ、混合物を、複合体形成の実施条件下（例えば塩およびpHにおける生理学的条件）でインキュベートする。インキュベート後、ビーズを洗浄して全ての非結合標識を除去し、マトリックスを固定し、放射標識を直接的にまたは複合体の解離した後に上清で決定する。別に、複合体をマトリックスから解離し、SDS-PAGEにより分離し、ビーズ画分に見られる受容体結合タンパク質のレベルを、標準的な電気泳動技法を使用してゲルから定量できる。例えば、ポリペプチドまたはその標的分子を、当分野で公知の技法を使用して、ビオチンおよびストレプトアビジンの複合体を利用して固定できる。別に、タンパク質と反応するがタンパク質のその標的分子への結合は妨害しない抗体を、プレートのウェルに誘導体化して、タンパク質を、抗体複合体によりウェルで捕獲できる。受容体結合タンパク質および候補化合物の調製物を、受容体タンパク質提示ウェル中でインキュベートし、ウェルに捕獲された複合体の量を定量できる。GST固定複合体について上記した複合体に加えて、かかる複合体を検出する方法は、受容体タンパク質標的分子と反応性であるか、または受容体タンパク質と反応性で標的分子と競合する、抗体を使用した複合体の検出；並びに、標的分子に関連した酵素活性の検出に依拠した酵素結合アッセイを含む。

【0142】

これらの薬物スクリーニングアッセイに従って同定した受容体タンパク質活性のモジュレーターを使用して、例えば、リンパ節、胸腺、脾臓、精巣、大腸、および末梢血リンパ球（Tヘルパー細胞（1および2）、CD3⁺（CD4およびCD8）細胞、B細胞および顆粒球を含むがこれに限定されない）などにおいて、2838タンパク質を発現する細胞を治療することにより、受容体経路により媒介される疾患に罹患した被検者を治療できる。これらの薬物スクリーニングアッセイに従って同定した受容体タンパク質活性のモジュレーターを使用して、例えば、乳房、骨格筋、脾臓および末梢血リンパ球並びにCD34⁺細胞および巨核球などにおいて、14618タンパク質を発現する細胞を治療することにより、受容体経路により媒介される疾患に罹患した被検者を治療できる。これらの薬物スクリーニングアッセイに従って同定した受容体タンパク質活性のモジュレーターを使用して、例えば、リンパ節、扁桃腺、脾臓、大腸、脾臓、末梢血細胞、胸腺、副腎および心臓並びに巨核球および赤芽球などにおいて、15334タンパク質を発現する細胞を治療することにより、受容体経路により媒介される疾患に罹患した被検者を治療できる。これらの治療法は、かかる治療の必要な被検者に、本明細書に記載した医薬組成物中のタンパク質活性のモジュレーターを投与する段階を含む。

【0143】

図13は、様々な正常ヒト組織における14618受容体の発現を示す。遺伝子は、乳房および骨格筋で高度に発現されている。従って、これらの組織に関与する疾患における遺伝子の発現は、特に関連性がある。有意な発現が、甲状腺、胎盤、胎児腎臓、胎児心臓、およびリンパ節でも生じる。従って、遺伝子の発現は、同様にこれらの組織に関与する疾患にも関連性がある。さらに、図13に示したように、より低いレベルの発現が、様々な他の組織で見られる。従って、遺伝子の発現は、これらの組織に関与する疾患にも関連性がある。

10

20

30

40

50

【0144】

14618受容体は、活性化を伴うまたは伴わない様々な造血細胞にも発現している。この遺伝子は、CD34⁺ 骨髄細胞にも高度に発現している。従って、この遺伝子の発現は、様々な血液細胞前駆体に関連性がある。それ故、この遺伝子の発現は、任意の主要な血液細胞型の欠損に關与する疾患、すなわち好中球減少症、血小板減少症または貧血に関連している。この遺伝子はまた、可動化末梢血細胞巨核球（G-CSFで可動化）にも高度に発現している。従って、この遺伝子の発現は、血小板減少症などの血小板機能に關与する疾患に関連がある。有意な発現が、可動化末梢血細胞リンパ球、可動化骨髄CD34⁺細胞、および臍帯血CD434⁺細胞にも見られる。従って、この遺伝子の発現は、これらの細胞の機能に関連があり、それ故、免疫機能または炎症に關与する疾患にも関連がある。さらに、遺伝子の発現は、活性化末梢血単核細胞にも生じる。活性化B細胞は、異な

10

【0145】

2838受容体は、図15に示した細胞および組織に発現している。高レベルの発現が、リンパ節および胸腺に示される。従って、遺伝子の発現は、特に、これらの組織に關与する疾患に関連している。極めて高い発現が、CD8細胞および活性化B細胞に見られる。高発現が、Tヘルパー細胞（1および2）、CD4細胞およびジャーカット細胞（T細胞系）にも生じている。発現は、活性化B細胞および活性化Tヘルパー細胞で異なる。発現は、これらの両方の細胞型の活性化時に増加する。従って、遺伝子の発現は、免疫機能および炎症に關与する疾患に関連がある。この遺伝子はまた、顆粒球で有意に発現している。従って、この遺伝子の発現は、これらの細胞に關与する疾患に関連がある。

20

【0146】

図16は、正常ヒト組織における15334受容体の発現を示す。この遺伝子は、リンパ節、扁桃腺および脾臓で高度に発現している。従って、この遺伝子の発現は、これらの組織に關与する疾患に特に関連している。この遺伝子の発現はまた、大腸、精巣、胎盤、胎児心臓、および脾臓でも高い。従って、この遺伝子の発現は、これらの組織に關与する疾患にも関連している。さらに、この図は、数個の他の組織における低い発現レベルを示す。従って、この遺伝子の発現は、これらの組織に關与する疾患に関連性があり得る。15334受容体の発現は、様々な造血細胞でも研究されている（図17～19）。極めて高い発現が、初期巨核球および赤芽球に生じる。従って、この遺伝子の発現は、赤血球分化および巨核球分化に関連し、従って、貧血および血小板減少症の治療に関連している。さらに、この遺伝子の発現は、活性化B細胞に比較して、休止B細胞で有意に増加している。従って、この遺伝子の発現は、B細胞免疫機能に関連している。さらに、低いレベルの発現が、様々な他の造血系統の細胞に見られる。図17参照。造血細胞における系統限定的様式でのこの遺伝子の発現により、発現は、系統細胞、赤血球／赤血球、または巨核球／血小板の発達の調節に関連している。

30

【0147】

脾臓に関連した疾患は、非特異的急性脾炎、鬱血性巨脾、および脾臓梗塞を含む脾腫大；新生物、先天性異常、および破裂を含むがこれに限定されない。脾腫大に関連した疾患は、非特異的脾炎、感染性単核球症、結核、腸チフス、ブルセラ症、サイトメガロウイルス、梅毒、マラリア、ヒストプラズマ症、トキソプラズマ症、カラアザール、トリパノソーマ症、住血吸虫症、レーシュマニア症、およびエキノコックス症などの感染；肝硬変、門脈または脾静脈血栓症、および心不全などの部分高血圧に関連した鬱血性状態；ホジキン病、非ホジキンリンパ腫／白血病、多発性骨髄腫、骨髄増殖疾患、溶血性貧血、および血小板減少性紫斑病などのリンパ血行性疾患；慢性関節リウマチおよび全身性エリテマトーデスなどの免疫－炎症疾患；ゴーシェ病、ニーマン・ピット病、およびムコ多糖症などの沈着病；および、アミロイド症、原発性新生物および嚢胞、および続発性新生物などの他の状態を含む。

40

【0148】

50

肺に関連した疾患は、先天性異常；無気肺；肺鬱血および浮腫（血行力学的肺浮腫および微小血管損傷、成人呼吸促迫症候群（び慢性肺胞傷害）、肺塞栓症、出血、および梗塞により引き起こされる浮腫を含む）、および肺性高血圧および血管硬化などの、血管を起源とする疾病；肺気腫、慢性気管支炎、気管支喘息、および気管支拡張症などの慢性閉塞性肺疾患；塵肺、サルコイドーシス、特発性肺繊維症、剥離性間質性肺炎、過敏症肺臓炎、肺好酸球増加症（好酸球増加症を伴う肺浸潤）、器質性肺炎を伴う閉塞性細気管支炎、び慢性肺出血症候群（グッドパスチャー症候群、特発性肺ヘモジデリン沈着症および他の出血症候群を含む）、コラーゲン血管疾患における肺併発、および肺胞蛋白症などの、散在性間質性（浸潤性、拘束性）疾病；薬物誘導肺疾患、放射線誘導肺疾患、および肺移植などの治療合併症；気管支原性癌（腫瘍随伴症候群、細気管支肺胞上皮癌、気管支カルチノイド、種々の腫瘍および転移性腫瘍などの神経内分泌腫瘍を含む）などの腫瘍；炎症性胸水、非炎症性胸水、気胸症および胸膜腫瘍（孤立性繊維症（胸膜繊維腫）および悪性中皮腫を含む）を含む、胸膜の病態を含むがこれに限定されない。

10

【0149】

大腸に関連した疾患は、閉塞症および狭窄症、メッケル憩室、先天性神経節欠損巨大大腸—ヒルシュスプルング病などの先天性異常；下痢および赤痢、感染性全腸炎（ウイルス性胃腸炎、細菌性全腸炎を含む）、壊死性全腸炎、抗生物質性大腸炎（偽膜性大腸炎）、および膠原性およびリンパ性大腸炎、種々の腸炎症疾患（寄生虫および原虫を含む）、後天性免疫不全症候群、移植、薬物誘導腸損傷、放射線全腸炎、好中球減少症大腸炎（盲腸炎）、および分流性大腸炎などの全腸炎；クローン病および潰瘍性大腸炎などの特発性炎症性腸疾患；非新生物性ポリープ、腺腫、家族性症候群、大腸直腸発癌、大腸直腸癌、およびカルチノイド腫瘍を含むがこれに限定されない。

20

【0150】

肝臓に関連した疾患は、肝臓損傷；黄疸および胆汁鬱滞、例えばビリルビンおよび胆汁形成；肝不全および肝硬変、例えば、硬変、門脈圧亢進症（腹水、門脈体静脈短絡を含む）および脾腫大；ウイルス性肝炎（A～E型肝炎感染および他の肝炎ウイルスによる肝臓を含む）、臨床病理学的症候群（例えば保因状態、無症候性肝臓、急性ウイルス肝炎、慢性ウイルス肝炎、および劇症肝炎）などの感染性疾患；自己免疫肝炎；薬物および毒素誘導肝疾患、例えばアルコール性肝疾患；先天的代謝異常および小児肝疾患、例えば血色素症、ウィルソン病、 α_1 -アンチトリプシン欠損症、および新生児肝炎；続発性胆汁性肝硬変、原発性胆汁性肝硬変、原発性硬化性胆道炎および胆道樹の異常などの肝内胆道疾患；肝臓への血流障害（肝動脈欠損および門脈閉塞および血栓症を含む）、肝臓を通る血流障害（受動鬱血および小葉中心性壊死および肝紫斑病を含む）、肝静脈流出閉塞（肝静脈血栓症（バッド・キアリ症候群）および静脈閉塞性疾患を含む）などの循環器疾患；子癰前症および子癰、妊娠の急性脂肪肝、および肝内胆汁鬱滞などの、妊娠に関連した肝疾患；骨髄移植後の薬物毒性、移植片宿主疾患および肝拒絶、および肝同種移植片への非免疫学的傷害などの、臓器または骨髄移植の肝合併症；結節過形成、腺腫、および悪性腫瘍（肝臓の原発性腫瘍および転移性腫瘍を含む）などの腫瘍および腫瘍状態を含むがこれに限定されない。

30

【0151】

子宮および子宮内膜に関連した疾患は、月経周期の子宮内膜組織学；無排卵周期、不適切な黄体期、経口避妊薬および誘発性子宮内膜変化、および閉経期および閉経後変化などの、機能的子宮内膜疾患；慢性子宮内膜炎などの炎症；腺筋症；子宮内膜ポリープ；子宮内膜過形成；子宮内膜癌などの悪性腫瘍；悪性混合ミューラー腫瘍などの、混合ミューラーおよび間葉腫瘍；平滑筋腫、平滑筋肉腫、および子宮内膜間質腫瘍を含む、子宮筋の腫瘍を含むがこれに限定されない。

40

【0152】

脳に関連した疾患は、ニューロンに関連した疾患、および膠細胞、星状膠細胞、乏突起膠細胞、上衣細胞、および小膠細胞などの膠細胞に関連した疾患；脳浮腫、頭蓋内圧上昇およびヘルニア形成、および水頭；神経管、前脳異常、後頭蓋窩以上、および脊髓空洞症お

50

よび脊髄水腫症などの、奇形および発達疾患；出世時脳損傷；低酸素、虚血、および梗塞（低血圧、低灌流、および局所的血液供給の閉塞に由来する低流動状態—全体的脳虚血および限局的脳虚血—の梗塞を含む）、頭蓋内出血（脳内（実質細胞内）出血、クモ膜下出血およびイチゴ状動脈瘤破裂を含む）、および血管形成異常、高血圧性脳疾患（ラクナ梗塞、細隙出血および高血圧性脳症を含む）；急性髄膜炎（急性化膿（細菌）髄膜炎および急性無菌（ウイルス）髄膜炎を含む）、急性焦点化膿性感染（脳膿瘍、硬膜下膿瘍、および硬膜外膿瘍を含む）、慢性細菌髄膜炎（結核およびマイコバクテリウム症、神経梅毒および神経ボレリア症（ライム病）を含む）、ウイルス髄膜炎（節足動物媒介（A r b o）ウイルス脳炎、単純ヘルペスウイルス1型、単純ヘルペスウイルス2型、水痘帯状疱疹ウイルス（帯状ヘルペス）、サイトメガロウイルス、灰白髄炎、狂犬病を含む）およびヒト免疫不全ウイルス1（H I V－1 髄膜炎（亜急性脳炎）、空胞性ミエロパシー、エイズ関連ミエロパシー、末梢神経障害、および小児エイズ、進行性多巣性白質脳症、亜急性硬化性全脳炎、真菌髄膜炎、神経系の他の感染症などの感染；伝染性海綿様脳症（プリオン病）；脱髄疾患（多発性硬化症、多発性硬化症変形、急性散在性脳脊髄炎および急性壊死性出血性脳脊髄炎、および脱髄を伴う他の疾病；大脳皮質に影響を及ぼす変性疾患（アルツハイマー病およびピック病を含む）、大脳基底核および脳幹の変性疾患（パーキンソン症、特発性パーキンソン病（振戦麻痺）、進行性核上麻痺、皮質基底変性を含む）、多系統萎縮（線条体黒質変性、シャイードレーガー症候群およびオリブ橋小脳萎縮症およびハンチントン病を含む）などの、変性疾患；脊髄小脳失調症（フリードライヒ運動失調症および毛細血管拡張性運動失調症を含む）を含む脊髄小脳変性、運動ニューロンに影響を及ぼす変性疾患（筋萎縮性側索硬化症（運動ニューロン疾患）、延髄脊髄萎縮症（ケネディー症候群）、および進行性脊髄性筋萎縮症を含む）；白質ジストロフィー（クラッペ病、異染性白質ジストロフィー、副腎白質ジストロフィー、ペリツェウス・メルツバッハー病およびカナバン病を含む）、ミトコンドリア脳筋症（リー病および他のミトコンドリア脳筋症を含む）などの先天性代謝障害；ビタミン欠乏症（例えばチアミン（ビタミンB₁）欠乏症およびビタミンB₁₂欠乏症）、神経学的代謝妨害後遺症（低血糖、高血糖、および肝性脳症を含む）、中毒疾患（一酸化炭素、メタノール、エタノールを含む）、および照射（メトトレキサートと照射の併用から誘発される損傷を含む）を含む、中毒および後天性代謝疾患；繊維性（び慢性）星状細胞腫および多型膠芽細胞腫、毛様性星状細胞腫、多形黄色星状細胞腫を含む神経膠腫、および脳幹神経膠腫、乏突起細胞腫、および上皮細胞腫および関連脳室周囲塊病変、神経腫瘍、低分化新生物（髄芽細胞腫を含む）、他の実質腫瘍（原発性脳リンパ腫を含む）、胚細胞腫瘍、および松果体実質腫瘍、髄膜腫、転移性腫瘍、腫瘍随伴症候群、末梢神経鞘腫瘍（シュワン細胞腫、神経線維腫症、および悪性末梢神経鞘腫瘍（悪性神経鞘腫）を含む）、および神経皮膚症候群（母斑症）（1型神経繊維腫症（N F 1）および2型神経線維腫症（N F 2）を含む神経線維腫症、結節性硬化、およびフォン・ヒッペル＝リンドウ病を含む）などの腫瘍を含むがこれに限定されない。

【0153】

T細胞に関連した疾患は、細胞性過敏症、例えば、遅延型過敏症およびT細胞性細胞毒性、および移植片拒絶；自己免疫疾患、例えば、全身性エリテマトーデス、シェーグレン症候群、全身性硬化症、炎症性ミオパシー、混合性結合性組織病、および結節性多発性動脈炎および他の脈管炎；胸腺形成不全、重症複合性免疫不全疾患、およびエイズなどの、原発性免疫不全を含むがこれに限定されない、免疫不全症候群；白血球減少症；白血球増多症、急性非特異的リンパ節炎、および慢性非特異的リンパ節炎を含むがこれに限定されない、反応性（炎症性）白血球増殖；急性リンパ芽球性白血病／リンパ腫、末梢T細胞および末梢T細胞リンパ腫を含むナチュラルキラー細胞新生物、分類不可能な成人T細胞白血病／リンパ腫、菌状息肉腫およびセザリー症候群、およびホジキン病などの、前駆T細胞新生物といったリンパ系新生物を含むがこれに限定されない、白血球の新生物による増殖を含むがこれに限定されない。

【0154】

皮膚の疾患は、色素沈着およびメラニン形成細胞の疾患（白斑、雀卵斑、肝斑、黒子、母斑細胞性母斑、異形成性母斑、および悪性黒色腫を含むがこれに限定されない）；上皮腫瘍（脂漏性角化症、黒色表皮腫、繊維上皮ポリープ、上皮嚢腫、角化棘細胞腫、および付属器（付属体）腫瘍を含むがこれに限定されない）；悪性前および悪性表皮腫瘍（光線性角化症、扁平上皮癌、基底細胞癌、およびメルケル細胞を含むがこれに限定されない）；真皮腫瘍（良性繊維性組織球腫、隆起性皮膚繊維肉腫、黄色腫、および真皮血管腫瘍を含むがこれに限定されない）；皮膚へ細胞移入する腫瘍（組織球症X、菌状息肉腫（皮膚T細胞リンパ腫）および肥満細胞症を含むがこれに限定されない）；表皮成熟疾患（魚鱗癬を含むがこれに限定されない）；急性炎症性皮膚疾患（蕁麻疹、急性湿疹皮膚炎、および多型紅斑を含むがこれに限定されない）；慢性炎症性皮膚疾患（乾癬、扁平苔せん、およびエリテマトーデスを含むがこれに限定されない）；水疱様膨隆（水疱性）疾患（天疱瘡、類天疱瘡、疱疹性皮膚炎および非炎症性水疱様膨隆疾患：表皮水疱症およびポルフィリン症を含むがこれに限定されない）；表皮付属器疾患（尋常性ざ瘡を含むがこれに限定されない）；皮下脂肪組織炎（結節性紅斑および硬結性紅斑を含むがこれに限定されない）；および感染および侵襲（例えば、いぼ、伝染性軟属腫、膿か疹、表在真菌感染、および節足動物咬合、刺針、および侵襲）を含むがこれに限定されない。

【0155】

正常な骨髄では、骨髄球シリーズ（多形核細胞）は、細胞成分の約60%を構成し、赤血球は20～30%を構成する。リンパ球、単球、細網細胞、形質細胞および巨核球は、共に10～20%を構成する。リンパ球は、正常な骨髄の5～15%を構成する。骨髄では、細胞型を、赤血球の前駆体（赤芽球）、マクロファージ（単芽球）、血小板（巨核球）、多形核白血球（骨髄芽球）、およびリンパ球（リンパ芽球）が1つの顕微鏡視野で見えるように加えて混合する。さらに、幹細胞が、異なる細胞系のために存在し、並びに、前駆幹細胞が、異なる系統の委任前駆細胞のために存在する。様々な種類の細胞および各段階は、当業者には公知であり、例えば、免疫学、免疫病理学および免疫、第5版、Se11ら、SimonおよびSchuster（1996）の42ページ（図2～8）に見られ、これは骨髄に見られる細胞型のその教義のために参考として取込む。従って、本発明は、これらの細胞から生じる疾患に関する。これらの疾患は、以下を含むがこれに限定されない：造血幹細胞に関与する疾患；委任リンパ系前駆細胞；BおよびT細胞を含むリンパ球；委任骨髄前駆体（単球、顆粒球、巨核球を含む）；および委任赤血球前駆体に関する疾患。これらは、白血病（Bリンパ系白血病、Tリンパ系白血病、未分化白血病を含む）；赤白血病、巨核芽球性白血病、単球性白血病；「白血病は、分化を伴いおよび伴わずに包含される」；慢性および急性骨髄芽球性白血病、慢性および急性リンパ急性白血病、慢性および急性骨髄性白血病、リンパ腫、骨髄異常形成症候群、慢性および急性骨髄性白血病、骨髄単球性白血病；慢性および急性骨髄芽球性白血病、慢性および急性骨髄性白血病、慢性および急性前骨髄球性白血病、慢性および急性骨髄球性白血病、単球-マクロファージ系統の血液学的悪性疾患、例えば、若年性慢性骨髄性白血病；続発性AML、前駆血液学的疾患；不応性貧血；再生不良性貧血；反応性皮膚血管内リンパ腫；樹状細胞の発現変化に関連した繊維性疾患、全身性硬化を含む疾患、E-M症候群、流行性毒性油症候群、好酸球性筋膜炎局在形の強皮症、ケロイド、および繊維性大腸疾患；血管腫様悪性繊維性組織球腫；原発性頭頸部扁平上皮癌を含む癌；カポジ肉腫を含む肉腫；哺乳動物繊維腺腫を含む、繊維腺腫および葉状腫瘍；間質腫瘍；組織球腫を含む葉状腫瘍；赤芽球症；神経線維腫症；血管内皮の疾患；脱髄、特に古い病巣において；血管原性浮腫、血管疾患、アルツハイマー病およびパーキンソン病；T細胞リンパ腫；B細胞リンパ腫を含むがこれに限定されない。

【0156】

心臓に関連した疾患は、心不全（心肥大、左側心不全および右側心不全を含むがこれに限定されない）；虚血性心疾患（狭心症、心筋梗塞、慢性虚血性心疾患、および急性心臓死を含むがこれに限定されない）；高血圧性心疾患（全身性（左側）高血圧性心疾患および肺性（右側）高血圧性心疾患を含むがこれに限定されない）；心臓弁膜症（石灰化大動脈

狭窄、先天性二尖大動脈弁の石灰化、および僧帽状輪状石灰化などの石灰化、および僧帽弁の粘液腫性変性（僧帽弁逸脱）、リウマチ熱およびリウマチ性心疾患、感染性心内膜炎、および非感染組織過形成、例えば、非細菌性血栓性心内膜炎および全身性エリテマトーデスの心内膜炎（リブマン・ザックス病）、カルチノイド心疾患、および人工弁の合併症により引き起こされる弁変性を含むがこれに限定されない）、心筋疾患（拡張型心筋症、肥大型心筋症、収縮性心筋症、および心筋炎を含むがこれに限定されない）；心膜疾患（心膜滲出および心膜血腫および心外膜炎（急性心外膜炎および治癒心外膜炎を含む）およびリウマチ様心疾患を含むがこれに限定されない）；新生物性心疾患（原発性心腫瘍、例えば、粘液腫、脂肪腫、乳頭繊維弾性腫、横紋筋腫、および肉腫、および非心性新生物の心臓効果を含むがこれに限定されない）；鬱血性心不全（左-右短絡-後期チアノーゼ、例えば、心房中隔欠損症、心室中隔欠損症、動脈管開存、および房室中隔欠損、右-左短絡-初期チアノーゼ、例えば、ファロー四徴症、大動脈転位、動脈幹、三尖弁閉鎖、および全肺静脈接続異常、閉塞性先天性異常、例えば、大動脈の縮窄、肺動脈弁狭窄および閉鎖症、および大動脈弁狭窄および閉鎖症、および心臓移植に関する疾患を含むがこれに限定されない）を含むがこれに限定されない。

【0157】

血管に關与する障害には、以下が含まれるが、これらに限定されない：損傷に対する血管壁の反応（内皮機能障害および内皮活性化ならびに内膜肥厚など）；先天性奇形を含む血管疾患（動静脈瘻、アテローム性動脈硬化症、および高血圧性血管疾患（高血圧など）など）（これらに限定されない）；炎症性疾患—脈管炎（巨細胞性（側頭）動脈炎、高安動脈炎、結節性多発性動脈炎（古典的）、川崎病（皮膚粘膜リンパ節症候群）、微視的多発性血管炎（微視的多発性動脈炎、過敏症、または白血球破碎性血管炎）、ヴェグナー肉芽腫症、閉塞性血栓血管炎（バーガー病）、他の障害に關連する脈管炎および感染性動脈炎など）；レイノー病；動脈瘤および解離（腹部大動脈瘤、梅毒性動脈瘤、および動脈解離など）；静脈およびリンパの疾患（拡張蛇行静脈、血栓性静脈炎および静脈血栓症、上大静脈の閉塞（上大静脈症候群）、下大静脈の閉塞（下大静脈症候群）、ならびにリンパ管炎およびリンパ水腫など）；良性腫瘍および腫瘍様病態を含む腫瘍（血管腫、リンパ管腫、グロムス腫瘍（グロムス血管腫）、血管拡張症、細菌性血管腫症など）および中程度（低い悪性度の境界）の腫瘍（カポジ肉腫および血管内皮腫など）、および悪性腫瘍（血管肉腫および血管外皮細胞腫など）；および血管疾患における治療介入の病理学（バルーン血管形成および關連技術）および血管置換（冠狀動脈バイパス術）が含まれる。

【0158】

赤血球に關連する障害には、以下の貧血が含まれるが、これらに限定されない：遺伝性球形赤血球症、赤血球酵素欠乏による溶血性疾患（グルコース-6-ホスフェートデヒドロゲナーゼ欠乏症）、鎌状赤血球病、サラセミア症候群、発作性夜間血色素尿症、免疫溶血性貧血、および赤血球に対する外傷に起因する溶血性貧血を含む溶血性貧血；および巨赤芽球性貧血を含む赤血球生成の減少による貧血（ビタミンB12欠乏による貧血（悪性貧血）；葉酸欠乏による貧血、鉄欠乏性貧血、慢性病による貧血、再生不良性貧血、赤芽球癆、および他の形態の骨髓不全など）。

【0159】

胸腺に關与する障害には、発達障害（胸腺形成不全症または無形成症を伴うディ・ジョージ症候群）；胸腺嚢胞；胸腺形成不全症（胸腺内のリンパ濾胞の出現に關与し、リンパ濾胞過形成を発症する）；胸腺腫（生殖細胞腫を含む）、リンパ腫、ホジキン病、およびカルチノイドなどが含まれる。胸腺腫には、良性または被膜を有する胸腺腫、および胸腺癌と呼ばれる悪性胸腺腫I型（浸潤性胸腺腫）またはII型を含むことができる。

【0160】

B細胞に關与する障害には、B細胞前駆体新生物（リンパ芽球性白血病／リンパ腫など）が含まれるがこれらに限定されない。末梢性B細胞新生物には、以下が含まれるが、これらに限定されない：慢性リンパ球性白血病／小リンパ球性リンパ腫、濾胞性リンパ腫、び慢性大B細胞型リンパ腫、バーキットリンパ腫、形質細胞性真性物、多発性骨髓腫および

関連する状態、リンパ形質細胞性リンパ腫（ヴァルデンストレームマクログロブリン血症）、外套細胞性リンパ腫、辺縁層リンパ腫（MALT腫）、および毛様細胞白血病。

【0161】

腎臓に関連する障害には、以下が含まれるがこれらに限定されない：先天性奇形（腎臓の嚢胞性疾患（嚢胞性腎形成異常、常染色体優性（成人）腎多嚢胞病、常染色体劣性（幼児）腎多嚢胞病）（これらに限定されない）および腎髄質の嚢胞性疾患（髄質性海綿腎、ネフロン肺結核－尿毒症性嚢胞性疾患複合体、）（これらに限定されない）；後天性（透析関連）嚢胞性疾患（孤立性嚢胞など））；糸球体損傷の病状を含む糸球体疾患（*in situ*免疫複合体蓄積を含むがこれに限定されない）（抗GBM腎炎、ヘイマン腎炎、移植抗原に対する抗体、循環免疫複合体腎炎、糸球体細胞に対する抗体、糸球体腎炎における細胞性免疫、第二経路の活性化、上皮細胞損傷、および細胞および可溶性メディエーターを含む糸球体損傷のメディエーターに関与する病態、急性糸球体腎炎（急性増殖性（連鎖球菌後、感染後）糸球体腎炎（連鎖球菌感染後糸球体腎炎および非連鎖球菌感染後急性糸球体腎炎を含むがこれらに限定されない）、急速進行性（半月体形成性）糸球体腎炎、ネフローゼ症候群、膜性糸球体腎炎（膜性腎症）、微小変化疾患（リポイドネフローゼ）、巣状分節状糸球体硬化症、膜性増殖性糸球体腎炎、IgA腎症（バージャー病）、層状増殖性および壊死性糸球体腎炎（層状糸球体腎炎）、遺伝性腎炎（アルポート症候群および薄膜疾患（両性家族性血尿）が含まれるがこれらに限定されない）など）、慢性糸球体腎炎、全身性疾患に関連する糸球体損傷（全身性紅斑性狼瘡、ヘーノホーシェーンライン紫斑病、細菌性心内膜炎、糖尿病性糸球体硬化症、アミロイドーシス、繊維性および免疫タクトイド糸球体腎炎、ならびに他の全身性障害が含まれるがこれらに限定されない））（これらに限定されない））；急性尿細管壊死および尿細管間質性腎炎を含む細管および間質に影響を与える疾患（腎盂腎炎および逆流性腎症、急性腎盂腎炎、慢性腎盂腎炎、および逆流性腎症が含まれるがこれらに限定されない）、薬物および毒素によって誘導される尿細管間質性腎炎（急性薬物誘導性間質性腎炎、鎮痛薬乱用による腎症、非ステロイド性抗炎症薬に関連する腎症が含まれるが、これらに限定されない）、および他の尿細管間質性疾患（尿酸腎症、高カルシウム血症および腎石灰沈着症ならびに多発性骨髄腫が含まれるが、これらに限定されない）；良性腎硬化症、悪性高血圧症、および促進性腎硬化症、腎動脈狭窄症、および血栓性微小血管症を含む血管の疾患（古典的（幼児期）溶血性尿毒症症候群、成人溶血性尿毒症症候群／血栓性血小板減少性紫斑病、特発性HUS／TTPが含まれるが、これらに限定されない）、他の血管疾患（アテローム硬化性虚血性腎疾患、鎌状赤血球腎症、汎発性皮質性壊死、および腎梗塞が含まれるが、これらに限定されない）；尿路閉塞（閉塞性尿路疾患）；尿路結石症（尿結石）；および腎臓腫瘍（良性腫瘍（腎乳頭腺種、腎線維腫または過誤腫（腎肥大性（*renomedullary*）間細胞腫瘍）、血管筋脂肪腫、および膨大細胞腫などならびに悪性腫瘍（腎盂の尿路上皮癌を含む腎細胞癌（副腎腫、腎臓腺癌）を含む））が含まれるが、これらに限定されない）。

【0162】

乳房の疾患には、以下が含まれるが、これらに限定されない：発育障害；炎症（急性乳腺炎、乳管周辺の乳腺炎、乳周辺の乳腺炎（再発性乳輪下膿瘍、乳管扁平上皮細胞化生）、乳管拡張症、脂肪壊死、肉芽腫性乳腺炎、およびシリコン乳房植込みに関連する病態が含まれるが、これらに限定されない）；繊維細胞の変化；増殖性乳房疾患（上皮性過形成、硬化性腺症、および小管乳頭腫が含まれるが、これらに限定されない）；腫瘍（間質腫瘍（線維腺種、葉状腫瘍、および肉腫など）および上皮性腫瘍（大管乳頭腫など））が含まれるが、これらに限定されない）；*in situ*での腺管癌（パジェット病を含む）および*in situ*での小葉癌を含む*in situ*（非浸潤性）癌および浸潤性（浸潤）癌（浸潤性腺管癌（非特異型）、浸潤性小葉癌、髄葉癌、膠様（粘液性）癌腫、管状腺癌、および浸潤性乳頭状癌が含まれるが、これらに限定されない）、および種々の悪性新生物を含む乳癌。

【0163】

男性の胸部の疾患には、女性化乳房および癌が含まれるが、これらに限定されない。

10

20

30

40

50

【0164】

精巣および精巣上体に関与する障害には、以下が含まれるが、これらに限定されない：奇形（停留睾丸など）、退行性病変（萎縮など）、炎症（（非特異性精巣上体炎および精巣炎など）、肉芽腫性（自己免疫）睾丸炎、および特異性炎症（淋病、おたふく風邪、結核、および梅毒を含むがこれらに限定されない））、ねじれを含む血管障害、生殖細胞のねじれを含む精巣腫瘍（精上皮腫、精母細胞精上皮腫、胚性癌腫、卵黄囊腫瘍絨毛腫、奇形腫、および混合腫瘍を含むが、これらに限定されない）、性索－性腺支質の腫瘍（ライディッヒ（間質性）細胞腫瘍およびセルトーリ細胞腫瘍（男性ホルモン細胞産生腫）ならびに精巣リンパ腫を含むがこれらに制限されない）、鞘膜の週の損傷。

【0165】

前立腺に関与する障害には、以下が含まれるが、これらに限定されない：炎症、良性肥厚（例えば、結節性再生（良性前立腺肥大または肥厚））、および腫瘍（癌など）。

【0166】

甲状腺に関与する障害には、以下が含まれるが、これらに限定されない：甲状腺機能亢進症；甲状腺機能低下症（クレチン症および粘液水腫が含まれるが、これらに限定されない）；甲状腺炎（橋本病、亜急性（肉芽腫性）甲状腺炎、亜急性リンパ球性（無痛）甲状腺炎が含まれるが、これらに限定されない）；グレーヴズ病；び慢性および多結節性甲状腺炎（び慢性無毒性（単一性）甲状腺炎および多結性甲状腺炎が含まれるが、これらに限定されない）；腺腫、他の良性腫瘍、および癌を含むがこれらに限定されない甲状腺新生物（乳頭状癌、濾胞腺癌、髄様癌、および退行性癌が含まれるが、これらに限定されない）

【0167】

骨格筋に関連する障害には、横紋筋肉腫などの腫瘍が含まれる。

【0168】

膵臓に関連する障害には、以下が含まれる：両性器奇形などの膵外分泌機能障害（異所性膵炎が含まれるがこれに限定されない）；嚢胞（偽性嚢胞が含まれるがこれらに限定されない）；腫瘍（嚢胞性腫瘍および膵臓癌が含まれるがこれらに限定されない）；内分泌性膵臓（真性糖尿病など）；島細胞腫瘍（膵島細胞腺種、ガストリン産生腫瘍、および他の稀な島細胞腫瘍が含まれるがこれらに限定されない）。

【0169】

小腸に関連する障害には、以下が含まれるが、これらに限定されない：吸収不良症候群（腹腔スプルー、熱帯性スプルー（感染後スプルー）など）、ホウィップル病、ジサッカリダーゼ（ラクターゼ）欠乏症、無βリポタンパク質血症、腺腫および腺癌を含む小腸の腫瘍。

【0170】

血小板減少（血小板減少症）に関連する障害には、特発性血小板減少性紫斑病（優勢特発性血小板減少性紫斑病を含む）、薬物誘導性血小板減少症、H I V 関連血小板減少症、および血栓性微小血管症；血栓性血小板減少性紫斑病および溶血性尿毒症症候群が含まれる。

【0171】

T細胞前駆体新生物に関連する障害には、Tリンパ芽球前駆体白血病／リンパ腫が含まれる。末梢性T細胞およびナチュラルキラー細胞新生物に関連する障害には、慢性Tリンパ球性白血病、大顆粒球リンパ性白血病、菌状息肉腫およびセザリー症候群、末梢性T細胞リンパ腫、不特定の血管免疫芽球性T細胞リンパ腫、血管中心性リンパ腫（NK／T細胞リンパ腫^{4 a}）、腸管T細胞リンパ腫、成人T細胞白血病／リンパ腫、および未分化大細胞リンパ腫が含まれる。

【0172】

卵巣に関連する疾患には、例えば、以下が含まれる：多嚢胞卵巣、スタイン－リーヴェンサール症候群、腹膜偽粘液腫、および間質性卵巣莢膜増殖症；卵巣腫瘍（体腔上皮腫瘍、漿液性腫瘍、粘液性腫瘍、子宮内腫瘍、明細胞腺癌、嚢腺線維腫、ブレンナー腫瘍、表面

10

20

30

40

50

上皮腫瘍など）；生殖細胞腫瘍（成熟（良性）奇形腫、単胚葉奇形腫、未熟型悪性奇形腫、未分化胚細胞腫、内胚葉洞腫瘍、絨毛癌など）；性索－吻合部腫瘍（顆粒膜細胞－卵胞膜細胞腫、テコーマー線維腫、男性ホルモン産生細胞腫、丘細胞腫瘍、および性腺芽腫など）；および転移性腫瘍（クルーケンベルク腫瘍など）。

【0173】

骨形成細胞には、骨芽前駆細胞、骨芽細胞、および骨細胞が含まれる。骨疾患は、これらの細胞が任意のその発達段階の間に骨格に影響を与え得るので複雑である。したがって、骨障害は、種々の症状を有し、身体の一部、複数、または全ての骨に関与し得る。このような障害には、以下が含まれる：先天性奇形、軟骨形成不全、および壊死性小人症、異常な基質に関連する疾患（I型コラーゲン疾患など）、骨粗鬆症、パジェット病、くる病、骨軟化症、高代謝回転骨形成異常、低代謝回転の再生不良疾患、骨壊死、化膿性骨髄炎、結核性骨髄炎、骨腫、類骨骨腫、骨芽細胞腫、骨肉腫、骨軟骨腫、軟骨腫、軟骨粘液線維腫、軟骨肉腫、線維性皮質欠損症、線維性骨異形成、線維肉腫、悪性線維性組織球腫、ユ－イング肉腫、未分化神経外胚葉性腫瘍、巨細胞腫、および転移性腫瘍。

10

【0174】

したがって、受容体ポリペプチドは、特に上記で開示のような受容体タンパク質の異常な発現または活性によって特徴づけられる受容体関連障害の治療に有用である。1つの実施形態では、本方法は、薬剤（例えば、本明細書中に記載のスクリーニングアッセイによって同定される薬剤）またはタンパク質の発現もしくは活性を調節する（例えば、アップレギュレートするかダウンレギュレートする）薬剤の組み合わせの投与を包含する。別の実施形態では、本発明は、タンパク質発現の減少または異常を補正するための治療として、タンパク質を投与するステップを包含する。

20

【0175】

タンパク質活性の刺激は、タンパク質が異常にダウンレギュレートされる、そして／またはタンパク質活性の増加がおそらく有利な効果を有する条件において望ましい。同様に、タンパク質活性の阻害は、タンパク質が異常にアップレギュレートされる、そして／またはタンパク質活性の減少がおそらく有利な効果を有する条件において望ましい。このような条件の一例では、被験体は、異常な発達または細胞分裂によって特徴づけられる障害を有する。このような条件の別の例では、被験体は、異常な造血反応によって特徴づけられる増殖疾患（例えば、癌）または障害を有する。このような条件の別の例では、患者の組織再生を達成することが望ましい（例えば、被験体が脳および脊髄損傷を有する場合、調節した様式で神経組を再生することが望ましい）。

30

【0176】

本発明のさらに別の態様では、本発明のタンパク質を「おとりタンパク質（bait protein）」として、2ハイブリッドアッセイまたは3ハイブリッドアッセイ（米国特許第5,283,317号、Zervosら、1993、Cell、72、223～232、Maduraら、1993、J. Biol. Chem.、268、12046～12054、Bartelら、1993、Biotechniques、14、920～924、Iwabuchiら、1993、Oncogene、8、1693～1696、およびBrent WO 94/10300を参照のこと）で使用する、本発明のタンパク質と結合するか相互作用する他のタンパク質（捕獲タンパク質）を同定し、その活性を調節することができる。

40

【0177】

2838受容体ポリペプチドはまた、特に、造血細胞（ヘルパーT細胞（1および2）、B細胞、および顆粒球など）において受容体タンパク質によって媒介される疾患または疾患の素因の診断用の標的を獲得するのに有用である。14618受容体ポリペプチドはまた、特に、乳房、骨格、筋肉、およびリンパ節ならびにCD34⁺前駆細胞およびこれらの前駆体によって産生される細胞（赤血球、リンパ球、および巨核球系）において受容体タンパク質によって媒介される疾患または疾患の素因の診断用の標的を獲得するのに有用である。15334受容体ポリペプチドはまた、特に、臍臓、脾臓、扁桃腺、およびリン

50

パ節ならびに巨核球および赤芽球系において受容体タンパク質によって媒介される疾患または疾患の素因の診断用の標的を獲得するのに有用である。したがって、本方法は、細胞、組織、または器官における受容体タンパク質の存在またはそのレベルの同定を提供する。本方法には、相互作用を検出することができるような、生物学的サンプルと受容体タンパク質との相互作用が可能な化合物との接触を含む。

【0178】

受容体タンパク質検出用の1つの薬剤は、受容体タンパク質に選択的に結合することができる抗体である。生物学的サンプルには、被験体から単離した組織、細胞、および生物学的体液、ならびに被験体に存在する組織、細胞、および体液が含まれる。

【0179】

受容体タンパク質はまた、種々の受容体タンパク質を有する患者における活動性疾患または疾患の素因の診断用の標的を提供する。したがって、受容体タンパク質を生物学的サンプルから単離し、異常な受容体タンパク質となる遺伝子変異の存在についてアッセイすることができる。これには、(異常なスプライシングの結果としての)アミノ酸の置換、欠失、挿入、再配列、および不適切な翻訳後修飾が含まれる。分析法には、電気泳動での移動度の変化、トリプティックペプチドの変化、細胞ベースのアッセイまたは無細胞アッセイにおける受容体活性の変化、リガンドまたは抗体結合パターンの変化、等電点の変化、直接的アミノ酸配列決定、およびタンパク質の変異の検出に有用な任意の他の公知のアッセイ技術が含まれる。

【0180】

受容体タンパク質検出用の *in vitro* 技術には、酵素結合免疫測定法 (ELISA)、ウェスタンブロット、免疫沈澱および免疫蛍光が含まれる。あるいは、タンパク質を、被験体への標識抗受容体抗体の移入によって、被験体において *in vivo* で検出することができる。例えば、抗体を、被験体中での存在および位置が標準的な画像技術で検出することができる放射性マーカーで標識することができる。被験体において発現する受容体タンパク質の対立遺伝子変異型を検出する方法およびサンプルにおける受容体タンパク質の断片を検出する方法が特に有用である。

【0181】

受容体ポリペプチドはまた、薬理ゲノム学的分析に有用である。薬理ゲノム学は、患者における薬物の性質の変化および異常な作用による、薬物反応における臨床的に有意な遺伝的変異型を取扱う。Eichelbaum、1997、Clin. Exp. Pharmacol. Physiol.、23 (10-11)、983~985、1996、およびLinder、1997、Clin. Chem.、43 (2)、254~266を参照のこと。これらの変異体の臨床的結果により、一定の個体において治療薬の毒性が増し、個体の代謝の変異型の結果として一定の個体において薬物治療が失敗する。したがって、個体の遺伝子型は、治療化合物が身体に作用する方法または身体が化合物を代謝する方法を同定することができる。さらに、薬物代謝酵素活性は、薬物作用の強度および持続時間の両方に影響を与える。したがって、個体の薬理ゲノム学により、個体の遺伝子型に基づいた予防または治療に有効な化合物およびこのような化合物の有効投薬量の選択が可能になる。いくつかの薬物代謝酵素における遺伝的多形の発見により、何故患者によって予想される薬物効果を得ることができないのか、薬物効果の悪化を示すのか、または標準的な薬物投薬量で毒性が強くなるのかが説明されている。強い代謝物の遺伝子型および弱い代謝物の遺伝子型において、多形を発現することができる。したがって、遺伝的多形により、1つの集団における1つまたは複数の受容体機能が別の集団の受容体機能と異なる受容体タンパク質の対立遺伝子タンパク質変異型を得ることができる。したがって、ポリペプチドで、標的が治療様式に影響を与えることができる遺伝的素因を確認できる。したがって、リガンドベースの治療において、多形により、アミノ末端細胞外ドメインおよび/またはリガンド結合において多少活性なリガンド結合領域、および受容体の活性化を得ることができる。したがって、リガンドの投薬量は、多形を含む所与の集団内での治療効果を最大にするために改変する必要があるだろう。

10

20

30

40

50

【0182】

受容体ポリペプチドはまた、臨床試験および他の治療中の治療効果の監視に有用である。したがって、遺伝子発現、タンパク質レベル、または受容体活性を増加または減少させるように設計された薬剤の治療効果を、終点標的として受容体ポリペプチドを用いて、治療クールにわたり監視することができる。監視ステップは、例えば、以下であり得る：(i) 薬物投与前に被験体から投与前サンプルを得るステップ、(ii) 投与前サンプルにおける特定のタンパク質の発現または活性レベルを検出するステップ、(iii) 被験体から1つまたは複数の投与後サンプルを得るステップ、(iv) 投与後サンプルにおけるタンパク質の発現または活性レベルを検出するステップ、(v) 投与前サンプルにおけるタンパク質と投与後サンプルにおけるタンパク質の発現または活性レベルを比較するステップ、(vi) 以上により、被験体に対する薬剤投与を増加または減少させるステップ。 10

【0183】

受容体ポリペプチドはまた、受容体関連疾患の治療に有用である。したがって、治療法には、リガンド結合を競合する受容体タンパク質の可溶性受容体または断片の使用が含まれる。これらの受容体または断片は、有効な競合を得るためにリガンドに対する高い親和性を有し得る。

【0184】

抗体

本発明はまた、2838、14618、および15334受容体タンパク質およびその変異型および断片に選択的に結合する抗体を提供する。抗体は、受容体タンパク質と実質的に相同ではない他のタンパク質にも結合するとしても、選択的に結合すると考える。これらの他のタンパク質は、受容体タンパク質断片またはドメインと相同性を共有する。特異的領域におけるこの保存により、相同的配列によって両タンパク質と結合する抗体が得られる。この場合、受容体タンパク質に結合する抗体は、依然として選択的であることが理解されるであろう。 20

【0185】

抗体を作製するために、ポリクローナル抗体およびモノクローナル抗体調製の標準的な技術を使用し、単離された受容体ポリペプチドを免疫原として使用して、抗体を作製する。全長タンパク質または抗原ペプチド断片のいずれかを使用することができる。高い抗原性指数を有する領域を、図2、図6、および図10に示す。 30

【0186】

抗体を、これらの領域またはこれらの領域の不連続な断片から調製することが好ましい。しかし、抗体を、本明細書中に記載のペプチドの任意の領域から調製することができる。好ましい断片は、リガンド結合を減少させるか完全に阻害する抗体を作製する。受容体の全てまたは受容体の一部（例えば、細胞内カルボキシ末端ドメイン、アミノ末端細胞外ドメイン、全膜貫通ドメインもしくは特異的セグメント、任意の細胞内もしくは細胞外ループ、または上記の任意の部分）に対する抗体を産生することができる。抗体を、特定の機能部位（リガンド結合部位、Gタンパク結合部位、またはリン酸化、グリコシル化、もしくはミリスチル化された部位）に対して産生することもできる。

【0187】

抗原性2838断片は、典型的には、少なくとも8つの連続したアミノ酸残基を含む。しかし、抗原ペプチドは、連続する少なくとも12、14アミノ酸残基、少なくとも15アミノ酸残基、少なくとも20アミノ酸残基、または少なくとも30アミノ酸残基を含むことができる。1つの実施形態では、断片は、タンパク質表面上に位置する領域（例えば、親水性領域）に対応する。しかし、これらの断片は、本発明の前に開示され得る任意の断片を包含すると解釈されない。 40

【0188】

抗原性14168断片は、典型的には、少なくとも9つの連続したアミノ酸残基を含む。しかし、抗原ペプチドは、連続する少なくとも12、14アミノ酸残基、少なくとも15アミノ酸残基、少なくとも20アミノ酸残基、または少なくとも30アミノ酸残基を含む 50

ことができる。1つの実施形態では、断片は、タンパク質表面上に位置する領域（例えば、親水性領域）に対応する。しかし、これらの断片は、本発明の前に開示され得る任意の断片を包含すると解釈されない。

【0189】

抗原性15334断片は、典型的には、少なくとも8つの連続したアミノ酸残基を含む。しかし、抗原ペプチドは、連続する少なくとも12、14アミノ酸残基、少なくとも15アミノ酸残基、少なくとも20アミノ酸残基、または少なくとも30アミノ酸残基を含むことができる。1つの実施形態では、断片は、タンパク質表面上に位置する領域（例えば、親水性領域）に対応する。しかし、これらの断片は、本発明の前に開示され得る任意の断片を包含すると解釈されない。

10

【0190】

抗体は、ポリクローナルまたはモノクローナルであり得る。インタクトな抗体またはその断片（例えば、FabまたはF(ab')₂）を使用することができる。

【0191】

抗体の検出可能な物質とのカップリング（すなわち、物理的結合）によって検出を容易にすることができる。検出可能な物質の例には、種々の酵素、置換基、蛍光物質、発光物質、生物発光物質、および放射性物質が含まれる。適切な酵素の例には、西洋ワサビペルオキシダーゼ、アルカリホスファターゼ、β-ガラクトシダーゼ、またはアセチルコリンエステラーゼが含まれる。適切な置換基複合体の例として、ストレプトアビジン／ビオチンおよびアビジン／ビオチンが含まれる。適切な蛍光物質の例として、ウンベリフェロン、フルオレセイン、フルオレセインイソチオシアネート、ローダミン、ジクロロトリアジニルアミンフルオレセイン、ダンシルクロリド、またはフィコエリトリンが含まれる。発光物質の例として、ルミノールが含まれる。生物発光物質の例として、ルシフェラーゼ、ルシフェリン、およびエクオリンが含まれる。適切な放射性物質の例として、¹²⁵I、¹³¹I、³⁵S、または³Hが含まれる。

20

【0192】

適切な免疫原性調製物は、天然もしくは組換えて発現させたタンパク質もしくは化学合成ペプチド由来であり得る。

【0193】

抗体の用途

30

抗体を使用して、標準的な技術（アフィニティークロマトグラフィーまたは免疫沈降）によって受容体タンパク質を単離することができる。抗体は、細胞由来の天然の受容体および宿主において発現させた組換えて産生させた受容体タンパク質の生成を容易にすることができる。

【0194】

抗体は、生物の種々の組織の間および正常な発達時の受容体の発現パターンを決定するために、細胞または組織における受容体タンパク質の存在の検出に有用である。

【0195】

発現量および発現パターンを評価するために、抗体を使用して、*in situ*、*in vivo*、または細胞溶解物中、または上清中の受容体タンパク質を検出することができる。

40

【0196】

抗体を使用して、発達中の異常な組織の分布または異常な発現を評価することができる。

【0197】

全長受容体タンパク質の循環断片の抗体検出を使用して、受容体代謝回転を同定することができる。

【0198】

さらに、抗体を使用して、疾患の活動段階などの疾患段階または受容体機能に関連する疾患に対する素因を有する個体における受容体発現を評価することができる。不適切な組織の分布、発達の発現、または受容体タンパク質の発現レベルによって障害が起こる場合、

50

正常な受容体タンパク質に対する抗体を調製することができる。障害が、受容体タンパク質における特異的変異によって特徴づけられる場合、この変異タンパク質に特異的な抗体を使用して、特異的変異受容体タンパク質の存在をアッセイすることができる。

【0199】

抗体を使用して、生物の種々の組織中の細胞の正常および異常な細胞下の局在性を評価することもできる。抗体は、全受容体または受容体の一部（例えば、アミノ末端の細胞外ドメインまたは細胞外ループの一部）に対して作製することができる。

【0200】

診断的用途を、遺伝子試験においてだけでなく、治療様式の監視においても適用することができる。したがって、治療が受容体発現レベルまたは異常な受容体の存在および異常な組織の分布または発達発現の回収を最終目的とする場合、受容体または関連断片に指向する抗体を使用して、治療効率を監視することができる。

【0201】

したがって、抗体を診断に使用して、例えば、所与の治療計画の効果を同定するために、臨床試験法の一部として組織におけるタンパク質レベルを監視することができる。

【0202】

さらに、抗体は、薬理ゲノム学的分析に有用である。したがって、多形受容体タンパク質に対して調製した抗体を使用して、治療計画の修正が必要な個体を同定することができる。

【0203】

抗体はまた、電気泳動での移動度、等電点、トリプティックペプチド消化、および当業者に公知の他の物理学的アッセイによって分析される異常な受容体タンパク質用の免疫学的マーカーとしての診断ツールとして有用である。

【0204】

抗体はまた、組織分類に有用である。したがって、特異的受容体タンパク質が特異的組織における発現に対応している場合、この受容体タンパク質に特異的な抗体を使用して組織の型を同定することができる。

【0205】

抗体はまた、法医学的同定に有用である。したがって、個体が特異的多形タンパク質が得られる特異的遺伝的多形性に対応している場合、多形タンパク質に特異的なタンパク質を、同定の補助として使用することができる。

【0206】

抗体はまた、受容体機能の阻害（例えば、リガンド結合の阻害）に有用である。

【0207】

これらの用途を、治療が受容体機能の阻害を含むような治療状況に適用することができる。抗体を使用して、例えば、リガンド結合を阻害することができる。機能に必要な部位を含む特異的断片または細胞に会合したインタクトな受容体に対して抗体を調製することができる。

【0208】

完全なヒト抗体は、ヒト患者の治療に特に望ましい。このヒト抗体作製技術の概要は、L on b e r g a n d H u z a r、1995、Int. Rev. Immunol.、13、65～93を参照のこと。このヒト抗体およびヒトモノクローナル抗体の作製技術およびこのような抗体の作製プロトコルの詳細な考察は、例えば、米国特許第5、625、126号、同第5、633、425号、同第5、569、825号、同第5、661、016号、および同第5、545、806号を参照のこと。

【0209】

本発明はまた、生物学的サンプル中の受容体タンパク質の存在を検出するための抗体を用いたキットを含む。キットは、生物学的サンプル中の受容体タンパク質検出用の抗体（標識抗体または標識可能な抗体など）および化合物または薬剤、サンプル中の受容体タンパク質の量を同定する手段、およびサンプル中の受容体タンパク質量と標準とを比較する手

段を含むことができる。化合物または薬剤を、適切な容器に封入することができる。さらに、キットは、受容体タンパク質を検出するためのキットを使用するための指示を含む。

【0210】

ポリヌクレオチド

配列番号2のヌクレオチド配列を、ヒト全長cDNAの配列決定によって得た。

配列番号4のヌクレオチド配列を、ヒト全長cDNAの配列決定によって得た。

配列番号6のヌクレオチド配列を、ヒト全長cDNAの配列決定によって得た。

【0211】

特別に開示したcDNAは、コード領域ならびに5'および3'非翻訳配列（配列番号2、配列番号4、および配列番号6）を含む。

10

【0212】

ヒト2838受容体cDNAは、約1617ヌクレオチド長であり、約319アミノ酸残基長である全長タンパク質をコードする。核酸は、脾臓、精巣、結腸、および末梢血リンパ球ならびに図15に記載の組織において発現する。配列番号1のアミノ酸の構造分析を、図3にヒドロパシープロットとして示す。図は、7回膜貫通セグメント、アミノ末端の細胞外ドメイン、およびカルボキシ末端の細胞内ドメインの推定構造を示す。

【0213】

ヒト14618受容体cDNAは、約1358ヌクレオチド長であり、約337アミノ酸残基長である全長タンパク質をコードする。核酸は、脾臓、および末梢血リンパ球、乳房、骨格筋、および図13に記載の他の組織ならびに図14に記載の種々の造血細胞、特に巨核球およびCD34⁺細胞において発現する。配列番号3のアミノ酸の構造分析を、図7にヒドロパシープロットとして示す。図は、7回膜貫通セグメント、アミノ末端の細胞外ドメイン、およびカルボキシ末端の細胞内ドメインの推定構造を示す。

20

【0214】

ヒト153340受容体cDNAは、約2559ヌクレオチド長であり、約372アミノ酸残基長である全長タンパク質をコードする。核酸は、リンパ節、扁桃腺、脾臓、結腸、脾臓、末梢血細胞、胸腺、副腎、および心臓ならびに図17～図19に記載の巨核球および赤血球系において発現する。配列番号5のアミノ酸の構造分析を、図11にヒドロパシープロットとして示す。図は、7回膜貫通セグメント、アミノ末端の細胞外ドメイン、およびカルボキシ末端の細胞内ドメインの推定構造を示す。

30

【0215】

本明細書中で使用される、用語「膜貫通セグメント」は、原形質膜に渡された疎水性ヘリックスを含む構造アミノ酸モチーフをいう。2838の全膜貫通ドメインは、アミノ酸約25～アミノ酸約292番目にわたる。14618の全膜貫通ドメインは、アミノ酸約29～アミノ酸約297番目にわたる。15334の全膜貫通ドメインは、アミノ酸約26～アミノ酸約299番目にわたる。セグメントが7回膜をわたり、このドメイン中に3つの細胞内ループおよび3つの細胞外ループが存在する。

【0216】

本発明は、2838受容体タンパク質をコードする単離ポリペプチドを提供する。用語「2838ポリペプチド」または「2838核酸」は、配列番号2に記載の配列をいう。

40

【0217】

本発明は、14618受容体タンパク質をコードする単離ポリペプチドを提供する。用語「14618ポリペプチド」または「14618核酸」は、配列番号4に記載の配列をいう。

【0218】

本発明は、15334受容体タンパク質をコードする単離ポリペプチドを提供する。用語「15334ポリペプチド」または「15334核酸」は、配列番号6に記載の配列をいう。

【0219】

用語「受容体ポリヌクレオチド」または「受容体核酸」は、さらに、2838、1461

50

8、および15334ポリヌクレオチドの変異型および断片を含む。

【0220】

「単離」受容体核酸は、天然の受容体核酸供給源に存在する他の核酸から分離したものである。好ましくは、「単離」核酸は、核酸が誘導される生物のゲノムDNA中の核酸（すなわち、核酸の5'および3'末端に位置する配列）に天然に隣接する配列を含まない。しかし、いくつかの隣接ヌクレオチド配列（例えば、5KBまで）を含むことができる。重要な点は、核酸を隣接配列から単離して、その核酸を本明細書中に記載の特定の操作（組換え発現、プローブおよびプライマーの調製など）および受容体核酸配列に特異的な他の用途に供することができることである。

【0221】

さらに、「単離」核酸分子（cDNAまたはRNA分子など）は、他の細胞物質、組換え技術によって作製した場合は培養培地、または化学合成した場合は化合物の前駆体もしくは他の薬品を実質的に含まない。しかし、核酸分子を他のコード配列または調節配列と融合し、依然として「単離」とみなすことができる。

【0222】

例えば、ベクター中に含まれる組換えDNA分子を、「単離」とみなす。単離DNA分子のさらなる例には、異種宿主細胞中に維持された組換えDNA分子または溶液中の（部分的または実質的に）精製されたDNA分子が含まれる。単離RNA分子には、本発明の単離DNA分子の*in vivo*または*in vitro* RNA転写物が含まれる。本発明の単離核酸分子には、さらに、このような合成分子が含まれる。

【0223】

いくつかの例では、単離物質は、組成物（例えば、他の物質を含む粗抽出物）、緩衝系または試薬混合物の一部を形成する。他の環境では、物質を、例えば、PAGEまたはカラムクロマトグラフィー（HPLCなど）で同定によって本質的に均質であるように精製することができる。好ましくは、単離核酸は全ての高分子種の少なくとも50、80、もしくは90%（モルベースに基づく）を含む。

【0224】

受容体ポリヌクレオチドは、成熟タンパク質+さらなるアミノまたはカルボキシル末端アミノ酸、または成熟ポリペプチド内のアミノ酸（例えば、成熟形態が1つを超えるポリペプチド鎖を有する場合）をコードすることができる。このような配列は、前駆体から成熟形態へのタンパク質のプロセッシングに役割を果たすか、タンパク質輸送を容易にするか、タンパク質の半減期を延長または短縮するか、アッセイまたは産生のためのタンパク質の操作を容易にすることなどができる。一般的に、*in situ*での場合のように、さらなるアミノ酸を、細胞酵素によって成熟タンパク質から離れてプロセッシングすることができる。

【0225】

受容体ポリヌクレオチドには、以下が含まれるが、これらに限定されない：成熟ポリペプチドのみをコードする配列、成熟ポリペプチドおよびさらなるコード配列（リーダー配列または分泌配列（例えば、プレプロまたはプロタンパク質配列）など）をコードする配列、さらなるコード配列を含むか含まない成熟配列をコードする配列+さらなる非コード配列（例えば、転写、mRNAプロセッシング（イントロンならびにスプライシングおよびポリアダニル化シグナルを含む）、リボゾーム結合およびmRNAの安定性の役割を果たす転写するが翻訳されない配列などの非コード5'および3'配列）。さらに、ポリヌクレオチドを、例えば、精製を容易にするペプチドをコードするマーカー配列に融合することができる。

【0226】

受容体ポリヌクレオチドは、クローニングまたは化学合成技術またはその組み合わせによって作製されるRNAの形態（mRNAなど）またはDNAの形態（cDNAおよびゲノムDNA）で存在し得る。核酸、特にDNAは、二本鎖または一本鎖であり得る。一本鎖核酸は、コード鎖（センス鎖）または非コード鎖（アンチセンス鎖）であり得る。

10

20

30

40

50

【0227】

1つの受容体核酸は、ヒト2838cDNAに対応する配列番号2に示したヌクレオチド配列を含む。

1つの受容体核酸は、ヒト14618cDNAに対応する配列番号4に示したヌクレオチド配列を含む。

1つの受容体核酸は、ヒト15334cDNAに対応する配列番号6に示したヌクレオチド配列を含む。

【0228】

1つの実施形態では、受容体核酸は、コード領域のみを含む。

【0229】

本発明は、さらに、遺伝コードの縮重のために配列番号2に記載のヌクレオチド配列と異なり、配列番号2に示すヌクレオチド配列によってコードされるタンパク質と同一のタンパク質をコードする、変異型受容体ポリヌクレオチドおよびその断片を提供する。

【0230】

本発明は、さらに、遺伝コードの縮重のために配列番号4に記載のヌクレオチド配列と異なり、配列番号4に示すヌクレオチド配列によってコードされるタンパク質と同一のタンパク質をコードする、変異型受容体ポリヌクレオチドおよびその断片を提供する。

【0231】

本発明は、さらに、遺伝コードの縮重のために配列番号6に記載のヌクレオチド配列と異なり、配列番号6に示すヌクレオチド配列によってコードされるタンパク質と同一のタンパク質をコードする、変異型受容体ポリヌクレオチドおよびその断片を提供する。

【0232】

本発明はまた、本明細書中に記載の変異型ポリペプチドをコードする受容体核酸分子を提供する。このようなポリヌクレオチドは、天然に存在し得るか（対立遺伝子変異型（同一の遺伝子座）、ホモログ（異なる遺伝子座）、およびオーソログ（異なる生物）など）、組換えDNA法または化学合成によって構築することができる。このような天然に存在しない変異型を、ポリヌクレオチド、細胞、または生物に適応する変異誘発技術によって作製することができる。したがって、上記のように、変異型は、ヌクレオチドの置換、欠失、逆位、および挿入を含むことができる。

【0233】

コード領域および非コード領域のいずれかまたは両方で変異体を得ることができる。変異体により、保存的および非保存的アミノ酸置換を得ることができる。

【0234】

典型的には、変異型は、配列番号2、4、および6ならびにその相補物からなる群から選択される核酸分子と実質的に同一である。

【0235】

オーソログ、ホモログ、および対立遺伝子変異型を、当該分野で公知の方法を用いて同定することができる。2838変異型は、配列番号2で示したヌクレオチド配列またはこの配列の断片に40～45%、45～50%、50～55%、55～60%、少なくとも約65%、典型的には少なくとも70～75%、より典型的には少なくとも80～85%、最も典型的には少なくとも約90～95%またはそれ以上相同な受容体をコードするヌクレオチド配列を含む。配列番号2に示すヌクレオチド配列またはその配列の断片にストリンジントな条件下でハイブリッド形成することができるこのような核酸分子を、容易に同定することができる。

【0236】

14618変異型は、配列番号4で示したヌクレオチド配列またはこの配列の断片に40～45%、45～50%、50～55%、55～60%、少なくとも約65%、典型的には少なくとも70～75%、より典型的には少なくとも80～85%、最も典型的には少なくとも約90～95%またはそれ以上相同な受容体をコードするヌクレオチド配列を含む。配列番号4に示すヌクレオチド配列またはその配列の断片にストリンジントな条件

10

20

30

40

50

下でハイブリッド形成することができるこのような核酸分子を、容易に同定することができる。

【0237】

15334変異型は、配列番号6で示したヌクレオチド配列またはこの配列の断片に、45～50%、50～55%、55～60%、少なくとも約65%、典型的には少なくとも70～75%、より典型的には少なくとも80～85%、最も典型的には少なくとも約90～95%またはそれ以上相同な受容体をコードするヌクレオチド配列を含む。配列番号6に示すヌクレオチド配列またはその配列の断片にストリンジェントな条件下でハイブリッド形成することができるこのような核酸分子を、容易に同定することができる。

【0238】

ストリンジェントなハイブリッド形成は、それが一般的な相同性（ポリA配列、または全てまたはほとんどのタンパク質、すべてのGPCR、全てのファミリーI GPCRまたは全てのプリン受容体に共通な配列など）による場合、実質的な相同性を示さないことが理解される。さらに、変異体には、本発明以前に開示されている可能性のあるいかなる核酸配列も含まれないことが理解される。

【0239】

本明細書中で使用される、用語「ストリンジェントな条件下でハイブリッド形成する」は、互いに少なくとも50～55%、55%相同な受容体ポリペプチドをコードするヌクレオチド配列が互いに典型的にハイブリッド形成したままであるハイブリッド形成および洗浄条件を記載することを意図する。これは、少なくとも約65%、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約90%、少なくとも約95%またはそれ以上互いに同一な配列が互いにハイブリッド形成したままのような条件であり得る。このようなストリンジェントな条件は、当業者に公知であり、Current Protocols in Molecular Biology、John Wiley & Sons、N. Y.、1989、6. 3. 1～6. 3. 6（参考として援用される）に見出すことができる。ストリンジェントなハイブリッド形成条件の一例は、約45℃の6×塩化ナトリウム／クエン酸ナトリウム（SSC）中でのハイブリッド形成およびその後の50～65℃の0. 2×SSC、0. 1% SDS中での1回または複数回の洗浄である。別の非限定的な例では、核酸分子を約45℃の6×塩化ナトリウム／クエン酸ナトリウム（SSC）中でのハイブリッド形成、その後室温での0. 2×SSC、0. 1% SDS中での低ストリンジェンシー洗浄、42℃での0. 2×SSC、0. 1% SDS中での中程度のストリンジェンシー洗浄、または65℃での0. 2×SSC、0. 1% SDS中での高ストリンジェンシー洗浄である。1つの実施形態では、ストリンジェントな条件下で配列番号2、4、および6の配列とハイブリッド形成する単離受容体核酸分子は、天然に存在する核酸分子に対応する。本明細書中で使用される、用語「天然に存在する」核酸分子は、天然に存在する核酸分子（例えば、天然のタンパク質をコードする）を有するRNAまたはDNAをいう。

【0240】

当業者に理解されるように、経験的および不安定化剤（ホルムアミドなど）または変性剤（SDSなど）のイオン強度、温度、および濃度に依存して、正確な条件を同定することができる。望ましいハイブリッド形成条件同定において考慮される他の因子には、核酸配列の長さ、塩基組成、ハイブリッド形成配列と他の同一配列内の配列のサブセットの発生頻度との間のミスマッチ%が含まれる。したがって、1つまたは複数のこれらのパラメータを変化させる一方で、2つの核酸分子について同程度の同一性および類似性を維持することによって、等価の条件を同定することができる。

【0241】

本発明はまた、ストリンジェントな条件下で、配列番号2、4、および6ならびに配列番号2、4、および6の相補物からなる群から選択される核酸配列とハイブリッド形成する一本鎖または二本鎖断片または部分を含む単離核酸を提供する。1つの実施形態では、核酸は、配列番号2、4、および6ならびにその相補物からなる群から選択されるヌクレオ

10

20

30

40

50

チド配列の一部からなる。本発明の核酸断片は、少なくとも約15、好ましくは少なくとも18、20、23、または25ヌクレオチド長であり、30、40、50、100、200、またはそれ以上のヌクレオチド長であり得る。本明細書中に記載の抗原タンパク質またはポリペプチドをコードするより長い断片（例えば、30またはそれ以上のヌクレオチド長）が有用である。

【0242】

さらに、本発明は、全長受容体ポリヌクレオチドの断片を含むポリヌクレオチドを提供する。断片は、一本鎖または二本鎖であり、DNAまたはRNAを含み得る。断片は、コード配列または非コード配列のいずれか由来であり得る。

【0243】

1つの実施形態では、1～およそヌクレオチド990由来の単離2838受容体核酸は、少なくとも16ヌクレオチド長であり、ストリンジェントな条件下で配列番号2のヌクレオチド配列を含む核酸分子とハイブリッド形成する。別の実施形態では、およそヌクレオチド1487～1617由来の核酸は、少なくとも20ヌクレオチドである。他の実施形態では、核酸は、少なくとも40、50、100、250、または500ヌクレオチド長またはそれ以上である。

【0244】

別の実施形態では、およそヌクレオチド1～およびヌクレオチド911のヌクレオチド由来の単離14618受容体核酸は、少なくとも8ヌクレオチド長であり、ストリンジェントな条件下で配列番号4のヌクレオチド配列を含む核酸分子とハイブリッド形成する。他の実施形態では、核酸は、少なくとも40、50、100、250、または500ヌクレオチド長またはそれ以上である。

【0245】

別の実施形態では、ヌクレオチド1～およそヌクレオチド1355由来の単離15334受容体核酸は、少なくとも18ヌクレオチド長であり、ストリンジェントな条件下で配列番号6のヌクレオチド配列を含む核酸分子とハイブリッド形成する。別の実施形態では、およそヌクレオチド868～およそヌクレオチド1355由来の核酸は、少なくとも11ヌクレオチドである。他の実施形態では、核酸は、少なくとも40、50、100、250、または500ヌクレオチド長またはそれ以上である。

【0246】

別の実施形態では、単離2838受容体核酸は、アミノ酸1～アミノ酸319の全コード領域をコードする。別の実施形態では、単離2838受容体核酸は、およそアミノ酸6～およそアミノ酸319の成熟タンパク質に対応する配列をコードする。別の実施形態では、単離14618受容体核酸は、アミノ酸1～アミノ酸337の全コード領域をコードする。別の実施形態では、単離14618受容体核酸は、およそアミノ酸6～およそアミノ酸337の成熟タンパク質に対応する配列をコードする。別の実施形態では、単離15334受容体核酸は、アミノ酸1～アミノ酸372の全コード領域をコードする。別の実施形態では、単離15334受容体核酸は、およそアミノ酸6～およそアミノ酸372の成熟タンパク質に対応する配列をコードする。

【0247】

3つ全ての受容体の他の断片は、本明細書中に記載のアミノ酸断片をコードするヌクレオチド配列を含む。さらに、断片は、特異的ドメインまたは本明細書中に記載の部位のサブ断片を含むことができる。断片はまた、上記の特異的アミノ酸配列またはその断片に対応する核酸配列を含む。本発明の核酸断片は、本発明以前に開示されている可能性のある断片を含むように構築されない。

【0248】

しかし、受容体断片には、全遺伝子を含まない任意の核酸配列を含むことが理解される。

【0249】

2838受容体核酸断片は、さらに、本明細書中に記載のドメイン、本明細書中にも記載のサブ領域、および特異的機能部位に対応する配列を含む。2838受容体核酸断片は、

10

20

30

40

50

1～およそ24のアミノ酸残基を含むアミノ末端細胞外ドメインを含むポリペプチド、膜貫通ドメインをわたす領域を含むポリペプチド（およそ25～およそ292のアミノ酸残基）、カルボキシ末端の細胞内ドメインを含むポリペプチド（およそ293～およそ319のアミノ酸残基）、およびGタンパク質受容体シグナルをコードするポリペプチド（118～120またはおよそ107～およそ123の周囲のアミノ酸残基）をコードする核酸、任意の7回膜貫通セグメント、細胞外または細胞内ループ、グリコシル化部位、リン酸化部位、ミリスチル化部位、およびアミド化部位をコードする核酸分子を含む。

【0250】

14618受容体核酸断片は、1～およそ28のアミノ酸残基を含むアミノ末端細胞外ドメインを含むポリペプチド、膜貫通ドメインをわたす領域を含むポリペプチド（およそ29～およそ297のアミノ酸残基）、カルボキシ末端の細胞内ドメインを含むペプチド（およそ298～およそ337のアミノ酸残基）、およびGタンパク質受容体シグナルをコードするポリペプチド（120～122またはおよそ110～およそ132の周囲のアミノ酸残基）をコードする核酸、任意の7回膜貫通セグメント、細胞外または細胞内ループ、グリコシル化部位、およびリン酸化部位をコードする核酸分子を含む。

10

【0251】

15334受容体核酸断片は、1～約25番目のアミノ酸残基を含むアミノ末端細胞外ドメインを含むポリペプチド、膜貫通ドメインをわたす領域を含むポリペプチド（およそ26～およそ299番目のアミノ酸残基）、カルボキシ末端の細胞内ドメインを含むペプチド（およそ300～およそ372番目のアミノ酸残基）、およびGタンパク質受容体シグナルをコードするポリペプチド（118～120番目またはおよそ110～およそ130番目の周囲のアミノ酸残基）をコードする核酸、任意の7回膜貫通セグメント、細胞外または細胞内ループ、グリコシル化部位、プロテインキナーゼC、cAMP、cGMP、およびカゼインキナーゼIIリン酸化部位、およびミリスチル化部位をコードする核酸分子を含む。

20

【0252】

ドメインの位置がコンピュータ分析によって予想されている場合、当業者は、これらのドメインを構築するアミノ酸残基が、ドメインを定義するために使用した基準に依存して変化し得ることを認識するであろう。

【0253】

受容体核酸断片はまた、ドメイン、セグメント、ループ、および上記の他の機能的部位の組み合わせを含む。したがって、例えば、受容体核酸は、アミノ末端細胞外ドメインおよび1つの膜貫通セグメントに対応する配列を含み得る。当業者は、多くの置換が可能であることを認識しているであろう。

30

【0254】

ドメインまたは部位の位置がコンピュータ分析によって予想されている場合、当業者は、これらのドメインを構築するアミノ酸残基が、ドメインを定義するために使用した基準に依存して変化し得ることを認識するであろう。

【0255】

本発明はまた、本明細書中に記載の受容体領域を保有するエピトープをコードする受容体核酸断片を提供する。

40

【0256】

単離受容体ポリヌクレオチド配列、特に断片は、DNAプローブおよびプライマーとして有用である。

【0257】

例えば、オリゴヌクレオチドプローブを合成するために、公知のヌクレオチド配列を用いて受容体遺伝子のコード領域を単離することができる。次いで、標識プローブを使用して、cDNAライブラリー、ゲノムDNAライブラリー、またはmRNAをスクリーニングし、コード領域に対応する核酸を単離することができる。さらに、PCR反応でプライマーを使用して、受容体遺伝子の特異的領域をクローニングすることができる。

50

【0258】

典型的には、プローブ／プライマーは、実質的に精製したオリゴヌクレオチドを含む。2838オリゴヌクレオチドは、典型的には、ストリンジェントな条件下で、配列番号2のセンス鎖またははアンチセンス鎖の少なくともそれぞれおよそ16および20個、典型的にはおよそ25、より典型的には40、50、または75個の連続したヌクレオチドとハイブリッド形成する1～990番目および1487～1617番目のヌクレオチド配列領域または他の受容体ポリヌクレオチドを含む。14618オリゴヌクレオチドは、典型的には、ストリンジェントな条件下で、配列番号4のセンス鎖またははアンチセンス鎖の少なくともおよそ9、12、典型的にはおよそ25、より典型的には40、50、または75個の連続したヌクレオチドとハイブリッド形成する1～911番目のヌクレオチド配列領域または他の受容体ポリヌクレオチドを含む。15334オリゴヌクレオチドは、典型的には、ストリンジェントな条件下で、配列番号6のセンス鎖またははアンチセンス鎖の少なくともおよそ11、12、典型的にはおよそ25、より典型的には40、50、または75個の連続したヌクレオチドとハイブリッド形成する868～1355番目のヌクレオチド配列領域または他の受容体ポリヌクレオチドを含む。15334オリゴヌクレオチドはまた、典型的には、ストリンジェントな条件下で、配列番号6のセンス鎖またははアンチセンス鎖の少なくともおよそ18、典型的にはおよそ25、より典型的には40、50、または75個の連続したヌクレオチドとハイブリッド形成する1～1355番目のヌクレオチド配列領域または他の受容体ポリヌクレオチドを含む。

10

【0259】

20

ポリヌクレオチドの使用

本発明の核酸配列を、さらに、「検索配列」として使用して、例えば、他のファミリーメンバーまたは関連配列の同定のために、公的なデータベースによって検索を行うことができる。このような検索を、Altschulら、1990、J. Mol. Biol.、215、403～10のNBLASTおよびXBLASTプログラム（バージョン2.0）を使用して行うことができる。本発明の核酸分子に相同なヌクレオチド配列を得るために、NBLASTプログラム（スコア＝100、長さ＝12）を具備したBLASTヌクレオチド検索を行うことができる。比較用のギャップアラインメントを得るために、Altschulら、1997、Nucleic Acids Res.、25（17）、3389～3402に記載のように、Gapped BLASTを利用することができる。BLASTおよびGapped BLASTプログラムを利用する場合、それぞれのプログラムのデフォルトパラメータ（例えば、XBLASTおよびNBLAST）を使用することができる。http://www.ncbi.nlm.nih.gov.を参照のこと。

30

【0260】

本発明の核酸断片により、アッセイ（以下に記載のアッセイなど）におけるプローブまたはプライマーが得られる。「プローブ」は、塩基特異的様式で、核酸の相補鎖とハイブリッド形成するオリゴヌクレオチドである。このようなプローブには、Nielsenら、1991、Science、254、1497～1500に記載のポリペプチド核酸が含まれる。典型的には、プローブは、配列番号2、4、および6ならびにその相補物からなる群から選択される核酸の少なくとも約15、典型的には、約20～25、より典型的には40、50、75の連続したヌクレオチドと高ストリンジェントな条件下でハイブリッド形成するヌクレオチド配列領域を含む。より典型的には、プローブは、さらに、標識（例えば、放射性同位体、蛍光化合物、酵素、または酵素の補因子）を含む。

40

【0261】

本明細書中で使用される、用語「プライマー」は、周知の方法（例えば、PCR、LCR）（本明細書中に記載されているが、これに限定されない）を使用するテンプレート指向性DNA合成の開始点として作用する一本鎖オリゴヌクレオチドをいう。プライマーの適切な長さは、特定の用途に依存するが、典型的には、約15～30ヌクレオチドの範囲である。用語「プライマー部位」は、プライマーがハイブリッド形成する標的DNA領域を

50

いう。用語「プライマー対」は、増幅される核酸配列の5'末端にハイブリッド形成する5'（上流）プライマーおよび増幅される配列の相補物にハイブリッド形成する3'（下流）プライマーのセットをいう。

【0262】

受容体ポリヌクレオチドは、プローブ、プライマー、および生物学的アッセイに有用である。

【0263】

本明細書中に記載のアッセイなどにおいて、ポリヌクレオチドを使用してGPCRの性質または機能を評価する場合、全長cDNAの全てまたはそれ未満が有用である。この場合、本発明以前に公知であり得る断片も含まれる。したがって、例えば、GPCR機能に特異的に指向するアッセイ（アゴニストまたはアンタゴニスト活性の評価）は、公知の断片の使用を含む。さらに、受容体機能の診断的アッセイ法を、任意の断片（本発明以前に公知であり得る断片を含む）を用いて行うこともできる。同様に、受容体機能不全の治療を含む方法では、当該分野で公知であり得る断片を含む全ての断片が含まれる。

10

【0264】

2838受容体ポリヌクレオチドは、配列番号1に記載のポリペプチドをコードする全長cDNAおよびゲノムクローンを単離し、配列番号1に示したものと同一のポリペプチドを産生する変異型または本明細書中に記載の他の変異型に対応するcDNAおよびゲノムクローンを単離するためのcDNAおよびゲノムDNA用のハイブリッド形成プローブとして有用である。変異体を、配列番号1に示したポリペプチドを単離したものと同一の組織および生物、同一の生物由来の異なる組織、または異なる生物由来の異なる組織から単離することができる。

20

【0265】

14628受容体ポリヌクレオチドは、配列番号3に記載のポリペプチドをコードする全長cDNAおよびゲノムクローンを単離し、配列番号3に示したものと同一のポリペプチドを産生する変異型または本明細書中に記載の他の変異型に対応するcDNAおよびゲノムクローンを単離するためのcDNAおよびゲノムDNA用のハイブリッド形成プローブとして有用である。変異体を、配列番号3に示したポリペプチドを単離したものと同一の組織および生物、同一の生物由来の異なる組織、または異なる生物由来の異なる組織から単離することができる。

30

【0266】

15334受容体ポリヌクレオチドは、配列番号5に記載のポリペプチドをコードする全長cDNAおよびゲノムクローンを単離し、配列番号5に示したものと同一のポリペプチドを産生する変異型または本明細書中に記載の他の変異型に対応するcDNAおよびゲノムクローンを単離するためのcDNAおよびゲノムDNA用のハイブリッド形成プローブとして有用である。変異体を、配列番号5に示したポリペプチドを単離したものと同一の組織および生物、同一の生物由来の異なる組織、または異なる生物由来の異なる組織から単離することができる。

【0267】

この方法は、発達段階で調節される遺伝子およびcDNAの単離に有用であるので、生物の発達における異なる点で同一の組織または異なる組織で発現させることができる。

40

【0268】

プローブは、受容体をコードする遺伝子の全長に沿った任意の配列に対応することができる。したがって、プローブは、5'非コード領域、コード領域、および3'非コード領域由来であり得る。しかし、プローブは、本発明以前に公知であり得る任意の配列に対応するように構築されない。

【0269】

2838核酸プローブは、例えば、配列番号1の全長cDNAまたはその断片（少なくとも11、16、18、20、30、50、100、250、または500ヌクレオチド長のオリゴヌクレオチドなど）であり、ストリンジェントな条件下でmRNAまたはDNA

50

と特異的にハイブリッド形成するのに十分であり得る。14618 核酸プローブは、例えば、配列番号3の全長cDNAまたはその断片（少なくとも8、12、15、30、50、100、250、または500ヌクレオチド長のオリゴヌクレオチドなど）であり、ストリンジェントな条件下でmRNAまたはDNAと特異的にハイブリッド形成するのに十分であり得る。15334 核酸プローブは、例えば、配列番号5の全長cDNAまたはその断片（少なくとも11、12、15、18、30、50、100、250、または500ヌクレオチド長のオリゴヌクレオチドなど）であり、ストリンジェントな条件下でmRNAまたはDNAと特異的にハイブリッド形成するのに十分であり得る。

【0270】

本明細書中に記載のポリヌクレオチドの断片はまた、本明細書中に記載のより長い断片または全長ポリヌクレオチドの合成に有用である。例えば、断片を、mRNAの任意の部分にハイブリッド形成することができ、より長いまたは全長cDNAを産生することができる。

10

【0271】

断片はまた、所望の長さおよび配列のアンチセンス分子の合成に有用である。

【0272】

本発明のアンチセンス核酸を、当該分野で公知の化学合成および酵素ライゲーション反応を用いて構築した配列番号2、4、および6のヌクレオチド配列を用いて設計することができる。例えば、アンチセンス核酸（例えば、アンチセンスオリゴヌクレオチド）を、天然に存在するヌクレオチドまたは分子の生物学的安定性を増加させるかアンチセンス核酸とセンス各案との間に形成された二重鎖の物理学的安定性を増加させるように設計された種々の修飾ヌクレオチド（例えば、ホスホロチオエート誘導体およびアクリジン置換ヌクレオチドを使用することができる）を用いて化学合成することができる。アンチセンスヌクレオチドの作製に使用することができる修飾ヌクレオチドの例には、以下が含まれる：5-フルオロウラシル、5-ブロモウラシル、5-クロロウラシル、5-ヨードウラシル、ヒポキサンチン、キサンチン、4-アセチルシトシン、5-（カルボキシヒドロキシルメチル）ウラシル、5-カルボキシメチルアミノメチル-2-チオウリジン、5-カルボキシメチルアミノメチルウラシル、ジヒドロウラシル、 β -D-ガラクトシルケオシン、イノシン、N6-イソペンテニルアデニン、1-メチルグアニン、1-メチルイノシン、2, 2-ジメチルグアニン、2-メチルアデニン、2-メチルグアニン、3-メチルシトシン、5-メチルシトシン、N6-アデニン、7-メチルグアニン、5-メチルアミノメチルウラシル、5-メトキシアミノメチル-2-チオウラシル、 β -D-マンノシルケオシン、5'-メトキシカルボキシメチルウラシル、5-メトキシウラシル、2-メトキシチオ-N6-イソペンテニルアデニン、ウラシル-5-オキシ酢酸（v）、ワイプトキソシン、プソイドウラシル、ケオシン、2-チオシトシン、5-メチル-2-チオウラシル、2-チオウラシル、4-チオウラシル、5-メチルウラシル、ウラシル-5-オキシ酢酸メチルエステル、ウラシル-5-オキサ酢酸（v）、5-メチル-2-チオウラシル、3-（3-アミノ-3-N-2-カルボキシプロピル）ウラシル、（acp3）w、および2, 6-ジアミノプリン。あるいは、核酸がアンチセンス方向（すなわち、挿入核酸から転写されたRNAは、目的の標的核酸に対してアンチセンス方向である）にサブクローニングされている発現ベクターを用いて、アンチセンス核酸を生物学的に作製することができる。

20

30

40

【0273】

さらに、本発明の核酸分子を、例えば、分子の安定性、ハイブリッド形成、または溶解性を改良するために、塩基部分、糖部分、またはリン酸塩バックボーンで改変することができる。例えば、核酸のデオキシリボースホスフェートバックボーンを、ペプチド核酸を作製するように改変することができる（Hyrupら、1996、Bioorganic & Medicinal Chemistry、4、5を参照のこと）。本明細書中で使用される、用語「ペプチド核酸」または「PNA」は、デオキシリボースホスフェートバックボーンが偽ペプチドバックボーンと置換され、たった4つの天然の核酸塩基が残され

50

るのみである核酸模倣物（例えば、DNA模倣物）をいう。PNAの中性バックボーンは、低イオン強度の条件下で、DNAおよびRNAと特異的にハイブリッド形成されることが示されている。PNAオリゴマー合成を、Hyrupら、1996、前出、Perry-O'Keefeら、1996、Proc. Natl. Acad. Sci. USA、93、14670に記載の標準的な固相ペプチド合成プロトコルを用いて行うことができる。さらに、親油性または他の補助基との結合、PNA-DNAキメラの形成、もしくはリポソームの使用、または当該分野で公知の他の薬物送達技術によって、PNAを、例えば、その安定性、特異性、または細胞取り込みを促進するように改変することができる。PNA-DNAキメラの合成を、Hyrup、1996、前出、Finnら、1996、Nucleic Acids Res.、24(17)、3357~63、Magら、1989、Nucleic Acids Res.、17、5973、およびPeterseerら、1975、Bioorganic Med. Chem. Lett.、5、1119に記載のように行うことができる。

【0274】

本発明の核酸分子および断片には、ペプチドなどの他の付加基（*in vivo*で宿主細胞受容体を標的するため）または細胞膜を通過する輸送を促進する因子（例えば、Letsingerら、1989、Proc. Natl. Acad. Sci. USA、86、6553~6556、Lemaîtreら、1987、Proc. Natl. Acad. Sci. USA、84、648~652、PCT公開番号WO 88/0918を参照のこと）または血液脳関門（例えば、PCT公開番号WO 89/10134を参照のこと）も含まれ得る。さらに、オリゴヌクレオチドを、ハイブリッド形成誘発切断剤（例えば、Krolら、1988、Bio-Techniques、6、958~976を参照のこと）または挿入剤（例えば、Zon、1988、Pharm Res.、5、539~549を参照のこと）を用いて修飾することができる。

【0275】

受容体ポリヌクレオチドはまた、受容体ポリヌクレオチドの任意の所与の領域を増幅させるためのPCRプライマーとして有用である。

【0276】

受容体ポリヌクレオチドはまた、組換えベクターの構築に有用である。このようなベクターには、受容体ポリペプチドの一部または全部を発現する発現ベクターが含まれる。ベクターはまた、受容体遺伝子および遺伝子産物の*in situ*発現を変更するための別のポリヌクレオチド配列（細胞ゲノムなど）への組込みに使用されている挿入ベクターを含む。例えば、内因性受容体コード配列を、1つまたは複数の特異的に組込んだ変異を含むコード配列の全部または一部と相同組換えを介して置換することができる。

【0277】

受容体ポリヌクレオチドはまた、抗原ペプチドの発現に有用である。高い抗原性指数を有するペプチド領域を、図2、図6、および図10に示す。

【0278】

受容体ポリヌクレオチドはまた、*in situ*ハイブリッド形成法（FISH（この技術の概説として、Vermaら、1988、Human Chromosomes: A Manual of Basic Techniques、Pergamon Press、New Yorkを参照のこと）など）および体細胞ハイブリッドのPCRマッピングによる受容体ポリヌクレオチドの染色体上の位置の同定用のプローブとして有用である。染色体に対するこの配列のマッピングは、疾患に関連する遺伝子を有するこれらの関連配列における重要な最初の段階である。

【0279】

染色体マッピング用の試薬をそれぞれ使用して、単一の染色体またはその染色体上の単一の部位をマークすることができ、複数の部位および／または複数の染色体のマーク用に試薬のパネルを使用することができる。マッピングには、遺伝子の非コード領域に対応する試薬が好ましい。コード配列は遺伝子ファミリー内で保存される傾向が強いので、染色体

マッピング中の交差ハイブリッド形成の機会が増加する。

【0280】

一旦配列が正確な染色体位置にマッピングされると、染色体上の配列の物理学的位置は、遺伝子マップデータに対応し得る。（このようなデータは、例えば、V. McKusick、Mendelian Inheritance in Man (Johns Hopkins University Welch Medical Libraryによりオンラインで利用可能である)に見出される)。次いで、同一の染色体領域にマッピングされた遺伝子および疾患との間の関係を、例えば、Egelandら、1987、Nature、325、783～787に記載の結合分析（物理的に隣接する遺伝子の同時遺伝）によって同定することができる。

10

【0281】

さらに、特定の遺伝子に関する疾患を罹患している個体と罹患していない個体との間のDNA配列の相違を同定することができる。罹患個体の一部または全部で変異が認められるが、罹患していない個体には全く認められない場合、おそらく、変異は特定の疾患の原因因子である。罹患および非罹患個体の比較には、一般に、染色体における構造の変化（可視的形態の染色体であるか、そのDNA配列に基づいたPCRを用いて検出可能な欠失または翻訳など）の最初の検索が含まれる。最終的に、何人かの個体由来の遺伝子の完全な配列決定を行って、変異の存在の確認および多形由来の変異の区別を行うことができる。

【0282】

受容体ポリヌクレオチドプローブはまた、組織分布に関する受容体およびその変異型をコードする遺伝子の存在パターン（例えば、遺伝子重複が存在するかどうかおよび重複が組織の全部または部分集合のみに存在するかどうか）の同定に有用である。遺伝子は天然存在するか、細胞、組織、または生物に外因的に移入され得る。

20

【0283】

受容体ポリヌクレオチドはまた、本明細書中に記載のポリヌクレオチドをコードする遺伝子から作製したmRNAの全部または一部に対応するリボザイムの設計に有用である。

【0284】

受容体ポリヌクレオチドはまた、受容体ポリヌクレオチドおよびポリペプチドの一部または全部を発現する宿主細胞の構築に有用である。

【0285】

受容体ポリヌクレオチドはまた、受容体ポリヌクレオチドおよびポリペプチドの一部または全部を発現するトランスジェニック動物の構築に有用である。

30

【0286】

受容体ポリヌクレオチドはまた、受容体ポリペプチドの一部または全部を発現するベクターの作製に有用である。

【0287】

受容体ポリヌクレオチドはまた、受容体核酸の発現レベルの同定用のハイブリッド形成プローブとして有用である。したがって、プローブを使用して、細胞、組織、および生物における受容体核酸の存在の検出またはレベルの同定を行うことができる。そのレベルが同定される核酸は、DNAまたはRNAであり得る。したがって、本明細書中に記載のポリペプチドに対応するプローブを使用して、所与の細胞、組織、または生物における遺伝子のコピー数を評価することができる。特に、これは受容体遺伝子の増幅が存在する場合に適切である。

40

【0288】

あるいは、プローブをin situハイブリッド形成でを使用して、受容体遺伝子の増加したコピーの位置、染色体エレメント上の位置、または受容体遺伝子が正常に見出されない染色体への組込み位置（例えば、均一に染色された領域）を評価することができる。

【0289】

これらの用途は、正常な結果と比較した受容体発現の増加または減少に関連する疾患（増殖疾患、分化もしくは発育疾患、または造血疾患など）の診断に関する。

50

【0290】

したがって、本発明は、試験サンプルを被験体から得て、核酸（例えば、mRNA、ゲノムDNA）を検出し、核酸の存在が核酸の異常な発現または活性に関連する疾患または障害の発症リスクを有する被験体を診断する、受容体核酸の異常な発現または活性に関連する疾患または障害の同定法を提供する。

【0291】

本発明の1つの態様は、個体が異常な核酸の発現または活性に関与する疾患または障害を有するかどうかまたは疾患または障害の発症のリスクを有するかどうかを同定するために、生物学的サンプル（例えば、血液、血清、細胞、組織）における核酸の発現ならびに活性を同定するための診断的アッセイに関する。このようなアッセイを、予知および予言目的で使用して、核酸分子の発現または活性に特徴づけられるか関連する障害の発症前に個体を予防的に治療することができる。

10

【0292】

mRNAの*in vitro*での検出技術には、ノーザンハイブリッド形成および*in situ*ハイブリッド形成が含まれる。DNAの*in vitro*での検出技術には、サザンハイブリッド形成および*in situ*ハイブリッド形成が含まれる。

【0293】

プローブを、被験体由来の細胞サンプル中の受容体をコードする核酸（例えば、mRNAまたはゲノムDNA）のレベルの測定または受容体に変異されている場合はその同定などによる、受容体タンパク質を発現する細胞または組織の同定用の診断試験キットの一部として使用することができる。

20

【0294】

核酸発現アッセイは、受容体核酸発現を調整する化合物（例えば、アンチセンス、ポリペプチド、ペプチド模倣物、小分子、または他の薬物）を同定するための薬物スクリーニングに有用である。細胞を、候補化合物と接触させ、mRNAの発現を同定する。候補化合物の存在下での受容体mRNAの発現レベルを、候補化合物の非存在下での受容体mRNAの発現レベルと比較する。次いで、候補化合物を、この比較に基づいて核酸発現のモジュレーターとして同定し、例えば、これを使用して、異常な核酸発現によって特徴づけられる障害を治療することができる。モジュレーターは、核酸に直接結合するか、核酸発現に影響を与える他の細胞成分との相互作用などによって間接的に発現を調整することができる。

30

【0295】

修飾法を、*in vitro*（例えば、薬剤と細胞との培養による）または患者もしくはトランスジェニック動物において*in vivo*（例えば、患者への薬剤投与による）で行うことができる。

【0296】

したがって、本発明は、受容体遺伝子の核酸発現に関連する障害の治療に使用することができる化合物の同定法を提供する。この方法は、典型的には、化合物の受容体核酸の発現調整能力をアッセイし、それにより、望ましくない受容体核酸発現によって特徴づけられる障害の治療に使用することができる化合物を同定するステップを包含する。

40

【0297】

アッセイを、細胞ベース系および無細胞系で行うことができる。細胞ベースのアッセイは、受容体核酸を天然に発現する細胞または特異的な核酸配列を発現するように遺伝子操作された組換え細胞を含む。

【0298】

あるいは、候補化合物を、患者またはトランスジェニック動物において*in vivo*でアッセイすることができる。

【0299】

受容体核酸発現アッセイには、核酸レベル（mRNAレベルなど）の直接的アッセイ、またはシグナル経路（サイクリックAMPまたはホスフェチジルイノシトールの代謝回転な

50

ど)に関連する付随的な化合物に対するアッセイを含むことができる。さらに、受容体タンパク質シグナル経路に対して上方またはダウンレギュレートされる遺伝発現もアッセイすることができる。この実施形態では、これらの遺伝子の調節領域を、ルシフェラーゼなどの受容体遺伝子に作動可能に連結させることができる。

【0300】

したがって、受容体遺伝子発現のモジュレーターを、細胞を候補化合物と接触させてmRNAの発現を同定する方法において同定することができる。候補化合物の存在下での受容体mRNAの発現レベルを、候補化合物の非存在下での受容体mRNAの発現レベルと比較する。次いで、候補化合物を、この比較を基本とした核酸発現のモジュレーターとして同定し、これを使用して異常な核酸発現によって特徴づけられる疾患の治療に使用することが10
できる。候補化合物の存在下でのmRNAの発現が、その非存在下での発現より統計的に有意に高い場合、候補化合物を核酸発現の刺激物質と同定する。候補化合物の存在下でのmRNAの発現が、その非損財質での発現より統計的に有意に低い場合、候補化合物を核酸発現のインヒビターと同定する。

【0301】

したがって、本発明は、標的として核酸を用い、受容体核酸発現を調整するための遺伝子モジュレーターとして薬物スクリーニングによって同定された化合物を用いた治療法を提供する。調整には、アップレギュレート（すなわち、活性化または作用）またはダウンレギュレート（抑制または拮抗）または核酸活性に対する効果（例えば、核酸が変異されて20
いるか不適当に修飾されている場合）が含まれる。治療は、核酸の異常な発現または活性によって特徴づけられる障害の治療である。

【0302】

あるいは、受容体核酸発現のモジュレーターは、薬物または小分子が受容体核酸発現を阻害する限り、本明細書中に記載のスクリーニングアッセイを用いて同定した小分子または薬物であり得る。

【0303】

受容体ポリヌクレオチドはまた、臨床試験または治療計画における受容体遺伝子の発現または活性に対する化合物の調整効果の監視に有用である。したがって、遺伝子発現パターンは、化合物（特に、患者が体性を示すことができる化合物）での継続的な治療効果のバロメーターとして役割を果たすことができる。遺伝子発現パターンはまた、化合物に対する30
罹患した細胞の生理学的反応の指標となるマーカーとしての役割を果たすことができる。したがって、このような監視により、この化合物または患者が体性を示さない別の化合物の投薬量の増加が可能である。同様に、核酸の発現レベルが所望のレベルより低下する場合、化合物の投与を適当に減少させることができる。

【0304】

監視ステップは、例えば、以下であり得る：（i）薬物投与前に被験体から投与前サンプルを得るステップ、（ii）投与前サンプルにおける本発明の特異的mRNAまたはゲノムDNAの発現レベルを検出するステップ、（iii）被験体から1つまたは複数の投与後サンプルを得るステップ、（iv）投与後サンプルにおけるmRNAまたはゲノムDNAの発現または活性レベルを検出するステップ、（v）投与前サンプル中のmRNAまたは40
はゲノムDNAと投与後サンプル中のmRNAまたはゲノムDNAの発現または活性レベルを比較するステップ、（vi）以上により、被験体に対する薬剤投与を増加または減少させるステップ。

【0305】

受容体ポリヌクレオチドはまた、受容体核酸における定量的変化、特に、病変を誘導する定量的変化の診断に有用である。ポリヌクレオチドを使用して、受容体遺伝子および遺伝発現産物（mRNAなど）の変異を検出することができる。ポリヌクレオチドを、受容体遺伝子の天然に存在する遺伝子変異を検出するためのハイブリッド形成プローブとして使用することができるので、変異を有する患者に変異による障害のリスクがあるかどうかを50
同定することができる。変異には、遺伝子中の1つまたは複数のヌクレオチドの欠失、付

加、または置換、染色体の再配列（逆位または転座など）、ゲノムDNAの修飾（異常なメチル化パターンなど）、または遺伝子コピー数の変化（増幅など）が含まれる。疾患がレポータータンパク質の過剰発現、発現不良、または発現の変化に起因する場合、機能不全に関連する受容体遺伝子の変異体態の検出により、活動性疾患または疾患の感受性についての診断ツールが得られる。

【0306】

受容体遺伝子の変異を、種々の技術によって核酸レベルで検出することができる。ゲノムDNAを、直接分析することも分析前にPCRを用いて増幅することもできる。RNAまたはcDNAを、同一の方法で使用する事ができる。

【0307】

一定の実施形態では、変異の検出には、ポリメラーゼ連鎖反応（PCR）（例えば、米国特許第4,683,195号および同第4,683,202号を参照のこと）（アンカーPCRまたはPACE PCRなど）またはライゲーション連鎖反応（LCR）（例えば、Landegranら、1988、Science、241、1077～1080およびNakazawaら、1994、PNAS、91、360～364を参照のこと）におけるプローブ／プライマーの使用が含まれる。後者は、遺伝子中の点変異の検出に特に有用であり得る（Abravayaら、1995、Nucleic Acids Res.、23、675～682を参照のこと）。この方法には、以下のステップを含むことができる：サンプルの細胞からの核酸（例えば、ゲノム、mRNAまたは両方）を単離するステップ、遺伝子のハイブリッド形成および増幅が起こる条件下で、核酸サンプルと1つまたは複数のプライマーとを接触させるステップ、および増幅産物の存在または非存在を検出するステップ、または増幅産物のサイズを検出してコントロールサンプルの長さと比較するステップ。検出および挿入を、正常な遺伝子型と比較した増幅産物のサイズの変化によって検出することができる。点変異を、増幅DNAの正常なRNAまたはアンチセンスDNA配列とのハイブリッド形成によって同定することができる。

【0308】

PCRおよび／またはLCRが本明細書中に記載の変異の検出に使用される任意の技術と組み合わせた予備増幅ステップとして望ましいことが意図される。

【0309】

代替増幅法には、以下が含まれる：自己持続配列複製（self sustained sequence replication）（Guatelliら、1990、Proc. Natl. Acad. Sci. USA、87、1874～1878）、転写増幅系（transcriptional amplification system）（Kwohら、1989、Proc. Natl. Acad. Sci. USA、86、1173～1177）、Q-βレプリカーゼ（Lizardiら、1988、Bio/Technology、6、1197）、または任意の他の核酸増幅法、およびその後の当業者に周知の技術を用いた増幅法の検出。これらの検出スキームは、核酸分子が非常に少数で存在する場合、核酸分子の検出に特に有用である。

【0310】

あるいは、受容体遺伝子の変異を、例えば、ゲル電気泳動によって同定された制限酵素消化パターンの変化によって直接同定することができる。

【0311】

さらに、配列特異的リボザイム（米国特許第5,498,531号）を使用して、リボザイム切断部位の発生または喪失による特異的変異の存在を得ることができる。

【0312】

完全に適合した配列を、ヌクレアーゼ切断消化アッセイまたは融点の相違によって区別することができる。

【0313】

特定の位置での配列変化を、ヌクレアーゼ保護アッセイ（RNaseおよびS1保護など）または化学的切断アッセイによって評価することもできる。

10

20

30

40

50

【0314】

さらに、変異株の受容体遺伝子と野生型遺伝子との間の配列の相違を、直接的DNA配列決定によって同定することができる。診断アッセイ（1995、Biotechniques、19、448）を行う場合、種々の自動化配列決定法（質量分析法による配列決定（例えば、PCT国際公開番号WO 94/16101、Cohenら、1996、Adv. Chromatogr.、36、127～162、およびGriffinら、1993、Appl. Biochem. Biotechnol.、38、147～159を参照のこと）を含む）を利用することができる。

【0315】

他の遺伝子変異の検出法には、以下が含まれる：切断因子からの保護を使用してRNA/RNAまたはRNA/DNA二重鎖における不適合塩基を検出する方法（Myersら、1985、Science、230、1242）、Cottonら、1988、PNAS、85、4397、Saleebaら、1992、Meth. Enzymol.、217、286～295、変異体および野生型核酸の電気泳動の移動度を比較する方法（Ortaら、1989、PNAS、86、2766、Cottonら、1993、Mutat. Res.、285、125～144、およびHayashiら、1992、Genet. Anal. Tech. Appl.、9、73～79）、および変性剤の勾配を含むポリアクリルアミドゲルにおける変異体または野生型断片の移動度を変性剤勾配ゲル電気泳動を用いて評価する方法（Myersら、1985、Nature、313、495）。アッセイの感度を、2次構造が配列の変化によってより感受性を有する（DNAよりも）RNAの使用によって増加させることができる。1つの実施形態では、本発明の方法は、ヘテロ二重鎖分析を利用して、電気泳動の移動度の変化を基本として二本鎖ヘテロ二重鎖を分離する（Keenら、1991、Trends Genet.、7、5）。他の点変異の検出法には、選択的オリゴヌクレオチドハイブリッド形成、選択的増幅、および選択的プライマー伸長が含まれる。

【0316】

1つの実施形態では、遺伝子変異を、サンプルおよびコントロール核酸（例えば、DNAまたはRNA）の数百または数千のオリゴヌクレオチドプローブ（Croninら、1996、Human Mutation 7：244～255；Kozaiら、1996、Nature Medicine 2：753～759）を含む高密度アレイとのハイブリッド形成によって同定することができる。例えば、遺伝子変異を、Croninら、前出に記載の発光DNAプローブを含む2次元アレイにおいて同定することができる。簡単に述べれば、プローブの第1のハイブリッド形成アレイを使用して、サンプルおよびコントロールにおけるDNAの長い鎖をスキャンして連続的重複プローブの直鎖状アレイの作製による配列間の塩基の変化を同定することができる。このステップにより、点変異の同定が可能である。このステップの後、検出される全ての変異型または変異体に相補的な小さな特異的プローブの使用による特異的変異の特徴づけを可能にする第2のハイブリッド形成アレイを行う。各変異アレイは、平行プローブ対、野生型遺伝子に相補的な遺伝子、および変異遺伝子に相補的な遺伝子からなる。

【0317】

受容体ヌクレオチドはまた、疾患の発症に必要ではないが治療様式に影響を与える遺伝子型についての個体の試験に有用である。したがって、ポリヌクレオチドを使用して、個体の遺伝子型と治療に使用される化合物に対する固体の反応との間の関係を研究することができる。この場合、例えば、リガンドに対する親和性を変化させる受容体遺伝子の変異により、受容体を活性化させる標準濃度のリガンドで薬物効果が過剰となり得るかまたは薬物効果が減少し得る。したがって、本明細書中に記載の受容体ポリヌクレオチドを使用して、治療に適切な化合物または投薬計画を選択するための個体における受容体遺伝子の変異の含有量を評価することができる。

【0318】

治療に影響を与える遺伝子変異を示すポリヌクレオチドは、個体の治療の変更に使用する

ことができる診断標的を提供する。したがって、これらの多形を含む組換え細胞および動物の産生により、治療化合物および投薬計画の効果的な臨床的設計が可能となる。

【0319】

本方法には、コントロール被験体由来のコントロール生物学的サンプルを得るステップと、mRNAまたはゲノムDNAの存在が生物学的サンプルにおいて検出されるように、コントロールサンプルとmRNAまたはゲノムDNAを検出することができる化合物または薬剤と接触させるステップと、コントロールサンプル中のmRNAまたはゲノムDNAの存在と試験サンプル中のmRNAまたはゲノムDNAの存在とを比較するステップとを含む。

【0320】

受容体ポリヌクレオチドはまた、配列が個々の染色体を用いて染色体上の特定の位置に同定される場合、染色体同定に有用である。第1に、DNA配列を、*in situ*または他の染色体特異的ハイブリッド形成によって染色体に適合させる。配列はまた、所望の種由来の個々の染色体を含む体細胞ハイブリッドのPCRスクリーニングに使用することができるPCRプライマーの調製によって特異的染色体に対応させることができる。プライマーに相同な遺伝子を含む染色体を含むハイブリッドのみが、増幅断片を産生する。染色体断片を使用して、半局在化を行うことができる。他のストラテジーには、標識流動分類染色体を用いた予備スクリーニングおよび染色体特異的ライブラリーとのハイブリッド形成による予備選択が含まれる。さらなるマッピングストラテジーには、伝統的に使用しているプローブよりも短いプローブを用いたハイブリッド形成が可能な蛍光*in situ*ハイブリッド形成が含まれる。染色体マッピング用の試薬をそれぞれ使用して、単一の染色体または染色体上の単一の部位を作製することができ、試薬のパネルを、複数の部位および／または複数の染色体の作製に使用することができる。実際に、遺伝子の非コード領域に対応する試薬は、マッピングでの使用に好ましい。コード配列は、遺伝子ファミリー内で保存される傾向が強いので、染色体マッピング中の交差ハイブリッド形成の機会を増加させる。

10

20

30

【0321】

受容体ポリヌクレオチドを使用して、小さな生物学的サンプル由来の個体を同定することもできる。これを、例えば、制限断片長の多形(RFLP)を用いて行って、個体を同定することができる。したがって、本明細書中に記載のポリヌクレオチドは、RFLPのDNAマーカーとして有用である(米国特許第5,272,057号を参照のこと)。

【0322】

さらに、受容体配列を使用して、個体のゲノム中の選択断片の実際のDNA配列を同定する別の技術を得ることができる。したがって、本明細書中に記載の受容体配列を使用して、配列の5'および3'末端由来の2つのPCRプライマーを調製することができる。次いで、これらのプライマーを使用して、その後の配列決定用の個体由来のDNAを増幅することができる。

【0323】

この様式で調製した個体由来の対応するDNA配列のパネルにより、個体がこのようなDNA配列の固有のセットを有するものとしての、固有の個体の同定を得ることができる。ヒトにおける対立遺伝子変異型は、各500塩基あたり約1つの頻度で起こると評価される。対立遺伝子変異型は、これらの配列のコード領域にいくらかの程度で起こり、非コード領域ではより高い程度で起こる。受容体配列を使用して、個体および組織由来のこのような配列を得ることができる。配列には、ヒトゲノムの固有の断片が存在する。本明細書中に記載の各配列を、いくらかの程度で、個体由来のDNAを同定目的で比較することができる標準として使用することができる。

40

【0324】

配列由来の試薬のパネルを使用して、個体についての固有の同定データベースを作成する場合、これらの試薬を後に使用して個体由来の組織を同定することができる。この固有の同定データベースを使用して、生きているか死んでいる個体の正の同定を、非常に小さな

50

組織サンプルから作製することができる。

【0325】

受容体ポリヌクレオチドを、法医学的同定目的に使用することもできる。PCR技術を使用して、非常に小さな生物学的サンプル（1本の毛髪、体液（例えば、血液、唾液、または精液））から得たDNA配列を増幅することができる。ついで、増幅配列を、サンプルの供給源の同定を可能にする標準と比較することができる。

【0326】

したがって、受容体ポリヌクレオチドを使用して、例えば、別の「同定マーカー」（すなわち、特定の個体に固有の別のDNA配列）の提供によってDNAベースの法医学的同定の信頼性を高めることができる、ヒトゲノム中の特異的遺伝子座に標的化されたポリヌクレオチド試薬（例えば、PCRプライマー）を得ることができる。上記のように、実際の塩基配列上方制限酵素によって得られた断片によって形成されたパターンとは別の正確なパターンとしての同定に使用することができる。非コード領域に標的化された配列は、より多数の多形が非コード領域で生じ、この技術を用いた個体の区別が容易になるので、特に有用である。断片は、少なくとも12塩基である。

10

【0327】

さらに、受容体ポリヌクレオチドを使用して、特異的組織を同定するためのポリヌクレオチド試薬（例えば、例えば、in situハイブリッド形成技術において使用することができる標識プローブまたは標識可能なプローブ）を得ることができる。これは、法病理学者が未知の起源の組織を与えられた場合、有用である。受容体プローブのパネルを使用して、種および／または器官の型によって組織を同定することができる。

20

【0328】

類似の様式で、プライマーおよびプローブを使用して、汚染について組織培養物をスクリーニングする（すなわち、培養物中の異なる細胞型の混合物の存在をスクリーニングする）ことができる。

【0329】

あるいは、受容体ポリヌクレオチドを使用してアンチセンスまたはリボザイム構築物によって受容体遺伝子配列の転写または翻訳を直接阻害することができる。したがって、異常に高いか望ましくない受容体遺伝子発現によって特徴づけられる障害では、核酸を治療に直接使用することができる。

30

【0330】

したがって、受容体ポリヌクレオチドは、細胞、組織、および生物における受容体遺伝子発現をコントロールするためのアンチセンスとして有用である。DNAアンチセンスポリヌクレオチドを、転写に関与する遺伝子領域に相補的になるように設計し、転写を阻害して受容体タンパク質の産生を阻害する。アンチセンスRNAまたはDNAポリヌクレオチドはmRNAとハイブリッド形成するので、受容体タンパク質へのmRNAの翻訳が阻害される。

【0331】

核酸発現の阻害に有用なアンチセンス分子の例として、配列番号2、4、または6の3'非翻訳領域に相補的な開始コドンおよびアンチセンス分子も含む、配列番号2、4、または6の5'非翻訳領域の断片に相補的なアンチセンス分子が含まれる。

40

【0332】

あるいは、アンチセンス分子のクラスを使用して、受容体核酸の発現を減少させるためにmRNAを不活化することができる。したがって、これらの分子により、異常なまたは望ましくない受容体核酸発現によって特徴づけられる障害を治療することができる。この技術には、mRNAの翻訳能力を減少させるmRNA中の1つまたは複数の領域に相補的なヌクレオチド配列を含むリボザイムによる切断が含まれる。可能な領域には、コード領域、特に、受容体タンパク質の触媒活性および他の機能的活性（リガンド結合など）に対応するコード領域が含まれる。

【0333】

50

受容体ポリヌクレオチドによれば、受容体遺伝子発現が異常な細胞を有する患者の遺伝子治療用のベクターも得られる。したがって、*ex vivo*で操作し、患者に戻した患者の細胞を含む組換え細胞を、患者を治療するための所望の受容体タンパク質を産生する個体に移入する。

【0334】

本発明はまた、生物学的サンプル中の受容体核酸の存在の検出用のキットを含む。例えば、キットは、生物学的サンプル中の受容体核酸を検出することができる標識されているか標識可能な核酸または薬剤などの試薬、サンプル中の受容体核酸の量を同定する手段、およびサンプル中の受容体核酸の量と標準とを比較する手段を含むことができる。化合物または薬剤を、適切な容器に封入することができる。キットは、さらに、受容体mRNAまたはDNAの検出用のキットの使用についての指示を含むことができる。 10

【0335】

コンピュータで読み込み可能な手段

本発明のヌクレオチド配列またはアミノ酸配列を、その使用を容易にするために種々の培地中에서도得られる。本明細書中で使用される、用語「提供された」は、本発明のヌクレオチド配列またはアミノ酸配列を含む、単離核酸またはアミノ酸分子以外の生産物をいう。このような生産物により、当業者が天然または精製形態で存在するヌクレオチド配列またはアミノ酸配列またはそのサブセットの試験（例えばオープンリーディングフレーム（ORF）のサブセットに直接適用できない手段を使用して生産物を試験することができる形態でのヌクレオチド配列またはアミノ酸配列が得られる。 20

【0336】

この実施形態の1つの適用では、本発明のヌクレオチド配列またはアミノ酸配列を、コンピュータで読み込み可能な媒体に記録することができる。本明細書中で使用される、用語「コンピュータ読み込み可能媒体」は、コンピュータで直接読み込みおよびアクセスすることができる任意の媒体をいう。このような媒体には、以下が含まれるが、これらに限定されない：磁気記憶媒体（フロッピーディスク、ハードディスク記憶媒体、および磁気テープなど）、光記憶媒体（CD-ROMなど）、電氣的記憶媒体（RAMおよびROMなど）、およびこれらのカテゴリーのハイブリッド（磁気／光記憶媒体など）。当業者は、現在公知の任意のコンピュータ媒体をいかに使用して媒体に記憶された本発明のヌクレオチド配列またはアミノ酸配列を有するコンピュータ読み込み可能媒体を含む生産物を製造することができるかを容易に認識する。 30

【0337】

本明細書中で使用される、用語「記録された」は、コンピュータ読み込み可能媒体上の情報の記憶プロセスをいう。当業者は、本発明のヌクレオチド配列またはアミノ酸配列情報を含む生産物の製造のために、コンピュータ読み込み可能媒体上の情報の現在公知の任意の読み込み法を容易に適用することができる。

【0338】

本発明のヌクレオチド配列またはアミノ酸配列情報を記憶させたコンピュータ読み込み可能媒体の製造のために、当業者は種々のデータ記憶構造を利用可能である。データ記憶構造の選択は、一般に、記憶情報のアクセスのために選択される手段に基づく。さらに、種々のデータ処理プログラムおよびフォーマットを使用して、コンピュータ読み込み可能媒体上の本発明のヌクレオチド配列情報を記憶することができる。配列情報を、市販のソフトウェア（Word PerfectおよびMicrosoft Wordなど）でフォーマットした文書処理テキストファイルとして処理するか、データベースアプリケーション（DB2、Sybase、Oracleなど）で記憶されたASCIIファイルの形態で処理することができる。当業者は、本発明のヌクレオチド配列情報が記録されたコンピュータ読み込み可能媒体を得るために、任意の数のデータ処理構築フォーマット（例えば、テキストファイルまたはデータベース）を容易に適用することができる。 40

【0339】

コンピュータ読み込み可能形態でのヌクレオチド配列またはアミノ酸配列の獲得によって 50

、当業者は、種々の目的に配列情報に日常的にアクセスすることができる。例えば、当業者は、データ記憶手段内に記憶された配列情報を用いて標的配列または標的構造モチーフを比較するために、コンピュータ読み込み可能形態の本発明のヌクレオチド配列またはアミノ酸配列を使用することができる。このような手段を使用して、特定の標的配列または標的モチーフに適合する本発明の配列の断片または領域を同定することができる。

【0340】

本明細書中で使用される、「標的配列」は、6つまたはそれ以上のヌクレオチドまたは2つまたはそれ以上のアミノ酸の任意のDNA配列またはアミノ酸配列であり得る。当業者は、標的配列が長いほど、標的配列はデータベースで無作為に発生しなくなることを容易に理解することができる。標的配列の最も好ましい配列長は、約10～100アミノ酸または約30～300ヌクレオチド残基である。しかし、遺伝子発現およびタンパク質のプロセッシングに關与する配列などの商業的に重要な断片はより短くても良いことが十分に認識される。 10

【0341】

本明細書中で使用される、「標的構造モチーフ」または「モチーフ」は、配列を標的モチーフの折りたたみの際に形成される3次元高次構造に基づいて選択された任意の合理的に選択された配列または配列の組み合わせをいう。当該分野で公知の種々の標的モチーフが存在する。タンパク質標的モチーフには、酵素活性部位およびシグナル配列が含まれるが、これらに限定されない。核酸標的配列には、プロモーター配列、ヘアピン構造および誘導発現エレメント（タンパク質結合配列）が含まれるが、これらに限定されない。 20

【0342】

分析および他の配列との比較用のコンピュータ読み取り可能媒体において得られる配列情報に、当業者が可能なコンピュータソフトウェアを公的に利用可能である。種々の公知のアルゴリズムが公的に開示されており、種々の市販の検索手段実行用のソフトウェアが本発明のコンピュータベースのシステムに使用するか、使用することができる。このようなソフトウェアには、Mac Pattern (EMBL)、BLASTNおよびBLASTX (NCBI)が含まれるが、これらに限定されない。

【0343】

例えば、SybaseシステムでBLAST (Altschulら、1990、J. Mol. Biol.、215、403～410) およびBLAZE (Brutlagら、1993、Comp. Chem.、17、203～207) 検索アルゴリズム実行ソフトウェアを使用して、ORFまたは他のライブラリー由来のタンパク質に相同性を有する本発明の配列のオープンリーディングフレーム (ORF) を同定することができる。このようなORFは、たんぱく質コード断片であり、種々の反応で使用される酵素などの商業的に重要なタンパク質の製造および商業的に有用な代謝産物の製造に有用である。 30

【0344】

ベクター／宿主細胞

本発明はまた、受容体ポリヌクレオチドを含むベクターも提供する。用語「ベクター」は、受容体ポリヌクレオチドを輸送することができる媒介物（好ましくは核酸分子）をいう。ベクターが核酸分子である場合、受容体ポリヌクレオチドをベクター核酸に共有結合させる。本発明のこの態様では、ベクターには、プラスミド、一本鎖または二本鎖ファージ、一本鎖または二本鎖RNAまたはDNAウイルスベクター、または人工染色体 (BAC、PAC、YAC、ORMACなど) が含まれる。 40

【0345】

ベクターを、受容体ポリヌクレオチドのさらなるコピーを複製および産生する染色体外エレメントとして宿主中に維持することができる。あるいは、ベクターを、宿主細胞ゲノムに組み込み、宿主細胞の複製時に受容体ポリヌクレオチドのさらなるコピーを産生することができる。

【0346】

本発明は、受容体ポリヌクレオチドの維持用のベクター（クローニングベクター）または 50

発現用のベクター（発現ベクター）を提供する。ベクターは、原核細胞または真核細胞またはその両方（シャトルベクター）で機能することができる。

【0347】

発現ベクターは、ポリヌクレオチドの転写が宿主中で行われるようにベクター中で受容体に作動可能に連結されたシス作用調節領域を含む。ポリヌクレオチドを、転写に影響を与えることができる別のポリヌクレオチドを有する宿主細胞に組込むことができる。したがって、第2のポリヌクレオチドは、ベクター由来の受容体ポリヌクレオチドの転写を可能にするシス調節制御コントロール領域と相互作用するトランス作用因子を提供することができる。あるいは、トランス作用因子を、宿主によって供給することができる。最終的にトランス作用因子を、ベクター自体から産生することができる。

10

【0348】

しかし、いくつかの実施形態では、受容体ポリヌクレオチドの転写および／または翻訳を無細胞系で得ることができることが理解される。

【0349】

本明細書中に記載のポリヌクレオチドを作動可能に連結することができる調節配列は、mRNA転写指向用のプロモーターを含む。これらには、λバクテリオファージ、lac、TRP、およびTAC由来のプロモーター、E. coli由来のプロモーター、SV40由来の初期および後期プロモーター、CMV媒介初期プロモーター、アデノウイルス初期および後期プロモーター、およびレトロウイルス長末端反復が含まれるが、これらに限定されない。

20

【0350】

転写を促進する調節領域に加えて、発現ベクターは、転写を調整する領域（リプレッサー結合部位およびエンハンサーなど）も含むことができる。例として、SV40エンハンサー、サイトメガロウイルス前初期エンハンサー、ポリオーマエンハンサー、アデノウイルスエンハンサー、およびレトロウイルスLTRエンハンサーが含まれる。

【0351】

転写開始および調節部位を含むことに加えて、発現ベクターはまた、転写終結に必要な配列および転写領域に翻訳のためのリボゾーム結合部位を含むことができる。他の発現用の調節制御エレメントは、開始および終結コドンならびにポリアデニル化シグナルを含む。当業者は、発現ベクターにおいて有用な多数の調節配列を認識しているであろう。このような調節配列は、例えば、Sambrookら、Molecular Cloning: A Laboratory Manual、第2版、Cold Spring Harbor Laboratory Press、Cold Spring Harbor、NY、1989に記載されている。

30

【0352】

種々の発現ベクターを使用して、受容体ポリヌクレオチドを発現させることができる。このようなベクターには、以下が含まれる：染色体、エピソーム、およびウイルス由来のベクター（例えば、細菌プラスミド、バクテリオファージ、酵母エピソーム、酵母染色体エレメント（酵母人工染色体を含む）、ウイルス（バキュロウイルス、パポバウイルス（SV40など）、ワクシニアウイルス、アデノウイルス、ポックスウイルス、仮性狂犬病ウイルス、およびレトロウイルス）由来のベクター）。ベクターはまた、これらの供給源の組み合わせ（プラスミドおよびバクテリオファージ遺伝子エレメント由来のベクター（例えば、コスミドおよびファージミド））由来であり得る。原核生物宿主および真核生物宿主用の適切なクローニングおよび発現ベクターは、Sambrookら、Molecular Cloning: A Laboratory Manual、第2版、Cold Spring Harbor Laboratory Press、Cold Spring Harbor、NY、1989に記載されている。

40

【0353】

調節配列は、温度、栄養素の添加、または外因性因子（ホルモンまたは他のリガンド）などにより1つまたは複数の宿主細胞における構築発現（すなわち、組織特異的）または1

50

つまたは複数の細胞型における誘導発現を提供することができる。原核生物宿主および真核生物宿主における構築発現および誘導発現を得るための種々のベクターは、当業者に周知である。

【0354】

受容体ポリヌクレオチドを、周知の方法によってベクター核酸に挿入することができる。一般に、最終的に発現されるDNA配列を、1つまたは複数の制限酵素でのDNA配列および発現ベクターの切断および得られた断片のライゲーションによって発現ベクターに連結する。制限酵素消化およびその後のライゲーションの手順は、当業者に周知である。

【0355】

周知の技術を用いて、適切なポリヌクレオチドを含むベクターを、増殖または発現用の適切な宿主細胞に移入することができる。細菌細胞には、*E. coli*、*Streptomyces*、および*Salmonella typhimurium*が含まれるが、これらに限定されない。真核細胞には、酵母、昆虫細胞(*Drosophila*など)、動物細胞(COSおよびCHO細胞など)、および植物細胞が含まれるが、これらに限定されない。

【0356】

本明細書中に記載のように、融合タンパク質としてポリペプチドを発現することが望ましい。したがって、本発明は、受容体ポリペプチドの産生が可能な融合ベクターを提供する。融合ベクターは、組換えタンパク質発現の増加、組換えタンパク質の溶解性の増加、および例えば、親和性精製のリガンドとしての作用によるタンパク質精製の補助が可能である。タンパク質分解部位を融合部分の結合部に移入して、所望のポリペプチドを最終的に融合部分から分離することができる。タンパク質分解酵素には、第Xa因子、トロンビン、およびエンテロキナーゼが含まれるが、これらに限定されない。典型的な融合発現ベクターには、標的組換えタンパク質を標的化するために、それぞれ、グルタチオンS-トランスフェラーゼ(GST)、マルトースE結合タンパク質、またはプロテインAに融合するpGEX(Smithら、1988、Gene、67、31~40)、pMAL(New England Biolabs、Beverly、MA)、およびpRIT5(Pharmacia、Piscataway、NJ)が含まれる。適切な誘導性非融合*E. coli*発現ベクターの例には、pTrc(Amannら、1988、Gene、69、301~315)およびpET 11d(Studierら、1990、Gene Expression Technology: Methods in Enzymology、185、60~89)が含まれる。

【0357】

宿主細胞の組換えタンパク質分解能力を阻害されている遺伝的背景を得ることによって、宿主細胞中の組換えタンパク質発現を最小にすることができる。(Gottesman、Gene Expression Technology: Methods in Enzymology、185、Academic Press、San Diego、California、1990、119~128)。あるいは、目的のポリヌクレオチド配列を、特定の宿主細胞(例えば、*E. coli*)に使用するための好ましいコドンを得るために変化させることができる。(Wadaら、1992、Nucleic Acids Res.、20、2111~2118)。

【0358】

受容体ポリヌクレオチドを、酵母で操作可能な発現ベクターによって発現させることもできる。酵母(例えば、*S. cerevisiae*)での発現用のベクターの例には、pYepSec1(Baldariら、1987、EMBO J.、6、229~234)、pMFa(Kurjanら、1982、Cell、30、933~943)、pJRY88(Schultzら、1987、Gene、54、113~123)、およびpYES2(Invitrogen Corporation、San Diego、CA)が含まれる。

【0359】

10

20

30

40

50

受容体ポリヌクレオチドを、例えば、バキュロウイルス発現ベクターを用いて、昆虫細胞中で発現させることもできる。培養昆虫細胞（例えば、Sf9細胞）におけるタンパク質発現に利用可能なバキュロウイルスベクターには、pAc系（Smithら、1983、Mol. Cell Biol.、3、2156～2165）およびpVL系（Lucklowら、1989、Virology、170、31～39）が含まれる。

【0360】

本発明の一定の実施形態では、本明細書中に記載のポリヌクレオチドを、哺乳動物の発現ベクターを用いて哺乳動物細胞中で発現させる。哺乳動物発現ベクターの例には、pCDM8（Seed、1987、Nature、329、840）およびpMT2PC（Kauffmanら、1987、EMBO J.、6、187～195）が含まれる。

10

【0361】

本明細書中に列挙した発現ベクターは、受容体ポリヌクレオチドの発現に有用な当業者が利用可能な周知のベクターのみによって得られる。当業者は、本明細書中に記載のポリヌクレオチドの増殖または発現の維持に適切な他のベクターを認識しているであろう。これらは、例えば、Sambrookら、Molecular Cloning: A Laboratory Manual、第2版、Cold Spring Harbor Laboratory、Cold Spring Harbor Press、Cold Spring Harbor、NY、1989に見出される。

【0362】

本発明はまた、本明細書中に記載の核酸配列を逆方向でベクターにクローニングされるが、アンチセンスRNAの転写が可能な調節配列に作動可能に連結されるベクターを含む。したがって、アンチセンス転写を、コード領域および非コード領域を含む、本明細書中に記載のポリヌクレオチド配列の全部または一部を作製することができる。このアンチセンスRNAの発現を、センスRNAの発現（調節配列、構造または誘導発現、組織特異的発現）に関連する上記の各パラメーターに供する。

20

【0363】

本発明はまた、本明細書中に記載のベクターを含む組換え宿主細胞に関する。したがって、宿主細胞には、原核細胞、下等真核細胞（酵母など）、他の真核細胞（昆虫など）、および高等真核細胞（哺乳動物細胞など）が含まれる。

【0364】

組換え宿主細胞を、当業者が容易に利用可能な技術によって、細胞への本明細書中に記載のベクター構築物の移入によって調製する。これらの技術には、以下が含まれるが、これらに限定されない：リン酸カリウムトランスフェクション、DEAEーデキストラン媒介トランスフェクション、陽イオン性脂質媒介トランスフェクション、エレクトロポレーション、形質導入、感染、リポフェクション、およびSambrookら、Molecular Cloning: A Laboratory Manual、第2版、Cold Spring Harbor Laboratory、Cold Spring Harbor Press、Cold Spring Harbor、NY、1989に記載の他の技術。

30

【0365】

宿主細胞は、1つを超えるベクターを含むことができる。したがって、異なるヌクレオチド配列を同一の細胞の異なるベクターに移入することができる。同様に、受容体ポリヌクレオチドを、単独または受容体ポリヌクレオチドと無関係の他のポリヌクレオチド（発現ベクターのトランス活性化を与えるもの）と共に移入することができる。1つを超えるベクターが細胞に移入された場合、ベクターを、単独で移入するか、同時に移入されるか、受容体ポリヌクレオチドベクターと結合することができる。

40

【0366】

バクテリオファージおよびウイルスベクターの場合、これらを、標準的な感染法および形質導入法によって封入またはカプセル化されたウイルスに移入することができる。ウイルスベクターは、複製コンピテントまたは複製欠損である。ウイルス複製が欠損している場

50

合、宿主中で欠損を補足する機能が得られるように複製される。

【0367】

ベクターは、一般に、組換えベクター構築物を含む細胞の亜集団の選択が可能な選択マーカーを含む。マーカーは、本明細書中に記載のポリヌクレオチドを含むベクター中に含まれるか、異なるベクター中に存在し得る。マーカーには、原核宿主細胞についてはテトラサイクリンまたはアンピシリン耐性遺伝子が含まれ、真核宿主細胞についてはジヒドロ葉酸レダクターゼまたはネオマイシン耐性遺伝子が含まれる。しかしながら、いずれのマーカーは表現型特徴が効果的であろう選択を提供する。

【0368】

成熟タンパク質を、適切な調節配列の調節下で細菌、酵母、哺乳動物細胞、および他の細胞中に産生することができる一方で、無細胞転写および翻訳系を用い、本明細書中に記載のDNA構築物由来のRNAを用いて、これらのタンパク質を産生することもできる。

【0369】

ポリペプチドの分泌を所望する場合、適切な分泌シグナルをベクターに組込む。シグナル配列は、受容体ポリペプチドに対して内因性であるか、これらのポリペプチドに対して異種であり得る。

【0370】

ポリペプチドが培地中に分泌されない場合、タンパク質を、標準的な破壊法（凍解、超音波処理、機械的破壊、溶解剤の使用などを含む）によって宿主細胞から単離することができる。次いで、ポリペプチドを回収して、周知の精製法（硫酸アンモニウム沈殿、酸抽出、陰イオンまたは陽イオン交換クロマトグラフィー、ホスホセルロースクロマトグラフィー、疎水性相互作用クロマトグラフィー、アフィニティークロマトグラフィー、ヒドロキシアパタイトクロマトグラフィー、レクチンクロマトグラフィー、または高速液体クロマトグラフィーを含む）によって精製することができる。

【0371】

本明細書中に記載のポリペプチドの組換え産生における宿主細胞に依存して、ポリペプチドは、細胞に依存して種々のグリコシル化パターンを有し、細菌で産生される場合はおそらく非グリコシル化パターンであり得ることも理解される。さらに、ポリペプチドは、宿主媒介プロセスの結果として、最初に修飾されたメチオニンを含み得る場合がある。

【0372】

ベクターおよび宿主細胞の使用

「宿主細胞」および「組換え宿主細胞」は、本発明の特定の細胞だけでなく、このような細胞の子孫または潜在的子孫をいう。変異または環境の影響によって後継世代に一定の修飾が起こり得るので、実際、このような子孫は、親細胞と同一ではないが、本明細書中で使用される用語の範囲内に依然として含まれることが理解される。

【0373】

本明細書中に記載のポリペプチドを発現する宿主細胞、特に組換え宿主細胞は、種々の用途がある。第1に、この細胞は、さらに精製して所望の量の受容体タンパク質または断片を産生することができる受容体タンパク質またはポリペプチドの産生に有用である。したがって、発現ベクターを含む宿主細胞は、ポリペプチド産生に有用である。

【0374】

宿主細胞はまた、受容体または受容体断片を含む細胞ベースのアッセイの構築に有用である。したがって、天然の受容体を発現する組換え宿主細胞は、受容体機能を刺激または阻害する化合物のアッセイに有用である。これには、リガンド結合、転写または翻訳レベルでの遺伝子発現、Gタンパク質相互作用、およびシグナル伝達経路の化合物が含まれる。

【0375】

宿主細胞はまた、これらの機能が影響を受ける受容体変異の同定に有用である。変異が自然発生して疾患を引き起こす場合、変異を含む宿主細胞は、天然の受容体に対するその効果によって示すことができる変異受容体に対する所望の効果（例えば、機能の刺激または阻害）を有する化合物のアッセイに有用である。

10

20

30

40

50

【0376】

組換え宿主細胞はまた、異種アミノ末端細胞外ドメイン（または他の結合領域）によって活性化を活性化または抑制する化合物を評価するための、本明細書中に記載のキメラポリペプチドの発現に有用である。あるいは、全膜貫通ドメイン（またはその一部）にわたる異種領域を使用して、任意の所与の宿主細胞に対する所望のアミノ末端細胞外ドメイン（または他の結合領域）の効果を評価することができる。この実施形態では、特異的宿主に適合する全膜貫通ドメイン（またはその一部）にわたる領域を使用してキメラベクターを作製する。あるいは、異種カルボキシ末端の細胞内（例えば、シグナル伝達）ドメインを、宿主細胞に移入することができる。

【0377】

さらに、1つまたは複数の種々の機能（例えば、リガンド結合またはGタンパク質結合）を増加させるかまたは減少させるように操作し、この機能を使用して個体中の受容体タンパク質を増加または置換するように変異受容体を設計することができる。したがって、宿主細胞により、異常な受容体の置換または治療効果が得られる異常な受容体の提供によって治療の利点を得ることができる。1つの実施形態では、細胞により異常に活性な受容体

10

【0378】

別の実施形態では、細胞により異常に不活性な受容体を得られる。これらの受容体は、個体における内因性受容体と競合することができる。

【0379】

別の実施形態では、リガンドの内因性受容体と競合させるために、活性化することができない受容体を発現する細胞を個体に移入する。例えば、過剰なリガンドが治療様式の一部である場合、このリガンドを治療の特定の点で不活性にする必要があり得る。リガンドと競合するが、受容体の活性化に影響を与え得ない細胞が有利であろう。

20

【0380】

宿主細胞ゲノムにおける内因性受容体ポリヌクレオチド配列の*in situ*での変化が可能な相同組換え宿主細胞を産生することもできる。宿主細胞には、安定な細胞株、*in vivo*での細胞、またはクローニング微生物が含まれるが、これらに限定されない。この技術は、WO 93/09222、WO 91/12650、WO 91/06667、米国特許第5,272,071号、および同第5,641,670号により完全に記載されている。簡単に述べれば、受容体遺伝子に対して近位または遠位の受容体ポリヌクレオチドまたは配列に対応する特異的ポリヌクレオチド配列を、遺伝子発現に影響を受け得る場合、相同組換えによって宿主細胞に組込む。1つの実施形態では、内因性配列の発現を増加させるか減少させる調節配列を移入する。したがって、受容体タンパク質を通常は産生しない細胞中に産生させることができる。あるいは、受容体タンパク質の発現の増加を、特定のレベルでタンパク質を正常に産生する細胞に影響を与えることができる。さらに、発現を、特異的調節配列の移入によって減少または消滅させることができる。調節配列は、受容体タンパク質配列と異種であるか、発現に影響を与える所望の変異を有する相同性配列であり得る。あるいは、全遺伝子を検出することができる。調節配列は、宿主細胞に特異的であるか、ある細胞型よりも機能し得る。なおさらに、特異的変異を、任意の所望の遺伝子領域に移入して変異受容体タンパク質を産生することができる。このような変異を、例えば、リガンド結合部位などの特異的な機能的領域に移入することができる。

30

40

【0381】

1つの実施形態では、宿主細胞は、受容体遺伝子の変化を含むトランスジェニック動物の作製に使用することができる、受精卵母細胞または胚幹細胞であり得る。あるいは、宿主細胞は、細胞の特異的サブセットを生じ、動物中のトランスジェニック組織の作製に使用することができる幹細胞または他の初期組織前駆体であり得る。相同組換えベクターについては、Thomasら、1987、Cell、51、503も参照のこと。ベクターを、胚幹細胞に移入し（例えば、エレクトロポレーションによる）、移入遺伝子が内因性受

50

容体遺伝子で相同組換えされている細胞を選択する（例えば、L i eら、1992、C e l l、69、915を参照のこと）。次いで、選択した細胞を、動物（マウス）の芽細胞に注射して、凝集キメラを形成させる（例えば、B r a d l e y、T e r a t o c a r c i n o m a s a n d E m b r y o m i c S t e m C e l l s : A p r a c t i c a l A p p r o a c h、E. J. R o b e r t s o n編、I R L、O x f o r d、1987、113～152を参照のこと）。次いで、キメラ胚を、適切な偽妊娠雌乳母に移植し、胚を出産させる。その生殖細胞に相同組換えDNAを保有する子孫を使用して、この動物の全細胞が導入遺伝子の生殖系列伝達によって相同組換えDNAを含む動物を繁殖させる。相同組換えベクターおよび相同組換え動物の構築法は、B r a d l e y、1991、C u r r e n t O p i n i o n i n B i o t e c h n o l o g y、2、82 10 3～829およびP C T国際公開番号W O 9 0 / 1 1 3 5 4、W O 9 1 / 0 1 1 4 0、およびW O 9 3 / 0 4 1 6 9にさらに記載されている。

【0382】

遺伝子操作した宿主細胞を使用して、非ヒトトランスジェニック動物を作成することができる。トランスジェニック動物は、動物の1つまたは複数の細胞が導入遺伝子を含む哺乳動物（例えば、ラットまたはマウスなどのげっ歯類）であることが好ましい。導入遺伝子は、トランスジェニック動物が発生し、トランスジェニック動物の1つまたは複数の細胞型または組織における成熟動物のゲノムに残存する細胞のゲノムに組込まれる外因性DNAである。これらの動物は、受容体タンパク質機能の研究ならびに受容体タンパク質活性のモジュレーターの同定および評価に有用である。 20

【0383】

トランスジェニック動物の他の例には、非ヒト霊長類、ヒツジ、イヌ、ウシ、ヤギ、ニワトリ、および両生類が含まれる。

【0384】

1つの実施形態では、宿主細胞は、受容体ポリヌクレオチド配列が移入されている受精卵母細胞または胚幹細胞である。

【0385】

トランスジェニック動物を、マイクロインジェクション、レトロウイルス感染などによる受精卵母細胞の雄性前核への核酸の移入および偽妊娠雌乳母動物中での卵母細胞の発生によって作製することができる。任意の受容体ヌクレオチド配列を、導入遺伝子として非ヒト動物（マウスなど）のゲノムに移入することができる。 30

【0386】

任意の調節配列または発現ベクターに有用な他の配列は、トランスジェニック配列の一部を形成することができる。これには、まだ含まれていない場合、イントロン配列およびポリアデニル化シグナルを含む。特定の細胞に受容体タンパク質発現を指向させるために、組織特異的調節配列を導入遺伝子に作動可能に連結させることができる。

【0387】

胚操作およびマイクロインジェクションを介したトランスジェニック動物（マウスなど）の作製法は、当該分野で慣習的であり、例えば、L e d e rらに付与された米国特許第4, 736, 866号、同第4, 870, 009号、W a g n e rらに付与された米国特許第4, 873, 191号、およびH o g a n、M a n i p u l a t i n g t h e M o u s e E m b r y o、C o l d S p r i n g H a r b o r L a b o r a t o r y P r e s s、C o l d S p r i n g H a r b o r、N Y、1986に記載されている。他のトランスジェニック動物作製用の類似の方法を使用する。初代トランスジェニック動物を、そのゲノムにおける導入遺伝子の存在および／または動物の組織または細胞中のトランスジェニックmRNAの発現に基づいて同定することができる。次いで、初代トランスジェニック動物を使用して、導入遺伝子を保有するさらなる動物を繁殖させることができる。さらに、導入遺伝子を保有するトランスジェニック動物を、他の導入遺伝子を保有する他のトランスジェニック動物とさらに繁殖させることができる。トランスジェニック動物には、全動物または動物中の組織が本明細書中に記載の相同組換え宿主細胞を用 50

いて作製した動物も含まれる。

【0388】

別の実施形態では、導入遺伝子の発現の調節が可能な選択された系を含む非ヒトトランスジェニック動物を作製することができる。このような系の一例は、バクテリオファージP1のcre/loxPリコンビナーゼ系である。cre/loxPリコンビナーゼ系の記載については、例えば、Laksora、1992、PNAS、89、6232～6236を参照のこと。リコンビナーゼ系の別の例は、S. cerevisiaeのFLPリコンビナーゼ系である(O'Gormanら、1991、Science、251、1351～1355)。cre/loxPリコンビナーゼ系を使用して導入遺伝子の発現を調節する場合、Creリコンビナーゼおよび選択されたタンパク質の両方をコードする導入遺伝子を含む動物が必要である。このような動物を、例えば、一方は選択されたタンパク質をコードする導入遺伝子を含み、他方はリコンビナーゼをコードする導入遺伝子を含む2つの導入遺伝子動物の交配によって、「二重」導入遺伝子動物の構築物から得ることができる。

【0389】

本明細書中に記載の非ヒト導入遺伝子動物のクローンを、Wilmutら、1997、Nature、385、810～813ならびにPCT国際公開番号WO 97/07668およびWO 97/07669に記載の方法にしたがって産生することもできる。簡単に述べれば、トランスジェニック動物由来の細胞(例えば、体細胞)を単離して、増殖サイクルを抜け出てG₀期に入るように誘導することができる。次いで、例えば、電気パルスの使用によって、静止期細胞が単離された同種の動物由来の除核卵母細胞に静止期細胞を融合させることができる。次いで、再構築した卵母細胞を培養して桑実胚または芽細胞まで発達させ、偽妊娠雌乳母動物に移す。この雌乳母から誕生した子孫は、細胞(例えば、体細胞)が単離された動物クローンである。

【0390】

本明細書中に記載のポリペプチドを発現する組換え細胞を含むトランスジェニック動物は、in vivoでの本明細書中に記載のアッセイの構築に有用である。したがって、in vivoで存在し、リガンド結合、受容体の活性化、およびシグナル伝達に影響を与え得る種々の生理学的因子は、in vitroでの無細胞または細胞ベースのアッセイからは明白ではない。したがって、in vivoでの受容体機能(リガンド相互作用を含む)、受容体機能およびリガンド相互作用に対する特異的変異受容体の影響、キメラ受容体の影響を評価するための非ヒトトランスジェニック動物の獲得に有用である。1つまたは複数の受容体機能を実質的または完全に除去する変異である無発現変異の影響を評価することも可能である。

【0391】

一般に、トランスジェニック動物の作製法は、本発明の核酸配列(トランスジェニック動物において受容体タンパク質を発現することができる核酸配列)を培養細胞またはin vivoで細胞に移入するステップを包含する。in vivoで移入した場合、核酸を、1つまたは複数の細胞型、すなわち、1つまたは複数の組織型が受容体タンパク質をコードする核酸を発現するようにインタクトな生物に移入する。あるいは、トランスジェニック動物の作製について本明細書中に記載のように、培養細胞(胚幹細胞など)のトランスフェクションによって生物中の実質的に全ての細胞に核酸を移入することができ、この細胞を用いて完全なトランスジェニック生物を作製することができる。記載のように、さらなる実施形態では、宿主細胞は、受精卵母細胞であり得る。次いで、このような細胞をメス乳母動物において発生させて、トランスジェニック生物を作製する。

【0392】

医薬組成物

受容体核酸分子、タンパク質(特に、アミノ末端細胞外ドメインなどの断片)、タンパク質のモジュレーター、および抗体(本明細書中で「活性化合物」ともいう)を、被験体(例えば、ヒト)への投与に適切な医薬組成物に組み込むことができる。このような組成物は

、典型的には、核酸分子、タンパク質、モジュレーター、または抗体および薬学的に受容可能なキャリアを含む。

【0393】

本明細書中で使用される、用語「薬学的に受容可能なキャリア」は、薬学的投与に適合する任意および全ての溶媒、分散媒、被覆物、抗菌薬および抗真菌薬、等張剤および吸収遅延剤などを含むことを意図する。薬学的有効成分用のこのような媒体および薬剤は、当該分野で周知である。任意の従来の媒体または薬剤が活性化合物と適合する範囲外で、このような媒体を、本発明の組成物中で使用することができる。補助的に活性な化合物を、この組成物に組込むこともできる。本発明の医薬組成物は、その意図する投与経路と適合するように処方される。投与経路の例には、非経口（例えば、静脈内）、皮内、皮下、経口（例えば、吸入）、経皮（局所）、経粘膜、および直腸投与が含まれる。非経口、皮内、または皮下への適用に使用される溶液または懸濁液には、以下の成分が含まれる：滅菌希釈剤（注射用の水、生理食塩水、不揮発油、ポリエチレングリコール、グリセリン、プロピレングリコール、または他の合成溶媒など）；抗菌薬（ベンジルアルコールまたはメチルパラベンなど）；抗酸化剤（アスコルビン酸または亜硫酸水素ナトリウムなど）；キレート剤（エチレンジアミン四酢酸など）；緩衝液（酢酸塩、クエン酸塩、またはリン酸塩など）および弾力性調整剤（塩化ナトリウムまたはデキストロースなど）。pHを、酸または塩基（塩酸または水酸化ナトリウム）で調整することができる。非経口用調製物を、ガラスまたはプラスチック製のアンプル、使い捨てシリンジ、または複数回投与用バイアル中に封入することができる。

10

20

【0394】

注射用に適切な医薬組成物には、滅菌水溶液（水溶性の場合）または分散液および滅菌注射用溶液または分散液の即席の調製用の滅菌粉末が含まれる。静脈内投与では、適切なキャリアには、生理食塩水、静菌水溶液、Cremophor ELTM（BASF、Parshippany、NJ）、またはリン酸緩衝化生理食塩水（PBS）が含まれる。全ての場合、組成物は滅菌されていなければならない、シリンジでの扱いやすい範囲の液体であるべきである。組成物は、製造および保存条件下で安定であり、微生物（細菌および真菌など）の汚染に反して保存されなければならない。キャリアは、例えば、水、エタノール、ポリオール（例えば、グリコール、プロピレングリコール、および液体ポリエチレングリコールなど）を含む溶媒または分散媒およびその適切な混合物であり得る。適切な流動性を、例えば、被覆物（レシチンなど）の使用、分散されている場合は必要な粒子サイズの維持、および界面活性剤の使用によって維持することができる。微生物作用の保護を、種々の抗菌薬および抗真菌薬（例えば、パラベン、クロロブタノール、フェノール、アスコルビン酸、チメロサルなど）によって達成することができる。多くの場合、組成物中に等張剤（例えば、糖、ポリアルコール（マンニトールなど）、ソルビトール、塩化ナトリウム）を含むことが好ましい。吸収遅延剤（例えば、モノステアリン酸アルミニウムおよびゼラチン）の添加によって、注射可能な組成物の吸収を延長させることができる。

30

【0395】

滅菌注射用溶液を、上記成分の1つまたは組み合わせを含む適切な溶液中への、必要量の活性化合物（例えば、受容体タンパク質または抗受容体抗体）の組込み、および必要ならばその後濾過滅菌することによって調製することができる。一般に、分散液を、ベースとなる分散媒および上記の必要な他の成分を含む滅菌賦形剤への活性化合物の組込みによって調製することができる。滅菌注射用溶液調製用の滅菌粉末の場合、好ましい調製法は、先に滅菌濾過した溶液から活性成分および任意のさらなる所望の成分の粉末を得るための減圧乾燥および凍結乾燥である。

40

【0396】

経口用組成物には、不活性希釈剤または食用キャリアが含まれる。これらを、ゼラチンカプセルに封入するか、錠剤に圧搾することができる。経口投与では、薬剤を、胃で残存するために腸溶形態に含め、さらに公知の方法でGI管の特定の領域で放出されるように被覆または混合することができる。経口治療投与の目的で、活性化合物を、補形薬と共に組

50

込んで、錠剤、トローチ、またはカプセルの形態で 사용할 수 있는。経口用組成物を、うがい薬としての用途で液体キャリアを用いて調製することもでき、液体キャリア中の化合物を経口で適用し、すすいで、吐き出すか飲み込む。薬学的に適合可能な結合剤および／またはアジュバント物質を、組成物の一部として含むことができる。錠剤、丸薬、カプセル、トローチなどは、以下の任意の成分または類似の性質の化合物を含むことができる：結合剤（微結晶性セルロース、トラガカントガム、またはゼラチンなど）；補形薬（デンプンまたはラクトース、崩壊薬（アルギン酸など）、Primogel、またはコーンスターチなど）；滑剤（ステアリン酸マグネシウムまたはSterotesなど）；潤滑剤（コロイド状シリコンジ옥シ드など）；甘味料（スクロースまたはサッカリンなど）；香料料（ペパーミント、サリチル酸メチル、またはレンジ香料など）。 10

【0397】

吸入投与では、化合物は、適切な噴霧剤（例えば、二酸化炭素などのガス）を含む圧縮コンテナまたはディスペンサーまたは噴霧器からエアゾールスプレーの形態で送達される。

【0398】

全身投与はまた、経粘膜または経皮的手段であり得る。経粘膜または経皮投与では、浸透させるバリアに適切な浸透剤を処方物に使用する。一般に、このような浸透剤は公知であり、例えば、経粘膜投与では、界面活性剤、胆汁酸塩、およびフシジン酸誘導体が含まれる。経粘膜投与を、鼻内噴霧または座薬の使用によって行うことができる。経皮投与では、活性化化合物を、一般に当該分野で公知の軟膏、軟膏、ゲル、またはクリームに処方する。 20

【0399】

化合物を、直腸送達用の座薬（例えば、ココアバターおよび他のグリセリドなどの従来の座薬基剤を用いる）または持続的浣腸の形態に調製することもできる。

【0400】

1つの実施形態では、活性化化合物を、移植およびマイクロカプセル化送達系を含む身体の迅速な排泄から化合物を保護するキャリア（放出調節処方物など）と共に調製する。生分解性で生体適合性のポリマー（エチレンビニルアセテート、ポリ無水物、ポリグリコール酸、コラーゲン、ポリオルソエステル、およびポリ乳酸）を使用することができる。このような処方物の調製法は、当業者に明らかである。物質を、Alza CorporationおよびNova Pharmaceuticals Inc. から購入することも 30
できる。リボソーム懸濁液（ウイルス抗原に対するモノクローナル抗体により感染細胞をターゲティングされたリボソームを含む）を、薬学的に受容可能なキャリアとして使用することもできる。これらを、当業者に周知の方法（例えば、米国特許第4,522,811号）にしたがって調製することができる。

【0401】

投与が容易で投薬量が均一な投薬単位形態での経口または非経口組成物を処方するのに特に有利である。本明細書中で使用される「投薬単位形態」は、治療を受ける被験体用の単位用量として適切な物理的分散単位をいう。予め秤量した量の活性化化合物を含む各単位は、必要な薬学的キャリアに関連する所望の治療効果が得られるように計算した。本発明の投薬単位形態の仕様は、活性化化合物の固有の特徴および達成すべき特定の治療効果、なら 40
びに個体治療用の活性化化合物のような複合材の当該分野での本来の制限に影響されるか依存する。

【0402】

本発明の核酸分子を、ベクターに挿入し、これを遺伝子治療のベクターとして使用することができる。遺伝子治療ベクターを、例えば、静脈内注射、局所投与（米国特許第5,328,470号）または定位固定注射（例えば、Chenら、1994、PNAS、91、3054～3057を参照のこと）によって被験体に送達させることができる。遺伝子治療ベクターの薬学的調製には条件を満たした希釈剤中に遺伝子治療ベクターを含むことができるか、遺伝子送達賦形剤が埋め込まれた放出遅延基質を含むことができる。あるいは、完全な遺伝子送達ベクターを組換え細胞（例えば、レトロウイルスベクター）からイ 50

ンタクトで産生することができる場合、薬学的調製物は、遺伝子送達系が得られる1つまたは複数の細胞を含むことができる。

【0403】

医薬組成物には、投与の支持と共にコンテナ、パック、またはディスペンサーを含むことができる。

【0404】

本明細書中で定義のように、治療有効量（すなわち、有効投薬量）のタンパク質またはポリペプチドは、約0.001～30mg/kg体重、好ましくは焼く0.01～25mg/kg体重、より好ましくは0.1～20mg/kg体重、さらにより好ましくは1～10mg/kg、2～9mg/kg、3～8mg/kg、4～7mg/kg、または5～6mg/kgの範囲である。 10

【0405】

当業者は、一定の因子が、患者の効果的な治療に必要な投薬量に影響を植え得ることを認識し、この因子には、以下が含まれるが、これらに限定されない：疾患または障害の重症度、薬歴、身体全体の健康および／または被験体の年齢。さらに、治療有効量のタンパク質、ポリペプチド、または抗体での被験体の治療を、一連の治療に含むことができる。好ましい例として、被験体を、約0.1～20mg/kg体重を1週間に1回約1週間～10週間、好ましくは2～8週間、よりこのましくは3～7週間、さらにより好ましくは約4、5、または6週間の範囲のタンパク質またはポリペプチドで治療する。治療に使用される有効量の抗体、タンパク質、またはポリペプチドを、特定の治療クールで増加または減少させることができることも認識される。投薬量に変化して、本明細書中に記載の診断アッセイの結果から明白になり得る。 20

【0406】

本発明は、発現または活性を調整する因子を含む。因子は、小分子であり得る。例えば、このような小分子には、ペプチド、ペプチド模倣物、アミノ酸、アミノ酸アナログ、ポリヌクレオチド、ポリヌクレオチドアナログ、ヌクレオチド、ヌクレオチドアナログ、または約10,000g/モル未満の分子量を有する無機化合物（すなわち、ヘテロ有機化合物および有機金属化合物を含む）、5,000g/モル未満の分子量を有する有機化合物または無機化合物、1,000g/モル未満の分子量を有する有機化合物または無機化合物、500g/モル未満の分子量を有する有機化合物または無機化合物、および塩、エステル、ならびにこのような化合物の薬学的に受容可能な他の形態が含まれるが、これらに限定されない。 30

【0407】

適切な用量の小分子が通常の技術を有する医師、獣医、または研究者の知識の範囲内で多数の因子に依存することが認識される。小分子の用量は、例えば、治療を受けるサンプルの同一性、サイズ、および条件に依存し、さらに、組成物の投与経路、および適用できるならば、本発明の核酸またはポリペプチドを有する小分子に実施者が所望する効果に依存して変化する。例として、投与量には、ミリグラムまたはマイクログラム量の小分子/キログラムの被験体サンプルの重さ（例えば、約1μg/kg～500mg/kg、約100μg/kg～5mg/kg、約1μg/kg～50mg/kg）が含まれる。小分子の適切な量は、調整される発現または活性に関連する小分子の効力に依存することがさらに理解される。このような適切な用量を、本明細書中に記載のアッセイを用いて同定することができる。本発明のポリペプチドまたは核酸の発現または活性を調整するために、1つまたは複数のこれらの小分子が動物（例えば、ヒト）に投与される場合、医師、獣医、または研究者は、例えば、最初は比較的低用量で処方して、その後適切な反応が得られるまで段階的に増加させる。さらに、任意の特定の動物についての特定の用量レベルは使用した特異的化合物の活性、被験体の年齢、体重、身体全体の健康、性、および被験体の食事、投与時間、投与経路、排泄速度、および任意の薬物の組み合わせ、および調整される発現または活性の程度を含む種々の因子に依存することが理解される。 40

【0408】

本発明を、多くの異なる形態で実施することができるが、本明細書中に記載の実施形態に制限されることを意図しない。むしろ、この開示が当業者に本発明を完全に伝達するようにこれらの実施形態を提供する。本発明の多数の修正および他の実施形態が当業者に予想されるが、本発明は、上記の記載に示した技術の利益を有する。特定の用語を使用したか、これらは、特記しない限り当該分野で使用されている。

【図面の簡単な説明】

【図 1】

図 1 は、Prosite データベースのタンパク質パターンに対する 2838 受容体の比較を示し、特に、7 回膜貫通セグメントロドプシンスーパーファミリーに対して高いスコアを示す。下線領域は、GPCR サイン、および特に GPCR に保存されたアルギニン残基の位置を示す。最もよく保存されている配列は、アスパラギン酸、アルギニンおよびチロシン (DRY) コドンである。DRY は、シグナル伝達に関与する。アルギニンは不変である。アスパラギン酸は、数個の GPCR に保存的に配置されている。今回の場合では、アルギニンが配列 DRF に見られ、これはロドプシンスーパーファミリー受容体の GPCR の DRY または不変アルギニンの位置に適合している。

10

【図 2】

図 2 は、2838 アミノ酸配列の解析： α β ターンおよびコイル領域；親水性；両親媒性領域；可動領域；抗原性係数；および表面確率プロットを示す。

【図 3】

図 3 は、2838 受容体疎水性プロットを示す。アミノ酸は、1 ~ 319 に対応し、7 回膜貫通セグメントを示す。

20

【図 4】

図 4 は、特定の機能的部位に対応するアミノ酸の 2838 オープンリーディングフレームの解析を示す。グリコシル化部位が、アミノ酸 5 ~ 8 番目に見られ、これはアミノ末端細胞外ドメインに相当する。第二のグリコシル化部位が、アミノ酸 171 ~ 174 に見られ、これは第二細胞外ループに相当する。プロテインキナーゼ C リン酸化部位が、アミノ酸 134 ~ 136 番目に見られ、これは第二細胞内ループにある。第二のプロテインキナーゼ C リン酸化部位が、アミノ酸 178 ~ 180 番目に見られ、これは第二細胞外ループにある。カゼインキナーゼ II リン酸化部位がアミノ酸 6 ~ 9 番目に見られ、これはカルボキシ末端細胞内ドメインにある。第二カゼインキナーゼ II リン酸化部位が、アミノ酸 95 ~ 98 番目に見られ、これは第一細胞外ループにある。N-ミリスチル化部位が、アミノ酸 34 ~ 39 番目に見られ、これは第一膜貫通部にある。第二の N-ミリスチル化部位がアミノ酸 107 ~ 112 番目に見られ、これは第三膜貫通部にある。第三の N-ミリスチル化部位がアミノ酸 140 ~ 145 番目に見られ、これは第四膜貫通部にある。アミド化部位が、アミノ酸 209 ~ 212 番目に見られ、これは第五膜貫通部および第三細胞内ループにおよぶ。さらに、位置が GPCR に対応し、不変アルギニンを含むアミノ酸が、アミノ酸 118 ~ 120 番目の配列 DRF に見られる。

30

【図 5】

図 5 は、Prosite データベースのタンパク質パターンに対する 14618 受容体の比較を示し、特に 7 回膜貫通セグメントロドプシンスーパーファミリーに対して高いスコアを示す。下線領域は、GPCR サイン、および特に GPCR に保存されたアルギニン残基の位置を示す。最もよく保存されている配列は、アスパラギン酸、アルギニン、チロシン (DRY) コドンである。DRY は、シグナル伝達に関与する。アルギニンは不変である。アスパラギン酸は、数個の GPCR に保存的に配置されている。今回の場合では、アルギニンが配列 DRF に見られ、これはロドプシンスーパーファミリー受容体の GPCR の DRY または不変アルギニンの位置に適合している。

40

【図 6】

図 6 は、14618 アミノ酸配列の解析： α β ターンおよびコイル領域；親水性；両親媒性領域；可動領域；抗原性係数；および表面確率プロットを示す。

【図 7】

50

図 7 は、1 4 6 1 8 受容体疎水性プロットを示す。アミノ酸は 1 ～ 3 3 7 番目に対応し、7 回膜貫通セグメントを示す。

【図 8】

図 8 は、特定の機能的部位に対応するアミノ酸の 1 4 6 1 8 オープンリーディングフレームの解析を示す。グリコシル化部位が、アミノ酸 6 ～ 9 番目に見られ、これはアミノ末端細胞外ドメインに相当する。第二のグリコシル化部位が、アミノ酸 1 6 9 ～ 1 7 2 番目に見られ、これは第二細胞外ループに相当する。第三のグリコシル化部位が、アミノ酸 1 8 0 ～ 1 8 3 番目に見られ、これは第二の細胞外ループに相当する。第四のグリコシル化部位が、アミノ酸 2 2 4 ～ 2 2 7 番目に見られ、これは第三の細胞内ループに相当する。第五のグリコシル化部位が、アミノ酸 2 6 2 ～ 2 6 5 番目に見られ、これは第三の細胞外ループに相当する。c A M P および c G M P 依存性プロテインキナーゼリン酸化部位が、アミノ酸 3 0 4 ～ 3 0 7、3 1 0 ～ 3 1 3 番目、および 3 2 3 ～ 3 2 6 に見られ、全てカルボキシ末端細胞内ドメインにある。プロテインキナーゼ C リン酸化部位が、アミノ酸 1 3 6 ～ 1 3 8 番目に見られ、これは第二の細胞内ループに相当する。第二および第三のプロテインキナーゼ C リン酸化部位が、アミノ酸 2 2 0 ～ 2 2 2 番目および 2 2 7 ～ 2 2 9 番目に見られ、両方共が、第三の細胞内ループに相当する。第四のプロテインキナーゼ C リン酸化部位が、アミノ酸 3 0 8 ～ 3 1 0 番目に見られ、これはカルボキシ末端細胞内ドメインに相当する。カゼインキナーゼ I I リン酸化部位がアミノ酸 1 3 ～ 1 6 番目および 1 7 ～ 2 0 番目に見られ、両方共がアミノ末端細胞外ドメインにある。第三のカゼインキナーゼ I I リン酸化部位が、カルボキシ末端細胞内ドメインに相当する、アミノ酸 3 2 6 ～ 3 2 9 番目に見られる。微小体 C 末端標的シグナルが、アミノ酸 3 3 5 ～ 3 3 8 番目に見られ、これはカルボキシ末端細胞内ドメインに相当する。さらに、位置が G P C R サインに対応し、不変アルギニンを含むアミノ酸が、アミノ酸 1 2 0 ～ 1 2 2 番目の配列 F R C に見られる。

【図 9】

図 9 は、P r o s i t e データベースのタンパク質パターンに対する 1 5 3 3 4 受容体の比較を示し、特に 7 回膜貫通セグメントロドプシンスーパーファミリーに対して高いスコアを示す。下線領域は、G P C R サイン、および特に G P C R に保存されたアルギニン残基の位置を示す。最もよく保存されている配列は、アスパラギン酸、アルギニン、チロシン (D R Y) コドンである。D R Y は、シグナル伝達に参与する。アルギニンは不変である。アスパラギン酸は、数個の G P C R に保存的に配置されている。今回の場合では、アルギニンが配列 D R F に見られ、これはロドプシンスーパーファミリー受容体の G P C R の D R Y または不変アルギニンの位置に適合する。

【図 10】

図 10 は、1 5 3 3 4 アミノ酸配列の解析：α β ターンおよびコイル領域；親水性；両親媒性領域；可動領域；抗原性係数；および表面確率プロットを示す。

【図 11】

図 11 は、1 5 3 3 4 受容体疎水性プロットを示す。アミノ酸は 7 回膜貫通セグメントを示す。

【図 12】

図 12 は、特定の機能的部位に対応するアミノ酸の 1 5 3 3 4 オープンリーディングフレームの解析を示す。グリコシル化部位が、アミノ酸 4 ～ 7 番目および 9 ～ 1 2 番目に見られ、これはアミノ末端細胞外ドメインにある。グリコシル化部位がまた、アミノ酸 2 5 1 ～ 2 5 4 番目に見られ、これは第六膜貫通部にある。さらなるグリコシル化部位が、アミノ酸 3 2 3 ～ 3 2 6 番目に見られ、これはカルボキシ末端ドメインにある。c A M P および c G M P 依存性プロテインキナーゼリン酸化部位が、アミノ酸 2 2 9 ～ 2 3 2 番目に見られ、これは第三の細胞内ループにある。プロテインキナーゼ C リン酸化部位が、アミノ末端ドメインに相当する、アミノ酸 2 1 ～ 2 3 番目、第三細胞内ループに相当する 2 1 1 ～ 2 1 3 番目、2 2 6 ～ 2 2 8 および 2 3 2 ～ 2 3 4、およびカルボキシ末端細胞内ドメインに相当する 3 0 7 ～ 3 0 9 番目および 3 3 2 ～ 3 3 4 番目に見られる。カゼインキナ

ーゼ I I リン酸化部位が、第二細胞外ループにあるアミノ酸 178～181 番目、および、カルボキシ末端細胞内ドメインにある 342～345 番目に見られる。N-ミリスチル化部位が、アミノ末端細胞外ドメインにあるアミノ酸 36～41 番目、第六膜貫通部および第三細胞外ループにおよぶ 258～263 番目、並びに、カルボキシ末端細胞内ドメインに相当する 324～329 番目に見られる。さらに、GPCR サイン、DRY は、アミノ酸 118～120 番目に見られ、これは第二の細胞内ループにある。

【図 13】

図 13 は、様々な正常なヒト組織における 14618 受容体の発現を示す。

【図 14】

図 14 は、様々な組織および造血系の細胞における 14618 受容体の発現を示す。

10

【図 15】

図 15 は、造血系の細胞を含む、様々な細胞および組織における 2838 受容体の発現を示す。

【図 16】

図 16 は、正常なヒト組織における 15334 受容体の発現を示す。

【図 17】

図 17 は、様々な組織および造血系細胞における 15334 受容体の発現を示す。

【図 18】

図 18 は、赤血球および巨核球を含む、様々な造血細胞／細胞系における 15334 受容体の発現を示す。

20

【図 19】

図 19 は、赤血球および巨核球を含む、様々な造血細胞または細胞系における 15334 受容体の発現を示す。

【図 1】

f1h2838 に関するタンパク質ファミリーノドメイン一致

>PF00001/7tm.17 膜貫通受容体 (ロドプシンファミリー)

スコア: 154.54 配列: 37 288 モデル: 1 269
 f1h2838 37 LNLAIHGFTFLKNRPDYAATSIYKINLAVDLLL-VLSLFPKMLVS 84
 vMgRWpFGdMGrIVwYFDYNNHYASIFLTCISIDRYLWACHPMYn
 f1h2838 85 QY-Q-SPPFS-LCTLVECLYFYSYMGVSFTICFISMDREL-AIRYPLLVS 130
 RWHTpRHRWVMIIIVwMSFIIISMPFLMFrWstyrDeneWnTWOnly
 f1h2838 131 HLRSPR-KIFGICCTIWLVTGSIPI-YSF-HGKY-EKY--H--CFH- 170
 DWPeWwVwYvILmtiimgFYIPH:IMIFCYwRIYRiaRIWHRMlpswQr
 f1h2838 171 NMSDDTWSAKYFFPLEVFGFLPMGIMGFCRS--I-HILLGRDRHTQD 217
 RRmSnRrERRIVKMIIIVwVFIIICWIPYFIvFMDTLM.HwFDefC.
 f1h2838 171 VVQQ-----KACIYSIAASLAVFVVSFLPWLGLFFLQFLVRNSFIVE-CR 261
 IwrrlWn.Y.IfeNLaYvNCpCINPIIY
 f1h2838 262 AKQSSIFFLQLSMCFSNVNC-CLDVFCY 288

FIG. 1.

【図 2】

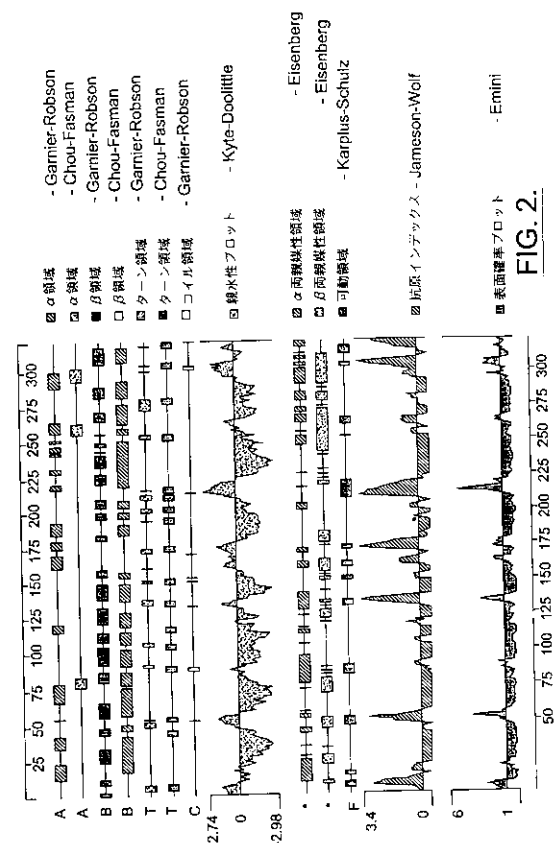


FIG. 2.

【 図 3 】

f 1 h 2 8 3 8 の解析 (320 aa)

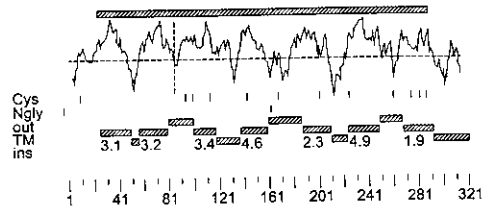


FIG. 3.

【 図 4 】

f1h2838のPositiveパターン一枚

>P000001/P0000001/ASN_GLYCOSYLATION Nグリコシル化部位

Query:	5	NTSG	8
Query:	171	NTSG	174

>PS00005/PIDC00005/PKC_PHOSP

Query: 134 SPR 136

Query: 178 SAK 180

>PS00006/PDOC00006/CK2_PHOSF

Query: 6 TSGD 9

Query: 95 TLVL 98

>P50Q0G8/P10C00008/MYRISTYL

Query: 34 GLLNL 39

Query: 107 GSVFTI 2.12
2 1.48 GICFTI 1.48

and

>P\$00009/P1000009/AMTBATL

Query: 209 LGRR 212

>PS00237/PDOC00210/G_PROTEI

Query: 107 GSVFTICFISMDRFLA1 123

>PS00339/PDOC00363/AA_TRNA_LIGASE_II_2 73/734

Query: 114 FISMDRFLAI 123

... ..

FIG. 4.

【図 5】

f1h14618のタンパク質ファミリー/ドメインHMM一致

>PF00001/7tm_1 7 膜貫通受容体 (ロドプシンファミリー)

スコア: 230.98 配列: 40 295 モデル: 1 269

[illegible]

FIG. 5.

【図 6】

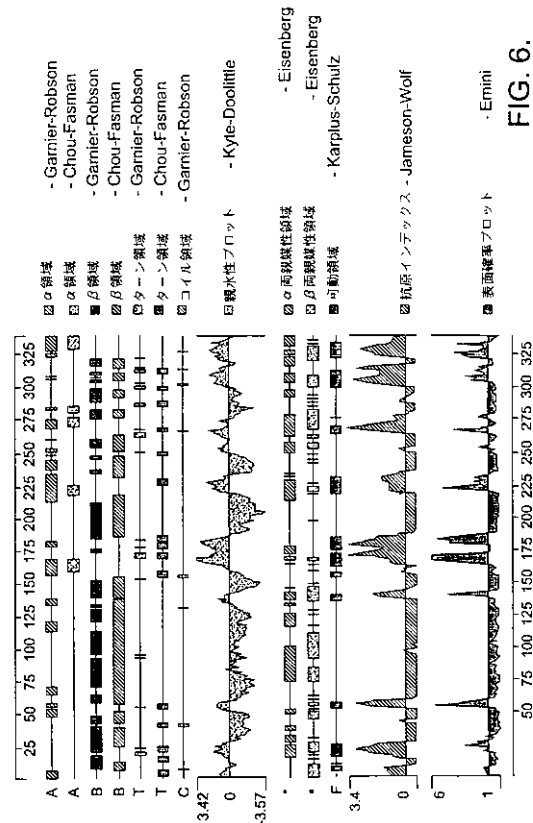
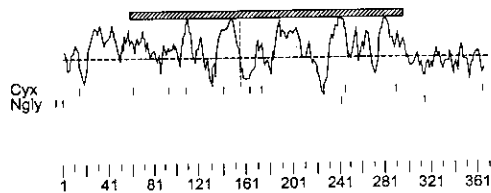


FIG. 6.

【図 1 1】

配列 28390 の解析 (372 アミノ酸)



>配列 28390
 MLANSSSTNSSVLPCEBYRPTRLHLVYSLVLAAGLPLNALALVWFRLRVHVSYSY
 MCNLAASDILLFTLSLPVRLSYIALHWPFPDLCQITGATFQHHVYSCIFMLINVERV
 AATVPLRLRLRPRVRLCLGVALLIVEAPARVHPSRCRYDLEVRIDFESFS
 DELVKGLPLVLLAEALGFLPLAAYYSRGVFTLRPDATDSQRRTYRLLANL
 VITLLGVYNSLAVYGLRSKLYAASVPARDYRGVLMYMLLAGANCLDPLVYYS
 AEGFRNLRGLGTPHRAITSATNGTRAAADRSASVTTDTRPDAAAGLLRPSDSHSL
 SSFTCCPDGAL

FIG. 10.

【図 1 2】

配列 28390 の Prosite パターン一致

Prosite パターン: 1995 年 2 月, 別表 12.2

>PS00001/PDOC0001/ASN_GLYCOSYLATION N-グリコシ化部位

Query: 4 NSSS 7
 Query: 9 NSSV 12
 Query: 251 NSTL 254
 Query: 323 NGTR 326

>PS00004/PDOC0004/CAMP_PHOSPHO_SITE cAMP および cGMP 依存性プロテインキナーゼリン酸化部位

Query: 229 RRKT 232

>PS00005/PDOC0005/PKC_PHOSPHO_SITE プロテインキナーゼ C リン酸化部位

Query: 21 THR 23
 Query: 211 SGR 213
 Query: 226 SQR 228
 Query: 232 TVR 234
 Query: 307 TLR 309
 Query: 332 SER 334

>PS00006/PDOC0006/CK2_PHOSPHO_SITE プロテインキナーゼ II リン酸化部位

Query: 178 SFSO 181
 Query: 342 TRPD 345

>PS00008/PDOC0008/MYRISTYL_Nミrist化部位

Query: 36 GLPLNA 41
 Query: 258 GLRSK 263
 Query: 324 GTRAL 329

>PS00237/PDOC00210/G_PROTEIN_RECEPTOR Gタンパク質共役受容体シグナル

Query: 107 GSCIFMLINVERIAAI 123

FIG. 12.

【図 1 3】



FIG. 13.

【図 1 4】

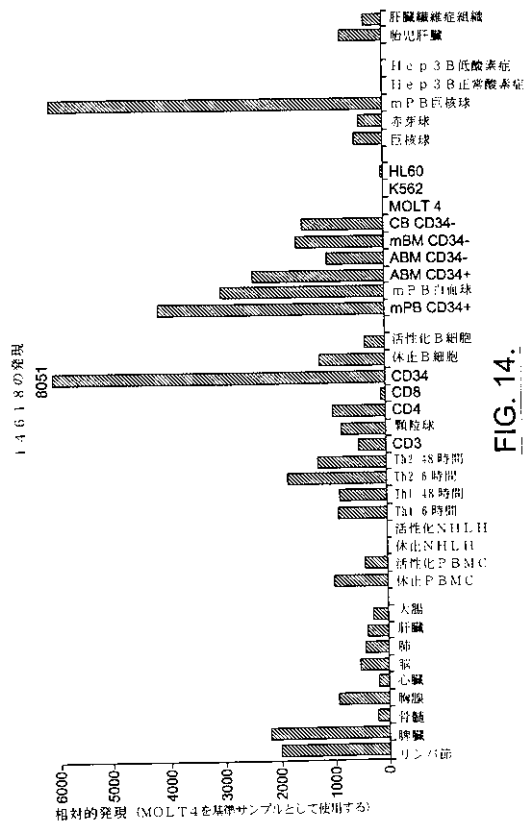


FIG. 14.

【图 16】

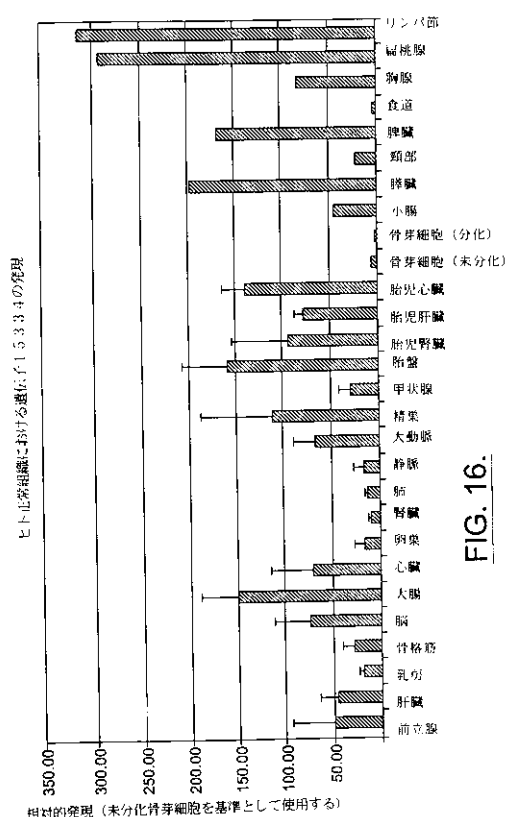
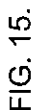


FIG. 16.

【图 18】

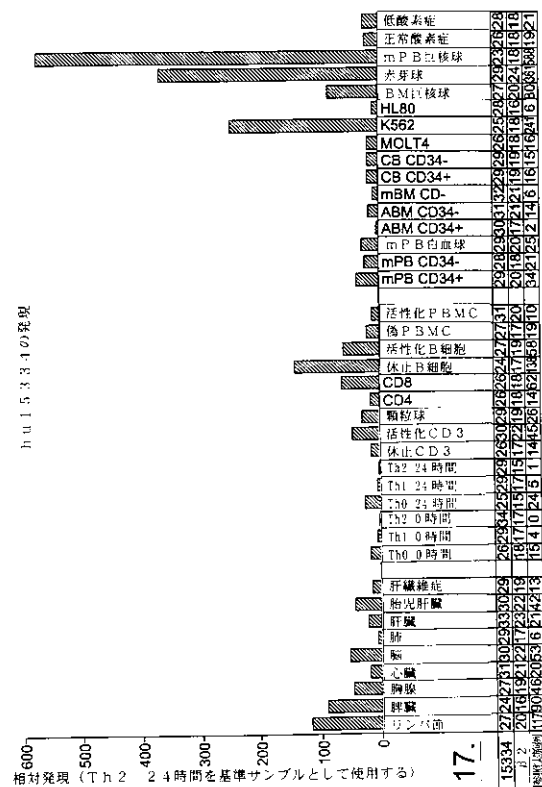
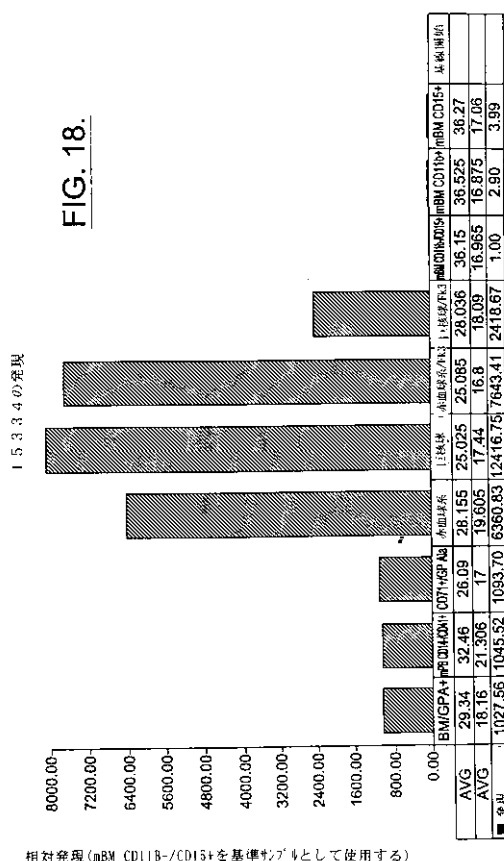


FIG. 17.



【国際公開パンフレット】

PCT		WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau	
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)			
(51) International Patent Classification 7: C12N 15/12, C07K 14/705, 16/28, G01N 33/53, C12Q 1/68, A61K 31/70, 38/17, A01K 67/027		A2	(11) International Publication Number: WO 00/23588 (43) International Publication Date: 27 April 2000 (27.04.00)
(21) International Application Number: PCT/US99/24368 (22) International Filing Date: 18 October 1999 (18.10.99) (30) Priority Data: 09/173,869 16 October 1998 (16.10.98) US Not furnished 18 October 1999 (18.10.99) US (71) Applicant (for all designated States except US): MILLEN- NIUM PHARMACEUTICALS, INC. [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US). (72) Inventors, and (73) Inventors/Applicants (for US only): GLUCKSMANN, Maria, Alexandra [AK/US]; 33 Summit Road, Lexington, MA 02173 (US); WEICEL, Nadine, S. [US/US]; 70 Park Street 853, Brookline, MA 02446 (US). (74) Agents: BROWN, Anne et al.; Alston & Bird LLP, P.O. Drawer 34009, Charlotte, NC 28234-4009 (US).		(81) Designated States: AE, AL, AM, AT, AU (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), DM, EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published Without international search report and to be republished upon receipt of that report.	
(54) Title: G-PROTEIN COUPLED RECEPTORS			
(57) Abstract The present invention relates to newly identified receptors belonging to the superfamily of G-protein-coupled receptors. The invention also relates to polynucleotides encoding the receptors. The invention further relates to methods using the receptor polypeptides and polynucleotides as a target for diagnosis and treatment in receptor-mediated disorders. The invention further relates to drug-screening methods using the receptor polypeptides and polynucleotides to identify agonists and antagonists for diagnosis and treatment. The invention further encompasses agonists and antagonists based on the receptor polypeptides and polynucleotides. The invention further relates to procedures for producing the receptor polypeptides and polynucleotides.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FJ	Fiji	LT	Lithuania	SK	Slovakia
AP	Apurimac	FR	France	LU	Luxembourg	SN	Senegal
AT	Austria	GA	Gabon	LV	Latvia	SZ	Swaziland
AU	Australia	GB	United Kingdom	MC	Monaco	TD	Chad
AZ	Azerbaijan	GE	Georgia	MD	Republic of Moldova	TC	Togo
BA	Bosnia and Herzegovina	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BB	Barbados	GN	Guinea	MK	The Former Yugoslav Republic of Macedonia	TM	Turkmenistan
BE	Belgium	GR	Greece	ML	Mali	TR	Turkey
BF	Burkina Faso	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BG	Bulgaria	IE	Ireland	MR	Mauritania	UA	Ukraine
BJ	Benin	IL	Israel	MW	Malawi	UG	Uganda
BK	Brazil	IS	Iceland	MX	Mexico	US	United States of America
BY	Belarus	IT	Italy	NE	Niger	UZ	Uzbekistan
CA	Canada	JP	Japan	NL	Netherlands	VN	Viet Nam
CF	Central African Republic	KE	Kenya	NO	Norway	YU	Yugoslavia
CG	Congo	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CH	Switzerland	KP	Democratic People's Republic of Korea	PL	Poland		
CI	Cote d'Ivoire	KR	Republic of Korea	PT	Portugal		
CM	Cameroun	KZ	Kazakhstan	RO	Romania		
CN	China	LC	Saint Lucia	RU	Russian Federation		
CU	Cuba	LI	Liechtenstein	SE	Sweden		
CZ	Czech Republic	LK	Sri Lanka	SG	Singapore		
DE	Germany	LR	Liberia				
DK	Denmark						
EE	Estonia						

WO 00/23585

PCT/US99/24368

G-PROTEIN COUPLED RECEPTORS

FIELD OF THE INVENTION

The present invention relates to newly identified receptors belonging to the superfamily of G-protein-coupled receptors. The invention also relates to polynucleotides encoding the receptors. The invention further relates to methods using the receptor polypeptides and polynucleotides as a target for diagnosis and treatment in receptor-mediated or -related disorders. The invention further relates to drug-screening methods using the receptor polypeptides and polynucleotides to identify agonists and antagonists for diagnosis and treatment. The invention further encompasses agonists and antagonists based on the receptor polypeptides and polynucleotides. The invention further relates to procedures for producing the receptor polypeptides and polynucleotides.

BACKGROUND OF THE INVENTION

15

G-protein coupled receptors

G-protein coupled receptors (GPCRs) constitute a major class of proteins responsible for transducing a signal within a cell. GPCRs have three structural domains: an amino terminal extracellular domain, a transmembrane domain containing seven transmembrane segments, three extracellular loops, and three intracellular loops, and a carboxy terminal intracellular domain. Upon binding of a ligand to an extracellular portion of a GPCR, a signal is transduced within the cell that results in a change in a biological or physiological property of the cell. GPCRs, along with G-proteins and effectors (intracellular enzymes and channels modulated by G-proteins), are the components of a modular signaling system that connects the state of intracellular second messengers to extracellular inputs.

GPCR genes and gene-products are potential causative agents of disease (Spiegel *et al.* (1993) *J. Clin. Invest.* 92:1119-1125; McKusick *et al.* (1993) *J. Med.*

WO 00/23588

PCT/US99/24368

Genet. 30:1-26). Specific defects in the rhodopsin gene and the V2 vasopressin receptor gene have been shown to cause various forms of retinitis pigmentosa (Nathans *et al.* (1992) *Annu. Rev. Genet.* 26:403-424), and nephrogenic diabetes insipidus (Holzman *et al.* (1993) *Hum. Mol. Genet.* 2:1201-1204). These receptors are of critical importance to both the central nervous system and peripheral physiological processes. Evolutionary analyses suggest that the ancestor of these proteins originally developed in concert with complex body plans and nervous systems.

The GPCR protein superfamily can be divided into five families: Family I, receptors typified by rhodopsin and the β_2 -adrenergic receptor and currently represented by over 200 unique members (Dohlman *et al.* (1991) *Annu. Rev. Biochem.* 60:653-688); Family II, the parathyroid hormone/calcitonin/secretin receptor family (Juppner *et al.* (1991) *Science* 254:1024-1026; Lin *et al.* (1991) *Science* 254:1022-1024); Family III, the metabotropic glutamate receptor family (Nakanishi (1992) *Science* 258:597-603); Family IV, the cAMP receptor family, important in the chemotaxis and development of *D. discoideum* (Klein *et al.* (1988) *Science* 241:1467-1472); and Family V, the fungal mating pheromone receptors such as STE2 (Kurjan (1992) *Annu. Rev. Biochem.* 61:1097-1129).

There are also a small number of other proteins which present seven putative hydrophobic segments and appear to be unrelated to GPCRs, they have not been shown to couple to G-proteins. *Drosophila* expresses a photoreceptor-specific protein, bride of sevenless (boss), a seven-transmembrane-segment protein which has been extensively studied and does not show evidence of being a GPCR (Hart *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:5047-5051). The gene frizzled (*fz*) in *Drosophila* is also thought to be a protein with seven transmembrane segments. Like boss, *fz* has not been shown to couple to G-proteins (Vinson *et al.* (1989) *Nature* 338:263-264).

G proteins represent a family of heterotrimeric proteins composed of α , β and γ subunits, that bind guanine nucleotides. These proteins are usually linked to cell surface receptors, e.g., receptors containing seven transmembrane segments. Following ligand binding to the GPCR, a conformational change is transmitted to the G protein, which causes the α -subunit to exchange a bound GDP molecule for a GTP molecule and to dissociate from the $\beta\gamma$ -subunits. The GTP-bound form of the α -subunit typically functions as an effector-modulating moiety, leading to the

WO 00/23588

PCT/US99/24368

production of second messengers, such as cAMP (e.g., by activation of adenylyl cyclase), diacylglycerol or inositol phosphates. Greater than 20 different types of α -subunits are known in humans. These subunits associate with a smaller pool of β and γ subunits. Examples of mammalian G proteins include Gi, Go, Gq, Gs and Gt. G proteins are described extensively in Lodish *et al. Molecular Cell Biology*, (Scientific American Books Inc., New York, N.Y., 1995), the contents of which are incorporated herein by reference. GPCRs, G proteins and G protein-linked effector and second messenger systems have been reviewed in *The G-Protein Linked Receptor Facts Book*, Watson *et al.* eds., Academic Press (1994).

10

Purinoreceptors

Purines, and especially adenosine and adenine nucleotides, have a broad range of pharmacological effects mediated through cell-surface receptors. For a general review, see Adenosine and Adenine Nucleotides in *The G-Protein Linked Receptor Facts Book*, Watson *et al.* (Eds.) Academic Press 1994, pp. 19-31.

15

Some effects of ATP include the regulation of smooth muscle activity, stimulation of the relaxation of intestinal smooth muscle and bladder contraction, stimulation of platelet activation by ADP when released from vascular endothelium, and excitatory effects in the central nervous system. Some effects of adenosine include vasodilation, bronchoconstriction, immunosuppression, inhibition of platelet aggregation, cardiac depression, stimulation of nociceptive afferents, inhibition of neurotransmitter release, pre- and postsynaptic depressant action, reducing motor activity, depressing respiration, inducing sleep, relieving anxiety, and inhibition of release of factors, such as hormones.

20

Distinct receptors exist for adenosine and adenine nucleotides. Clinical actions of such analogs as methylxanthines, for example, theophylline and caffeine, are thought to achieve their effects by antagonizing adenosine receptors. Adenosine has a low affinity for adenine nucleotide receptors, while adenine nucleotides have a low affinity for adenosine receptors.

25

30

There are four accepted subtypes of adenosine receptors, designated A₁, A_{2A}, A_{2B}, and A₃. In addition, an A₄ receptor has been proposed based on labeling by 2-phenylaminoadenosine (Cornfield *et al.* (1992) *Mol. Pharmacol.* 42:552-561).

WO 00/23588

PCT/US99/24368

P_{2X} receptors are ATP-gated cation channels (See *Neuropharmacology* 36 (1977)). The proposed topology for P_{2X} receptors is two transmembrane regions, a large extracellular loop, and intracellular N and C-termini.

Numerous cloned receptors designated P_{2Y} have been proposed to be members of the G-protein coupled family. UDP, UTP, ADP, and ATP have been identified as agonists. To date, P_{2Y(1-7)} have been characterized although it has been proposed that P_{2Y7} may be a leukotriene B₄ receptor (Yokomizo *et al.* (1997) *Nature* 387:620-624). It is widely accepted, however, that P_{2Y 1, 2, 4, and 6} are members of the G-protein coupled family of P_{2Y} receptors.

At least three P₂ purinoceptors from the hematopoietic cell line IIIb1, have been identified by intracellular calcium mobilization and by photoaffinity labeling (Akbar *et al.* (1996) *J. Biochem.* 271:18363-18567).

The A₁ adenosine receptor was designated in view of its ability to inhibit adenylyclase. The receptors are distributed in many peripheral tissues such as heart, adipose, kidney, stomach and pancreas. They are also found in peripheral nerves, for example intestine and vas deferens. They are present in high levels in the central nervous system, including cerebral cortex, hippocampus, cerebellum, thalamus, and striatum, as well as in several cell lines. Agonists and antagonists can be found on page 22 of *The G-Protein Linked Receptor Facts Book* cited above, herein incorporated by reference. These receptors are reported to inhibit adenylyclase and voltage-dependent calcium channels and to activate potassium channels through a pertussis-toxin-sensitive G-protein suggested to be of the G_i/G_o class. A₁ receptors have also been reported to induce activation of phospholipase C and to potentiate the ability of other receptors to activate this pathway.

The A_{2A} adenosine receptor has been found in brain, such as striatum, olfactory tubercle and nucleus accumbens. In the periphery, A₂ receptors mediate vasodilation, immunosuppression, inhibition of platelet aggregation, and gluconeogenesis. Agonists and antagonists are found in *The G-Protein Linked Receptor Facts Book* cited above on page 25, herein incorporated by reference. This receptor mediates activation of adenylyclase through G_s.

The A_{2B} receptor has been shown to be present in human brain and in rat intestine and urinary bladder. Agonists and antagonists are discussed on page 27 of *The G-Protein Linked Receptor Facts Book* cited above, herein incorporated by reference. This receptor mediates the stimulation of cAMP through G_s.

WO 00/23588

PCT/US99/24368

The A₃ adenosine receptor is expressed in testes, lung, kidney, heart, central nervous system, including cerebral cortex, striatum, and olfactory bulb. A discussion of agonists and antagonists can be found on page 28 of *The G-Protein Linked Receptor Facts Book* cited above, herein incorporated by reference. The receptor

5 mediates the inhibition of adenylyl cyclase through a pertussis-toxin-sensitive G-protein, suggested to be of the G_i/G_o class.

The P_{2y} purinoceptor shows a similar affinity for ATP and ADP with a lower affinity for AMP. The receptor has been found in smooth muscle, for example, taeni caeci and in vascular tissue where it induces vasodilation through endothelium-

10 dependent release of nitric oxide. It has also been shown in avian erythrocytes. Agonists and antagonists are discussed on page 30 of *The G-Protein Linked Receptor Facts Book* cited above, herein incorporated by reference. The receptor function through activation of phosphoinositide metabolism through a pertussis-toxin-insensitive G-protein, suggested to be of the G_q/G₁₂ class.

15 Accordingly, GPCRs, and especially purinoceptors, are a major target for drug action and development. Accordingly, it is valuable to the field of pharmaceutical development to identify and characterize previously unknown GPCRs. The present invention advances the state of the art by providing a previously unidentified human GPCR having homology to purinoceptors.

20

SUMMARY OF THE INVENTION

It is an object of the invention to identify novel GPCRs.

25 It is a further object of the invention to provide novel GPCR polypeptides that are useful as reagents or targets in receptor assays applicable to treatment and diagnosis of GPCR-mediated disorders.

It is a further object of the invention to provide polynucleotides corresponding to the novel GPCR receptor polypeptides that are useful as targets and reagents in receptor assays applicable to treatment and diagnosis of GPCR-mediated disorders

30 and useful for producing novel receptor polypeptides by recombinant methods.

A specific object of the invention is to identify compounds that act as agonists and antagonists and modulate the expression of the novel receptors.

WO 00/23588

PCT/US99/24368

A further specific object of the invention is to provide compounds that modulate expression of the receptors for treatment and diagnosis of GPCR-related disorders.

The invention is thus based on the identification of novel GPCRs, designated the 2838, 14618, and 15334 receptors.

The invention provides isolated 2838 receptor polypeptides including a polypeptide having the amino acid sequence shown in SEQ ID NO:1.

The invention provides isolated 14618 receptor polypeptides including a polypeptide having the amino acid sequence shown in SEQ ID NO:3.

The invention provides isolated 15334 receptor polypeptides including a polypeptide having the amino acid sequence shown in SEQ ID NO:5.

The invention also provides isolated 2838 receptor nucleic acid molecules having the sequence shown in SEQ ID NO:2.

The invention also provides isolated 14618 receptor nucleic acid molecules having the sequence shown in SEQ ID NO:4.

The invention also provides isolated 15334 receptor nucleic acid molecules having the sequence shown in SEQ ID NO:6.

The invention also provides variant polypeptides having an amino acid sequence that is substantially homologous to the amino acid sequence shown in SEQ ID NO:1.

The invention also provides variant polypeptides having an amino acid sequence that is substantially homologous to the amino acid sequence shown in SEQ ID NO:3.

The invention also provides variant polypeptides having an amino acid sequence that is substantially homologous to the amino acid sequence shown in SEQ ID NO:5.

The invention also provides variant nucleic acid sequences that are substantially homologous to the nucleotide sequence shown in SEQ ID NO:2.

The invention also provides variant nucleic acid sequences that are substantially homologous to the nucleotide sequence shown in SEQ ID NO:4.

The invention also provides variant nucleic acid sequences that are substantially homologous to the nucleotide sequence shown in SEQ ID NO:6.

WO 00/23598

PCT/US99/24368

The invention also provides fragments of the polypeptide shown in SEQ ID NO:1 and nucleotide sequence shown in SEQ ID NO:2, as well as substantially homologous fragments of the polypeptide or nucleic acid.

- 5 The invention also provides fragments of the polypeptide shown in SEQ ID NO:3 and nucleotide sequence shown in SEQ ID NO:4, as well as substantially homologous fragments of the polypeptide or nucleic acid.

The invention also provides fragments of the polypeptide shown in SEQ ID NO:5 and nucleotide sequence shown in SEQ ID NO:6, as well as substantially homologous fragments of the polypeptide or nucleic acid.

- 10 The invention further provides nucleic acid constructs comprising the nucleic acid molecules described above. In a preferred embodiment, the nucleic acid molecules of the invention are operatively linked to a regulatory sequence.

- 15 The invention also provides vectors and host cells for expressing the receptor nucleic acid molecules and polypeptides and particularly recombinant vectors and host cells.

The invention also provides methods of making the vectors and host cells and methods for using them to produce the receptor nucleic acid molecules and polypeptides.

- 20 The invention also provides antibodies or antigen-binding fragments thereof that selectively bind the receptor polypeptides and fragments.

The invention also provides methods of screening for compounds that modulate expression or activity of the receptor polypeptides or nucleic acid (RNA or DNA).

- 25 The invention also provides a process for modulating receptor polypeptide or nucleic acid expression or activity, especially using the screened compounds. Modulation may be used to treat conditions related to aberrant activity or expression of the receptor polypeptides or nucleic acids.

The invention also provides assays for determining the presence or absence of and level of the receptor polypeptides or nucleic acid molecules in a biological sample, including for disease diagnosis.

- 30 The invention also provides assays for determining the presence of a mutation in the receptor polypeptides or nucleic acid molecules, including for disease diagnosis.

In still a further embodiment, the invention provides a computer readable means containing the nucleotide and/or amino acid sequences of the nucleic acids and polypeptides of the invention, respectively

WO 00/23588

PCT/US99/24368

DESCRIPTION OF THE DRAWINGS

Figure 1 shows a comparison of the 2838 receptor against the Prosite data base of protein patterns, specifically showing a high score against the seven transmembrane segment rhodopsin superfamily. The underlined area shows a GPCR signature, and specifically the position of an arginine residue, conserved in GPCRs. The most commonly conserved sequence is an aspartate, arginine, tyrosine (DRY) triplet. DRY is implicated in signal transduction. Arginine is invariant. Aspartate is conservatively placed in several GPCRs. In the present case, the arginine is found in the sequence DRF, which matches the position of DRY or invariant arginine in GPCRs of the rhodopsin superfamily of receptors.

Figure 2 shows an analysis of the 2838 amino acid sequence: α -turn and coil regions; hydrophilicity; amphipathic regions; flexible regions; antigenic index; and surface probability plot.

Figure 3 shows a 2838 receptor hydrophobicity plot. The amino acids correspond to 1-319 and show the seven transmembrane segments.

Figure 4 shows an analysis of the 2838 open reading frame for amino acids corresponding to specific functional sites. A glycosylation site is found at amino acids 5-8 which corresponds to the amino terminal extracellular domain. A second glycosylation site is found at amino acids 171-174 which corresponds to the second extracellular loop. A protein kinase C phosphorylation site is found at amino acids 134-136 which is in the second intracellular loop. A second protein kinase C phosphorylation site is found at amino acids 178-180 which is in the second extracellular loop. A casein kinase II phosphorylation site is found at amino acids 6-9 which is in the carboxy terminal intracellular domain. A second casein kinase II phosphorylation site is found at amino acids 95-98 which is in the first extracellular loop. An N-myristoylation site is found at amino acids 34-39 which is in the first transmembrane segment. A second N-myristoylation site is found at amino acids 107-112 which is in the third transmembrane segment. A third N-myristoylation site is found at amino acids 140-145 which is in the fourth transmembrane segment. An

WO 00/23588

PCT/US99/24368

amidation site is found at amino acids 209-212 which spans the fifth transmembrane segment and third intracellular loop. In addition, amino acids corresponding in position to the GPCR signature and containing the invariant arginine are found in the sequence DRF at amino acids 118-120.

5

Figure 5 shows a comparison of the 14618 receptor against the Prosite data base of protein patterns, specifically showing a high score against the seven transmembrane segment rhodopsin superfamily. The underlined area shows a GPCR signature, and specifically the position of an arginine residue, conserved in GPCRs.

10 The most commonly conserved sequence is an aspartate, arginine, tyrosine (DRY) triplet. DRY is implicated in signal transduction. Arginine is invariant. Aspartate is conservatively placed in several GPCRs. In the present case, the arginine is found in the sequence DRF, which matches the position of DRY or invariant arginine in GPCRs of the rhodopsin superfamily of receptors.

15

Figure 6 shows an analysis of the 14618 amino acid sequence: α turn and coil regions; hydrophilicity; amphipathic regions; flexible regions; antigenic index; and surface probability plot.

20

Figure 7 shows a 14618 receptor hydrophobicity plot. The amino acids correspond to 1-337 and show the seven transmembrane segments.

Figure 8 shows an analysis of the 14618 open reading frame for amino acids corresponding to specific functional sites. A glycosylation site is found at amino acids 6-9 which corresponds to the amino terminal extracellular domain. A second glycosylation site is found at amino acids 169-172 which corresponds to the second extracellular loop. A third glycosylation site is found at amino acids 180-183 which also corresponds to the second extracellular loop. A fourth glycosylation site found at amino acids 224-227 which corresponds to the third intracellular loop. A fifth glycosylation site is found at amino acids 262-265 which corresponds to the third extracellular loop. A cAMP- and cGMP-dependent protein kinase phosphorylation site is found at amino acids 304-307, 310-313, and 323-326, all in the carboxy terminal intracellular domain. A protein kinase C phosphorylation site is found at amino acids 136-138 which corresponds to the second intracellular loop. Second and

25

30

WO 00/23588

PCT/US99/24368

third protein kinase C phosphorylation sites are found at amino acids 220-222 and 227-229, both corresponding to the third intracellular loop. A fourth protein kinase C phosphorylation site is found at amino acids 308-310, corresponding to the carboxy terminal intracellular domain. Casein kinase II phosphorylation sites are found at amino acids 13-16 and 17-20, both in the amino terminal extracellular domain. A third casein kinase II phosphorylation site is found at amino acids 326-329 corresponding to the carboxy terminal intracellular domain. A microbodies C-terminal targeting signal is found at amino acids 335-338, corresponding to the carboxy terminal intracellular domain. In addition, amino acids corresponding in position to the GPCR signature and containing the invariant arginine are found in the sequence FRC at amino acids 120-122.

Figure 9 shows a comparison of the 15334 receptor against the Prosite data base of protein patterns, specifically showing a high score against the seven transmembrane segment rhodopsin superfamily. The underlined area shows a GPCR signature, and specifically the position of an arginine residue, conserved in GPCRs. The most commonly conserved sequence is an aspartate, arginine, tyrosine (DRY) triplet. DRY is implicated in signal transduction. Arginine is invariant. Aspartate is conservatively placed in several GPCRs. In the present case, the arginine is found in the sequence DRY, which matches the position of DRY or invariant arginine in GPCRs of the rhodopsin superfamily of receptors.

Figure 10 shows an analysis of the 15334 amino acid sequence: alpha turn and coil regions; hydrophilicity; amphipathic regions; flexible regions; antigenic index; and surface probability plot.

Figure 11 shows a 15334 receptor hydrophobicity plot. The amino acids show the seven transmembrane segments.

Figure 12 shows an analysis of the 15334 open reading frame for amino acids corresponding to specific functional sites. Glycosylation sites are found at amino acids 4-7 and 9-12, which are in the amino terminal extracellular domain. A glycosylation site is also found at amino acids 251-254, which is in the sixth transmembrane segment. A further glycosylation site is found from amino acid 323-

WO 00/23588

PCT/US99/24368

326, which is in the carboxy terminal domain. A cAMP- and cGMP-dependent protein kinase phosphorylation site is found at amino acids 229-232, which is in the third intracellular loop. Protein kinase C phosphorylation sites are found from amino acids 21-23, which corresponds to the amino terminal domain, 211-213, 226-228, and 232-234, which corresponds to the third intracellular loop, and 307-309 and 332-334, which corresponds to the carboxy terminal intracellular domain. A casein kinase II phosphorylation site is found from amino acids 178-181, which is in the second extracellular loop and from 342-345, which is in the carboxy terminal intracellular domain. N-myristoylation sites are found at amino acids 36-41, which is in the amino terminal extracellular domain, 258-263, which spans the sixth transmembrane segment and third extracellular loop, and 324-329, which corresponds to the carboxy terminal intracellular domain. In addition, a GPCR signature, DRY, is found at amino acids 118-120, which is in the second intracellular loop.

15 **Figure 13** shows expression of the 14618 receptor in various normal human tissues.

Figure 14 shows the 14618 receptor expression in various tissues and in cells of the hematopoietic system

20 **Figure 15** shows expression of the 2838 receptor in various cells and tissues, including cells of the hematopoietic system.

Figure 16 shows expression of the 15334 receptor in normal human tissues.

25 **Figure 17** shows expression of the 15334 receptor in various tissues and in cells of the hematopoietic system.

Figure 18 shows expression of the 15334 receptor in various hematopoietic cells/cell lines, including erythroid and megakaryocyte cells.

30 **Figure 19** shows expression of the 15334 receptor in various hematopoietic cells or cell lines, including erythroid and megakaryocyte cells.

WO 00/23588

PCT/US99/24368

DETAILED DESCRIPTION OF THE INVENTION

Receptor function/signal pathway

The 2838, 14618, and 15334 receptor proteins are GPCRs that participate in signaling pathways. As used herein, a "signaling pathway" refers to the modulation (e.g., stimulation or inhibition) of a cellular function/activity upon the binding of a ligand to the GPCR (2838, 14618, or 15334 protein). Examples of such functions include mobilization of intracellular molecules that participate in a signal transduction pathway, e.g., phosphatidylinositol 4,5-bisphosphate (PIP₂), inositol 1,4,5-triphosphate (IP₃) and adenylate cyclase; polarization of the plasma membrane; production or secretion of molecules; alteration in the structure of a cellular component; cell proliferation, e.g., synthesis of DNA; cell migration; cell differentiation; and cell survival. Since the 2838 receptor protein is expressed in thymus, lymph node, spleen, testes, colon, and peripheral blood lymphocytes including but not limited to activated T helper cells-1, activated T-helper cells-2, CD3(both CD4 and CD8), activated B cells, and granulocytes, cells participating in a 2838 receptor protein signaling pathway include, but are not limited to, these tissues and cells. Since the 14618 receptor protein is expressed in breast, skeletal muscle, thyroid, lymph node, spleen, and peripheral blood lymphocytes including, but not limited to, CD34⁺ cells, resting B cells, and megakaryocytes, cells participating in a 14618 receptor protein signaling pathway include, but are not limited to, cells derived from these tissues and cells. Since the 15334 receptor protein is expressed in colon, pancreas, tonsil, lymph node, spleen, thymus, adrenal gland, heart, and peripheral blood cells, including but not limited to those shown in Figure 17, megakaryocytes, and erythroblasts, cells participating in a 15334 receptor protein signaling pathway include, but are not limited to, cells derived from these tissues and cells.

The response mediated by the receptor protein depends on the type of cell. For example, in some cells, binding of a ligand to the receptor protein may stimulate an activity such as release of compounds, gating of a channel, cellular adhesion, migration, differentiation, etc., through phosphatidylinositol or cyclic AMP metabolism and turnover while in other cells, the binding of the ligand will produce a different result. Regardless of the cellular activity/response modulated by the receptor protein, it is universal that the protein is a GPCR and interacts with G proteins to produce one or more secondary signals, in a variety of intracellular signal

WO 00/23588

PCT/US99/24368

transduction pathways, e.g., through phosphatidylinositol or cyclic AMP metabolism and turnover, in a cell.

As used herein, "phosphatidylinositol turnover and metabolism" refers to the molecules involved in the turnover and metabolism of phosphatidylinositol 4,5-bisphosphate (PIP₂) as well as to the activities of these molecules. PIP₂ is a phospholipid found in the cytosolic leaflet of the plasma membrane. Binding of ligand to the receptor activates, in some cells, the plasma-membrane enzyme phospholipase C that in turn can hydrolyze PIP₂ to produce 1,2-diacylglycerol (DAG) and inositol 1,4,5-triphosphate (IP₃). Once formed IP₃ can diffuse to the endoplasmic reticulum surface where it can bind an IP₃ receptor, e.g., a calcium channel protein containing an IP₃ binding site. IP₃ binding can induce opening of the channel, allowing calcium ions to be released into the cytoplasm. IP₃ can also be phosphorylated by a specific kinase to form inositol 1,3,4,5-tetraphosphate (IP₄), a molecule which can cause calcium entry into the cytoplasm from the extracellular medium. IP₃ and IP₄ can subsequently be hydrolyzed very rapidly to the inactive products inositol 1,4-bisphosphate (IP₂) and inositol 1,3,4-triphosphate, respectively. These inactive products can be recycled by the cell to synthesize PIP₂. The other second messenger produced by the hydrolysis of PIP₂, namely 1,2-diacylglycerol (DAG), remains in the cell membrane where it can serve to activate the enzyme protein kinase C. Protein kinase C is usually found soluble in the cytoplasm of the cell, but upon an increase in the intracellular calcium concentration, this enzyme can move to the plasma membrane where it can be activated by DAG. The activation of protein kinase C in different cells results in various cellular responses such as the phosphorylation of glycogen synthase, or the phosphorylation of various transcription factors, e.g., NF- κ B. The language "phosphatidylinositol activity", as used herein, refers to an activity of PIP₂ or one of its metabolites.

Another signaling pathway in which the receptor may participate is the cAMP turnover pathway. As used herein, "cyclic AMP turnover and metabolism" refers to the molecules involved in the turnover and metabolism of cyclic AMP (cAMP) as well as to the activities of these molecules. Cyclic AMP is a second messenger produced in response to ligand-induced stimulation of certain G protein coupled receptors. In the cAMP signaling pathway, binding of a ligand to a GPCR can lead to the activation of the enzyme adenylyl cyclase, which catalyzes the synthesis of cAMP. The newly synthesized cAMP can in turn activate a cAMP-dependent protein kinase.

WO 00/23588

PCT/US99/24368

This activated kinase can phosphorylate a voltage-gated potassium channel protein, or an associated protein, and lead to the inability of the potassium channel to open during an action potential. The inability of the potassium channel to open results in a decrease in the outward flow of potassium, which normally repolarizes the membrane of a neuron, leading to prolonged membrane depolarization.

Polypeptides

The invention is based on the discovery of novel G-coupled protein receptors. Specifically, an expressed sequence tag (EST) was selected based on homology to G-protein-coupled receptor sequences. This EST was used to design primers based on sequences that it contains and used to identify a 2838 cDNA from a B cell cDNA library, a 14618 cDNA from a liver and spleen cDNA library, and a 15334 cDNA from a spleen cDNA library. Positive clones were sequenced and the overlapping fragments were assembled. Analysis of the assembled sequences revealed that the three cloned cDNA molecules encode G-protein coupled receptors.

The invention thus relates to a novel GPCR having the deduced amino acid sequence shown in SEQ ID NO:1.

The invention also thus relates to a novel GPCR having the deduced amino acid sequence shown in SEQ ID NO:3.

The invention also thus relates to a novel GPCR having the deduced amino acid sequence shown in SEQ ID NO:5.

The Sequence Listing shows the 2838 nucleotide sequence (SEQ ID NO:2) and the deduced 2838 amino acid sequence (SEQ ID NO:1). It is predicted that amino acids 1 to about 24 constitute the amino terminal extracellular domain, amino acids about 25-292 constitute the region spanning the transmembrane domain, and amino acids about 293-319 constitute the carboxy terminal intracellular domain. The transmembrane domain contains seven transmembrane segments, three extracellular loops and three intracellular loops. The transmembrane segments are found from about amino acid 25 to about amino acid 49, from about amino acid 56 to about amino acid 79, from about amino acid 100 to about amino acid 117, from about amino acid 138 to about amino acid 159, from about amino acid 187 to about amino acid 210, from about amino acid 224 to about amino acid 248, and from about amino acid 268 to about amino acid 292. Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular

WO 00/23588

PCT/US99/24368

loops are found from about amino acid 50 to about amino acid 55, from about amino acid 118 to about amino acid 137, and from about amino acid 211 to about amino acid 223. The three extracellular loops are found at from about amino acid 80 to about amino acid 99, from about amino acid 160 to about amino acid 186, and from about amino acid 249 to about amino acid 267.

The Sequence Listing shows the 14618 nucleotide sequence (SEQ ID NO:4) and the deduced 14618 amino acid sequence (SEQ ID NO:3). It is predicted that amino acids 1 to about 28 constitute the amino terminal extracellular domain, amino acids about 29-297 constitute the region spanning the transmembrane domain, and amino acids about 298-337 constitute the carboxy terminal intracellular domain. The transmembrane domain contains seven transmembrane segments, three extracellular loops and three intracellular loops. The transmembrane segments are found from about amino acid 29 to about amino acid 49, from about amino acid 60 to about amino acid 84, from about amino acid 103 to about amino acid 127, from about amino acid 142 to about amino acid 161, from about amino acid 194 to about amino acid 217, from about amino acid 231 to about amino acid 247, and from about amino acid 276 to about amino acid 297. Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular loops are found from about amino acid 50 to about amino acid 59, from about amino acid 128 to about amino acid 141, and from about amino acid 218 to about amino acid 230. The three extracellular loops are found at from about amino acid 85 to about amino acid 102, from about amino acid 162 to about amino acid 193, and from about amino acid 248 to about amino acid 275.

The Sequence Listing shows the 15334 nucleotide sequence (SEQ ID NO:6) and the deduced 15334 amino acid sequence (SEQ ID NO:5). It is predicted that amino acids 1 to about 25 constitute the amino terminal extracellular domain, amino acids about 26-299 constitute the region spanning the transmembrane domain, and amino acids about 300-372 constitute the carboxy terminal intracellular domain. The transmembrane domain contains seven transmembrane segments, three extracellular loops and three intracellular loops. The transmembrane segments are found from about amino acid 26 to about amino acid 48, from about amino acid 56 to about amino acid 77, from about amino acid 99 to about amino acid 115, from about amino acid 140 to about amino acid 157, from about amino acid 188 to about amino acid 209, from about amino acid 235 to about amino acid 259, and from about amino acid

WO 00/23588

PCT/US99/24368

277 to about amino acid 299. Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular loops are found from about amino acid 49 to about amino acid 55, from about amino acid 116 to about amino acid 139, and from about amino acid 210 to about amino acid 234. The three extracellular loops are found at from about amino acid 78 to about amino acid 98, from about amino acid 158 to about amino acid 187, and from about amino acid 260 to about amino acid 276.

The "2838 receptor polypeptide" or "2838 receptor protein" refers to the polypeptide in SEQ ID NO:1. The "14618 receptor polypeptide" or "14618 receptor protein" refers to the polypeptide in SEQ ID NO:3. The "15334 receptor polypeptide" or "15334 receptor protein" refers to the polypeptide in SEQ ID NO:5. The term "receptor protein" or "receptor polypeptide", however, further includes the numerous variants of 2838, 14618, or 15334 polypeptides described herein, as well as fragments derived from the full length 2838, 14618, or 15334 polypeptides and variants.

The present invention thus provides isolated or purified 2838, 14618, and 15334 receptor polypeptides and variants and fragments thereof.

The 2838 polypeptide is a 319 residue protein exhibiting three main structural domains. The amino terminal extracellular domain is identified to be within residues 1 to about 24 in SEQ ID NO:1. The transmembrane domain is identified to be within residues from about 25 to about 292 in SEQ ID NO:1. The carboxy terminal intracellular domain is identified to be within residues from about 293 to 319 in SEQ ID NO:1. The transmembrane domain contains seven segments that span the membrane. The transmembrane segments are found from about amino acid 25 to about amino acid 49, from about amino acid 56 to about amino acid 79, from about amino acid 100 to about amino acid 117, from about amino acid 138 to about amino acid 159, from about amino acid 187 to about amino acid 210, from about amino acid 224 to about amino acid 248, and from about amino acid 268 to about amino acid 292. Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular loops are found from about amino acid 50 to about amino acid 55, from about amino acid 118 to about amino acid 137, and from about amino acid 211 to about amino acid 223. The three extracellular loops are found at from about amino acid 80 to about amino acid 99, from about amino acid 160 to about amino acid 186, and from about amino acid 249 to about amino acid 267.

WO 00/23588

PCT/US99/24368

The transmembrane domain includes a GPCR signal transduction signature, DRP, at residues 118-120. The sequence includes an arginine at residue 119, an invariant amino acid in GPCRs.

Based on a BLAST search, highest homology was shown to purinoceptors.

5 The 14618 polypeptide is a 337 residue protein exhibiting three main structural domains. The amino terminal extracellular domain is identified to be within residues 1 to about 28 in SEQ ID NO:3. The transmembrane domain is identified to be within residues from about 29 to about 297 in SEQ ID NO:3. The carboxy terminal intracellular domain is identified to be within residues from about 298 to 337 in SEQ ID NO:3. The transmembrane domain contains seven segments that span the membrane. The transmembrane segments are found from about amino acid 29 to about amino acid 49, from about amino acid 84 to about amino acid 60, from about amino acid 103 to about amino acid 127, from about amino acid 142 to about amino acid 161, from about amino acid 194 to about amino acid 217, from about amino acid 231 to about amino acid 247, and from about amino acid 276 to about amino acid 297. Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular loops are found from about amino acid 50 to about amino acid 59, from about amino acid 128 to about amino acid 141, and from about amino acid 218 to about amino acid 230. The three extracellular loops are found at from about amino acid 85 to about amino acid 102, from about amino acid 162 to about amino acid 193, and from about amino acid 248 to about amino acid 275.

20 The transmembrane domain includes a GPCR signal transduction signature, FRC, at residues 121-123. The sequence includes an arginine at residue 122, an invariant amino acid in GPCRs.

25 Based on a BLAST search, highest homology was shown to purinoceptors.

The 15334 polypeptide is a 372 residue protein exhibiting three main structural domains. The amino terminal extracellular domain is identified to be within residues 1 to about 25 in SEQ ID NO:5. The transmembrane domain is identified to be within residues from about 26 to about 299 in SEQ ID NO:5. The carboxy terminal intracellular domain is identified to be within residues from about 300 to 372 in SEQ ID NO:5. The transmembrane domain contains seven segments that span the membrane. The transmembrane segments are found from about amino acid 26 to about amino acid 48, from about amino acid 56 to about amino acid 77, from about

WO 00/23588

PCT/US99/24368

amino acid 99 to about amino acid 115, from about amino acid 140 to about amino acid 157, from about amino acid 188 to about amino acid 209, from about amino acid 235 to about amino acid 259, and from about amino acid 277 to about amino acid 299.

Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular loops are found from about amino acid 49 to about amino acid 55, from about amino acid 116 to about amino acid 139, and from about amino acid 210 to about amino acid 234. The three extracellular loops are found at from about amino acid 78 to about amino acid 98, from about amino acid 158 to about amino acid 187, and from about amino acid 260 to about amino acid 276.

The transmembrane domain includes a GPCR signal transduction signature, DRY, at residues 118-120. The sequence includes an arginine at residue 119, an invariant amino acid in GPCRs.

Based on a BLAST search, highest homology was shown to purinoceptors.

As used herein, a polypeptide is said to be "isolated" or "purified" when it is substantially free of cellular material when it is isolated from recombinant and non-recombinant cells, or free of chemical precursors or other chemicals when it is chemically synthesized. A polypeptide, however, can be joined to another polypeptide with which it is not normally associated in a cell and still be considered "isolated" or "purified."

The receptor polypeptides can be purified to homogeneity. It is understood, however, that preparations in which the polypeptide is not purified to homogeneity are useful and considered to contain an isolated form of the polypeptide. The critical feature is that the preparation allows for the desired function of the polypeptide, even in the presence of considerable amounts of other components. Thus, the invention encompasses various degrees of purity.

In one embodiment, the language "substantially free of cellular material" includes preparations of the receptor polypeptide having less than about 30% (by dry weight) other proteins (*i.e.*, contaminating protein), less than about 20% other proteins, less than about 10% other proteins, or less than about 5% other proteins. When the receptor polypeptide is recombinantly produced, it can also be substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, less than about 10%, or less than about 5% of the volume of the protein preparation.

WO 00/23588

PCT/US99/24368

A receptor polypeptide is also considered to be isolated when it is part of a membrane preparation or is purified and then reconstituted with membrane vesicles or liposomes.

The language "substantially free of chemical precursors or other chemicals" includes preparations of the receptor polypeptide in which it is separated from chemical precursors or other chemicals that are involved in its synthesis. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of the polypeptide having less than about 30% (by dry weight) chemical precursors or other chemicals, less than about 20% chemical precursors or other chemicals, less than about 10% chemical precursors or other chemicals, or less than about 5% chemical precursors or other chemicals.

In one embodiment, the receptor polypeptide comprises the amino acid sequence shown in SEQ ID NO:1. However, the invention also encompasses sequence variants. Variants include a substantially homologous protein encoded by the same genetic locus in an organism, *i.e.*, an allelic variant. The receptor maps to chromosome 2, in close proximity to W1-7921. Variants also encompass proteins derived from other genetic loci in an organism, but having substantial homology to the 2838 receptor protein of SEQ ID NO:1. Variants also include proteins substantially homologous to the 2838 receptor protein but derived from another organism, *i.e.*, an ortholog. Variants also include proteins that are substantially homologous to the 2838 receptor protein that are produced by chemical synthesis. Variants also include proteins that are substantially homologous to the 2838 receptor protein that are produced by recombinant methods. It is understood, however, that variants exclude any amino acid sequences disclosed prior to the invention.

In another embodiment, the receptor polypeptide comprises the amino acid sequence shown in SEQ ID NO:3. However, the invention also encompasses sequence variants. Variants include allelic variants. Variants also encompass proteins derived from other genetic loci in an organism, but having substantial homology to the 14618 receptor protein of SEQ ID NO:3. Variants also include substantially homologous orthologs. Variants also include proteins that are substantially homologous to the 14618 receptor protein that are produced by chemical synthesis. Variants also include proteins that are substantially homologous to the 14618 receptor protein that are produced by recombinant methods. It is understood,

WO 00/23588

PCT/US99/24368

however, that variants exclude any amino acid sequences disclosed prior to the invention.

In another embodiment, the receptor polypeptide comprises the amino acid sequence shown in SEQ ID NO:5. However, the invention also encompasses sequence variants. Variants include allelic variants. The receptor maps to chromosome 12 in close proximity to SHGC-30262. Variants also encompass proteins derived from other genetic loci in an organism, but having substantial homology to the 15334 receptor protein of SEQ ID NO:5. Variants also include substantially homologous orthologs. Variants also include proteins that are substantially homologous to the 15334 receptor protein that are produced by chemical synthesis. Variants also include proteins that are substantially homologous to the 15334 receptor protein that are produced by recombinant methods. It is understood, however, that variants exclude any amino acid sequences disclosed prior to the invention.

As used herein, two proteins (or a region of the proteins) are substantially homologous to the 2838 protein when the amino acid sequences are at least about 40-45%, 45-50%, 50-55%, 55-60%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous. A substantially homologous amino acid sequence, according to the present invention, will be encoded by a nucleic acid sequence hybridizing to the nucleic acid sequence, or portion thereof, of the sequence shown in SEQ ID NO:2 under stringent conditions as more fully described below.

As used herein, two proteins (or a region of the proteins) are substantially homologous to the 14618 protein when the amino acid sequences are at least about 35-40%, 40-45%, 45-50%, 55-60%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous. A substantially homologous amino acid sequence, according to the present invention, will be encoded by a nucleic acid sequence hybridizing to the nucleic acid sequence, or portion thereof, of the sequence shown in SEQ ID NO:4 under stringent conditions as more fully described below.

As used herein, two proteins (or a region of the proteins) are substantially homologous to the 15334 protein when the amino acid sequences are at least about 40-45%, 45-50%, 50-55%, 55-60%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous. A

WO 00/23588

PCT/US99/24368

substantially homologous amino acid sequence, according to the present invention, will be encoded by a nucleic acid sequence hybridizing to the nucleic acid sequence, or portion thereof, of the sequence shown in SEQ ID NO 6 under stringent conditions as more fully described below.

5 To determine the percent homology of two amino acid sequences, or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of one protein or nucleic acid for optimal alignment with the other protein or nucleic acid). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When
10 a position in one sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the other sequence, then the molecules are homologous at that position. As used herein, amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity". The percent homology between the two sequences is a function of the number of identical positions shared
15 by the sequences (*i.e.*, per cent homology equals the number of identical positions/total number of positions times 100).

The invention also encompasses polypeptides having a lower degree of identity but having sufficient similarity so as to perform one or more of the same functions performed by the 2838, 14618, or 15334 polypeptides. Similarity is
20 determined by conserved amino acid substitution. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Conservative substitutions are likely to be phenotypically silent. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu, and Ile; interchange of the hydroxyl
25 residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr. Guidance concerning which amino acid changes are likely to be phenotypically silent are found in Bowie *et al.* (1990) *Science* 247: 1306-1310.

30

WO 00/23588

PCT/US99/24368

TABLE 1. Conservative Amino Acid Substitutions.

Aromatic	Phenylalanine
	Tryptophan
	Tyrosine
Hydrophobic	Leucine
	Isoleucine
	Valine
Polar	Glutamine
	Asparagine
Basic	Arginine
	Lysine
	Histidine
Acidic	Aspartic Acid
	Glutamic Acid
Small	Alanine
	Serine
	Threonine
	Methionine
	Glycine

- Both identity and similarity can be readily calculated (*Computational*
- 5 *Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part 1*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987, and *Sequence Analysis Primer*,
- 10 Gribskov, M. and Devereux, J., eds., M. Stockton Press, New York, 1991).

WO 00/23588

PCT/US99/24368

A preferred, non-limiting example of such a mathematical algorithm is described in Karlin *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402. When utilizing

5 BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>. In one embodiment, parameters for sequence comparison can be set at score= 100, wordlength=12, or can be varied (e.g., W=5 or W=20).

In a preferred embodiment, the percent identity between two amino acid

10 sequences is determined using the Needleman *et al.* (1970) (*J. Mol. Biol.* 48:444-453) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a BLOSUM 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between

15 two nucleotide sequences is determined using the GAP program in the GCG software package (Devereux *et al.* (1984) *Nucleic Acids Res.* 12(1):387) (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6.

Another preferred, non-limiting example of a mathematical algorithm utilized

20 for the comparison of sequences is the algorithm of Myers and Miller, CABIOS (1989). Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Additional

25 algorithms for sequence analysis are known in the art and include ADVANCE and ADAM as described in Torellis *et al.* (1994) *Comput. Appl. Biosci.* 10:3-5; and FASTA described in Pearson *et al.* (1988) *PNAS* 85:2444-8.

A variant polypeptide can differ in amino acid sequence by one or more substitutions, deletions, insertions, inversions, fusions, and truncations or a

30 combination of any of these.

Variant polypeptides can be fully functional or can lack function in one or more activities. Thus, in the present case, variations can affect the function, for example, of one or more of the regions corresponding to ligand binding, membrane association, G-protein binding and signal transduction.

WO 00/23568

PCT/US99/24368

Fully functional variants typically contain only conservative variation or variation in non-critical residues or in non-critical regions. Functional variants can also contain substitution of similar amino acids which result in no change or an insignificant change in function. Alternatively, such substitutions may positively or negatively affect function to some degree

Non-functional variants typically contain one or more non-conservative amino acid substitutions, deletions, insertions, inversions, or truncation or a substitution, insertion, inversion, or deletion in a critical residue or critical region.

As indicated, variants can be naturally-occurring or can be made by recombinant means or chemical synthesis to provide useful and novel characteristics for the receptor polypeptide. This includes preventing immunogenicity from pharmaceutical formulations by preventing protein aggregation.

Useful variations further include alteration of ligand binding characteristics. For example, one embodiment involves a variation at the binding site that results in binding but not release, or slower release, of ligand. A further useful variation at the same sites can result in a higher affinity for ligand. Useful variations also include changes that provide for affinity for another ligand. Another useful variation includes one that allows binding but which prevents activation by the ligand. Another useful variation includes variation in the transmembrane G-protein-binding/signal transduction domain that provides for reduced or increased binding by the appropriate G-protein or for binding by a different G-protein than the one with which the receptor is normally associated. Another useful variation provides a fusion protein in which one or more domains or subregions is operationally fused to one or more domains or subregions from another G-protein coupled receptor.

Amino acids that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham *et al.* (1989) *Science* 244:1081-1085). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as receptor binding or *in vitro*, or *in vitro* proliferative activity. Sites that are critical for ligand-receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al.* (1992) *J. Mol. Biol.* 224:899-904; de Vos *et al.* (1992) *Science* 255:306-312).

WO 00/23588

PCT/US99/24368

Substantial homology can be to the entire nucleic acid or amino acid sequence or to fragments of these sequences.

The invention thus also includes polypeptide fragments of the 2838 receptor protein. Fragments can be derived from the amino acid sequence shown in SEQ ID NO.1. However, the invention also encompasses fragments of the variants of the 2838 receptor protein as described herein.

The invention thus also includes polypeptide fragments of the 14618 receptor protein. Fragments can be derived from the amino acid sequence shown in SEQ ID NO.3. However, the invention also encompasses fragments of the variants of the 14618 receptor protein as described herein.

The invention thus also includes polypeptide fragments of the 15334 receptor protein. Fragments can be derived from the amino acid sequence shown in SEQ ID NO.5. However, the invention also encompasses fragments of the variants of the 15334 receptor protein as described herein.

The fragments to which the invention pertains, however, are not to be construed as encompassing fragments that may be disclosed prior to the present invention.

Fragments can retain one or more of the biological activities of the protein, for example the ability to bind to a G-protein or ligand, as well as fragments that can be used as an immunogen to generate receptor antibodies.

Biologically active fragments of 2838 receptor (peptides which are, for example, 8, 12, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) can comprise a domain or motif, e.g., an extracellular or intracellular domain or loop, one or more transmembrane segments, or parts thereof, G-protein binding site, or GPCR signature, glycosylation sites, protein kinase C and casein kinase II phosphorylation sites, N-myristoylation and amidation sites. Such domains or motifs can be identified by means of routine computerized homology searching procedures.

Possible fragments include, but are not limited to: 1) soluble peptides comprising the entire amino terminal extracellular domain from amino acid 1 to about amino acid 24 of SEQ ID NO.1, or parts thereof; 2) peptides comprising the entire carboxy terminal intracellular domain from about amino acid 293 to amino acid 319 of SEQ ID NO.1, or parts thereof; 3) peptides comprising the region spanning the entire transmembrane domain from about amino acid 25 to about amino acid 292, or parts thereof; 4) any of the specific transmembrane segments, or parts thereof, from

WO 00/23588

PCT/US99/24368

- about amino acid 25 to about amino acid 49, from about amino acid 56 to about amino acid 79, from about amino acid 100 to about amino acid 117, from about amino acid 138 to about amino acid 159, from about amino acid 187 to about amino acid 210, from about amino acid 224 to about amino acid 248, and from about amino acid 268 to about amino acid 292, 5) any of the three intracellular or three extracellular loops, or parts thereof, from about amino acid 50 to about amino acid 55, from about amino acid 118 to about amino acid 137, from about amino acid 211 to about amino acid 223, from about amino acid 80 to about amino acid 99, from about amino acid 160 to about amino acid 186, and from about amino acid 249 to about amino acid 267.
- 10 Fragments further include combinations of the above fragments, such as an amino terminal domain combined with one or more transmembrane segments and the attendant extra or intracellular loops or one or more transmembrane segments, and the attendant intra or extracellular loops, plus the carboxy terminal domain. Thus, any of the above fragments can be combined. Other fragments include the mature protein
- 15 from about amino acid 6 to 319. Other fragments contain the various functional sites described herein, such as phosphorylation sites, glycosylation sites, and myristoylation sites and a sequence containing the GPCR signature sequence. Fragments, for example, can extend in one or both directions from the functional site to encompass 5, 10, 15, 20, 30, 40, 50, or up to 100 amino acids. Further, fragments
- 20 can include sub-fragments of the specific domains mentioned above, which sub-fragments retain the function of the domain from which they are derived. Fragments also include amino acid sequences greater than 7 amino acids from amino acid 1 to about amino acid 264. Fragments also include antigenic fragments and specifically those shown to have a high antigenic index in Figure 2.
- 25 Accordingly, possible fragments include fragments defining a ligand-binding site, fragments defining a glycosylation site, fragments defining membrane association, fragments defining phosphorylation sites, fragments defining interaction with G proteins and signal transduction, and fragments defining myristoylation sites. By this is intended a discrete fragment that provides the relevant function or allows
- 30 the relevant function to be identified. In a preferred embodiment, the fragment contains the ligand-binding site.

Biologically active fragments of 14618 receptor (peptides which are, for example, 9, 12, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) can comprise a domain or motif, e.g., an extracellular or intracellular domain

WO 00/23588

PCT/US99/24368

or loop, one or more transmembrane segments, or parts thereof, G-protein binding site, or GPCR signature, glycosylation sites, and cAMP- and cGMP-dependent, protein kinase C, and casein kinase II phosphorylation sites.

Possible fragments include, but are not limited to: 1) soluble peptides

- 5 comprising the entire amino terminal extracellular domain about amino acid 1 to about amino acid 28 of SEQ ID NO:3, or parts thereof; 2) peptides comprising the entire carboxy terminal intracellular domain from about amino acid 298 to amino acid 337 of SEQ ID NO:3, or parts thereof; 3) peptides comprising the region spanning the entire transmembrane domain from about amino acid 29 to about amino acid 297, or
- 10 parts thereof; 4) any of the specific transmembrane segments, or parts thereof, from about amino acid 29 to about amino acid 49, from about amino acid 60 to about amino acid 84, from about amino acid 103 to about amino acid 127, from about amino acid 142 to about amino acid 161, from about amino acid 194 to about amino acid 217, from about amino acid 231 to about amino acid 247, and from about amino acid
- 15 276 to about amino acid 297; 5) any of the three intracellular or three extracellular loops, or parts thereof, from about amino acid 50 to about amino acid 59, from about amino acid 128 to about amino acid 141, from about amino acid 218 to about amino acid 230, from about amino acid 85 to about amino acid 102, from about amino acid 162 to about amino acid 193, and from about amino acid 248 to about amino acid 275.
- 20 Fragments further include combinations of the above fragments, such as an amino terminal domain combined with one or more transmembrane segments and the attendant extra or intracellular loops or one or more transmembrane segments, and the attendant intra or extracellular loops, plus the carboxy terminal domain. Thus, any of the above fragments can be combined. Other fragments include the mature protein
- 25 from about amino acid 6 to 337. Other fragments contain the various functional sites described herein, such as phosphorylation sites, glycosylation sites, and a sequence containing the GPCR signature sequence. Fragments, for example, can extend in one or both directions from the functional site to encompass 5, 10, 15, 20, 30, 40, 50, or up to 100 amino acids. Further, fragments can include sub-fragments of the specific domains mentioned above, which sub-fragments retain the function of the domain
- 30 from which they are derived. Fragments also include amino acid sequences greater than 8 amino acids from amino acid 1 to about amino acid 244.

Fragments also include antigenic fragments and specifically those shown to have a high antigenic index in Figure 2.

WO 00/23588

PCT/US99/24368

Accordingly, possible fragments include fragments defining a ligand-binding site, fragments defining a glycosylation site, fragments defining membrane association, fragments defining phosphorylation sites, fragments defining interaction with G proteins and signal transduction, and fragments defining myristoylation sites.

- 5 By this is intended a discrete fragment that provides the relevant function or allows the relevant function to be identified. In a preferred embodiment, the fragment contains the ligand-binding site.

- Biologically active fragments (peptides which are, for example, 8, 12, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) can comprise a
10 domain or motif, e.g., an extracellular or intracellular domain or loop, one or more transmembrane segments, or parts thereof; G-protein binding site, or GPCR signature, glycosylation sites, cAMP, cGMP, protein kinase C, and casein kinase II phosphorylation sites, and N-myristoylation sites.

- Possible fragments include, but are not limited to: 1) soluble peptides
15 comprising the entire amino terminal extracellular domain about amino acid 1 to about amino acid 23 of SEQ ID NO.5, or parts thereof; 2) peptides comprising the entire carboxy terminal intracellular domain from about amino acid 300 to amino acid 372 of SEQ ID NO.5, or parts thereof; 3) peptides comprising the region spanning the entire transmembrane domain from about amino acid 26 to about amino acid 299, or
20 parts thereof; 4) any of the specific transmembrane segments, or parts thereof, from about amino acid 26 to about amino acid 48, from about amino acid 56 to about amino acid 77, from about amino acid 99 to about amino acid 115, from about amino acid 140 to about amino acid 157, from about amino acid 188 to about amino acid 209, from about amino acid 235 to about amino acid 259, and from about amino acid
25 277 to about amino acid 299; 5) any of the three intracellular or three extracellular loops, or parts thereof, from about amino acid 49 to about amino acid 55, from about amino acid 78 to about amino acid 98, from about amino acid 116 to about amino acid 139, from about amino acid 158 to about amino acid 187, from about amino acid 210 to about amino acid 234, and from about amino acid 260 to about amino acid 276.
30 Fragments further include combinations of the above fragments, such as an amino terminal domain combined with one or more transmembrane segments and the attendant extra or intracellular loops or one or more transmembrane segments, and the attendant intra or extracellular loops, plus the carboxy terminal domain. Thus, any of the above fragments can be combined. Other fragments include the mature protein

WO 00/23588

PCT/US99/24368

from about amino acid 6 to 372. Other fragments contain the various functional sites described herein, such as phosphorylation sites, glycosylation sites, and myristoylation sites and a sequence containing the GPCR signature sequence.

Fragments, for example, can extend in one or both directions from the functional site to encompass 5, 10, 15, 20, 30, 40, 50, or up to 100 amino acids. Further, fragments can include sub-fragments of the specific domains mentioned above, which sub-fragments retain the function of the domain from which they are derived. Fragments also include amino acid sequences greater than 7 amino acids.

Fragments also include antigenic fragments and specifically those shown to have a high antigenic index in Figure 2.

These regions can be identified by well-known methods involving computerized homology analysis.

Accordingly, possible fragments include fragments defining a ligand-binding site, fragments defining a glycosylation site, fragments defining membrane association, fragments defining phosphorylation sites, fragments defining interaction with G proteins and signal transduction, and fragments defining myristoylation sites. By this is intended a discrete fragment that provides the relevant function or allows the relevant function to be identified. In a preferred embodiment, the fragment contains the ligand-binding site.

The invention also provides 2838 receptor fragments with immunogenic properties. These contain an epitope-bearing portion of the 2838 receptor protein and variants. These epitope-bearing peptides are useful to raise antibodies that bind specifically to a receptor polypeptide or region or fragment. These peptides can contain at least 8, 12, at least 14, or between at least about 15 to about 30 amino acids.

The invention also provides 14618 receptor fragments with immunogenic properties. These contain an epitope-bearing portion of the 14618 receptor protein and variants. These epitope-bearing peptides are useful to raise antibodies that bind specifically to a receptor polypeptide or region or fragment. These peptides can contain at least 9, 12, at least 14, or between at least about 15 to about 30 amino acids.

The invention also provides 15334 receptor fragments with immunogenic properties. These contain an epitope-bearing portion of the 15334 receptor protein and variants. These epitope-bearing peptides are useful to raise antibodies that bind specifically to a receptor polypeptide or region or fragment. These peptides can contain at least 8, 12, at least 14, or between at least about 15 to about 30 amino acids.

WO 00/23588

PCT/US99/24368

Non-limiting examples of antigenic polypeptides that can be used to generate antibodies include peptides derived from the amino terminal extracellular domain or any of the extracellular loops. Regions having a high antigenicity index are shown in Figures 2, 6, and 10.

- 5 The epitope-bearing receptor and polypeptides may be produced by any conventional means (Houghten (1985) *Proc. Natl. Acad. Sci. USA* 82:5131-5135). Simultaneous multiple peptide synthesis is described in U.S. Patent No. 4,631,211.

- Fragments can be discrete (not fused to other amino acids or polypeptides) or can be within a larger polypeptide. Further, several fragments can be comprised
10 within a single larger polypeptide. In one embodiment a fragment designed for expression in a host can have heterologous pre- and pro-polypeptide regions fused to the amino terminus of the receptor fragment and an additional region fused to the carboxyl terminus of the fragment.

- The invention thus provides chimeric or fusion proteins. These comprise a
15 receptor protein operatively linked to a heterologous protein having an amino acid sequence not substantially homologous to the receptor protein. "Operatively linked" indicates that the receptor protein and the heterologous protein are fused in-frame. The heterologous protein can be fused to the N-terminus or C-terminus of the receptor protein.

- 20 In one embodiment the fusion protein does not affect receptor function *per se*. For example, the fusion protein can be a GST-fusion protein in which the receptor sequences are fused to the C-terminus of the GST sequences. Other types of fusion proteins include, but are not limited to, enzymatic fusion proteins, for example beta-galactosidase fusions, yeast two-hybrid GAL fusions, poly-His fusions and Ig fusions.
25 Such fusion proteins, particularly poly-His fusions, can facilitate the purification of recombinant receptor protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of a protein can be increased by using a heterologous signal sequence. Therefore, in another embodiment, the fusion protein contains a heterologous signal sequence at its N-terminus.

- 30 EP-A-O 464 533 discloses fusion proteins comprising various portions of immunoglobulin constant regions. The Fc is useful in therapy and diagnosis and thus results, for example, in improved pharmacokinetic properties (EP-A 0232 262). In drug discovery, for example, human proteins have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists. Bennett *et al.*

WO 00/23588

PCT/US99/24368

(1995) *J. Mol. Recog.* 8:52-58 and Johanson *et al.* (1995) *J. Biol. Chem.* 270, 16:9459-9471. Thus, this invention also encompasses soluble fusion proteins containing a receptor polypeptide and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclass (IgG, IgM, IgA, IgE). Preferred as immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. For some uses it is desirable to remove the Fc after the fusion protein has been used for its intended purpose, for example when the fusion protein is to be used as antigen for immunizations. In a particular embodiment, the Fc part can be removed in a simple way by a cleavage sequence which is also incorporated and can be cleaved with factor Xa.

A chimeric or fusion protein can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different protein sequences are ligated together in-frame in accordance with conventional techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see Ausubel *et al.*, (1992) *Current Protocols in Molecular Biology*). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST protein). A receptor protein-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the receptor protein.

Another form of fusion protein is one that directly affects receptor functions. Accordingly, a receptor polypeptide is encompassed by the present invention in which one or more of the receptor domains (or parts thereof) has been replaced by homologous domains (or parts thereof) from another G-protein coupled receptor or other type of receptor. Accordingly, various permutations are possible. The amino terminal extracellular domain, or subregion thereof, (for example, ligand-binding) can be replaced with the domain or subregion from another ligand-binding receptor protein. Alternatively, the entire transmembrane domain, or any of the seven segments or loops, or parts thereof, for example, G-protein-binding/signal transduction, can be replaced. Finally, the carboxy terminal intracellular domain or

WO 80/23588

PCT/US99/24368

subregion can be replaced. Thus, chimeric receptors can be formed in which one or more of the native domains or subregions has been replaced.

The isolated 2838 receptor protein can be purified from cells that naturally express it, such as from lymph node, thymus, spleen, testes, colon, and peripheral blood lymphocytes, and from those cells in which it is significantly expressed as shown in Figure 15, such as activated T-helper cells (1 and 2), hypoxic Hep 3B cells, CD3 cells (both CD4 and CD8), activated B cells, Jurkat cells, among others, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods.

The isolated 14618 receptor protein can be purified from cells that naturally express it, such as from breast, skeletal muscle, lymph node, spleen and blood peripheral lymphocytes, as well as those tissues in which the gene is significantly expressed as shown in Figures 13 and 14 including, but not limited to, CD34⁺ cells and megakaryocytes, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods.

The isolated 15334 receptor protein can be purified from cells that naturally express it, such as from colon, placenta, pancreas, tonsil, lymph node, spleen, peripheral blood cells, thymus, adrenal gland and heart, as well as those tissues shown in Figures 16-19, including K562 cells, erythroblasts, and megakaryocytes, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods.

In one embodiment, the protein is produced by recombinant DNA techniques. For example, a nucleic acid molecule encoding the receptor polypeptide is cloned into an expression vector, the expression vector introduced into a host cell and the protein expressed in the host cell. The protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques.

Polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally-occurring amino acids. Further, many amino acids, including the terminal amino acids, may be modified by natural processes, such as processing and other post-translational modifications, or by chemical modification techniques well known in the art. Common modifications that occur naturally in polypeptides are described in basic texts, detailed monographs, and the research literature, and they are well known to those of skill in the art.

WO 00/23588

PCT/US99/24368

Accordingly, the polypeptides also encompass derivatives or analogs in which a substituted amino acid residue is not one encoded by the genetic code, in which a substituent group is included, in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence for purification of the mature polypeptide or a pro-protein sequence.

Known modifications include, but are not limited to, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well-known to those of skill in the art and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as *Proteins - Structure and Molecular Properties*, 2nd Ed., T.E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as by Wold, F., *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York 1-12 (1983); Seifter *et al.* (1990) *Meth. Enzymol.* 182: 626-646 and Rattan *et al.* (1992) *Ann. N.Y. Acad. Sci.* 663:48-62).

As is also well known, polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of post-translation events, including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, branched and branched circular polypeptides may be synthesized by non-translational natural processes and by synthetic methods.

WO 00/23588

PCT/US99/24368

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. Blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally-occurring and synthetic polypeptides. For instance, the amino

5 terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost invariably will be N-formylmethionine.

The modifications can be a function of how the protein is made. For recombinant polypeptides, for example, the modifications will be determined by the host cell posttranslational modification capacity and the modification signals in the

10 polypeptide amino acid sequence. Accordingly, when glycosylation is desired, a polypeptide should be expressed in a glycosylating host, generally a eukaryotic cell. Insect cells often carry out the same posttranslational glycosylations as mammalian cells and, for this reason, insect cell expression systems have been developed to efficiently express mammalian proteins having native patterns of glycosylation.

15 Similar considerations apply to other modifications.

The same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain more than one type of modification.

20 Polypeptide uses

The receptor polypeptides are useful for producing antibodies specific for the 2838, 14618, and 15334 receptor protein, regions, or fragments. Regions having a high antigenicity index score are shown in Figures 2, 6, and 10.

The receptor polypeptides, variants, and fragments (including those which

25 may have been disclosed prior to the present invention) are useful for biological assays related to GPCRs. Such assays involve any of the known GPCR functions or activities or properties useful diagnosis and treatment of GPCR-related conditions.

The receptor polypeptides are also useful in drug screening assays, in cell-based or cell-free systems. Cell-based systems can be native, *i.e.*, cells that normally

30 express the receptor protein, as a biopsy or expanded in cell culture. In one embodiment, however, cell-based assays involve recombinant host cells expressing the receptor protein.

Determining the ability of the test compound to interact with the polypeptide can also comprise determining the ability of the test compound to preferentially bind

WO 00/23588

PCT/US99/24368

to the polypeptide as compared to the ability of the ligand, or a biologically active portion thereof, to bind to the polypeptide.

The polypeptides can be used to identify compounds that modulate receptor activity. Such compounds, for example, can increase or decrease affinity or rate of binding to a known ligand, compete with ligand for binding to the receptor, or displace ligand bound to the receptor. 2838, 14618, and 15334 protein and appropriate variants and fragments can be used in high-throughput screens to assay candidate compounds for the ability to bind to the receptor. These compounds can be further screened against a functional receptor to determine the effect of the compound on the receptor activity. Compounds can be identified that activate (agonist) or inactivate (antagonist) the receptor to a desired degree. Modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). Examples include purine analogs such as those discussed above.

The receptor polypeptides can be used to screen a compound for the ability to stimulate or inhibit interaction between the receptor protein and a target molecule that normally interacts with the receptor protein. The target can be ligand or a component of the signal pathway with which the receptor protein normally interacts (for example, a G-protein or other interactor involved in cAMP or phosphatidylinositol turnover and/or adenylate cyclase, or phospholipase C activation). The assay includes the steps of combining the receptor protein with a candidate compound under conditions that allow the receptor protein or fragment to interact with the target molecule, and to detect the formation of a complex between the protein and the target or to detect the biochemical consequence of the interaction with the receptor protein and the target, such as any of the associated effects of signal transduction such as G-protein phosphorylation, cyclic AMP or phosphatidylinositol turnover, and adenylate cyclase or phospholipase C activation.

Determining the ability of the protein to bind to a target molecule can also be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (e.g., BIAcore™). Changes in the optical phenomenon surface

WO 00/23588

PCT/US99/24368

plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including:
 5 biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide
 10 oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:6909; Lirb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994) *J. Med.*
 15 *Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233. Libraries of compounds may be presented in solution (e.g., Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Iodor (1993) *Nature* 364:555-556),
 20 bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390), (Devlin (1990) *Science* 249:404-406), (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 97:6378-6382), (Felici (1991) *J. Mol. Biol.* 222:301-310), (Ladner *supra*).

Candidate compounds include, for example, 1) purine analogs, 2) peptides
 such as soluble peptides, including Ig-tailed fusion peptides and members of random peptide libraries (see, e.g., Lam *et al.* (1991) *Nature* 354:82-84; Houghten *et al.*
 (1991) *Nature* 354:84-86) and combinatorial chemistry-derived molecular libraries
 made of D- and/or L- configuration amino acids; 3) phosphopeptides (e.g., members
 30 of random and partially degenerate, directed phosphopeptide libraries, see, e.g., Songyang *et al.* (1993) *Cell* 72:767-778); 4) antibodies (e.g., polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, and single chain antibodies as well as Fab, F(ab')₂, Fab expression library fragments, and epitope-binding fragments of

WO 00/23588

PCT/US99/24368

antibodies); and 5) small organic and inorganic molecules (e.g., molecules obtained from combinatorial and natural product libraries).

One candidate compound is a soluble full-length receptor or fragment that competes for ligand binding. Other candidate compounds include mutant receptors or appropriate fragments containing mutations that affect receptor function and thus compete for ligand. Accordingly, a fragment that competes for ligand, for example with a higher affinity, or a fragment that binds ligand but does not allow release, is encompassed by the invention.

The invention provides other end points to identify compounds that modulate (stimulate or inhibit) receptor activity. The assays typically involve an assay of events in the signal transduction pathway that indicate receptor activity. Thus, the expression of genes that are up- or down-regulated in response to the receptor protein dependent signal cascade can be assayed. In one embodiment, the regulatory region of such genes can be operably linked to a marker that is easily detectable, such as luciferase. Alternatively, phosphorylation of the receptor protein, or a receptor protein target, could also be measured.

Binding and/or activating compounds can also be screened by using chimeric receptor proteins in which the amino terminal extracellular domain, or parts thereof, the entire transmembrane domain or subregions, such as any of the seven transmembrane segments or any of the intracellular or extracellular loops and the carboxy terminal intracellular domain, or parts thereof, can be replaced by heterologous domains or subregions. For example, a G-protein-binding region can be used that interacts with a different G-protein than that which is recognized by the native receptor. Accordingly, a different set of signal transduction components is available as an end-point assay for activation. Alternatively, the entire transmembrane portion or subregions (such as transmembrane segments or intracellular or extracellular loops) can be replaced with the entire transmembrane portion or subregions specific to a host cell that is different from the host cell from which the amino terminal extracellular domain and/or the G-protein-binding region are derived. This allows for assays to be performed in other than the specific host cell from which the receptor is derived. Alternatively, the amino terminal extracellular domain (and/or other ligand-binding regions) could be replaced by a domain (and/or other binding region) binding a different ligand, thus, providing an assay for test compounds that interact with the heterologous amino terminal extracellular domain

WO 00/23588

PCT/US99/24368

(or region) but still cause signal transduction. Finally, activation can be detected by a reporter gene containing an easily detectable coding region operably linked to a transcriptional regulatory sequence that is part of the native signal transduction pathway.

5 The receptor polypeptides are also useful in competition binding assays in methods designed to discover compounds that interact with the receptor. Thus, a compound is exposed to a receptor polypeptide under conditions that allow the compound to bind or to otherwise interact with the polypeptide. Soluble receptor polypeptide is also added to the mixture. If the test compound interacts with the
10 soluble receptor polypeptide, it decreases the amount of complex formed or activity from the receptor target. This type of assay is particularly useful in cases in which compounds are sought that interact with specific regions of the receptor. Thus, the soluble polypeptide that competes with the target receptor region is designed to contain peptide sequences corresponding to the region of interest.

15 To perform cell free drug screening assays, it is desirable to immobilize either the receptor protein, or fragment, or its target molecule to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay.

Techniques for immobilizing proteins on matrices can be used in the drug
20 screening assays. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be bound to a matrix. For example, glutathione-S-transferase/2838, 14618, and 15334 fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates (e.g., ³⁵S-labeled) and
25 the candidate compound, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly, or in the supernatant after the complexes are dissociated. Alternatively, the complexes can be dissociated from the
30 matrix, separated by SDS-PAGE, and the level of receptor-binding protein found in the bead fraction quantitated from the gel using standard electrophoretic techniques. For example, either the polypeptide or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin using techniques well known in the art. Alternatively, antibodies reactive with the protein but which do not interfere with

WO 00/23588

PCT/US99/24368

binding of the protein to its target molecule can be derivatized to the wells of the plate, and the protein trapped in the wells by antibody conjugation. Preparations of a receptor-binding protein and a candidate compound are incubated in the receptor protein-presenting wells and the amount of complex trapped in the well can be quantitated. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the receptor protein target molecule, or which are reactive with receptor protein and compete with the target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

Modulators of receptor protein activity identified according to these drug screening assays can be used to treat a subject with a disorder mediated by the receptor pathway, by treating cells that express the 2838 protein, such as in lymph node, thymus, spleen, testes, colon, and peripheral blood lymphocytes including, but not limited to, T-helper cells (1 and 2), CD3⁺ (CD4 and CD8) cells, B cells and granulocytes. Modulators of receptor protein activity identified according to these drug screening assays can be used to treat a subject with a disorder mediated by the receptor pathway, by treating cells that express the 14618 protein, such as in breast, skeletal muscle, spleen and peripheral blood lymphocytes as well as CD34⁺ cells and megakaryocytes. Modulators of receptor protein activity identified according to these drug screening assays can be used to treat a subject with a disorder mediated by the receptor pathway, by treating cells that express the 15334 protein, such as in lymph node, tonsil, pancreas, colon, spleen, peripheral blood cells, thymus, adrenal gland and heart as well as megakaryocytes and erythroblasts. These methods of treatment include the steps of administering the modulators of protein activity in a pharmaceutical composition as described herein, to a subject in need of such treatment.

Figure 13 shows expression of the 14618 receptor in a variety of normal human tissues. The gene is highly expressed in breast and skeletal muscle. Accordingly, expression of the gene in disorders involving these tissues is particularly relevant. Significant expression also occurs in the thyroid, placenta, fetal kidney, fetal heart, and lymph node. Accordingly, expression of the gene is relevant to disorders involving these tissues as well. Furthermore, as shown in Figure 13, lower levels of

WO 00/23588

PCT/US99/24368

expression are seen in a variety of other tissues. Accordingly, expression of the gene is also relevant in disorders involving these tissues.

The 14618 receptor is also expressed in various hematopoietic cells with and without activation. The gene is highly expressed in CD34⁺ bone marrow cells.

- 5 Accordingly, expression of the gene is relevant in a variety of blood cell progenitors. Expression of this gene is therefore relevant to disorders involving deficiencies in any of the major blood cell types, *i.e.* neutropenia, thrombocytopenia or anemia. The gene is also highly expressed in mobilized peripheral blood cell megakaryocytes (mobilized with G-CSF). Accordingly, expression of this gene is relevant to disorders
10 involving platelet function, such as thrombocytopenia. Significant expression is also seen in mobilized peripheral blood cell leukocytes, mobilized bone marrow CD34⁺ cells, and cord blood CD434⁺ cells. Accordingly, expression of the gene is relevant to function of these cells, and therefore relevant to disorders involving immune function or inflammation. Further, expression of the gene occurs in activated peripheral blood
15 mononuclear cells. Activated B cells differentially express the gene. Accordingly, expression of the gene is relevant to disorders involving immune function and/or inflammation.

The 2338 receptor is expressed in the cells and tissues shown in Figure 15.

- High levels of expression are shown in lymph node and thymus. Accordingly,
20 expression of the gene is especially relevant to disorders involving these tissues. Extremely high expression is found in CD8 cells and activated B cells. High expression also occurs in activated T-helper cells (1 and 2), CD4 cells and Jurkat cells (a T-cell line). Expression is differential in activated B cells and activated T-helper cells. Expression increases upon activation in both of these cell types. Accordingly,
25 expression of the gene is relevant to disorders involving immune function and inflammation. The gene is also significantly expressed in granulocytes. Accordingly, expression of the gene is relevant to disorders involving these cells.

Figure 16 shows expression of the 15334 receptor in normal human tissues.

- The gene is highly expressed in lymph node, tonsil, and pancreas. Accordingly,
30 expression of the gene is particularly relevant to disorders involving these tissues. Expression of the gene is also high in colon, testis, placenta, fetal heart, and spleen. Accordingly, expression of the gene is also relevant to disorders involving these tissues. In addition, the figure shows low levels of expression in several other tissues. Accordingly, expression of the gene may be relevant to disorders involving these

WO 00/23588

PCT/US99/24368

tissues. Expression of the 15334 receptor has also been studied in various hematopoietic cells (Figures 17-19). Extremely high expression occurs in primary megakaryocytes and erythroblasts. Accordingly, expression of the gene is relevant to erythrocyte differentiation and megakaryocyte differentiation and thus is relevant to treating anemia and thrombocytopenia. Further, expression of the gene is significantly increased in resting B cells compared to activated B cells. Accordingly, expression of the gene is relevant to B cell immune function. Further, lower levels of expression are found in various other cells of the hematopoietic lineage. See Figure 17. The expression of this gene in hematopoietic cells in a lineage-restricted manner indicates that expression is relevant in regulating the development of the lineage cells, erythrocytes/red blood cells, or megakaryocytes/platelets.

Disorders involving the spleen include, but are not limited to, splenomegaly, including nonspecific acute splenitis, congestive splenomegaly, and splenic infarcts, neoplasms, congenital anomalies, and rupture. Disorders associated with splenomegaly include infections, such as nonspecific splenitis, infectious mononucleosis, tuberculosis, typhoid fever, brucellosis, cytomegalovirus, syphilis, malaria, histoplasmosis, toxoplasmosis, kala-azar, trypanosomiasis, schistosomiasis, leishmaniasis, and echinococcosis; congestive states related to partial hypertension, such as cirrhosis of the liver, portal or splenic vein thrombosis, and cardiac failure; lymphohematogenous disorders, such as Hodgkin disease, non-Hodgkin lymphomas/leukemia, multiple myeloma, myeloproliferative disorders, hemolytic anemias, and thrombocytopenic purpura; immunologic-inflammatory conditions, such as rheumatoid arthritis and systemic lupus erythematosus; storage diseases such as Gaucher disease, Niemann-Pick disease, and mucopolysaccharidoses; and other conditions, such as amyloidosis, primary neoplasms and cysts, and secondary neoplasms.

Disorders involving the lung include, but are not limited to, congenital anomalies; atelectasis; diseases of vascular origin, such as pulmonary congestion and edema, including hemodynamic pulmonary edema and edema caused by microvascular injury, adult respiratory distress syndrome (diffuse alveolar damage), pulmonary embolism, hemorrhage, and infarction, and pulmonary hypertension and vascular sclerosis; chronic obstructive pulmonary disease, such as emphysema, chronic bronchitis, bronchial asthma, and bronchiectasis; diffuse interstitial (infiltrative, restrictive) diseases, such as pneumoconioses, sarcoidosis, idiopathic

WO 00/23588

PCT/US99/24368

pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia (pulmonary infiltration with eosinophilia), *Bronchiolitis obliterans*-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, including Goodpasture syndrome, idiopathic pulmonary hemosiderosis and other hemorrhagic syndromes, pulmonary involvement in collagen vascular disorders, and pulmonary alveolar proteinosis; complications of therapies, such as drug-induced lung disease, radiation-induced lung disease, and lung transplantation; tumors, such as bronchogenic carcinoma, including paraneoplastic syndromes, bronchioloalveolar carcinoma, neuroendocrine tumors, such as bronchial carcinoid, miscellaneous tumors, and metastatic tumors; pathologies of the pleura, including inflammatory pleural effusions, noninflammatory pleural effusions, pneumothorax, and pleural tumors, including solitary fibrous tumors (pleural fibroma) and malignant mesothelioma.

Disorders involving the colon include, but are not limited to, congenital anomalies, such as atresia and stenosis, Meckel diverticulum, congenital aganglionic megacolon-Hirschsprung disease; enterocolitis, such as diarrhea and dysentery, infectious enterocolitis, including viral gastroenteritis, bacterial enterocolitis, necrotizing enterocolitis, antibiotic-associated colitis (pseudomembranous colitis), and collagenous and lymphocytic colitis, miscellaneous intestinal inflammatory disorders, including parasites and protozoa, acquired immunodeficiency syndrome, transplantation, drug-induced intestinal injury, radiation enterocolitis, neutropenic colitis (typhilitis), and diversion colitis; idiopathic inflammatory bowel disease, such as Crohn disease and ulcerative colitis; tumors of the colon, such as non-neoplastic polyps, adenomas, familial syndromes, colorectal carcinogenesis, colorectal carcinoma, and carcinoid tumors.

Disorders involving the liver include, but are not limited to, hepatic injury; jaundice and cholestasis, such as bilirubin and bile formation, hepatic failure and cirrhosis, such as cirrhosis, portal hypertension, including ascites, portosystemic shunts, and splenomegaly; infectious disorders, such as viral hepatitis, including hepatitis A-E infection and infection by other hepatitis viruses, clinicopathologic syndromes, such as the carrier state, asymptomatic infection, acute viral hepatitis, chronic viral hepatitis, and fulminant hepatitis; autoimmune hepatitis; drug- and toxin-induced liver disease, such as alcoholic liver disease; inborn errors of metabolism and pediatric liver disease, such as hemochromatosis, Wilson disease, α_1 -

WO 00/23588

PCT/US99/24368

- antitrypsin deficiency, and neonatal hepatitis; intrahepatic biliary tract disease, such as secondary biliary cirrhosis, primary biliary cirrhosis, primary sclerosing cholangitis, and anomalies of the biliary tree, circulatory disorders, such as impaired blood flow into the liver, including hepatic artery compromise and portal vein obstruction and
- 5 thrombosis, impaired blood flow through the liver, including passive congestion and centrilobular necrosis and peliosis hepatis, hepatic vein outflow obstruction, including hepatic vein thrombosis (Budd-Chiari syndrome) and veno-occlusive disease; hepatic disease associated with pregnancy, such as preeclampsia and eclampsia, acute fatty liver of pregnancy, and intrahepatic cholestasis of pregnancy; hepatic complications
- 10 of organ or bone marrow transplantation, such as drug toxicity after bone marrow transplantation, graft-versus-host disease and liver rejection, and nonimmunologic damage to liver allografts; tumors and tumorous conditions, such as nodular hyperplasias, adenomas, and malignant tumors, including primary carcinoma of the liver and metastatic tumors.
- 15 Disorders involving the uterus and endometrium include, but are not limited to, endometrial histology in the menstrual cycle; functional endometrial disorders, such as anovulatory cycle, inadequate luteal phase, oral contraceptives and induced endometrial changes, and menopausal and postmenopausal changes; inflammations, such as chronic endometritis; adenomyosis; endometriosis; endometrial polyps;
- 20 endometrial hyperplasia; malignant tumors, such as carcinoma of the endometrium; mixed Mullerian and mesenchymal tumors, such as malignant mixed Mullerian tumors, tumors of the myometrium, including leiomyomas, leiomyosarcomas, and endometrial stromal tumors.
- Disorders involving the brain include, but are not limited to, disorders
- 25 involving neurons, and disorders involving glia, such as astrocytes, oligodendrocytes, ependymal cells, and microglia; cerebral edema, raised intracranial pressure and herniation, and hydrocephalus; malformations and developmental diseases, such as neural tube defects, forebrain anomalies, posterior fossa anomalies, and syringomyelia and hydromyelia; perinatal brain injury; cerebrovascular diseases, such as those
- 30 related to hypoxia, ischemia, and infarction, including hypotension, hypoperfusion, and low-flow states--global cerebral ischemia and focal cerebral ischemia--infarction from obstruction of local blood supply, intracranial hemorrhage, including intracerebral (intraparenchymal) hemorrhage, subarachnoid hemorrhage and ruptured berry aneurysms, and vascular malformations, hypertensive cerebrovascular disease,

WO 00/23588

PCT/US99/24368

- including lacunar infarcts, slit hemorrhages, and hypertensive encephalopathy; infections, such as acute meningitis, including acute pyogenic (bacterial) meningitis and acute aseptic (viral) meningitis, acute focal suppurative infections, including brain abscess, subdural empyema, and extradural abscess, chronic bacterial
- 5 meningoencephalitis, including tuberculosis and mycobacterioses, neurosyphilis, and neuroborreliosis (Lyme disease), viral meningoencephalitis, including arthropod-borne (Arbo) viral encephalitis, *Herpes simplex* virus Type 1, *Herpes simplex* virus Type 2, *Varicella-zoster* virus (*Herpes zoster*), cytomegalovirus, poliomyelitis, rabies, and human immunodeficiency virus 1, including HIV-1 meningoencephalitis
- 10 (subacute encephalitis), vacuolar myelopathy, AIDS-associated myopathy, peripheral neuropathy, and AIDS in children, progressive multifocal leukoencephalopathy, subacute sclerosing panencephalitis, fungal meningoencephalitis, other infectious diseases of the nervous system; transmissible spongiform encephalopathies (prion diseases); demyelinating diseases, including multiple sclerosis, multiple sclerosis variants, acute disseminated encephalomyelitis and acute necrotizing hemorrhagic
- 15 encephalomyelitis, and other diseases with demyelination; degenerative diseases, such as degenerative diseases affecting the cerebral cortex, including Alzheimer disease and Pick disease, degenerative diseases of basal ganglia and brain stem, including Parkinsonism, idiopathic Parkinson disease (paralysis agitans), progressive
- 20 supranuclear palsy, corticobasal degeneration, multiple system atrophy, including striatonigral degeneration, Shy-Drager syndrome, and olivopontocerebellar atrophy, and Huntington disease, spinocerebellar degenerations, including spinocerebellar ataxias, including Friedreich ataxia, and ataxia-telangiectasia, degenerative diseases affecting motor neurons, including amyotrophic lateral sclerosis (motor neuron
- 25 disease), bulbospinal atrophy (Kennedy syndrome), and spinal muscular atrophy; inborn errors of metabolism, such as leukodystrophies, including Krabbe disease, metachromatic leukodystrophy, adrenoleukodystrophy, Pelizaeus-Merzbacher disease, and Canavan disease, mitochondrial encephalomyopathies, including Leigh disease and other mitochondrial encephalomyopathies; toxic and acquired metabolic
- 30 diseases, including vitamin deficiencies such as thiamine (vitamin B₁) deficiency and vitamin B₁₂ deficiency, neurologic sequelae of metabolic disturbances, including hypoglycemia, hyperglycemia, and hepatic encephatopathy, toxic disorders, including carbon monoxide, methanol, ethanol, and radiation, including combined methotrexate and radiation-induced injury, tumors, such as gliomas, including astrocytoma,

WO 00/23588

PCT/US99/24368

- including fibrillary (diffuse) astrocytoma and glioblastoma multiforme, pilocytic astrocytoma, pleomorphic xanthoastrocytoma, and brain stem glioma, oligodendroglioma, and ependymoma and related paraventricular mass lesions, neuronal tumors, poorly differentiated neoplasms, including medulloblastoma, other
- 5 parenchymal tumors, including primary brain lymphoma, germ cell tumors, and pineal parenchymal tumors, meningiomas, metastatic tumors, paraneoplastic syndromes, peripheral nerve sheath tumors, including schwannoma, neurofibroma, and malignant peripheral nerve sheath tumor (malignant schwannoma), and neurocutaneous syndromes (phakomatosis), including neurofibromatosis, including
- 10 Type 1 neurofibromatosis (NF1) and TYPE 2 neurofibromatosis (NF2), tuberous sclerosis, and Von Hippel-Lindau disease.

- Disorders involving T-cells include, but are not limited to, cell-mediated hypersensitivity, such as delayed type hypersensitivity and T-cell-mediated cytotoxicity, and transplant rejection; autoimmune diseases, such as systemic lupus
- 15 erythematosus, Sjögren syndrome, systemic sclerosis, inflammatory myopathies, mixed connective tissue disease, and polyarteritis nodosa and other vasculitides; immunologic deficiency syndromes, including but not limited to, primary immunodeficiencies, such as thymic hypoplasia, severe combined immunodeficiency diseases, and AIDS; leukopenia; reactive (inflammatory) proliferations of white cells,
- 20 including but not limited to, leukocytosis, acute nonspecific lymphadenitis, and chronic nonspecific lymphadenitis; neoplastic proliferations of white cells, including but not limited to lymphoid neoplasms, such as precursor T-cell neoplasms, such as acute lymphoblastic leukemia/lymphoma, peripheral T-cell and natural killer cell neoplasms that include peripheral T-cell lymphoma, unspecified, adult T-cell
- 25 leukemia/lymphoma, mycosis fungoides and Sézary syndrome, and Hodgkin disease.

- Diseases of the skin, include but are not limited to, disorders of pigmentation and melanocytes, including but not limited to, vitiligo, freckle, melasma, lentigo, nevocellular nevus, dysplastic nevi, and malignant melanoma; benign epithelial
- 30 tumors, including but not limited to, seborrheic keratoses, acanthosis nigricans, fibroepithelial polyp, epithelial cyst, keratoacanthoma, and adnexal (appendage) tumors; premalignant and malignant epidermal tumors, including but not limited to, actinic keratosis, squamous cell carcinoma, basal cell carcinoma, and merkel cell carcinoma; tumors of the dermis, including but not limited to, benign fibrous histiocytoma, dermatofibrosarcoma protuberans, xanthomas, and dermal vascular

WO 00/23588

PCT/US99/24368

tumors; tumors of cellular immigrants to the skin, including but not limited to, histiocytosis X, mycosis fungoides (cutaneous T-cell lymphoma), and mastocytosis; disorders of epidermal maturation, including but not limited to, ichthyosis; acute inflammatory dermatoses, including but not limited to, urticaria, acute eczematous dermatitis, and erythema multiforme; chronic inflammatory dermatoses, including but not limited to, psoriasis, lichen planus, and lupus erythematosus; blistering (bullous) diseases, including but not limited to, pemphigus, bullous pemphigoid, dermatitis herpetiformis, and noninflammatory blistering diseases; epidermolysis bullosa and porphyria, disorders of epidermal appendages, including but not limited to, acne vulgaris; panniculitis, including but not limited to, erythema nodosum and erythema induratum; and infection and infestation, such as verrucae, molluscum contagiosum, impetigo, superficial fungal infections, and arthropod bites, stings, and infestations.

In normal bone marrow, the myelocytic series (polymorphonuclear cells) make up approximately 60% of the cellular elements, and the erythrocytic series, 20-30%. Lymphocytes, monocytes, reticular cells, plasma cells and megakaryocytes together constitute 10-20%. Lymphocytes make up 5-15% of normal adult marrow. In the bone marrow, cell types are added mixed so that precursors of red blood cells (erythroblasts), macrophages (monoblasts), platelets (megakaryocytes), polymorphonuclear leucocytes (myeloblasts), and lymphocytes (lymphoblasts) can be visible in one microscopic field. In addition, stem cells exist for the different cell lineages, as well as a precursor stem cell for the committed progenitor cells of the different lineages. The various types of cells and stages of each would be known to the person of ordinary skill in the art and are found, for example, on page 42 (Figure 2-8) of *Immunology, Immunopathology and Immunity*, Fifth Edition, Sell *et al.* Simon and Schuster (1996), incorporated by reference for its teaching of cell types found in the bone marrow. According, the invention is directed to disorders arising from these cells. These disorders include but are not limited to the following: diseases involving hematopoietic stem cells; committed lymphoid progenitor cells; lymphoid cells including B and T-cells; committed myeloid progenitors, including monocytes, granulocytes, and megakaryocytes; and committed erythroid progenitors. These include but are not limited to the leukemias, including B-lymphoid leukemias, T-lymphoid leukemias, undifferentiated leukemias, erythroleukemia, megakaryoblastic leukemia, monocytic, [leukemias are encompassed with and without differentiation]; chronic and acute lymphoblastic leukemia, chronic and acute lymphocytic leukemia,

WO 00/23588

PCT/US99/24368

- chronic and acute myelogenous leukemia, lymphoma, myelo dysplastic syndrome, chronic and acute myeloid leukemia, myelomonocytic leukemia; chronic and acute myeloblastic leukemia, chronic and acute myelogenous leukemia, chronic and acute promyelocytic leukemia, chronic and acute myelocytic leukemia, hematologic malignancies of monocyte-macrophage lineage, such as juvenile chronic myelogenous leukemia, secondary AML, antecedent hematological disorder; refractory anemia; aplastic anemia; reactive cutaneous angioendotheliomatosis; fibrosing disorders involving altered expression in dendritic cells, disorders including systemic sclerosis, E-M syndrome, epidemic toxic oil syndrome, eosinophilic fasciitis
- 10 localized forms of scleroderma, keloid, and fibrosing colonopathy; angiomatoid malignant fibrous histiocytoma; carcinoma, including primary head and neck squamous cell carcinoma; sarcoma, including kaposi's sarcoma; fibroadenoma and phyllodes tumors, including mammary fibroadenoma; stromal tumors; phyllodes tumors, including histiocytoma, erythroblastosis; neurofibromatosis; diseases of the
- 15 vascular endothelium; demyelinating, particularly in old lesions; gliosis, vasogenic edema, vascular disease, Alzheimer's and Parkinson's disease; T-cell lymphomas; B cell lymphomas.

- Disorders involving the heart, include but are not limited to, heart failure, including but not limited to, cardiac hypertrophy, left-sided heart failure, and right-
- 20 sided heart failure; ischemic heart disease, including but not limited to angina pectoris, myocardial infarction, chronic ischemic heart disease, and sudden cardiac death; hypertensive heart disease, including but not limited to, systemic (left-sided) hypertensive heart disease and pulmonary (right-sided) hypertensive heart disease; valvular heart disease, including but not limited to, valvular degeneration caused by
- 25 calcification, such as calcific aortic stenosis, calcification of a congenitally bicuspid aortic valve, and mitral annular calcification, and myxomatous degeneration of the mitral valve (mitral valve prolapse), rheumatic fever and rheumatic heart disease, infective endocarditis, and noninfected vegetations, such as nonbacterial thrombotic endocarditis and endocarditis of systemic lupus erythematosus (Libman-Sacks
- 30 disease), carcinoid heart disease, and complications of artificial valves; myocardial disease, including but not limited to dilated cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy, and myocarditis; pericardial disease, including but not limited to, pericardial effusion and hemopericardium and pericarditis, including acute pericarditis and healed pericarditis, and rheumatoid heart

WO 00/23588

PCT/US99/24368

disease; neoplastic heart disease, including but not limited to, primary cardiac tumors, such as myxoma, lipoma, papillary fibroelastoma, rhabdomyoma, and sarcoma, and cardiac effects of noncardiac neoplasms; congenital heart disease, including but not limited to, left-to-right shunts--late cyanosis, such as atrial septal defect, ventricular septal defect, patent ductus arteriosus, and atrioventricular septal defect, right-to-left shunts--early cyanosis, such as tetralogy of fallot, transposition of great arteries, truncus arteriosus, tricuspid atresia, and total anomalous pulmonary venous connection, obstructive congenital anomalies, such as coarctation of aorta, pulmonary stenosis and atresia, and aortic stenosis and atresia, and disorders involving cardiac transplantation.

Disorders involving blood vessels include, but are not limited to, responses of vascular cell walls to injury, such as endothelial dysfunction and endothelial activation and intimal thickening, vascular diseases including, but not limited to, congenital anomalies, such as arteriovenous fistula, atherosclerosis, and hypertensive vascular disease, such as hypertension; inflammatory disease--the vasculitides, such as giant cell (temporal) arteritis, Takayasu arteritis, polyarteritis nodosa (classic), Kawasaki syndrome (mucocutaneous lymph node syndrome), microscopic polyangiitis (microscopic polyarteritis, hypersensitivity or leukocytoclastic angitis), Wegener granulomatosis, thromboangiitis obliterans (Buerger disease), vasculitis associated with other disorders, and infectious arteritis; Raynaud disease; aneurysms and dissection, such as abdominal aortic aneurysms, syphilitic (lentic) aneurysms, and aortic dissection (dissecting hematoma); disorders of veins and lymphatics, such as varicose veins, thrombophlebitis and phlebotrombosis, obstruction of superior vena cava (superior vena cava syndrome), obstruction of inferior vena cava (inferior vena cava syndrome), and lymphangitis and lymphedema; tumors, including benign tumors and tumor-like conditions, such as hemangioma, lymphangioma, glomus tumor (glomangioma), vascular ectasias, and bacillary angiomatosis, and intermediate-grade (borderline low-grade malignant) tumors, such as Kaposi sarcoma and hemangioendothelioma, and malignant tumors, such as angiosarcoma and hemangiopericytoma; and pathology of therapeutic interventions in vascular disease, such as balloon angioplasty and related techniques and vascular replacement, such as coronary artery bypass graft surgery.

Disorders involving red cells include, but are not limited to, anemias, such as hemolytic anemias, including hereditary spherocytosis, hemolytic disease due to

WO 00/23588

PCT/US99/24368

erythrocyte enzyme defects: glucose-6-phosphate dehydrogenase deficiency, sickle cell disease, thalassemia syndromes, paroxysmal nocturnal hemoglobinuria, immunohemolytic anemia, and hemolytic anemia resulting from trauma to red cells; and anemias of diminished erythropoiesis, including megaloblastic anemias, such as

5 anemias of vitamin B12 deficiency: pernicious anemia, and anemia of folate deficiency; iron deficiency anemia, anemia of chronic disease, aplastic anemia, pure red cell aplasia, and other forms of marrow failure

Disorders involving the thymus include developmental disorders, such as DiGeorge syndrome with thymic hypoplasia or aplasia; thymic cysts; thymic

10 hypoplasia, which involves the appearance of lymphoid follicles within the thymus, creating thymic follicular hyperplasia; and thymomas, including germ cell tumors, lymphomas, Hodgkin disease, and carcinoids. Thymomas can include benign or encapsulated thymoma, and malignant thymoma Type I (invasive thymoma) or Type II, designated thymic carcinoma.

Disorders involving B cells include, but are not limited to precursor B cell neoplasms, such as lymphoblastic leukemia/lymphoma. Peripheral B cell neoplasms include, but are not limited to, chronic lymphocytic leukemia/small lymphocytic lymphoma, follicular lymphoma, diffuse large B cell lymphoma, Burkitt lymphoma, plasma cell neoplasms, multiple myeloma, and related entities, lymphoplasmacytic

20 lymphoma (Waldenström macroglobulinemia), mantle cell lymphoma, marginal zone lymphoma (MAL.Toma), and hairy cell leukemia.

Disorders involving the kidney include, but are not limited to, congenital anomalies including, but not limited to, cystic diseases of the kidney, that include but are not limited to, cystic renal dysplasia, autosomal dominant (adult) polycystic kidney

25 disease, autosomal recessive (childhood) polycystic kidney disease, and cystic diseases of renal medulla, which include, but are not limited to, medullary sponge kidney, and nephronophthisis-uremic medullary cystic disease complex, acquired (dialysis-associated) cystic disease, such as simple cysts, glomerular diseases including pathologies of glomerular injury that include, but are not limited to, in situ immune

30 complex deposition, that includes, but is not limited to, anti-GBM nephritis, Heymann nephritis, and antibodies against planted antigens, circulating immune complex nephritis, antibodies to glomerular cells, cell-mediated immunity in glomerulonephritis, activation of alternative complement pathway, epithelial cell injury, and pathologies involving mediators of glomerular injury including cellular and soluble mediators, acute

WO 00/23588

PCT/US99/24368

- glomerulonephritis, such as acute proliferative (poststreptococcal, postinfectious) glomerulonephritis, including but not limited to, poststreptococcal glomerulonephritis and nonstreptococcal acute glomerulonephritis, rapidly progressive (crescentic) glomerulonephritis, nephrotic syndrome, membranous glomerulonephritis (membranous nephropathy), minimal change disease (lipoid nephrosis), focal segmental glomerulosclerosis, membranoproliferative glomerulonephritis, IgA nephropathy (Berger disease), focal proliferative and necrotizing glomerulonephritis (focal glomerulonephritis), hereditary nephritis, including but not limited to, Alport syndrome and thin membrane disease (benign familial hematuria), chronic glomerulonephritis,
- 10 glomerular lesions associated with systemic disease, including but not limited to, systemic lupus erythematosus, Henoch-Schönlein purpura, bacterial endocarditis, diabetic glomerulosclerosis, amyloidosis, fibrillary and immunotactoid glomerulonephritis, and other systemic disorders, diseases affecting tubules and interstitium, including acute tubular necrosis and tubulointerstitial nephritis, including
- 15 but not limited to, pyelonephritis and urinary tract infection, acute pyelonephritis, chronic pyelonephritis and reflux nephropathy, and tubulointerstitial nephritis induced by drugs and toxins, including but not limited to, acute drug-induced interstitial nephritis, analgesic abuse nephropathy, nephropathy associated with nonsteroidal anti-inflammatory drugs, and other tubulointerstitial diseases including, but not limited to,
- 20 urate nephropathy, hypercalcemia and nephrocalcinosis, and multiple myeloma; diseases of blood vessels including benign nephrosclerosis, malignant hypertension and accelerated nephrosclerosis, renal artery stenosis, and thrombotic microangiopathies including, but not limited to, classic (childhood) hemolytic-uremic syndrome, adult hemolytic-uremic syndrome/thrombotic thrombocytopenic purpura, idiopathic
- 25 HUS/TTP, and other vascular disorders including, but not limited to, atherosclerotic ischemic renal disease, atheroembolic renal disease, sickle cell disease nephropathy, diffuse cortical necrosis, and renal infarcts; urinary tract obstruction (obstructive uropathy), urolithiasis (renal calculi, stones); and tumors of the kidney including, but not limited to, benign tumors, such as renal papillary adenoma, renal fibroma or hamartoma
- 30 (renomedullary interstitial cell tumor), angiomyolipoma, and oncocytoma, and malignant tumors, including renal cell carcinoma (hypernephroma, adenocarcinoma of kidney), which includes urothelial carcinomas of renal pelvis.

Disorders of the breast include, but are not limited to, disorders of development, inflammations, including but not limited to, acute mastitis, periductal mastitis, periductal

WO 00/23588

PCT/US99/24368

- mastitis (recurrent subareolar abscess, squamous metaplasia of lactiferous ducts), mammary duct ectasia, fat necrosis, granulomatous mastitis, and pathologies associated with silicone breast implants; fibrocystic changes; proliferative breast disease including, but not limited to, epithelial hyperplasia, sclerosing adenosis, and small duct papillomas;
- 5 tumors including, but not limited to, stromal tumors such as fibroadenoma, phyllodes tumor, and sarcomas, and epithelial tumors such as large duct papilloma; carcinoma of the breast including in situ (noninvasive) carcinoma that includes ductal carcinoma in situ (including Paget's disease) and lobular carcinoma in situ, and invasive (infiltrating) carcinoma including, but not limited to, invasive ductal carcinoma, no special type,
- 10 invasive lobular carcinoma, medullary carcinoma, colloid (mucinous) carcinoma, tubular carcinoma, and invasive papillary carcinoma, and miscellaneous malignant neoplasms.

Disorders in the male breast include, but are not limited to, gynecomastia and carcinoma

- Disorders involving the testis and epididymis include, but are not limited to,
- 15 congenital anomalies such as cryptorchidism, regressive changes such as atrophy, inflammations such as nonspecific epididymitis and orchitis, granulomatous (autoimmune) orchitis, and specific inflammations including, but not limited to, gonorrhea, mumps, tuberculosis, and syphilis, vascular disturbances including torsion, testicular tumors including germ cell tumors that include, but are not limited to,
- 20 seminoma, spermatocytic seminoma, embryonal carcinoma, yolk sac tumor, choriocarcinoma, teratoma, and mixed tumors, tumors of sex cord-gonadal stroma including, but not limited to, Leydig (interstitial) cell tumors and Sertoli cell tumors (androblastoma), and testicular lymphoma, and miscellaneous lesions of tunica vaginalis.

- Disorders involving the prostate include, but are not limited to, inflammations,
- 25 benign enlargement, for example, nodular hyperplasia (benign prostatic hypertrophy or hyperplasia), and tumors such as carcinoma.

- Disorders involving the thyroid include, but are not limited to, hyperthyroidism; hypothyroidism including, but not limited to, cretinism and myxedema; thyroiditis including, but not limited to, Hashimoto thyroiditis, subacute (granulomatous) thyroiditis,
- 30 and subacute lymphocytic (painless) thyroiditis; Graves disease; diffuse and multinodular goiter including, but not limited to, diffuse nontoxic (simple) goiter and multinodular goiter; neoplasms of the thyroid including, but not limited to, adenomas, other benign tumors, and carcinomas, which include, but are not limited to, papillary

WO 00/23588

PCT/US99/24368

carcinoma, follicular carcinoma, medullary carcinoma, and anaplastic carcinoma, and congenital anomalies.

Disorders involving the skeletal muscle include tumors such as rhabdomyosarcoma.

- 5 Disorders involving the pancreas include those of the exocrine pancreas such as congenital anomalies, including but not limited to, ectopic pancreas; pancreatitis, including but not limited to, acute pancreatitis; cysts, including but not limited to, pseudocysts; tumors, including but not limited to, cystic tumors and carcinoma of the pancreas; and disorders of the endocrine pancreas such as, diabetes mellitus; islet cell
10 tumors, including but not limited to, insulinomas, gastrinomas, and other rare islet cell tumors.

- Disorders involving the small intestine include the malabsorption syndromes such as, celiac sprue, tropical sprue (postinfectious sprue), whipple disease, disaccharidase (lactase) deficiency, abetalipoproteinemia, and tumors of the small
15 intestine including adenomas and adenocarcinoma.

- Disorders related to reduced platelet number, thrombocytopenia, include idiopathic thrombocytopenic purpura, including acute idiopathic thrombocytopenic purpura, drug-induced thrombocytopenia, HIV-associated thrombocytopenia, and thrombotic microangiopathies: thrombotic thrombocytopenic purpura and hemolytic-
20 uremic syndrome.

- Disorders involving precursor T-cell neoplasms include precursor T lymphoblastic leukemia/lymphoma. Disorders involving peripheral T-cell and natural killer cell neoplasms include T-cell chronic lymphocytic leukemia, large granular lymphocytic leukemia, mycosis fungoides and Sézary syndrome, peripheral T-cell
25 lymphoma, unspecified, angioimmunoblastic T-cell lymphoma, angiocentric lymphoma (NK/T-cell lymphoma^{4a}), intestinal T-cell lymphoma, adult T-cell leukemia/lymphoma, and anaplastic large cell lymphoma.

- Disorders involving the ovary include, for example, polycystic ovarian disease, Stein-Leventhal syndrome, Pseudomyxoma peritonei and stromal
30 hyperthecosis; ovarian tumors such as, tumors of coelomic epithelium, serous tumors, mucinous tumors, endometrioid tumors, clear cell adenocarcinoma, cystadenofibroma, Brenner tumor, surface epithelial tumors; germ cell tumors such as mature (benign) teratomas, monodermal teratomas, immature malignant teratomas, dysgerminoma, endodermal sinus tumor, choriocarcinoma; sex cord-stromal tumors

WO 00/23588

PCT/US99/24368

such as, granulosa-theca cell tumors, thecoma-fibromas, androblastomas, hill cell tumors, and gonadoblastoma; and metastatic tumors such as Krukenberg tumors.

Bone-forming cells include the osteoprogenitor cells, osteoblasts, and osteocytes. The disorders of the bone are complex because they may have an impact on the skeleton during any of its stages of development. Hence, the disorders may have variable manifestations and may involve one, multiple or all bones of the body. Such disorders include, congenital malformations, achondroplasia and thanatophoric dwarfism, diseases associated with abnormal matrix such as type I collagen disease, osteoporosis, paget disease, rickets, osteomalacia, high-turnover osteodystrophy, low-turnover of aplastic disease, osteonecrosis, pyogenic osteomyelitis, tuberculous osteomyelitis, osteoma, osteoid osteoma, osteoblastoma, osteosarcoma, osteochondroma, chondromas, chondroblastoma, chondromyxoid fibroma, chondrosarcoma, fibrous cortical defects, fibrous dysplasia, fibrosarcoma, malignant fibrous histiocytoma, ewing sarcoma, primitive neuroectodermal tumor, giant cell tumor, and metastatic tumors.

The receptor polypeptides are thus useful for treating a receptor-associated disorder characterized by aberrant expression or activity of a receptor protein, especially as disclosed above. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) expression or activity of the protein. In another embodiment, the method involves administering a protein as therapy to compensate for reduced or aberrant expression or activity of the protein.

Stimulation of protein activity is desirable in situations in which the protein is abnormally downregulated and/or in which increased protein activity is likely to have a beneficial effect. Likewise, inhibition of protein activity is desirable in situations in which the protein is abnormally upregulated and/or in which decreased protein activity is likely to have a beneficial effect. In one example of such a situation, a subject has a disorder characterized by aberrant development or cellular differentiation. In another example of such a situation, the subject has a proliferative disease (e.g., cancer) or a disorder characterized by an aberrant hematopoietic response. In another example of such a situation, it is desirable to achieve tissue regeneration in a subject (e.g., where a subject has undergone brain or spinal cord injury and it is desirable to regenerate neuronal tissue in a regulated manner).

WO 00/23588

PCT/US99/24368

In yet another aspect of the invention, the proteins of the invention can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brem WO 94/10300), to identify other proteins (captured proteins) which bind to or interact with the proteins of the invention and modulate their activity.

The 2838 receptor polypeptides also are useful to provide a target for diagnosing a disease or predisposition to disease mediated by the receptor protein, especially in hematopoietic cells such as T-helper cells (1 and 2), B cells, and granulocytes. The 14618 receptor polypeptides also are useful to provide a target for diagnosing a disease or predisposition to disease mediated by the receptor protein, especially in breast, skeletal, muscle and lymphnode, as well as CD34⁺ progenitor cells and cells generated from these progenitors, which include erythroid, lymphoid, and megakaryocytic lineages. The 15334 receptor polypeptides also are useful to provide a target for diagnosing a disease or predisposition to disease mediated by the receptor protein, especially in pancreas, spleen, tonsils and lymphnodes, as well as in cells of the megakaryocyte and erythroblast lineages. Accordingly, methods are provided for detecting the presence, or levels of, the receptor protein in a cell, tissue, or organism. The method involves contacting a biological sample with a compound capable of interacting with the receptor protein such that the interaction can be detected.

One agent for detecting receptor protein is an antibody capable of selectively binding to receptor protein. A biological sample includes tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject.

The receptor protein also provides a target for diagnosing active disease, or predisposition to disease, in a patient having a variant receptor protein. Thus, receptor protein can be isolated from a biological sample, assayed for the presence of a genetic mutation that results in aberrant receptor protein. This includes amino acid substitution, deletion, insertion, rearrangement, (as the result of aberrant splicing events), and inappropriate post-translational modification. Analytic methods include altered electrophoretic mobility, altered tryptic peptide digest, altered receptor activity in cell-based or cell-free assay, alteration in ligand or antibody-binding pattern,

WO 00/23588

PCT/US99/24368

altered isoelectric point, direct amino acid sequencing, and any other of the known assay techniques useful for detecting mutations in a protein.

In vitro techniques for detection of receptor protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. Alternatively, the protein can be detected *in vivo* in a subject by introducing into the subject a labeled anti-receptor antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques. Particularly useful are methods which detect the allelic variant of a receptor protein expressed in a subject and methods which detect fragments of a receptor protein in a sample.

The receptor polypeptides are also useful in pharmacogenomic analysis. Pharmacogenomics deal with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, e.g., Eichelbaum (1997) *Clin. Exp. Pharmacol. Physiol.* 23(10-11):983-985 (1996), and Linder (1997) *Clin. Chem.* 43(2):254-266. The clinical outcomes of these variations result in severe toxicity of therapeutic drugs in certain individuals or therapeutic failure of drugs in certain individuals as a result of individual variation in metabolism. Thus, the genotype of the individual can determine the way a therapeutic compound acts on the body or the way the body metabolizes the compound. Further, the activity of drug metabolizing enzymes effects both the intensity and duration of drug action. Thus, the pharmacogenomics of the individual permit the selection of effective compounds and effective dosages of such compounds for prophylactic or therapeutic treatment based on the individual's genotype. The discovery of genetic polymorphisms in some drug metabolizing enzymes has explained why some patients do not obtain the expected drug effects, show an exaggerated drug effect, or experience serious toxicity from standard drug dosages. Polymorphisms can be expressed in the phenotype of the extensive metabolizer and the phenotype of the poor metabolizer. Accordingly, genetic polymorphism may lead to allelic protein variants of the receptor protein in which one or more of the receptor functions in one population is different from those in another population. The polypeptides thus allow a target to ascertain a genetic predisposition that can affect treatment modality. Thus, in a ligand-based treatment, polymorphism may give rise to amino terminal extracellular domains and/or other ligand-binding regions that are more or less active in ligand binding, and receptor activation. Accordingly, ligand dosage would

WO 00/23588

PCT/US99/24368

necessarily be modified to maximize the therapeutic effect within a given population containing a polymorphism. As an alternative to genotyping, specific polymorphic polypeptides could be identified.

The receptor polypeptides are also useful for monitoring therapeutic effects during clinical trials and other treatment. Thus, the therapeutic effectiveness of an agent that is designed to increase or decrease gene expression, protein levels or receptor activity can be monitored over the course of treatment using the receptor polypeptides as an end-point target. The monitoring can be, for example, as follows: (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression or activity of a specified protein in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the protein in the post-administration samples; (v) comparing the level of expression or activity of the protein in the pre-administration sample with the protein in the post-administration sample or samples; and (vi) increasing or decreasing the administration of the agent to the subject accordingly.

The receptor polypeptides are also useful for treating a receptor-associated disorder. Accordingly, methods for treatment include the use of soluble receptor or fragments of the receptor protein that compete for ligand binding. These receptors or fragments can have a higher affinity for the ligand so as to provide effective competition.

Antibodies

The invention also provides antibodies that selectively bind to the 2838, 14618, and 15334 receptor proteins and variants and fragments. An antibody is considered to selectively bind, even if it also binds to other proteins that are not substantially homologous with the receptor protein. These other proteins share homology with a fragment or domain of the receptor protein. This conservation in specific regions gives rise to antibodies that bind to both proteins by virtue of the homologous sequence. In this case, it would be understood that antibody binding to the receptor protein is still selective.

To generate antibodies, an isolated receptor polypeptide is used as an immunogen to generate antibodies using standard techniques for polyclonal and monoclonal antibody preparation. Either the full-length protein or antigenic peptide

WO 00/23588

PCT/US99/24368

fragment can be used. Regions having a high antigenicity index are shown in Figures 2, 6, and 10.

Antibodies are preferably prepared from these regions or from discrete fragments in these regions. However, antibodies can be prepared from any region of the peptide as described herein. A preferred fragment produces an antibody that diminishes or completely prevents ligand-binding. Antibodies can be developed against the entire receptor or portions of the receptor, for example, the intracellular carboxy terminal domain, the amino terminal extracellular domain, the entire transmembrane domain or specific segments, any of the intra or extracellular loops, or any portions of the above. Antibodies may also be developed against specific functional sites, such as the site of ligand-binding, the site of G protein coupling, or sites that are phosphorylated, glycosylated, or myristoylated.

An antigenic 2838 fragment will typically comprise at least 8 contiguous amino acid residues. The antigenic peptide can comprise, however, a contiguous sequence of at least 12, 14 amino acid residues, at least 15 amino acid residues, at least 20 amino acid residues, or at least 30 amino acid residues. In one embodiment, fragments correspond to regions that are located on the surface of the protein, e.g., hydrophilic regions. These fragments are not to be construed, however, as encompassing any fragments which may be disclosed prior to the invention.

An 14618 antigenic fragment will typically comprise at least 9 contiguous amino acid residues. The antigenic peptide can comprise, however, a contiguous sequence of at least 12, 14 amino acid residues, at least 15 amino acid residues, at least 20 amino acid residues, or at least 30 amino acid residues. In one embodiment, fragments correspond to regions that are located on the surface of the protein, e.g., hydrophilic regions. These fragments are not to be construed, however, as encompassing any fragments which may be disclosed prior to the invention.

An antigenic 15334 fragment will typically comprise at least 8 contiguous amino acid residues. The antigenic peptide can comprise, however, a contiguous sequence of at least 12, 14 amino acid residues, at least 15 amino acid residues, at least 20 amino acid residues, or at least 30 amino acid residues. In one embodiment, fragments correspond to regions that are located on the surface of the protein, e.g., hydrophilic regions. These fragments are not to be construed, however, as encompassing any fragments which may be disclosed prior to the invention.

WO 00/23588

PCT/US99/24368

Antibodies can be polyclonal or monoclonal. An intact antibody, or a fragment thereof (e.g. Fab or F(ab')₂) can be used.

- Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various
- 5 enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials
- 10 include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin, an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ¹²⁵I, ¹³¹I, ³⁵S or ³H.
- 15 An appropriate immunogenic preparation can be derived from native, recombinantly expressed, protein or chemically synthesized peptides.

Antibody Uses

- The antibodies can be used to isolate a receptor protein by standard
- 20 techniques, such as affinity chromatography or immunoprecipitation. The antibodies can facilitate the purification of the natural receptor protein from cells and recombinantly produced receptor protein expressed in host cells.

- The antibodies are useful to detect the presence of receptor protein in cells or tissues to determine the pattern of expression of the receptor among various tissues in
- 25 an organism and over the course of normal development.

The antibodies can be used to detect receptor protein *in situ*, *in vitro*, or in a cell lysate or supernatant in order to evaluate the abundance and pattern of expression.

The antibodies can be used to assess abnormal tissue distribution or abnormal expression during development.

- 30 Antibody detection of circulating fragments of the full length receptor protein can be used to identify receptor turnover.

Further, the antibodies can be used to assess receptor expression in disease states such as in active stages of the disease or in an individual with a predisposition toward disease related to receptor function. When a disorder is caused by an

WO 00/23588

PCT/US99/24368

inappropriate tissue distribution, developmental expression, or level of expression of the receptor protein, the antibody can be prepared against the normal receptor protein. If a disorder is characterized by a specific mutation in the receptor protein, antibodies specific for this mutant protein can be used to assay for the presence of the specific mutant receptor protein.

The antibodies can also be used to assess normal and aberrant subcellular localization of cells in the various tissues in an organism. Antibodies can be developed against the whole receptor or portions of the receptor, for example, portions of the amino terminal extracellular domain or extracellular loops.

The diagnostic uses can be applied, not only in genetic testing, but also in monitoring a treatment modality. Accordingly, where treatment is ultimately aimed at correcting receptor expression level or the presence of aberrant receptors and aberrant tissue distribution or developmental expression, antibodies directed against the receptor or relevant fragments can be used to monitor therapeutic efficacy.

Antibodies accordingly can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen.

Additionally, antibodies are useful in pharmacogenomic analysis. Thus, antibodies prepared against polymorphic receptor proteins can be used to identify individuals that require modified treatment modalities.

The antibodies are also useful as diagnostic tools as an immunological marker for aberrant receptor protein analyzed by electrophoretic mobility, isoelectric point, tryptic peptide digest, and other physical assays known to those in the art.

The antibodies are also useful for tissue typing. Thus, where a specific receptor protein has been correlated with expression in a specific tissue, antibodies that are specific for this receptor protein can be used to identify a tissue type.

The antibodies are also useful in forensic identification. Accordingly, where an individual has been correlated with a specific genetic polymorphism resulting in a specific polymorphic protein, an antibody specific for the polymorphic protein can be used as an aid in identification.

The antibodies are also useful for inhibiting receptor function, for example, blocking ligand binding.

These uses can also be applied in a therapeutic context in which treatment involves inhibiting receptor function. An antibody can be used, for example, to block

WO 00/23588

PCT/US99/24368

ligand binding. Antibodies can be prepared against specific fragments containing sites required for function or against intact receptor associated with a cell.

Completely human antibodies are particularly desirable for therapeutic treatment of human patients. For an overview of this technology for producing

5 human antibodies, see Lonberg and Huszar (1995, *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, e.g., U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806.

10 The invention also encompasses kits for using antibodies to detect the presence of a receptor protein in a biological sample. The kit can comprise antibodies such as a labeled or labelable antibody and a compound or agent for detecting receptor protein in a biological sample; means for determining the amount of receptor protein in the sample; and means for comparing the amount of receptor protein in the
15 sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect receptor protein.

Polynucleotides

20 The nucleotide sequence in SEQ ID NO 2 was obtained by sequencing the human full length cDNA.

The nucleotide sequence in SEQ ID NO 4 was obtained by sequencing the human full length cDNA.

25 The nucleotide sequence in SEQ ID NO 6 was obtained by sequencing the human full length cDNA.

The specifically disclosed cDNAs comprise the coding region and 5' and 3' untranslated sequences (SEQ ID NO 2, SEQ ID NO 4, and SEQ ID NO 6).

The human 2838 receptor cDNA is approximately 1617 nucleotides in length and encodes a full length protein that is approximately 319 amino acid residues in
30 length. The nucleic acid is expressed in spleen, testes, colon, and peripheral blood lymphocytes as well as those tissues shown in Figure 15. Structural analysis of the amino acid sequence of SEQ ID NO.1 is provided in Figure 3, a hydropathy plot. The figure shows the putative structure of the seven transmembrane segments, the amino terminal extracellular domain and the carboxy terminal intracellular domain.

WO 00/23588

PCT/US99/24368

The human 14618 receptor cDNA is approximately 1358 nucleotides in length and encodes a full length protein that is approximately 337 amino acid residues in length. The nucleic acid is expressed in spleen, peripheral blood lymphocytes, breast, skeletal muscle, and other tissues shown in Figure 13, as well as various hematopoietic cells shown in Figure 14, especially megakaryocytes and CD34⁺ cells. Structural analysis of the amino acid sequence of SEQ ID NO:3 is provided in Figure 7, a hydropathy plot. The figure shows the putative structure of the seven transmembrane segments, the amino terminal extracellular domain and the carboxy terminal intracellular domain.

The human 153340 receptor cDNA is approximately 2559 nucleotides in length and encodes a full length protein that is approximately 372 amino acid residues in length. The nucleic acid is expressed in lymphnode, tonsils, pancreas, colon, spleen, peripheral blood cells, thymus, adrenal gland and heart as well as cells of the megakaryocytic and erythroid lineages as shown in Figures 17-19. Structural analysis of the amino acid sequence of SEQ ID NO:5 is provided in Figure 11, a hydropathy plot. The figure shows the putative structure of the seven transmembrane segments, the amino terminal extracellular domain and the carboxy terminal intracellular domain.

As used herein, the term "transmembrane segment" refers to a structural amino acid motif which includes a hydrophobic helix that spans the plasma membrane. The entire transmembrane domain of 2838 spans from about amino acid 25 to about amino acid 292. The entire transmembrane domain of 14618 spans from about amino acid 29 to about amino acid 297. The entire transmembrane domain of 15334 spans from about amino acid 26 to about amino acid 299. Seven segments span the membrane and there are three intracellular and three extracellular loops in this domain.

The invention provides isolated polynucleotides encoding a 2838 receptor protein. The term "2838 polynucleotide" or "2838 nucleic acid" refers to the sequence shown in SEQ ID NO:2.

The invention provides isolated polynucleotides encoding a 14618 receptor protein. The term "14618 polynucleotide" or "14618 nucleic acid" refers to the sequence shown in SEQ ID NO:4.

The invention provides isolated polynucleotides encoding a 15334 receptor protein. The term "15334 polynucleotide" or "15334 nucleic acid" refers to the sequence shown in SEQ ID NO:6.

WO 00/23588

PCT/US99/24368

The term "receptor polynucleotide" or "receptor nucleic acid" further includes variants and fragments of the 2838, 14618, and 15334 polynucleotides.

An "isolated" receptor nucleic acid is one that is separated from other nucleic acid present in the natural source of the receptor nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. However, there can be some flanking nucleotide sequences, for example up to about 5KB. The important point is that the nucleic acid is isolated from flanking sequences such that it can be subjected to the specific manipulations described herein such as recombinant expression, preparation of probes and primers, and other uses specific to the receptor nucleic acid sequences.

Moreover, an "isolated" nucleic acid molecule, such as a cDNA or RNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. However, the nucleic acid molecule can be fused to other coding or regulatory sequences and still be considered isolated.

For example, recombinant DNA molecules contained in a vector are considered isolated. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the isolated DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

In some instances, the isolated material will form part of a composition (for example, a crude extract containing other substances), buffer system or reagent mix. In other circumstances, the material may be purified to essential homogeneity, for example as determined by PAGE or column chromatography such as HPLC. Preferably, an isolated nucleic acid comprises at least about 50, 80 or 90 % (on a molar basis) of all macromolecular species present.

The receptor polynucleotides can encode the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor

WO 00/23588

PCT/US99/24368

to a mature form, facilitate protein trafficking, prolong or shorten protein half-life or facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in situ*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

5 The receptor polynucleotides include, but are not limited to, the sequence encoding the mature polypeptide alone, the sequence encoding the mature polypeptide and additional coding sequences, such as a leader or secretory sequence (e.g., a pre-pro or pro-protein sequence), the sequence encoding the mature polypeptide, with or without the additional coding sequences, plus additional non-coding sequences, for
10 example introns and non-coding 5' and 3' sequences such as transcribed but non-translated sequences that play a role in transcription, mRNA processing (including splicing and polyadenylation signals), ribosome binding and stability of mRNA. In addition, the polynucleotide may be fused to a marker sequence encoding, for example, a peptide that facilitates purification.

15 Receptor polynucleotides can be in the form of RNA, such as mRNA, or in the form DNA, including cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The nucleic acid, especially DNA, can be double-stranded or single-stranded. Single-stranded nucleic acid can be the coding strand (sense strand) or the non-coding strand (anti-sense
20 strand).

One receptor nucleic acid comprises the nucleotide sequence shown in SEQ ID NO:2, corresponding to human 2838 cDNA.

One receptor nucleic acid comprises the nucleotide sequence shown in SEQ ID NO:4, corresponding to human 14618 cDNA.

25 One receptor nucleic acid comprises the nucleotide sequence shown in SEQ ID NO:6, corresponding to human 15334 cDNA.

In one embodiment, the receptor nucleic acid comprises only the coding region.

The invention further provides variant receptor polynucleotides, and fragments
30 thereof, that differ from the nucleotide sequence shown in SEQ ID NO:2 due to degeneracy of the genetic code and thus encode the same protein as that encoded by the nucleotide sequence shown in SEQ ID NO:2.

The invention further provides variant receptor polynucleotides, and fragments thereof, that differ from the nucleotide sequence shown in SEQ ID NO:4 due to

WO 00/23588

PCT/US99/24368

degeneracy of the genetic code and thus encode the same protein as that encoded by the nucleotide sequence shown in SEQ ID NO 4.

The invention further provides variant receptor polynucleotides, and fragments thereof, that differ from the nucleotide sequence shown in SEQ ID NO 6 due to degeneracy of the genetic code and thus encode the same protein as that encoded by the nucleotide sequence shown in SEQ ID NO 6.

The invention also provides receptor nucleic acid molecules encoding the variant polypeptides described herein. Such polynucleotides may be naturally occurring, such as allelic variants (same locus), homologs (different locus), and orthologs (different organism), or may be constructed by recombinant DNA methods or by chemical synthesis. Such non-naturally occurring variants may be made by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. Accordingly, as discussed above, the variants can contain nucleotide substitutions, deletions, inversions and insertions.

Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions.

Typically, variants have a substantial identity with a nucleic acid molecule selected from the group consisting of SEQ ID NOS: 2, 4, and 6 and its complements.

Orthologs, homologs, and allelic variants can be identified using methods well known in the art. 2838 variants comprise a nucleotide sequence encoding a receptor that is 40-45%, 45-50%, 50-55%, 55-60%, at least about 65%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous to the nucleotide sequence shown in SEQ ID NO 2 or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions, to the nucleotide sequence shown in SEQ ID NO 2 or a fragment of the sequence.

14618 variants comprise a nucleotide sequence encoding a receptor that is 40-45%, 45-50%, 50-55%, 55-60%, at least about 65%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous to the nucleotide sequence shown in SEQ ID NO 4 or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions, to the nucleotide sequence shown in SEQ ID NO 4 or a fragment of the sequence.

WO 00/23588

PCT/US99/24368

15334 variants comprise a nucleotide sequence encoding a receptor that is 45-50%, 50-55%, 55-60%, at least about 65%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous to the nucleotide sequence shown in SEQ ID NO:6 or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions, to the nucleotide sequence shown in SEQ ID NO:6 or a fragment of the sequence.

It is understood that stringent hybridization does not indicate substantial homology where it is due to general homology, such as poly A sequences, or sequences common to all or most proteins, all GPCRs, or all family I GPCRs or even all purinoceptors. Moreover, it is understood that variants do not include any of the nucleic acid sequences that may have been disclosed prior to the invention.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences encoding a receptor polypeptide at least 50-55%, 55% homologous to each other typically remain hybridized to each other. The conditions can be such that sequences at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 90%, at least about 95% or more identical to each other remain hybridized to one another. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6, incorporated by reference. One example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. In another non-limiting example, nucleic acid molecules are allowed to hybridize in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more low stringency washes in 0.2X SSC/0.1% SDS at room temperature, or by one or more moderate stringency washes in 0.2X SSC/0.1% SDS at 42°C, or washed in 0.2X SSC/0.1% SDS at 65°C for high stringency. In one embodiment, an isolated receptor nucleic acid molecule that hybridizes under stringent conditions to the sequence of SEQ ID NOS 2, 4, and 6 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

WO 00/23588

PCT/US99/24369

As understood by those of ordinary skill, the exact conditions can be determined empirically and depend on ionic strength, temperature and the concentration of destabilizing agents such as formamide or denaturing agents such as SDS. Other factors considered in determining the desired hybridization conditions include the length of the nucleic acid sequences, base composition, percent mismatch between the hybridizing sequences and the frequency of occurrence of subsets of the sequences within other non-identical sequences. Thus, equivalent conditions can be determined by varying one or more of these parameters while maintaining a similar degree of identity or similarity between the two nucleic acid molecules.

The present invention also provides isolated nucleic acids that contain a single or double stranded fragment or portion that hybridizes under stringent conditions to a nucleotide sequence selected from the group consisting of SEQ ID NOS: 2, 4, and 6 and the complements of SEQ ID NOS: 2, 4, and 6. In one embodiment, the nucleic acid consists of a portion of a nucleotide sequence selected from the group consisting of SEQ ID NOS: 2, 4, and 6 and the complements. The nucleic acid fragments of the invention are at least about 15, preferably at least about 18, 20, 23 or 25 nucleotides, and can be 30, 40, 50, 100, 200 or more nucleotides in length. Longer fragments, for example, 30 or more nucleotides in length, which encode antigenic proteins or polypeptides described herein are useful.

Furthermore, the invention provides polynucleotides that comprise a fragment of the full length receptor polynucleotides. The fragment can be single or double stranded and can comprise DNA or RNA. The fragment can be derived from either the coding or the non-coding sequence.

In one embodiment, an isolated 2838 receptor nucleic acid from nucleotide 1 to around nucleotide 990 is at least 16 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:2. In another embodiment, the nucleic acid from around nucleotide 1487-1617 is at least 20 nucleotides. In other embodiments, the nucleic acid is at least 40, 50, 100, 250 or 500 nucleotides in length or greater.

In another embodiment, an isolated 14168 receptor nucleic acid from around nucleotide 1 to around nucleotide 911 is at least 8 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:4. In other embodiments, the nucleic acid is at least 40, 50, 100, 250, or 500 nucleotides in length or greater.

WO 00/23588

PCT/US99/24368

In another embodiment, an isolated 15334 receptor nucleic acid from nucleotide 1 to around nucleotide 1355 is at least 18 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:6. In another embodiment, the nucleic acid from
 5 around nucleotide 868 to around 1355 is at least 11 nucleotides. In other embodiments, the nucleic acid is at least 40, 50, 100, 250, or 500 nucleotides in length or greater.

In another embodiment, an isolated 2838 receptor nucleic acid encodes the entire coding region from amino acid 1 to amino acid 319. In another embodiment,
 10 the isolated 2838 receptor nucleic acid encodes a sequence corresponding to the mature protein from about amino acid 6 to about amino acid 319. In another embodiment, an isolated 14618 receptor nucleic acid encodes the entire coding region from amino acid 1 to amino acid 337. In another embodiment, the isolated 14618 receptor nucleic acid encodes a sequence corresponding to the mature protein from
 15 about amino acid 6 to about amino acid 337. In another embodiment, an isolated 15334 receptor nucleic acid encodes the entire coding region from amino acid 1 to amino acid 372. In another embodiment, the isolated 15334 receptor nucleic acid encodes a sequence corresponding to the mature protein from about amino acid 6 to about amino acid 372.

20 Other fragments of all three receptors include nucleotide sequences encoding the amino acid fragments described herein. Further, fragments can include subfragments of the specific domains or sites described herein. Fragments also include nucleic acid sequences corresponding to specific amino acid sequences described above or fragments thereof. Nucleic acid fragments, according to the
 25 present invention, are not to be construed as encompassing those fragments that may have been disclosed prior to the invention.

However, it is understood that a receptor fragment includes any nucleic acid sequence that does not include the entire gene.

2838 receptor nucleic acid fragments further include sequences corresponding
 30 to the domains described herein, subregions also described, and specific functional sites. 2838 receptor nucleic acid fragments include nucleic acid molecules encoding a polypeptide comprising the amino terminal extracellular domain including amino acid residues from 1 to about 24, a polypeptide comprising the region spanning the transmembrane domain (amino acid residues from about 25 to about 292), a

WO 99/23588

PCT/US99/24368

polypeptide comprising the carboxy terminal intracellular domain (amino acid residues from about 293 to about 319), and a polypeptide encoding the G-protein receptor signature (118-120 or surrounding amino acid residues from about 107 to about 123), nucleic acid molecules encoding any of the seven transmembrane segments, extracellular or intracellular loops, glycosylation sites, phosphorylation sites, myristoylation sites, and amidation site.

14618 receptor nucleic acid fragments include nucleic acid molecules encoding a polypeptide comprising the amino terminal extracellular domain including amino acid residues from 1 to about 28, a polypeptide comprising the region spanning the transmembrane domain (amino acid residues from about 29 to about 297), a polypeptide comprising the carboxy terminal intracellular domain (amino acid residues from about 298 to about 337), and a polypeptide encoding the G-protein receptor signature (120-122 or surrounding amino acid residues from about 110 to about 132), nucleic acid molecules encoding any of the seven transmembrane segments, extracellular or intracellular loops, glycosylation sites and phosphorylation sites.

15334 receptor nucleic acid fragments include nucleic acid molecules encoding a polypeptide comprising the amino terminal extracellular domain including amino acid residues from 1 to about 25, a polypeptide comprising the region spanning the transmembrane domain (amino acid residues from about 26 to about 299), a polypeptide comprising the carboxy terminal intracellular domain (amino acid residues from about 300 to about 372), and a polypeptide encoding the G-protein receptor signature (118-120 or surrounding amino acid residues from about 110 to about 130), nucleic acid molecules encoding any of the seven transmembrane segments, extracellular or intracellular loops, glycosylation sites, protein kinase C, cAMP, cGMP, and casein kinase II phosphorylation sites, and myristoylation sites.

Where the location of the domains have been predicted by computer analysis, one of ordinary skill would appreciate that the amino acid residues constituting these domains can vary depending on the criteria used to define the domains.

Receptor nucleic acid fragments also include combinations of the domains, segments, loops, and other functional sites described above. Thus, for example, a receptor nucleic acid could include sequences corresponding to the amino terminal extracellular domain and one transmembrane segment. A person of ordinary skill in the art would be aware of the many permutations that are possible.

WO 00/23588

PCT/US99/24368

Where the location of the domains or sites have been predicted by computer analysis, one of ordinary skill would appreciate that the amino acid residues constituting these domains can vary depending on the criteria used to define the domains.

The invention also provides receptor nucleic acid fragments that encode epitope bearing regions of the receptor proteins described herein.

The isolated receptor polynucleotide sequences, and especially fragments, are useful as DNA probes and primers.

For example, the coding region of a receptor gene can be isolated using the known nucleotide sequence to synthesize an oligonucleotide probe. A labeled probe can then be used to screen a cDNA library, genomic DNA library, or mRNA to isolate nucleic acid corresponding to the coding region. Further, primers can be used in PCR reactions to clone specific regions of receptor genes.

A probe/primer typically comprises substantially purified oligonucleotide. The 2838 oligonucleotide typically comprises a region of nucleotide sequence from 1-990 and from 1487-1617 that hybridizes under stringent conditions to at least about 16 and 20, respectively, typically about 25, more typically about 40, 50 or 75 consecutive nucleotides of SEQ ID NO:2 sense or anti-sense strand or other receptor polynucleotides. The 14618 oligonucleotide typically comprises a region of nucleotide sequence from 1-911 that hybridizes under stringent conditions to at least about 9, 12, typically about 25, more typically about 40, 50 or 75 consecutive nucleotides of SEQ ID NO:4 sense or anti-sense strand or other receptor polynucleotides. The 15334 oligonucleotide typically comprises a region of nucleotide sequence from 868-1355 that hybridizes under stringent conditions to at least about 11, 12, typically about 25, more typically about 40, 50 or 75 consecutive nucleotides of SEQ ID NO:6 sense or anti-sense strand or other receptor polynucleotides. The 15334 oligonucleotide also typically comprises a region of nucleotide sequence from nucleotide 1-1355 that hybridizes under stringent conditions to at least about 18, typically about 25, more typically about 40, 50, or 75 consecutive nucleotides of SEQ ID NO:6 sense or anti-sense strand or other receptor polynucleotides.

Polynucleotide Uses

The nucleic acid sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example,

WO 00/23588

PCT/US99/24368

- identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences
- 5 homologous to the nucleic acid molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.
- 10 The nucleic acid fragments of the invention provide probes or primers in assays such as those described below. "Probes" are oligonucleotides that hybridize in a base-specific manner to a complementary strand of nucleic acid. Such probes include polypeptide nucleic acids, as described in Nielsen *et al.* (1991) *Science* 254:1497-1500. Typically, a probe comprises a region of nucleotide sequence that
- 15 hybridizes under highly stringent conditions to at least about 15, typically about 20-25, and more typically about 40, 50 or 75 consecutive nucleotides of a nucleic acid selected from the group consisting of SEQ ID NOS: 2, 4, and 6 and the complements thereof. More typically, the probe further comprises a label, e.g., radioisotope, fluorescent compound, enzyme, or enzyme co-factor.
- 20 As used herein, the term "primer" refers to a single-stranded oligonucleotide which acts as a point of initiation of template-directed DNA synthesis using well-known methods (e.g., PCR, LCR) including, but not limited to those described herein. The appropriate length of the primer depends on the particular use, but typically ranges from about 15 to 30 nucleotides. The term "primer site" refers to the area of
- 25 the target DNA to which a primer hybridizes. The term "primer pair" refers to a set of primers including a 5' (upstream) primer that hybridizes with the 5' end of the nucleic acid sequence to be amplified and a 3' (downstream) primer that hybridizes with the complement of the sequence to be amplified.
- The receptor polynucleotides are useful for probes, primers, and in biological
- 30 assays.
- Where the polynucleotides are used to assess GPCR properties or functions, such as in the assays described herein, all or less than all of the entire cDNA can be useful. In this case, even fragments that may have been known prior to the invention are encompassed. Thus, for example, assays specifically directed to GPCR functions.

WO 00/23588

PCT/US99/24368

such as assessing agonist or antagonist activity, encompass the use of known fragments. Further, diagnostic methods for assessing receptor function can also be practiced with any fragment, including those fragments that may have been known prior to the invention. Similarly, in methods involving treatment of receptor dysfunction, all fragments are encompassed including those which may have been known in the art

The 2838 receptor polynucleotides are useful as a hybridization probe for cDNA and genomic DNA to isolate a full-length cDNA and genomic clones encoding the polypeptide described in SEQ ID NO:1 and to isolate cDNA and genomic clones that correspond to variants producing the same polypeptide shown in SEQ ID NO:1 or the other variants described herein. Variants can be isolated from the same tissue and organism from which the polypeptide shown in SEQ ID NO:1 was isolated, different tissues from the same organism, or from different organisms.

The 14618 receptor polynucleotides are useful as a hybridization probe for cDNA and genomic DNA to isolate a full-length cDNA and genomic clones encoding the polypeptide described in SEQ ID NO:3 and to isolate cDNA and genomic clones that correspond to variants producing the same polypeptide shown in SEQ ID NO:3 or the other variants described herein. Variants can be isolated from the same tissue and organism from which the polypeptide shown in SEQ ID NO:3 was isolated, different tissues from the same organism, or from different organisms.

The 15334 receptor polynucleotides are useful as a hybridization probe for cDNA and genomic DNA to isolate a full-length cDNA and genomic clones encoding the polypeptide described in SEQ ID NO:5 and to isolate cDNA and genomic clones that correspond to variants producing the same polypeptide shown in SEQ ID NO:5 or the other variants described herein. Variants can be isolated from the same tissue and organism from which the polypeptide shown in SEQ ID NO:5 was isolated, different tissues from the same organism, or from different organisms.

This method is useful for isolating genes and cDNA that are developmentally-controlled and therefore may be expressed in the same tissue or different tissues at different points in the development of an organism.

The probe can correspond to any sequence along the entire length of the gene encoding the receptor. Accordingly, it could be derived from 5' noncoding regions, the coding region, and 3' noncoding regions. Probes, however, are not to be construed as corresponding to any sequences that may be known prior to the invention.

WO 00/23588

PCT/US99/24368

The 2838 nucleic acid probe can be, for example, the full-length cDNA of SEQ ID NO:1, or a fragment thereof, such as an oligonucleotide of at least 11, 16, 18, 20, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to mRNA or DNA. The 14618 nucleic acid probe can be, for example, the full-length cDNA of SEQ ID NO:3, or a fragment thereof, such as an oligonucleotide of at least 8, 12, 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to mRNA or DNA. The 15334 nucleic acid probe can be, for example, the full-length cDNA of SEQ ID NO:5, or a fragment thereof, such as an oligonucleotide of at least 11, 12, 15, 18, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to mRNA or DNA.

Fragments of the polynucleotides described herein are also useful to synthesize larger fragments or full-length polynucleotides described herein. For example, a fragment can be hybridized to any portion of an mRNA and a larger or full-length cDNA can be produced.

The fragments are also useful to synthesize antisense molecules of desired length and sequence.

Antisense nucleic acids of the invention can be designed using the nucleotide sequences of SEQ ID NOS: 2, 4, and 6 constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylquosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxycarbonylmethyl-2-thiouracil, beta-D-mannosylquosine, 5-methoxycarbonylmethyluracil, 5-methoxyuracil, 2-methylthio-

WO 00/23588

PCT/US99/24368

N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-riouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest).

Additionally, the nucleic acid molecules of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4:5). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996), *supra*; Perry-O'Keefe *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:14670. PNAs can be further modified, e.g., to enhance their stability, specificity or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996), *supra*; Finn *et al.* (1996) *Nucleic Acids Res.* 24(17):3357-63; Mag *et al.* (1989) *Nucleic Acids Res.* 17:5973, and Peterser *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5:1119.

The nucleic acid molecules and fragments of the invention can also include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Leisinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. WO 88/0918) or the blood brain barrier (see, e.g., PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (see, e.g., Krol *et al.*

WO 00/23588

PCT/US99/24368

(1988) *Bio-Techniques* 6:958-976) or intercalating agents (see, e.g., Zon (1988) *Pharm Res.* 5:539-549).

The receptor polynucleotides are also useful as primers for PCR to amplify any given region of a receptor polynucleotide.

5 The receptor polynucleotides are also useful for constructing recombinant vectors. Such vectors include expression vectors that express a portion of, or all of, the receptor polypeptides. Vectors also include insertion vectors, used to integrate into another polynucleotide sequence, such as into the cellular genome, to alter *in situ* expression of receptor genes and gene products. For example, an endogenous
10 receptor coding sequence can be replaced via homologous recombination with all or part of the coding region containing one or more specifically introduced mutations.

The receptor polynucleotides are also useful for expressing antigenic peptides. Peptide regions having a high antigenicity index are shown in Figures 2, 6, and 10.

The receptor polynucleotides are also useful as probes for determining the
15 chromosomal positions of the receptor polynucleotides by means of *in situ* hybridization methods, such as FISH (for a review of this technique, see Verma *et al.* (1988) *Human Chromosomes: A Manual of Basic Techniques* (Pergamon Press, New York), and PCR mapping of somatic cell hybrids. The mapping of the sequences to chromosomes is an important first step in correlating these sequences with genes
20 associated with disease.

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to
25 noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, *Mendelian Inheritance
30 in Man*, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland *et al.* (1987) *Nature* 325:783-787.

WO 00/23588

PCT/US99/24368

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with a specified gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

The receptor polynucleotide probes are also useful to determine patterns of the presence of the gene encoding the receptors and their variants with respect to tissue distribution, for example, whether gene duplication has occurred and whether the duplication occurs in all or only a subset of tissues. The genes can be naturally occurring or can have been introduced into a cell, tissue, or organism exogenously.

The receptor polynucleotides are also useful for designing ribozymes corresponding to all, or a part, of the mRNA produced from genes encoding the polynucleotides described herein.

The receptor polynucleotides are also useful for constructing host cells expressing a part, or all, of the receptor polynucleotides and polypeptides.

The receptor polynucleotides are also useful for constructing transgenic animals expressing all, or a part, of the receptor polynucleotides and polypeptides.

The receptor polynucleotides are also useful for making vectors that express part, or all, of the receptor polypeptides.

The receptor polynucleotides are also useful as hybridization probes for determining the level of receptor nucleic acid expression. Accordingly, the probes can be used to detect the presence of, or to determine levels of, receptor nucleic acid in cells, tissues, and in organisms. The nucleic acid whose level is determined can be DNA or RNA. Accordingly, probes corresponding to the polypeptides described herein can be used to assess gene copy number in a given cell, tissue, or organism. This is particularly relevant in cases in which there has been an amplification of the receptor genes.

Alternatively, the probe can be used in an *in situ* hybridization context to assess the position of extra copies of the receptor genes, as on extrachromosomal

WO 00/23588

PCT/US99/24368

elements or as integrated into chromosomes in which the receptor gene is not normally found, for example as a homogeneously staining region.

These uses are relevant for diagnosis of disorders involving an increase or decrease in receptor expression relative to normal results, such as a proliferative disorder, a differentiative or developmental disorder, or a hematopoietic disorder.

Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant expression or activity of receptor nucleic acid, in which a test sample is obtained from a subject and nucleic acid (e.g., mRNA, genomic DNA) is detected, wherein the presence of the nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant expression or activity of the nucleic acid.

One aspect of the invention relates to diagnostic assays for determining nucleic acid expression as well as activity in the context of a biological sample (e.g., blood, serum, cells, tissue) to determine whether an individual has a disease or disorder, or is at risk of developing a disease or disorder, associated with aberrant nucleic acid expression or activity. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with expression or activity of the nucleic acid molecules.

In vitro techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detecting DNA includes Southern hybridizations and *in situ* hybridization.

Probes can be used as a part of a diagnostic test kit for identifying cells or tissues that express a receptor protein, such as by measuring a level of a receptor-encoding nucleic acid in a sample of cells from a subject e.g., mRNA or genomic DNA, or determining if a receptor gene has been mutated.

Nucleic acid expression assays are useful for drug screening to identify compounds that modulate receptor nucleic acid expression (e.g., antisense, polypeptides, peptidomimetics, small molecules or other drugs). A cell is contacted with a candidate compound and the expression of mRNA determined. The level of expression of receptor mRNA in the presence of the candidate compound is compared to the level of expression of receptor mRNA in the absence of the candidate compound. The candidate compound can then be identified as a modulator of nucleic acid expression based on this comparison and be used, for example to treat a disorder.

WO 00/23588

PCT/US99/24368

characterized by aberrant nucleic acid expression. The modulator can bind to the nucleic acid or indirectly modulate expression, such as by interacting with other cellular components that affect nucleic acid expression.

Modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject) in patients or in transgenic animals.

The invention thus provides a method for identifying a compound that can be used to treat a disorder associated with nucleic acid expression of the receptor gene. The method typically includes assaying the ability of the compound to modulate the expression of the receptor nucleic acid and thus identifying a compound that can be used to treat a disorder characterized by undesired receptor nucleic acid expression.

The assays can be performed in cell-based and cell-free systems. Cell-based assays include cells naturally expressing the receptor nucleic acid or recombinant cells genetically engineered to express specific nucleic acid sequences.

Alternatively, candidate compounds can be assayed *in vivo* in patients or in transgenic animals.

The assay for receptor nucleic acid expression can involve direct assay of nucleic acid levels, such as mRNA levels, or on collateral compounds involved in the signal pathway (such as cyclic AMP or phosphatidylinositol turnover). Further, the expression of genes that are up- or down-regulated in response to the receptor protein signal pathway can also be assayed. In this embodiment the regulatory regions of these genes can be operably linked to a reporter gene such as luciferase.

Thus, modulators of receptor gene expression can be identified in a method wherein a cell is contacted with a candidate compound and the expression of mRNA determined. The level of expression of receptor mRNA in the presence of the candidate compound is compared to the level of expression of receptor mRNA in the absence of the candidate compound. The candidate compound can then be identified as a modulator of nucleic acid expression based on this comparison and be used, for example to treat a disorder characterized by aberrant nucleic acid expression. When expression of mRNA is statistically significantly greater in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of nucleic acid expression. When nucleic acid expression is statistically significantly less in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of nucleic acid expression.

WO 00/23588

PCT/US99/24368

Accordingly, the invention provides methods of treatment, with the nucleic acid as a target, using a compound identified through drug screening as a gene modulator to modulate receptor nucleic acid expression. Modulation includes both up-regulation (*i.e.* activation or agonization) or down-regulation (suppression or antagonization) or effects on nucleic acid activity (e.g. when nucleic acid is mutated or improperly modified). Treatment is of disorders characterized by aberrant expression or activity of the nucleic acid.

Alternatively, a modulator for receptor nucleic acid expression can be a small molecule or drug identified using the screening assays described herein as long as the drug or small molecule inhibits the receptor nucleic acid expression.

The receptor polynucleotides are also useful for monitoring the effectiveness of modulating compounds on the expression or activity of the receptor gene in clinical trials or in a treatment regimen. Thus, the gene expression pattern can serve as a barometer for the continuing effectiveness of treatment with the compound, particularly with compounds to which a patient can develop resistance. The gene expression pattern can also serve as a marker indicative of a physiological response of the affected cells to the compound. Accordingly, such monitoring would allow either increased administration of the compound or the administration of alternative compounds to which the patient has not become resistant. Similarly, if the level of nucleic acid expression falls below a desirable level, administration of the compound could be commensurately decreased.

Monitoring can be, for example, as follows: (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a specified mRNA or genomic DNA of the invention in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the mRNA or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the mRNA or genomic DNA in the pre-administration sample with the mRNA or genomic DNA in the post-administration sample or samples; and (vi) increasing or decreasing the administration of the agent to the subject accordingly.

The receptor polynucleotides are also useful in diagnostic assays for qualitative changes in receptor nucleic acid, and particularly in qualitative changes that lead to pathology. The polynucleotides can be used to detect mutations in receptor genes and gene expression products such as mRNA. The polynucleotides

WO 00/23588

PCT/US99/24368

can be used as hybridization probes to detect naturally-occurring genetic mutations in the receptor gene and thereby to determine whether a subject with the mutation is at risk for a disorder caused by the mutation. Mutations include deletion, addition, or substitution of one or more nucleotides in the gene, chromosomal rearrangement, such as inversion or transposition, modification of genomic DNA, such as aberrant methylation patterns or changes in gene copy number, such as amplification. Detection of a mutated form of the receptor gene associated with a dysfunction provides a diagnostic tool for an active disease or susceptibility to disease when the disease results from overexpression, underexpression, or altered expression of a receptor protein.

Mutations in the receptor gene can be detected at the nucleic acid level by a variety of techniques. Genomic DNA can be analyzed directly or can be amplified by using PCR prior to analysis. RNA or cDNA can be used in the same way.

In certain embodiments, detection of the mutation involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.*, (1988) *Science* 241:1077-1080, and Nakazawa *et al.*, (1994) *PNAS* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the gene (see Abravaya *et al.* (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a gene under conditions such that hybridization and amplification of the gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. Deletions and insertions can be detected by a change in size of the amplified product compared to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to normal RNA or antisense DNA sequences.

It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional

WO 00/23588

PCT/US99/24368

amplification system (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al.* (1988) *BioTechnology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well-known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

Alternatively, mutations in a receptor gene can be directly identified, for example, by alterations in restriction enzyme digestion patterns determined by gel electrophoresis.

Further, sequence-specific ribozymes (U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

Perfectly matched sequences can be distinguished from mismatched sequences by nuclease cleavage digestion assays or by differences in melting temperature.

Sequence changes at specific locations can also be assessed by nuclease protection assays such as RNase and S1 protection or the chemical cleavage method.

Furthermore, sequence differences between a mutant receptor gene and a wild-type gene can be determined by direct DNA sequencing. A variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, e.g., PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA duplexes (Myers *et al.* (1985) *Science* 230:1242; Cotton *et al.* (1988) *PNAS* 85:4397; Saleeba *et al.* (1992) *Meth. Enzymol.* 217:286-295), electrophoretic mobility of mutant and wild type nucleic acid is compared (Orita *et al.* (1989) *PNAS* 86:2766; Cotton *et al.* (1993) *Mutat. Res.* 285:125-144, and Hayashi *et al.* (1992) *Genet. Anal. Tech. Appl.* 9:73-79), and movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (Myers *et al.* (1985) *Nature* 313:495). The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex

WO 00/23588

PCT/US99/24368

molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet.* 7:5). Examples of other techniques for detecting point mutations include, selective oligonucleotide hybridization, selective amplification, and selective primer extension.

5 In other embodiments, genetic mutations can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotide probes (Cronin *et al.* (1996) *Human Mutation* 7:244-255; Kozal *et al.* (1996) *Nature Medicine* 2:753-759). For example, genetic mutations can be identified in two dimensional arrays containing

10 light-generated DNA probes as described in Cronin *et al. supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the

15 characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

The receptor polynucleotides are also useful for testing an individual for a

20 genotype that while not necessarily causing the disease, nevertheless affects the treatment modality. Thus, the polynucleotides can be used to study the relationship between an individual's genotype and the individual's response to a compound used for treatment (pharmacogenomic relationship). In the present case, for example, a mutation in the receptor gene that results in altered affinity for ligand could result in

25 an excessive or decreased drug effect with standard concentrations of ligand that activates the receptor. Accordingly, the receptor polynucleotides described herein can be used to assess the mutation content of the receptor gene in an individual in order to select an appropriate compound or dosage regimen for treatment.

Thus polynucleotides displaying genetic variations that affect treatment

30 provide a diagnostic target that can be used to tailor treatment in an individual. Accordingly, the production of recombinant cells and animals containing these polymorphisms allow effective clinical design of treatment compounds and dosage regimens.

WO 00/23588

PCT/US99/24368

The methods can involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting mRNA, or genomic DNA, such that the presence of mRNA or genomic DNA is detected in the biological sample, and comparing the presence of mRNA or genomic DNA in the control sample with the presence of mRNA or genomic DNA in the test sample

The receptor polynucleotides are also useful for chromosome identification when the sequence is identified with an individual chromosome and to a particular location on the chromosome. First, the DNA sequence is matched to the chromosome by *in situ* or other chromosome-specific hybridization. Sequences can also be correlated to specific chromosomes by preparing PCR primers that can be used for PCR screening of somatic cell hybrids containing individual chromosomes from the desired species. Only hybrids containing the chromosome containing the gene homologous to the primer will yield an amplified fragment. Sublocalization can be achieved using chromosomal fragments. Other strategies include prescreening with labeled flow-sorted chromosomes and preselection by hybridization to chromosome-specific libraries. Further mapping strategies include fluorescence *in situ* hybridization which allows hybridization with probes shorter than those traditionally used. Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on the chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

The receptor polynucleotides can also be used to identify individuals from small biological samples. This can be done for example using restriction fragment-length polymorphism (RFLP) to identify an individual. Thus, the polynucleotides described herein are useful as DNA markers for RFLP (See U.S. Patent No. 5,272,057).

Furthermore, the receptor sequence can be used to provide an alternative technique which determines the actual DNA sequence of selected fragments in the genome of an individual. Thus, the receptor sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify DNA from an individual for subsequent sequencing.

WO 00/23588

PCT/US99/24368

Panels of corresponding DNA sequences from individuals prepared in this manner can provide unique individual identifications, as each individual will have a unique set of such DNA sequences. It is estimated that allelic variation in humans occurs with a frequency of about once per each 500 bases. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. The receptor sequences can be used to obtain such identification sequences from individuals and from tissue. The sequences represent unique fragments of the human genome. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes.

If a panel of reagents from the sequences is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

The receptor polynucleotides can also be used in forensic identification procedures. PCR technology can be used to amplify DNA sequences taken from very small biological samples, such as a single hair follicle, body fluids (e.g. blood, saliva, or semen). The amplified sequence can then be compared to a standard allowing identification of the origin of the sample.

The receptor polynucleotides can thus be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e. another DNA sequence that is unique to a particular individual). As described above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to the noncoding region are particularly useful since greater polymorphism occurs in the noncoding regions, making it easier to differentiate individuals using this technique. Fragments are at least 12 bases.

The receptor polynucleotides can further be used to provide polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue. This is useful in cases in

WO 00/23588

PCT/US99/24368

which a forensic pathologist is presented with a tissue of unknown origin. Panels of receptor probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these primers and probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

Alternatively, the receptor polynucleotides can be used directly to block transcription or translation of receptor gene sequences by means of antisense or ribozyme constructs. Thus, in a disorder characterized by abnormally high or undesirable receptor gene expression, nucleic acids can be directly used for treatment.

The receptor polynucleotides are thus useful as antisense constructs to control receptor gene expression in cells, tissues, and organisms. A DNA antisense polynucleotide is designed to be complementary to a region of the gene involved in transcription, preventing transcription and hence production of receptor protein. An antisense RNA or DNA polynucleotide would hybridize to the mRNA and thus block translation of mRNA into receptor protein.

Examples of antisense molecules useful to inhibit nucleic acid expression include antisense molecules complementary to a fragment of the 5' untranslated region of SEQ ID NOS 2, 4, or 6 which also includes the start codon and antisense molecules which are complementary to a fragment of the 3' untranslated region of SEQ ID NOS 2, 4, or 6.

Alternatively, a class of antisense molecules can be used to inactivate mRNA in order to decrease expression of receptor nucleic acid. Accordingly, these molecules can treat a disorder characterized by abnormal or undesired receptor nucleic acid expression. This technique involves cleavage by means of ribozymes containing nucleotide sequences complementary to one or more regions in the mRNA that attenuate the ability of the mRNA to be translated. Possible regions include coding regions and particularly coding regions corresponding to the catalytic and other functional activities of the receptor protein, such as ligand binding.

The receptor polynucleotides also provide vectors for gene therapy in patients containing cells that are aberrant in receptor gene expression. Thus, recombinant cells, which include the patient's cells that have been engineered *ex vivo* and returned to the patient, are introduced into an individual where the cells produce the desired receptor protein to treat the individual.

WO 00/23588

PCT/US99/24368

The invention also encompasses kits for detecting the presence of a receptor nucleic acid in a biological sample. For example, the kit can comprise reagents such as a labeled or labelable nucleic acid or agent capable of detecting receptor nucleic acid in a biological sample, means for determining the amount of receptor nucleic acid in the sample; and means for comparing the amount of receptor nucleic acid in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect receptor mRNA or DNA.

10 Computer Readable Means

The nucleotide or amino acid sequences of the invention are also provided in a variety of mediums to facilitate use thereof. As used herein, "provided" refers to a manufacture, other than an isolated nucleic acid or amino acid molecule, which contains a nucleotide or amino acid sequence of the present invention. Such a manufacture provides the nucleotide or amino acid sequences, or a subset thereof (e.g., a subset of open reading frames (ORFs)) in a form which allows a skilled artisan to examine the manufacture using means not directly applicable to examining the nucleotide or amino acid sequences, or a subset thereof, as they exist in nature or in purified form.

In one application of this embodiment, a nucleotide or amino acid sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium that can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. The skilled artisan will readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide or amino acid sequence of the present invention.

As used herein, "recorded" refers to a process for storing information on computer readable medium. The skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide or amino acid sequence information of the present invention.

WO 00/23588

PCT/US99/24368

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide or amino acid sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. The skilled artisan can readily adapt any number of data processor structuring formats (e.g., text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing the nucleotide or amino acid sequences of the invention in computer readable form, the skilled artisan can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the nucleotide or amino acid sequences of the invention in computer readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

As used herein, a "target sequence" can be any DNA or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences.

WO 00/23588

PCT/US99/24368

Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium for analysis and comparison to other sequences. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, MacPattern (EMBL), BLASTN and BLASTX (NCBI).

For example, software which implements the BLAST (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410) and BLAZE (Brutlag *et al.* (1993) *Comp. Chem.* 17:203-207) search algorithms on a Sybase system can be used to identify open reading frames (ORFs) of the sequences of the invention which contain homology to ORFs or proteins from other libraries. Such ORFs are protein encoding fragments and are useful in producing commercially important proteins such as enzymes used in various reactions and in the production of commercially useful metabolites.

Vectors/host cells

The invention also provides vectors containing the receptor polynucleotides. The term "vector" refers to a vehicle, preferably a nucleic acid molecule, that can transport the receptor polynucleotides. When the vector is a nucleic acid molecule, the receptor polynucleotides are covalently linked to the vector nucleic acid. With this aspect of the invention, the vector includes a plasmid, single or double stranded phage, a single or double stranded RNA or DNA viral vector, or artificial chromosome, such as a BAC, PAC, YAC, OR MAC.

A vector can be maintained in the host cell as an extrachromosomal element where it replicates and produces additional copies of the receptor polynucleotides. Alternatively, the vector may integrate into the host cell genome and produce additional copies of the receptor polynucleotides when the host cell replicates.

The invention provides vectors for the maintenance (cloning vectors) or vectors for expression (expression vectors) of the receptor polynucleotides. The vectors can function in procaryotic or eukaryotic cells or in both (shuttle vectors).

Expression vectors contain cis-acting regulatory regions that are operably linked in the vector to the receptor polynucleotides such that transcription of the

WO 00/23588

PCT/US99/24368

polynucleotides is allowed in a host cell. The polynucleotides can be introduced into the host cell with a separate polynucleotide capable of affecting transcription. Thus, the second polynucleotide may provide a trans-acting factor interacting with the cis-regulatory control region to allow transcription of the receptor polynucleotides from the vector. Alternatively, a trans-acting factor may be supplied by the host cell. Finally, a trans-acting factor can be produced from the vector itself.

It is understood, however, that in some embodiments, transcription and/or translation of the receptor polynucleotides can occur in a cell-free system.

The regulatory sequence to which the polynucleotides described herein can be operably linked include promoters for directing mRNA transcription. These include, but are not limited to, the left promoter from bacteriophage λ , the lac, TRP, and TAC promoters from *E. coli*, the early and late promoters from SV40, the CMV immediate early promoter, the adenovirus early and late promoters, and retrovirus long-terminal repeats.

In addition to control regions that promote transcription, expression vectors may also include regions that modulate transcription, such as repressor binding sites and enhancers. Examples include the SV40 enhancer, the cytomegalovirus immediate early enhancer, polyoma enhancer, adenovirus enhancers, and retrovirus LTR enhancers.

In addition to containing sites for transcription initiation and control, expression vectors can also contain sequences necessary for transcription termination and, in the transcribed region a ribosome binding site for translation. Other regulatory control elements for expression include initiation and termination codons as well as polyadenylation signals. The person of ordinary skill in the art would be aware of the numerous regulatory sequences that are useful in expression vectors. Such regulatory sequences are described, for example, in Sambrook *et al. Molecular Cloning: A Laboratory Manual*, 2nd. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

A variety of expression vectors can be used to express a receptor polynucleotide. Such vectors include chromosomal, episomal, and virus-derived vectors, for example vectors derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, including yeast artificial chromosomes, from viruses such as baculoviruses, papovaviruses such as SV40, Vaccinia viruses, adenoviruses, poxviruses, pseudorabies viruses, and retroviruses

WO 00/23588

PCT/US99/24368

Vectors may also be derived from combinations of these sources such as those derived from plasmid and bacteriophage genetic elements, e.g. cosmids and phagemids. Appropriate cloning and expression vectors for prokaryotic and eukaryotic hosts are described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

The regulatory sequence may provide constitutive expression in one or more host cells (*i.e.* tissue specific) or may provide for inducible expression in one or more cell types such as by temperature, nutrient additive, or exogenous factor such as a hormone or other ligand. A variety of vectors providing for constitutive and inducible expression in prokaryotic and eukaryotic hosts are well known to those of ordinary skill in the art.

The receptor polynucleotides can be inserted into the vector nucleic acid by well-known methodology. Generally, the DNA sequence that will ultimately be expressed is joined to an expression vector by cleaving the DNA sequence and the expression vector with one or more restriction enzymes and then ligating the fragments together. Procedures for restriction enzyme digestion and ligation are well known to those of ordinary skill in the art.

The vector containing the appropriate polynucleotide can be introduced into an appropriate host cell for propagation or expression using well-known techniques. Bacterial cells include, but are not limited to, *E. coli*, *Streptomyces*, and *Salmonella typhimurium*. Eukaryotic cells include, but are not limited to, yeast, insect cells such as *Drosophila*, animal cells such as COS and CHO cells, and plant cells.

As described herein, it may be desirable to express the polypeptide as a fusion protein. Accordingly, the invention provides fusion vectors that allow for the production of the receptor polypeptides. Fusion vectors can increase the expression of a recombinant protein, increase the solubility of the recombinant protein, and aid in the purification of the protein by acting for example as a ligand for affinity purification. A proteolytic cleavage site may be introduced at the junction of the fusion moiety so that the desired polypeptide can ultimately be separated from the fusion moiety. Proteolytic enzymes include, but are not limited to, factor Xa, thrombin, and enterokinase. Typical fusion expression vectors include pGEX (Smith *et al.* (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST).

WO 00/23588

PCT/US99/24368

maltose E binding protein, or protein A, respectively, to the target recombinant protein. Examples of suitable inducible non-fusion *E. coli* expression vectors include pTc (Amann *et al.* (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.* (1990) *Gene Expression Technology: Methods in Enzymology* 185:60-89).

5 Recombinant protein expression can be maximized in a host bacteria by providing a genetic background wherein the host cell has an impaired capacity to proteolytically cleave the recombinant protein. (Gottesman *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Alternatively, the sequence of the polynucleotide of interest can be
10 altered to provide preferential codon usage for a specific host cell, for example *E. coli*. (Wada *et al.* (1992) *Nucleic Acids Res.* 20:2111-2118).

The receptor polynucleotides can also be expressed by expression vectors that are operative in yeast. Examples of vectors for expression in yeast e.g., *S. cerevisiae* include pYepSec1 (Baldari, *et al.* (1987) *EMBO J.* 6:229-234), pM1a (Kurjan *et al.*
15 (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.* (1987) *Gene* 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA).

The receptor polynucleotides can also be expressed in insect cells using, for example, baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf9 cells) include the pAc series
20 (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pV1 series (Lucklow *et al.* (1989) *Virology* 170:31-39).

In certain embodiments of the invention, the polynucleotides described herein are expressed in mammalian cells using mammalian expression vectors. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and
25 pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195).

The expression vectors listed herein are provided by way of example only of the well-known vectors available to those of ordinary skill in the art that would be useful to express the receptor polynucleotides. The person of ordinary skill in the art would be aware of other vectors suitable for maintenance propagation or expression
30 of the polynucleotides described herein. These are found for example in Sambrook, *et al. T. Molecular Cloning: A Laboratory Manual*, 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

The invention also encompasses vectors in which the nucleic acid sequences described herein are cloned into the vector in reverse orientation, but operably linked

WO 00/23588

PCT/US99/24368

to a regulatory sequence that permits transcription of antisense RNA. Thus, an antisense transcript can be produced to all, or to a portion, of the polynucleotide sequences described herein, including both coding and non-coding regions.

Expression of this antisense RNA is subject to each of the parameters described above

5 in relation to expression of the sense RNA (regulatory sequences, constitutive or inducible expression, tissue-specific expression).

The invention also relates to recombinant host cells containing the vectors described herein. Host cells therefore include prokaryotic cells, lower eukaryotic cells such as yeast, other eukaryotic cells such as insect cells, and higher eukaryotic
10 cells such as mammalian cells.

The recombinant host cells are prepared by introducing the vector constructs described herein into the cells by techniques readily available to the person of ordinary skill in the art. These include, but are not limited to, calcium phosphate transfection, DEAE-dextran-mediated transfection, cationic lipid-mediated
15 transfection, electroporation, transduction, infection, lipofection, and other techniques such as those found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual*, 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Host cells can contain more than one vector. Thus, different nucleotide
20 sequences can be introduced on different vectors of the same cell. Similarly, the receptor polynucleotides can be introduced either alone or with other polynucleotides that are not related to the receptor polynucleotides such as those providing trans-acting factors for expression vectors. When more than one vector is introduced into a cell, the vectors can be introduced independently, co-introduced or joined to the
25 receptor polynucleotide vector.

In the case of bacteriophage and viral vectors, these can be introduced into cells as packaged or encapsulated virus by standard procedures for infection and transduction. Viral vectors can be replication-competent or replication-defective. In the case in which viral replication is defective, replication will occur in host cells
30 providing functions that complement the defects.

Vectors generally include selectable markers that enable the selection of the subpopulation of cells that contain the recombinant vector constructs. The marker can be contained in the same vector that contains the polynucleotides described herein or may be on a separate vector. Markers include tetracycline or ampicillin-resistance

WO 00/23588

PCT/US99/24368

genes for prokaryotic host cells and dihydrofolate reductase or neomycin resistance for eukaryotic host cells. However, any marker that provides selection for a phenotypic trait will be effective.

While the mature proteins can be produced in bacteria, yeast, mammalian cells, and other cells under the control of the appropriate regulatory sequences, cell-free transcription and translation systems can also be used to produce these proteins using RNA derived from the DNA constructs described herein.

Where secretion of the polypeptide is desired, appropriate secretion signals are incorporated into the vector. The signal sequence can be endogenous to the receptor polypeptides or heterologous to these polypeptides.

Where the polypeptide is not secreted into the medium, the protein can be isolated from the host cell by standard disruption procedures, including freeze thaw, sonication, mechanical disruption, use of lysing agents and the like. The polypeptide can then be recovered and purified by well-known purification methods including ammonium sulfate precipitation, acid extraction, anion or cationic exchange chromatography, phosphocellulose chromatography, hydrophobic-interaction chromatography, affinity chromatography, hydroxylapatite chromatography, lectin chromatography, or high performance liquid chromatography.

It is also understood that depending upon the host cell in recombinant production of the polypeptides described herein, the polypeptides can have various glycosylation patterns, depending upon the cell, or maybe non-glycosylated as when produced in bacteria. In addition, the polypeptides may include an initial modified methionine in some cases as a result of a host-mediated process.

25 Uses of vectors and host cells

It is understood that "host cells" and "recombinant host cells" refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

The host cells expressing the polypeptides described herein, and particularly recombinant host cells, have a variety of uses. First, the cells are useful for producing receptor proteins or polypeptides that can be further purified to produce desired

WO 00/23588

PCT/US99/24368

amounts of receptor protein or fragments. Thus, host cells containing expression vectors are useful for polypeptide production.

Host cells are also useful for conducting cell-based assays involving the receptor or receptor fragments. Thus, a recombinant host cell expressing a native
5 receptor is useful to assay for compounds that stimulate or inhibit receptor function. This includes ligand binding, gene expression at the level of transcription or translation, G-protein interaction, and components of the signal transduction pathway.

Host cells are also useful for identifying receptor mutants in which these functions are affected. If the mutants naturally occur and give rise to a pathology,
10 host cells containing the mutations are useful to assay compounds that have a desired effect on the mutant receptor (for example, stimulating or inhibiting function) which may not be indicated by their effect on the native receptor.

Recombinant host cells are also useful for expressing the chimeric polypeptides described herein to assess compounds that activate or suppress activation
15 by means of a heterologous amino terminal extracellular domain (or other binding region). Alternatively, a heterologous region spanning the entire transmembrane domain (or parts thereof) can be used to assess the effect of a desired amino terminal extracellular domain (or other binding region) on any given host cell. In this embodiment, a region spanning the entire transmembrane domain (or parts thereof)
20 compatible with the specific host cell is used to make the chimeric vector. Alternatively, a heterologous carboxy terminal intracellular, e.g., signal transduction, domain can be introduced into the host cell.

Further, mutant receptors can be designed in which one or more of the various functions is engineered to be increased or decreased (e.g., ligand binding or G-protein
25 binding) and used to augment or replace receptor proteins in an individual. Thus, host cells can provide a therapeutic benefit by replacing an aberrant receptor or providing an aberrant receptor that provides a therapeutic result. In one embodiment, the cells provide receptors that are abnormally active

In another embodiment, the cells provide receptors that are abnormally
30 inactive. These receptors can compete with endogenous receptors in the individual.

In another embodiment, cells expressing receptors that cannot be activated, are introduced into an individual in order to compete with endogenous receptors for ligand. For example, in the case in which excessive ligand is part of a treatment modality, it may be necessary to inactivate this ligand at a specific point in treatment

WO 00/23588

PCT/US99/24368

Providing cells that compete for the ligand, but which cannot be affected by receptor activation would be beneficial.

Homologously recombinant host cells can also be produced that allow the *in situ* alteration of endogenous receptor polynucleotide sequences in a host cell genome. The host cell includes, but is not limited to, a stable cell line, cell *in vivo*, or cloned microorganism. This technology is more fully described in WO 93/09222, WO 91/12650, WO 91/06667, U.S. 5,272,071, and U.S. 5,641,670. Briefly, specific polynucleotide sequences corresponding to the receptor polynucleotides or sequences proximal or distal to a receptor gene are allowed to integrate into a host cell genome by homologous recombination where expression of the gene can be affected. In one embodiment, regulatory sequences are introduced that either increase or decrease expression of an endogenous sequence. Accordingly, a receptor protein can be produced in a cell not normally producing it. Alternatively, increased expression of receptor protein can be effected in a cell normally producing the protein at a specific level. Further, expression can be decreased or eliminated by introducing a specific regulatory sequence. The regulatory sequence can be heterologous to the receptor protein sequence or can be a homologous sequence with a desired mutation that affects expression. Alternatively, the entire gene can be deleted. The regulatory sequence can be specific to the host cell or capable of functioning in more than one cell type. Still further, specific mutations can be introduced into any desired region of the gene to produce mutant receptor proteins. Such mutations could be introduced, for example, into the specific functional regions such as the ligand-binding site.

In one embodiment, the host cell can be a fertilized oocyte or embryonic stem cell that can be used to produce a transgenic animal containing the altered receptor gene. Alternatively, the host cell can be a stem cell or other early tissue precursor that gives rise to a specific subset of cells and can be used to produce transgenic tissues in an animal. See also Thomas *et al.* (1987) *Cell* 51:503 for a description of homologous recombination vectors. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced gene has homologously recombined with the endogenous receptor gene is selected (see e.g., Li *et al.* (1992) *Cell* 69:915). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras (see e.g., Bradley in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, F.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into

WO 00/23588

PCT/US99/24368

a suitable pseudopregnant female foster animal and the embryo brought to term.

Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos. WO 90/11354; WO 91/01140; and WO 93/04169.

The genetically engineered host cells can be used to produce non-human transgenic animals. A transgenic animal is preferably a mammal, for example a rodent, such as a rat or mouse, in which one or more of the cells of the animal include a transgene. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal in one or more cell types or tissues of the transgenic animal. These animals are useful for studying the function of a receptor protein and identifying and evaluating modulators of receptor protein activity.

Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, and amphibians.

In one embodiment, a host cell is a fertilized oocyte or an embryonic stem cell into which receptor polynucleotide sequences have been introduced

A transgenic animal can be produced by introducing nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Any of the receptor nucleotide sequences can be introduced as a transgene into the genome of a non-human animal, such as a mouse.

Any of the regulatory or other sequences useful in expression vectors can form part of the transgenic sequence. This includes intronic sequences and polyadenylation signals, if not already included. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the receptor protein to particular cells.

Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.* U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan

WO 00/23588

PCT/US99/24368

Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the transgene in its genome and/or expression of transgenic mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene can further be bred to other transgenic animals carrying other transgenes. A transgenic animal also includes animals in which the entire animal or tissues in the animal have been produced using the homologously recombinant host cells described herein.

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso *et al.* (1992) *PNAS* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *S. cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355). If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein is required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyst and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, e.g., the somatic cell, is isolated.

WO 00/23588

PCT/US99/24368

Transgenic animals containing recombinant cells that express the polypeptides described herein are useful to conduct the assays described herein in an *in vivo* context. Accordingly, the various physiological factors that are present *in vivo* and that could effect ligand binding, receptor activation, and signal transduction, may not be evident from *in vitro* cell-free or cell-based assays. Accordingly, it is useful to provide non-human transgenic animals to assay *in vivo* receptor function, including ligand interaction, the effect of specific mutant receptors on receptor function and ligand interaction, and the effect of chimeric receptors. It is also possible to assess the effect of null mutations, that is mutations that substantially or completely eliminate one or more receptor functions.

In general, methods for producing transgenic animals include introducing a nucleic acid sequence according to the present invention, the nucleic acid sequence capable of expressing the receptor protein in a transgenic animal, into a cell in culture or *in vivo*. When introduced *in vivo*, the nucleic acid is introduced into an intact organism such that one or more cell types and, accordingly, one or more tissue types, express the nucleic acid encoding the receptor protein. Alternatively, the nucleic acid can be introduced into virtually all cells in an organism by transfecting a cell in culture, such as an embryonic stem cell, as described herein for the production of transgenic animals, and this cell can be used to produce an entire transgenic organism. As described, in a further embodiment, the host cell can be a fertilized oocyte. Such cells are then allowed to develop in a female foster animal to produce the transgenic organism.

Pharmaceutical compositions

The receptor nucleic acid molecules, protein (particularly fragments such as the amino terminal extracellular domain), modulators of the protein, and antibodies (also referred to herein as "active compounds") can be incorporated into pharmaceutical compositions suitable for administration to a subject, e.g., a human. Such compositions typically comprise the nucleic acid molecule, protein, modulator, or antibody and a pharmaceutically acceptable carrier.

As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for

WO 00/23588

PCT/US99/24368

pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, such media can be used in the compositions of the invention. Supplementary active compounds can also be incorporated into the compositions. A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as

WO 00/23588

PCT/US99/24368

mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

- Sterile injectable solutions can be prepared by incorporating the active
- 5 compound (e.g., a receptor protein or anti-receptor antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above.
- 10 In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

- Oral compositions generally include an inert diluent or an edible carrier. They
- 15 can be enclosed in gelatin capsules or compressed into tablets. For oral administration, the agent can be contained in enteric forms to survive the stomach or further coated or mixed to be released in a particular region of the GI tract by known methods. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules.
- 20 Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or
- 25 compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin, an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin, or a flavoring agent such as peppermint, methyl salicylate, or
- 30 orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

WO 00/23588

PCT/US99/24368

Systemic administration can also be by transmucosal or transdermal means

For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems.

Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

WO 00/23588

PCT/US99/24368

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (U.S. 5,328,470) or by stereotactic injection (see e.g., Chen *et al.* (1994) *PNAS* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g. retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

As defined herein, a therapeutically effective amount of protein or polypeptide (i.e., an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight.

The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments. In a preferred example, a subject is treated with antibody, protein, or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody, protein, or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids,

WO 00/23588

PCT/US99/24368

amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention. Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (e.g., about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram). It is furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (e.g., a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

This invention may be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments

WO 00/23588

PCT/US99/24368

are provided so that this disclosure will fully convey the invention to those skilled in the art. Many modifications and other embodiments of the invention will come to mind in one skilled in the art to which this invention pertains having the benefit of the teachings presented in the foregoing description. Although specific terms are
5 employed, they are used as in the art unless otherwise indicated.

WO 00/23588

PCT/US99/24368

THAT WHICH IS CLAIMED:

1. An isolated polypeptide having an amino acid sequence selected from the group consisting of:
 - 5 (a) The amino acid sequence shown in SEQ ID NO 1, the amino acid sequence shown in SEQ ID NO 3, and the amino acid sequence shown in SEQ ID NO 5;
 - (b) The amino acid sequence of an allelic variant of the amino acid sequence shown in SEQ ID NO 1, the amino acid sequence of an allelic variant of the amino acid sequence shown in SEQ ID NO 3, and the amino acid sequence of an allelic variant of the amino acid sequence shown in SEQ ID NO 5;
 - 10 (c) The amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 1, wherein the sequence variant is encoded by a nucleic acid molecule hybridizing to the nucleic acid molecule shown in SEQ ID NO 2 under stringent conditions, the amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 3, wherein the sequence variant is encoded by a nucleic acid molecule hybridizing to the nucleic acid molecule shown in SEQ ID NO 4 under stringent conditions, and the amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 5, wherein the sequence variant is encoded by a nucleic acid molecule hybridizing to the nucleic acid molecule shown in SEQ ID NO 6 under stringent conditions;
 - 20 (d) A fragment of the amino acid sequence shown in SEQ ID NO 1, wherein the fragment comprises at least 8 contiguous amino acids from 1-264, a fragment of the amino acid sequence shown in SEQ ID NO 3, wherein the fragment comprises at least 9 contiguous amino acids from 1-244, and a fragment of the amino acid sequence shown in SEQ ID NO 5, wherein the fragment comprises at least 8 contiguous amino acids;
 - 25 (e) The amino acid sequence of the mature receptor polypeptide from about amino acid 6 to about amino acid 319, shown in SEQ ID NO 1, the amino acid sequence of the mature receptor polypeptide from about amino acid 6 to about amino acid 337, shown in SEQ ID NO 3, and the amino acid sequence of the mature receptor polypeptide from about amino acid 6 to about amino acid 372, shown in SEQ ID NO 5;
 - 30 (f) The amino acid sequence of the polypeptide shown in SEQ ID NO 1, from about amino acid 1 to about amino acid 24, the amino acid sequence of the polypeptide shown in SEQ ID NO 3, from about amino acid 1 to about amino acid 28,

WO 00/23588

PCT/US99/24368

and the amino acid sequence of the polypeptide shown in SEQ ID NO 5, from about amino acid 1 to about amino acid 25;

(g) The amino acid sequence of an epitope bearing region of any one of the polypeptides of (a)-(f).

5

2 An isolated antibody that selectively binds to a polypeptide of claim 1, (a)-(g).

3. An isolated nucleic acid molecule having a nucleotide sequence selected from the group consisting of:

10 (a) The nucleotide sequence shown in SEQ ID NO 2, the nucleotide sequence shown in SEQ ID NO 4, and the nucleotide sequence shown in SEQ ID NO 6;

(b) A nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO 1, a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO 3, and a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO 5;

15 (c) A nucleotide sequence complementary to any of the nucleotide sequences in (a) or (b).

20 4. An isolated nucleic acid molecule having a nucleotide sequence selected from the group consisting of:

(a) A nucleotide sequence encoding an amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 1 that hybridizes to the nucleotide sequence shown in SEQ ID NO 2 under stringent conditions, a nucleotide sequence encoding an amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 3 that hybridizes to the nucleotide sequence shown in SEQ ID NO 4 under stringent conditions, and a nucleotide sequence encoding an amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 5 that hybridizes to the nucleotide sequence shown in SEQ ID NO 6 under stringent conditions;

25

30 (b) A nucleotide sequence complementary to the nucleotide sequences in (a).

WO 00/23588

PCT/US99/24368

5. An isolated nucleic acid molecule a polynucleotide having a nucleotide sequence selected from the group consisting of:
- (a) A nucleotide sequence encoding a fragment of the amino acid sequence shown in SEQ ID NO 1, wherein the fragment comprises at least 8 contiguous amino acids from 1-264, a nucleotide sequence encoding a fragment of the amino acid sequence shown in SEQ ID NO 3, wherein the fragment comprises at least 9 contiguous amino acids from 1-244, and a nucleotide sequence encoding a fragment of the amino acid sequence shown in SEQ ID NO 5, wherein the fragment comprises at least 8 contiguous amino acids;
- (b) A nucleotide sequence complementary to the nucleotide sequences in (a).
6. A nucleic acid vector comprising the nucleic acid sequences in any of claims 3-5.
7. A host cell containing the vector of claim 6.
8. A method for producing any of the polypeptides in claim 1 comprising introducing a nucleotide sequence encoding any of the polypeptide sequences in (a)-(f) into a host cell, and culturing the host cell under conditions in which the proteins are expressed from the nucleic acid.
9. A method for detecting the presence of any of the polypeptides in claim 1 in a sample, said method comprising contacting said sample with an agent that specifically allows detection of the presence of the polypeptide in the sample and then detecting the presence of the polypeptide.
10. The method of claim 9, wherein said agent is capable of selective physical association with said polypeptide.
11. The method of claim 10, wherein said agent binds to said polypeptide.
12. The method of claim 11, wherein said agent is an antibody.

WO 00/23588

PCT/US99/24368

13. The method of claim 11, wherein said agent is a ligand.
14. A kit comprising reagents used for the method of claim 9, wherein the reagents comprise an agent that specifically binds to said polypeptide.
- 5 15. A method for detecting the presence of any of the nucleic acid sequences in any of claims 3-5 in a sample, the method comprising contacting the sample with an oligonucleotide that hybridizes to the nucleic acid sequences under stringent conditions and determining whether the oligonucleotide binds to the nucleic acid sequence in the sample.
- 10 16. The method of claim 15, wherein the nucleic acid, whose presence is detected, is mRNA.
- 15 17. A kit comprising reagents used for the method of claim 15, wherein the reagents comprise a compound that hybridizes under stringent conditions to any of the nucleic acid molecules.
18. A method for identifying an agent that interacts with any of the polypeptides in claim 1, said method comprising contacting said agent with a cell capable of allowing an interaction between said polypeptide and said agent such that said polypeptide can interact with said agent and measuring the interaction.
- 20 19. A method of screening a cell to identify an agent that interacts with any of the polypeptides in claim 1, said method comprising contacting said agent with a cell capable of allowing an interaction between said polypeptide and said agent such that said polypeptide can interact with said agent and measuring the interaction.
- 25 20. A method for identifying an agent that modulates the level or activity of any of the polypeptides of claim 1 in a cell, said method comprising contacting said agent with a cell capable of expressing said polypeptide such that said polypeptide level or activity can be modulated in said cell by said agent and measuring said polypeptide level or activity.
- 30

WO 00/23588

PCT/US99/24368

21. A method of screening a cell to identify an agent that modulates the level or activity of any of the polypeptides of claim 1 in a cell, said method comprising contacting said agent with a cell capable of expressing said polypeptide such that said polypeptide level or activity can be modulated in said cell by said agent and measuring said polypeptide level or activity.
22. The method of either of claims 19 or 20 wherein said interaction is binding
23. The method of claim 22, wherein a fragment of the polypeptide is contacted.
24. A method for modulating the activity of any of the polypeptides in claim 1, the method comprising contacting any of the polypeptides of claim 1 with an agent under conditions that allow the agent to modulate the activity of the polypeptide.
25. A method for identifying an agent that interacts with any of the nucleic acid molecules of claims 3-5, said method comprising contacting said agent with a cell capable of allowing an interaction between said nucleic acid molecule and said agent such that said nucleic acid molecule can interact with said agent and measuring the interaction.
26. A method of screening a cell to identify an agent that interacts with any of the nucleic acid molecules of claims 3-5, said method comprising contacting said agent with a cell capable of allowing an interaction between said nucleic acid molecule and said agent such that said nucleic acid molecule can interact with said agent and measuring the interaction.
27. A method for identifying an agent that modulates the level or activity of any of the nucleic acid molecules of claims 3-5 in a cell, said method comprising contacting said agent with a cell capable of expressing said nucleic acid molecule such that said nucleic acid molecule level or activity can be modulated in said cell by said agent and measuring said nucleic acid molecule level or activity

WO 00/23588

PCT/US99/24368

28. A method of screening a cell to identify an agent that modulates the level or activity of any of the nucleic acid molecules of claims 3-5 in a cell, said method comprising contacting said agent with a cell capable of expressing said nucleic acid molecule such that said nucleic acid molecule level or activity can be modulated in said cell by said agent and measuring said nucleic acid molecule level or activity.
29. The method of either of claims 25 or 26 wherein said interaction is binding.
30. A method for modulating the activity of any of the nucleic acid molecules in claim 3-5, the method comprising contacting any of the nucleic acid molecules of claim 1 with an agent under conditions that allow the agent to modulate the activity of the nucleic acid molecule.
31. A pharmaceutical composition containing any of the polypeptides in claim 1 in a pharmaceutically acceptable carrier.
32. A pharmaceutical composition containing any of the nucleic acid molecules of claims 3-5 in a pharmaceutically acceptable carrier.
33. A non-human transgenic animal wherein one or more of the cells of said animal contains any of the nucleic acid sequences of claims 3-5.
34. A non-human transgenic animal wherein one or more of cells of said animal contains any of the nucleic acid sequences of claims 3-5 wherein said cell expresses any of the polypeptides of claim 1.
35. A method for producing a transgenic animal according to claims 32 and 33, said method comprising introducing any of the nucleic acid sequences of claim 1 into a cell, wherein said cell is present in said animal or gives rise to said animal.

WO 00/23588

PCT/US99/24368

36. An agent identified using any of the methods of claims 19-22 or 25-28.
37. A method of treating a disorder in a subject comprising administering the pharmaceutical compositions of either of claims 31 or 32.
- 5 38. A method of treating a disorder in a subject by modulating the level or expression of any of the polypeptides of claim 1 in said subject.
- 10 39. A method of treating a disorder in a subject by modulating the level or activity of any of the nucleic acid molecules of claims 3-5 in said subject.
40. The method of either of claims 9 or 15 wherein said detecting is in a cell derived from a subject having a disorder involving said cell.
- 15 41. The method of either of claims 24 or 30 wherein said modulation is in a subject having a disorder involving said cell.
42. The method of claim 39 wherein said disorder is anemia, neutropenia or thrombocytopenia.
- 20 43. The method of claim 41 wherein said disorder is anemia, neutropenia or thrombocytopenia.

WO 00/23588

1 / 19

PCT/US99/24368

Protein Family/Domain HMM Matches for flh2838

>PF00001:tm_1 7 transmembrane receptor (rhodopsin family)

Score: 154.54 Seq: 37/288 Model: 1/269
 *GNILVIWVLCRYRRMR...TPMNYFIVNLAVADLLFslfcmPFWMvyy
 N+L I ++ + R +++++NLAV DIL+ +++++PF MV
 flh2838 37 LNLLAIHGPFSTFLXNRWPDYAATSIYHINLAVFDLLL-VLSLEFKMVL 84
 vMqgRUpFGdfMCrTWaYFDYMNMYASIFELTcISIDRYLWAICHPrYm
 PF++ +C ++ +++++MY S+F++++IS+DR+L AI +P+
 flh2838 95 QV-Q-SFFPS-LCTLVECLYPVSMYGSVPTICFISMDEFL-AIRYPLVS 130
 RWHtpRHIAWvMiiIiWvMSFLISMPPFLMFrWetyrDeneWnmTWcmTy
 +++PR + + ++ IWV+ S+P+ + F + E + M C+
 flh2838 131 HLRSPR-KIFGICCTIWVLVWTGSIFL-YSP-HGKV--ERY--M--CFH 170
 DWPeWwRwYvILmtiimgFYIPMILMFCYwRIYRJaRlWMMNipswQr
 + ++W + V + + GP++FM IM EC +R I ++ + ++ Q
 flh2838 171 NMSDDTWSARVFFPLEVPGLLENGIMSPCCSRG--I-HILGKRDRYQD 217
 RRrmSmRrERRivKMLiIiHvVPIICWIPYFIvmfKSTLM.Mw+FCofC.
 + + + + + + + + V8++++LE ++ +F++ L + + + R C
 flh2838 218 WVQD-----KACLYGIANSLAVFVVSFLPVHLGFFLQFLVRNSFIWE-CR 261
 IwzrlWm.Y.IfeWLaYvMcpCinPIIY*
 + + + + + + + + VNC C+ + Y
 flh2838 262 AKQISFFLQLSMCFNNVC-CLDVFCY 288

FIGURE 2

WO 00/23588

2 / 19

PCT/US99/24368

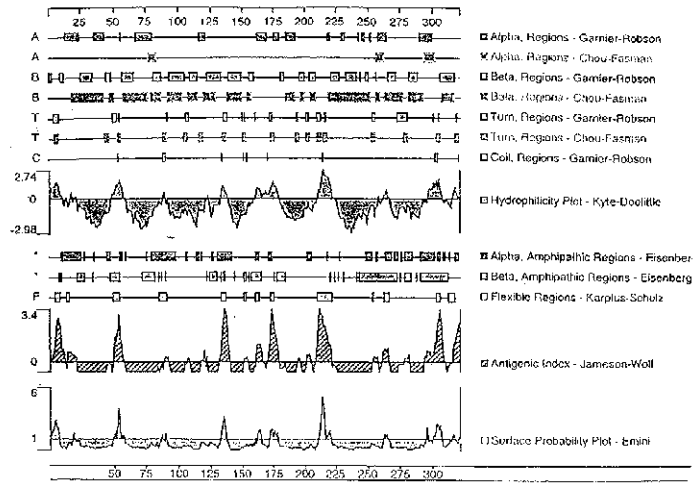


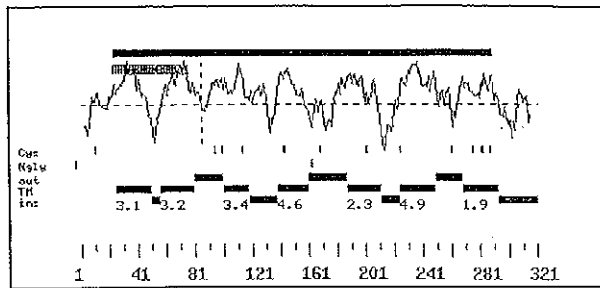
FIGURE 2

WO 00/23588

3 / 19

PCT/US99/24368

Analysis of flh2838 (320 aa)



```

>flh2838
MSQQNTSGDCLFDGVNGLMKTLOFAVHIPTFVLGLLNLGAHGFSTELKNRMPDYAATS
TYMINLAVFDLLVLVLSLFFKMVLSQVQSPFSLCTLVECLYFVSHYGVFTICFISMDF
LAIRYPLLVSHLRSPRKIFGICCTIWLWVWGSIPYSPHGKVEKYMCFHMSDDTWSAK
VFPLEVFGFLPMGIMGPCCSRSHILGRRDHTODWVQKACTIYSIAASLAVFVVSFL
PVHLGFFLOFLVRNSFIVECRAKQSIISFLQLSMCFSNVNCCLDVPCYYPVIRKFRMNIR
AHRFSRVQLVLQDTTISRQ*

```

FIGURE 3

WO 00/23588

PCT/US99/24368

4 / 19

Prosit Pattern Matches for flh2838

>PS00001|PDOC00001|ASN_GLYCOSYLATION N-glycosylation site.

Query: 5 NTSG 8

Query: 171 RNSD 174

>PS00005|PDOC00005|PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 134 SPR 136

Query: 178 SAK 180

>PS00006|PDOC00006|CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query: 6 TSGD 9

Query: 95 TLVE 98

>PS00008|PDOC00008|MYRISTYL N-myristoylation site.

Query: 34 GLLNL 39

Query: 107 GSVFTI 112

Query: 140 GICCTI 145

>PS00009|PDOC00009|AMIDATION Amidation site.

Query: 209 LGRR 212

>PS00227|PDOC00210|G_PROTEIN_RECEPTOR G-protein coupled receptors signature.

Query: 107 GSVFTICFISMDRFLA1 123

>PS00119|PDOC00363|AA_TRNA_LIGASE_T1_2 Aminoacyl-transfer RNA synthetases class-II sign

Query: 114 FISMDRFLA1 123

FIGURE 4

WO 00/23588

5 / 19

PCT/US99/24368

Protein Family/Domain HMM Matches for flh14618

>PF00001|7tm_1 7 transmembrane receptor (rhodopsin family)

Score: 210.98 Seq: 40 295 Model: 1 269

*GNILVIWVtCryRRMRTFKNYFIvNLAvADLLF=lfcmPFWRvYyvWqg
 GN++V++V++++ + ++ + +++++NLAVADLL + T+P+ VYIV+ G
 flh14618 40 GNGFVLVLIKTYHKKSFAQVYMINLAVADLLC-VCTLPLRVVYVYHKG 87

RmpFGDfNCrIwmYFDYMMMYASTFELTcISIDRYLWAICHfMcrYmRRMT
 W+FGDF+CR+++Y Y+N+Y SIFF+T++S+ R +Al+ P++ ++ T
 flh14618 88 IWLFGDFLCRLSTYALYVNLVCSIFFMTAMSEFRC-IATVFPVQINLVT 136

pRRRAWVMIIiIWVMSFLISMPEFLMFrWstYrDENEwNmTWCMlyDWPe
 + +A+++++IW++ +L S P F ++ +DE + T C++ ++
 flh14618 137 QK-KAREVCVGIWIFVILTSSP-F-LM-AKPKDEKN--NTKCFE-PPQD 179

...wMWrWYVILMtiingFYIPMIIMlFCYwRIYRIarLWMRMlpswQrR
 +++ +++ +++++ GF+IP+ I+++CY++I I L++++ S+++
 flh14618 180 NQTKNHVLVLHVVSLVGGFIIPFVIIVCYTMl--I--LTLLK-KSMKK- 223

RrmSmrFRERRiVkmIiIIMvVFIICWlPYFIvmEHUTLMMwvF.....
 ++++++ M++++ +P++++PY+I + +++++
 flh14618 224 ---NLSSHKAIGNIMVVTAALVSVFMPYHIQRT-----IHLHFLHNETK 265

.CeFCIwrriWmY.IfeWLaYvNCpCinPIIY*
 C+ +++R++ I++ LA +NC C++P++Y
 flh14618 266 PCD-SVLRMQKSVVITLSLAASNC-CFDELLY 295

FIGURE 5

WO 00/23588

6 / 19

PCT/US99/24368

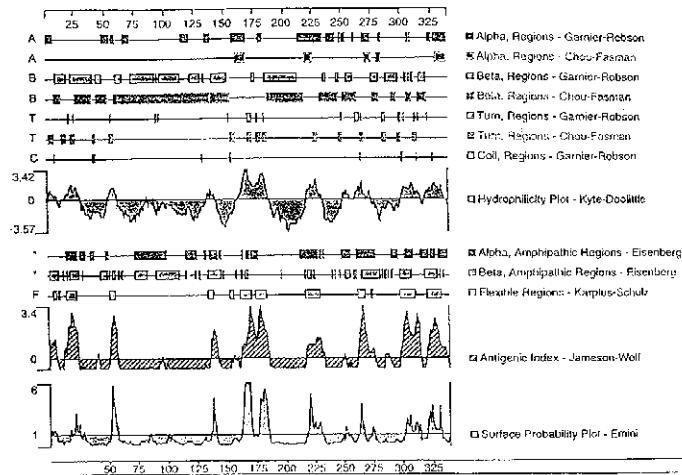


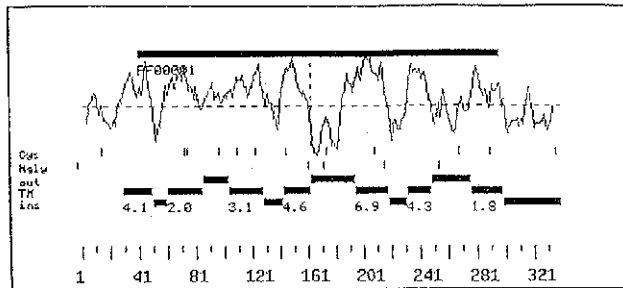
FIGURE 6

WO 00/23588

7 / 19

PCT/US99/24368

Analysis of flh14618 (338 aa)



```

>flh14618
MDETGNLTVSSATCHDITDOFRNQVYSTLYSMISVVGFPNGFVLYVLIKTYHKKSAPQV
YKINLAVADLLCVCTLPKRVVYVHKGWLFQDFLCRLSTYALYVNLVCSIFFNTAMSEF
RCIAIVFPVQNIINLTOKKAREVCVGWIFVILTSSPFLNAKPOKDEKQNTKCEPFPQDN
QTRNHVVLVHVSLSVGGFIIEFVLIIVCYTHIILTLLKSKMKMLSSHKKATGMWVTA
AFLVSFMPYHIQRTIHLHFLHNETKPCDSVLRHQKSVVITLSLAASNCCFDPLLYFFSGG
NFRKRLSTFRKHSLSSTVTVPRKKASLPEKGDEICKV*

```

FIGURE 7

WO 00/23588

8 / 19

PCT/US99/24368

Prosite Pattern Matches for fh14618

>PS00001|PDOC00001|ASN_GLYCOSYLATION N-glycosylation site.

Query: 6 NLTV 9
 Query: 169 NMTK 172
 Query: 180 NQTK 181
 Query: 224 NLSS 227
 Query: 262 NETH 265

>PS00004|PDOC00004|CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation site.

Query: 304 KRLG 307
 Query: 310 RKHS 313
 Query: 323 KKAS 326

>PS00005|PDOC00005|PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 136 TOR 138
 Query: 220 SMK 222
 Query: 229 GIK 229
 Query: 308 TFR 310

>PS00006|PDOC00006|CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query: 13 TCHD 16
 Query: 17 TIDD 20
 Query: 326 SLPE 329

>PS00142|PDOC00299|MICROBODIES_CTER Microbodies C-terminal targeting signal.

Query: 335 CKV 338

FIGURE 8

PCT/US99/24368

FIGURE 9

WO 00/23588

10 / 19

PCT/US99/24368

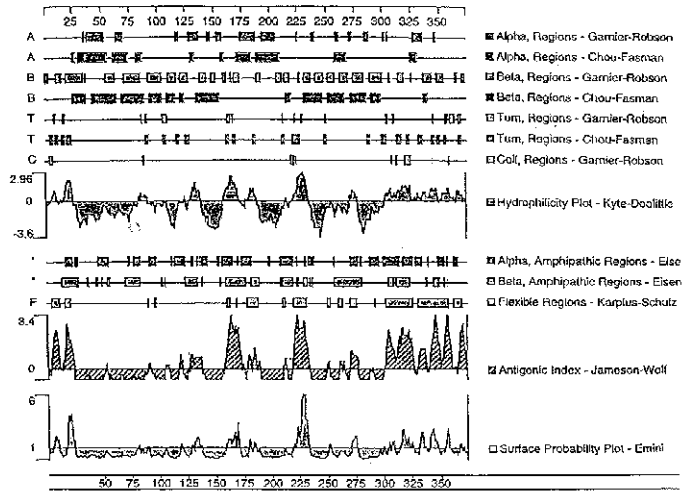


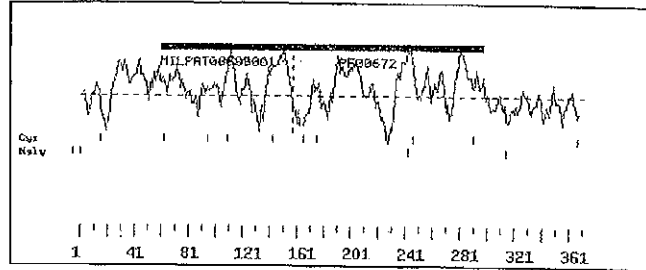
FIGURE 10

WO 00/23588

11 / 19

PCT/US99/24368

Analysis of sequence28390 (372 aa)



```

>sequence28390
MLANSSSTNSVLPDVRPHRLHVVSLVLAAGLPINALAHVFLRALTVHSVSVY
MCNLAASDLPLTSLPVRLEYYALHWPFPDLCCQTG.IIQMNNYSSCIFLMLINVDY
AATVHPLRLRLRRPRVARLCLGVWALTLVFAVPAARVHPSPRCRYRDLVRLCUESPS
DEINWGLLELLPLVLLAPALGFLPLAAVVYSSGRVPWTLARPDATCSQRRKTVRELLANT
VIFLACFPVYNSTLAVYGLLRSLVAASVFARDRVSGVLAVMVLGAGNCVLDPLVYVPS
ABGPRSTLEGLSTPHNARTSATNGTRALAQSERSAVTTDATRPDAASQGLRPFSDSHL
SSFTQCPQDSAL

```

FIGURE 11

WO 00/23588

PCT/US99/24368

12 / 19

Prosite Pattern Matches for sequence28390

Prosite version: Release 12.2 of February 1995

>PS00001|PDOC00001|ASN_GLYCOSYLATION N-glycosylation site.

Query: 4 NSSS 7

Query: 9 NSSV 12

Query: 251 NSTL 254

Query: 323 NQTR 326

>PS00004|PDOC00004|CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation site.

Query: 229 RPKT 232

>PS00005|PDOC00005|PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 21 THR 23

Query: 211 SGR 213

Query: 226 SQE 228

Query: 232 TVR 234

Query: 307 TLR 309

Query: 332 SER 334

>PS00006|PDOC00006|CK2_PHOSPHO_SITE Casein Kinase II phosphorylation site.

Query: 178 SFSI 181

Query: 342 TRFD 345

>PS00008|PDOC00008|MYRISTYL N-myristoylation site.

Query: 36 GLPLDA 41

Query: 258 GLDRSK 263

Query: 324 GTRAAI 329

>PS00217|PDOC00210|G_PROTEIN_RECEPTOR G-protein coupled receptors signature.

Query: 107 GSCXFLHLINVDRYAAI 123

FIGURE 12

WO 00/23588

13 / 19

PCT/US99/24368

Gene 14618 Expression in Normal Human Tissues

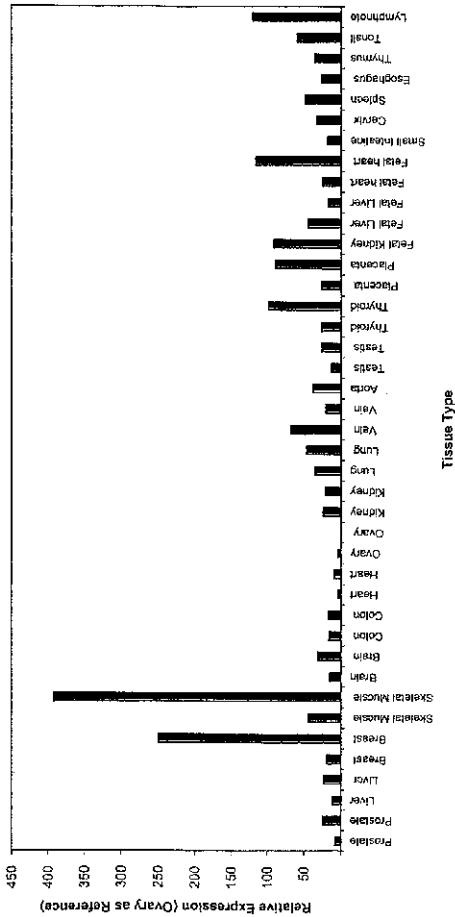
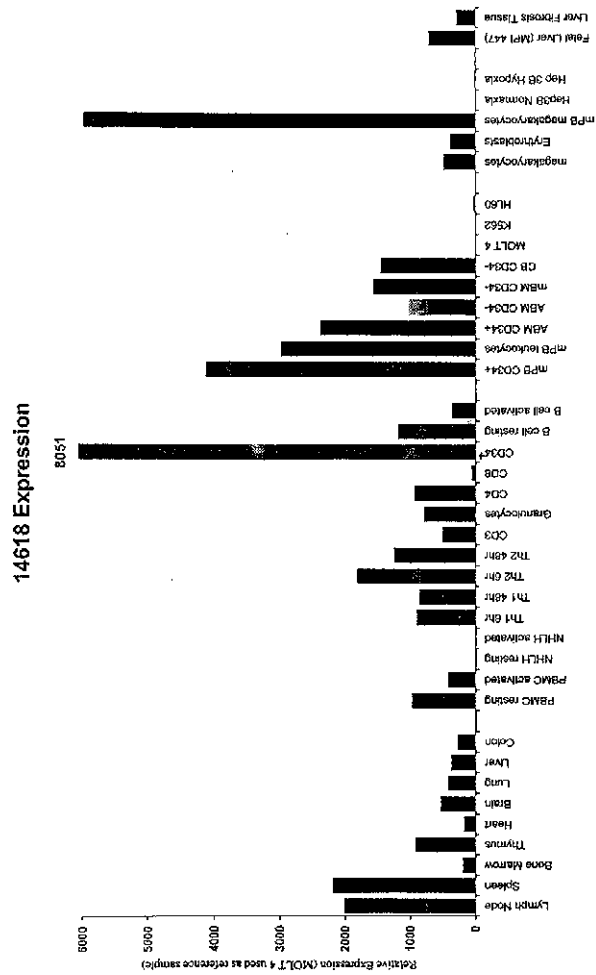


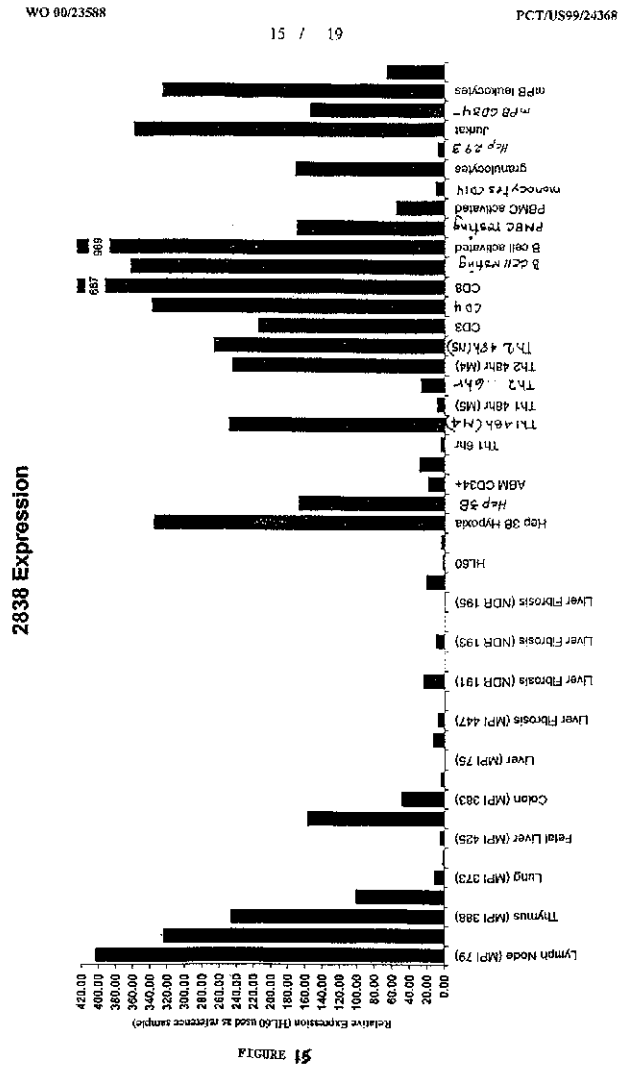
FIGURE 3

WO 00/23588

14 / 19

PCT/US99/24368





WO 00/23588

16 / 19

PCT/US99/24368

Gene 15334 Expression in Normal Human Tissues

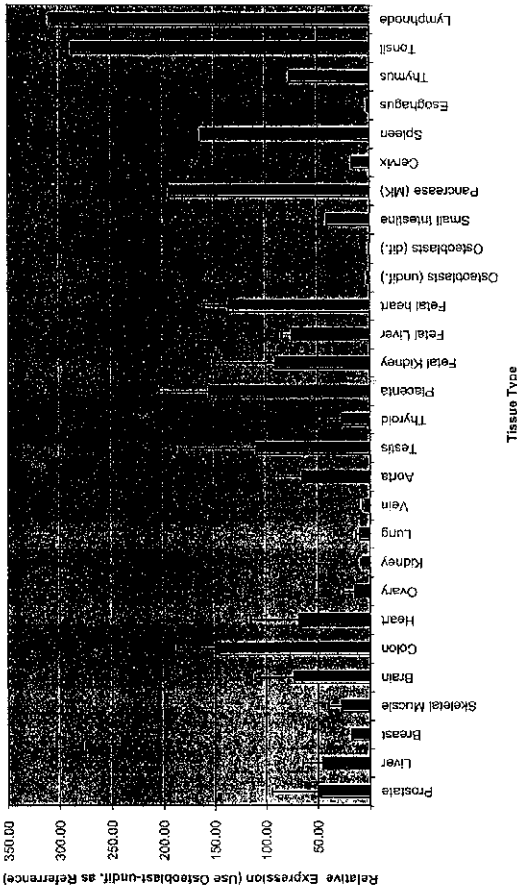


FIGURE 16

WO 00/23588

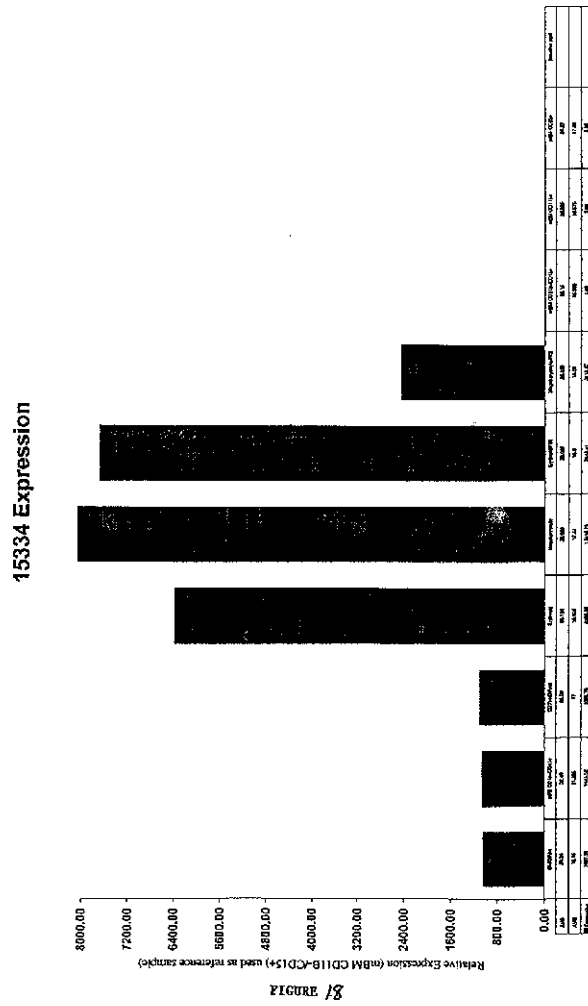
17 / 19

PCT/US99/24368

hu15334 Expression



FIGURE 17





WO 00/23588

PCT/US99/24368

SEQUENCE LISTING

<110> Glucksmann, Marc A.

<120> Novel G-Protein Coupled Receptors

<130> 5800-11-1

<140>

<141>

<150> 7

<170> PatentIn Ver. 2.0

<210> 1

<211> 319

<212> PRT

<213> Homo sapiens

<400> 1

Met Ser Gln Gln Asn Thr Ser Gly Asp Cys Leu Phe Asp Gly Val Asn
 1 5 10 15

Glu Leu Met Lys Thr Leu Gln Phe Ala Val His Ile Pro Thr Phe Val
 20 25 30

Leu Gly Leu Leu Leu Asn Leu Leu Ala Ile His Gly Phe Ser Thr Phe
 35 40 45

Leu Lys Asn Arg Trp Pro Asp Tyr Ala Ala Thr Ser Ile Tyr Met Ile
 50 55 60

Asn Leu Ala Val Phe Asp Leu Leu Leu Val Leu Ser Leu Pro Phe Lys
 65 70 75 80

Met Val Leu Ser Gln Val Gln Ser Pro Phe Pro Ser Leu Cys Thr Leu
 85 90 95

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

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

Val Ser His Leu Arg Ser Pro Arg Lys Ile Phe Gly Ile Cys Cys Thr
 130 135 140

WO 00/23588 PCT/US99/24368

Ile Trp Val Leu Val Trp Thr Gly Ser Ile Pro Ile Tyr Ser Phe His
 145 150 155 160

Gly Lys Val Glu Lys Tyr Met Cys Phe His Asn Met Ser Asp Asp Thr
 165 170 175

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

Pro Met Gly Ile Met Gly Phe Cys Cys Ser Arg Ser Ile His Ile Leu
 195 200 205

Leu Gly Arg Arg Asp His Thr Gln Asp Trp Val Gln Gln Lys Ala Cys
 210 215 220

Ile Tyr Ser Ile Ala Ala Ser Leu Ala Val Phe Val Val Ser Phe Leu
 225 230 235 240

Pro Val His Leu Gly Phe Phe Leu Gln Phe Leu Val Arg Asn Ser Phe
 245 250 255

Ile Val Glu Cys Arg Ala Lys Gln Ser Ile Ser Phe Phe Leu Gln Leu
 260 265 270

Ser Met Cys Phe Ser Asn Val Asn Cys Cys Leu Asp Val Phe Cys Tyr
 275 280 285

Tyr Phe Val Ile Lys Glu Phe Arg Met Asn Ile Arg Ala His Arg Pro
 290 295 300

Ser Arg Val Gln Leu Val Leu Gln Asp Thr Thr Ile Ser Arg Gly
 305 310 315

<210> 2
 <211> 1617
 <212> DNA
 <213> Homo sapiens

<400> ?
 aggtacagcc ttggccatt aagaaacta ggcaaggacc tacaactga ccttgctctt 60
 gtggactga gttggtatc aatgggatg aattgctgtg tcatgctggg naggctttt 120
 tggattcttga caaagtcatt tgaatccctc actcaagag agtgaaggga gccccgtctg 180
 atctgttggg gctgtaggaa gaacacagag tcaagcaaac accagtggg acitgctgtg 240
 tgaaggtgtc aacagctga tgaacacat acagtttga gtccacatcc ccaacttctg 300
 cctgggcttg ctctcacc tcctggccat ccatgggttt agcaatttc ttaagaacag 360
 gtggcccgat tatgtgcca cctcaatcra catgataaac ctgucagctt ttgaactgct 420
 acitgctctc tccctccat tcaagatggt cctgtccag gcaagctcc cctccctga 480

WO 00/23588

PCT/US99/24368

ctgttgcaac ctggtggagt gctttacttt cgtcagcatg tacggagcgg tcttcacat 540
 ctgttcacat acaatggacc ggtttatggc catccgttac cctgtactgg tgaagcaact 600
 cccgtccccc aggaagatct ttgggactct ctgcacacct tgggtcttgg tgtggaccgg 660
 aagcatccct atctacagtl tccatgggaa agtggaaaaa tccatgtgct tccacaacat 720
 gtctgatgat acctggagcg caaagglett ctctccgcct gagtgcttgg gcttctctcl 780
 tccatggggc atcatggcct lctgctgctc caggagcacc cactctctgc tggggccggc 840
 agaccacacc caggactggg tgcagcagaa agcctgcctc tccagcatcg cagccagcct 900
 ggtctctctc gttgtctctt tctctcagct ccaactgggg tctctctctc agttctctgt 960
 gaggaaacag tttatcttag aatgcagagc caagcagagc cccagcttct tcttccaatt 1020
 gtccatgtgt tcttccacc tcaactgtct cctggatgtt tctgtcact accttctcat 1080
 caaagaattc agcatgaaa tccgggctca caggctctcc aggttccagc tcttctctca 1140
 ggacacacag atctccgggg gataacggaa ggacatcttg tccaggggaa gaaagccctg 1200
 gccctgaatt ctggtacggg atttccgctc cagggttttt gatgtggagg gatgatccgc 1260
 accatcttca ctgctgtgct tcccttctgt gccatctgag tgcagctctt gctcattata 1320
 ccccaagac cttttttcca ctgcacagac agcttatncc acccagtggt cagggtcttc 1380
 tgaagacccc acagaccagg tgaattactg atttctttag ccaaaaacta tagagcagaa 1440
 gaatttgaaa agagaatgag acctgttcaa caaggtgctc tccactcttc cctattttcc 1500
 tgttctactg ggaagttctg gaaagaaaaa gagagagaga aaagaggtta aggaaggagc 1560
 caagagagtc agttatlggg gagagtgtct tgggcagagg tgggtctgta gggatga 1617

<210> 3

<211> 337

<212> PRT

<213> Homo sapiens

<400> 3

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

Thr Ile Asp Asp Phe Arg Asn Gln Val Tyr Ser Thr Leu Tyr Ser Met
 20 25 30

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

Ile Lys Thr Tyr His Lys Lys Ser Ala Phe Gln Val Tyr Met Ile Asn
 50 55 60

Leu Ala Val Ala Asp Leu Leu Cys Val Cys Thr Leu Pro Leu Arg Val
 65 70 75 80

Val Tyr Tyr Val His Lys Gly Ile Trp Leu Phe Gly Asp Phe Leu Cys
 85 90 95

Arg Leu Ser Thr Tyr Ala Leu Tyr Val Asn Leu Tyr Cys Ser Ile Phe
 100 105 110

Phe Met Thr Ala Met Ser Phe Phe Arg Cys Ile Ala Ile Val Phe Pro

WO 00/23588 PCT/US99/24368

115 120 125

Val Gln Asn Ile Asn Leu Val Thr Gln Lys Lys Ala Arg Phe Val Cys
130 135 140

Val Gly Ile Trp Ile Phe Val Ile Leu Thr Ser Ser Pro Phe Leu Met
145 150 155 160

Ala Lys Pro Gln Lys Asp Glu Lys Asn Asn Thr Lys Cys Phe Glu Pro
165 170 175

Pro Gln Asp Asn Gln Thr Lys Asn His Val Leu Val Leu His Tyr Val
180 185 190

Ser Leu Val Gly Gly Phe Ile Ile Pro Phe Val Ile Ile Ile Val Cys
195 200 205

Tyr Thr Met Ile Ile Leu Thr Asn Leu Lys Lys Ser Met Lys Lys Asn
210 215 220

Leu Ser Ser His Lys Lys Ala Ile Gly Met Ile Met Val Val Thr Ala
225 230 235 240

Ala Phe Leu Val Ser Phe Met Pro Tyr His Ile Gln Arg Thr Ile His
245 250 255

Leu His Phe Leu His Asn Glu Thr Lys Pro Cys Asp Ser Val Leu Arg
260 265 270

Met Gln Lys Ser Val Val Ile Thr Leu Ser Leu Ala Ala Ser Asn Cys
275 280 285

Cys Phe Asp Pro Leu Leu Tyr Phe Phe Ser Gly Gly Asn Phe Arg Lys
290 295 300

Arg Leu Ser Thr Phe Arg Lys His Ser Leu Ser Ser Val Thr Tyr Val
305 310 315 320

Pro Arg Lys Lys Ala Ser Leu Pro Glu Lys Gly Glu Glu Ile Cys Lys
325 330 335

Val

<210> 4
<211> 1358
<212> DNA

W018023588
<213> Homo sapiens

PCT/US99/24368

<400> 4
aatttcaggc cagaatttgc cccgagggctg gtatgacgaa ttttcggaag acttgagcct 60
tgcttctgag aacaaacgca naagacacgt aaactgtgga ccttcaagtt agcagcgtgg 120
gcttcctcta atallacacg gtaaaaggca ttgacacca taagaaggaa catttgtaga 180
ggtaaccagc tgcagaaag agcacaagc cagacattcg tagagaacaa tggatgaac 240
aggaatctc acagtatctt ctgcacacg ccatgacac attgatgact tccgcaatca 300
agtgtatccc aactgtact ctatgactc tgttgtagc tctctggca atggcttctg 360
gctctatgtc ctctlaaaa cctctacaa yaagtacgc utccaaqtat acatgattca 420
tttagcagtc gcagatctcc ttgtctgtg nacatgcct ctccgctggg tctattatgt 480
tcacaaagcg atttggctct ttggtagctt ctgtgccc ctcagcacct atgctttgta 540
tgtcaacctc tattgtagca tctctcttat gacggccatg agcttctcc ggtgcattgc 600
aattgtcttt ccagtcacga acattaatc ggttacacag aaaaaagca ggtttgtgtg 660
tgtaggtatt tggatttttg tgattttgac cagttctcca ttctaatgg caaaacaca 720
aaaaatgag aaaaataata caaagtctt tgagcccca caagacaac aactlaaaa 780
tcattgtttg gcttgcatt atgtgtcatt ggttgggg ttatcctcc cttttgttat 840
tataatgttc tctacacaa tctacattt gacctaata aaaaatcca tgaaaaaaa 900
ctgtcaagt catlaaagc ctataggat gatcatggtc gtagccctg cttttttagt 960
cagttctatg ccatatcaca tcaacgtac cattaactt catttttca acaatgaac 1020
taaacctcgt gatctgccc ttagaatgca gaagtcggtc gtcataact tgcctctgga 1080
tgatccaat tgttccttg accctctcc atattcttt tctggggta attttagaa 1140
aaggctgtct acatttaga agcatttttt gtcagcgtg acttatgac caagaagaa 1200
ggctctttt ccagaaaag gagaagaa atgtaaagta tgglttaac ccatctccag 1260
tcacaaacaa tgaataatg tcccaataa agtatltgt caatcatt acaaaaaaa 1320
taaaatctt acataaaaa aaaaaaaa aggaaaa 1380

<210> 5
<211> 372
<212> FR1
<213> Homo sapiens

<400> 5
Met Leu Ala Asn Ser Ser Ser Thr Asn Ser Ser Val Leu Pro Cys Pro
1 5 10 15
Asp Tyr Arg Pro Thr His Arg Leu His Leu Val Val Tyr Ser Leu Val
20 25 30
Leu Ala Ala Gly Leu Pro Leu Asn Ala Leu Ala Leu Thr Val Phe Leu
35 40 45
Arg Ala Leu Arg Val His Ser Val Val Ser Val Tyr Met Cys Asn Leu
50 55 60
Ala Ala Ser Asp Leu Leu Phe Thr Leu Ser Leu Pro Val Arg Leu Ser
65 70 75 80

WO 00/23588 PCT/US99/24368

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

Met Leu Ile Asn Val Asp Arg Tyr Ala Gly Ile Val His Pro Leu Arg
115 120 125

Leu Arg His Leu Arg Arg Ala 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

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

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

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

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 Gln 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

Ala Thr Asn Gly Thr Arg Ala Ala Leu Ala Gln Ser Glu Arg Ser Ala
325 330 335

WO 00/23588 PCT/US99/24368
 Val Thr Thr Asp Ala Thr Arg Pro Asp Ala Ala Ser Val Gly Leu Leu
 340 345 350

Arg Pro Ser Asp Ser His Ser Leu Ser Ser Phe Thr Gln Cys Pro Gln
 355 360 365

Asp Ser Ala Leu
 370

<210> 6
 <211> 2359
 <212> DNA
 <213> Homo sapiens

<400> 6
 caatgaactc cctctgcttg ttttgggacc atgtctgac agcctctagg cccagcccc 60
 gtaggtgaat gccatgcac gattctggty tgcctcatalg cctcccccag ctactccca 120
 atcccaattt ggcacgatgt tagcacaacg cctctcaacc aacagctcttg ttctcccgta 180
 tcttgactac cgaactacc accgcctgca ctgggtggcc taagacttgg tctgggtgc 240
 cgggtccccc ttaaacggca tagcctcttg ggtcttcttg cggcggctga ggtgacac 300
 ggtgctgagc ggtacatgt qtaacctggc ggcacggac ctgctctca cctctctgc 360
 gccgttcgt ctctctact acgaactga ccactggccc ttcccggac tctgtgcca 420
 gacgacgggc gccatctcc agatgaacat gtacggcagg tgcattctcc tcatgctcat 480
 caacgtgqac cgtacggccg gcatctgca cccgtgga ctggccacc tgggggggg 540
 ccgcctggcg cggctgctct gctcggggt gtggggtct atctctgggt ttgcccgtgc 600
 cgcgcgcgcg gtgcacaggc cctcgcgtt cgcctacgg gacctcagg tgcctctatg 660
 ctctgagagc ttacggagcg agctgtgga aggnaggctg ctgcctclog tctgtctggg 720
 cgaaggcgtg ggttctctgc tgcctctggt ggcgtggtc tactctctgg gccagctct 780
 ctggagctg cgcgcgcgcg acgcaacgca gaggcagcg cggcggaag cgtcgggtc 840
 cctgctggtt aacctctga cttctctgt gtgctctgt cctacaaca gacgtctg 900
 ggtctacggc ctgctcgga gcaagctggt ggcggcagc gtgcctgccc gccatcgct 960
 ggcgggggtg ctgctggtga tggctctgt ggcgggggc aactgctgc tggacccgt 1020
 ggtgtactac tttagggcg agggcttccg caacacctg agcggtctg gcactccga 1080
 ccggggcagg acctcggcca caacgggac ggcggggcg ctgcgcact ccgaaggta 1140
 cgcctcacc accgacgcca ccaggccgga tgcgcaggt cgggggtgc tccgaccta 1200
 cgaetccacc tctctgctt ccttcacaca gtgtccccag gattccgct tctgaacaca 1260
 catgcccatt cgtctcctg gccgactac caacgctct cgtctctgga ggtctacag 1320
 gtgtacacac aagaaggtg gctgggact tggaccttg ggtggcaat ccagcttag 1380
 aaogcagag agtacaaat gtggaagcca gggccaggg aaggcagtc tctgggaal 1440
 ggtctcttta aactgtgag accgagaga ccccttccc agcggtggg agtgatgag 1500
 aaagccacc cgtgcagag gcagagagg acgaatgcc ttgggtggg cagggcatta 1560
 aactgtana agctggttag atgaccaga aaatgggcat tctgattttt aaccogcacc 1620
 aggggcttga ggttgaaga gcaacagggt tgggggcaa agctactgag atgctgttc 1680
 atctgctgac ttctgtctag gctcatggat gccacccctt tctatttgg cctaggcttc 1740
 cctgtctcac cactgagggc taatcacaga gtctctatgg ccgaactac attctctct 1800
 gcatagtgac ttgtgcaat ttgacttgg cctccagcat gggatagtg gggcaaggca 1860
 aaactaacct agatgtccc cctcaaacac atccagttc aaacctttt taggttatcc 1920

WO 00/23588

PCT/US99/24368

```

ttttttccat caccatccct ttccagggc tntccattt taggtctta acattcttc 1980
ttttttctc Ltctctgttt ctctctctc tctctctct tctctctct 2040
ctctctctct ctctctctt gtcctagta aggtctctt ctctctctt aagctctgt 2100
ctctctctt ttctctctc gctctctct ttctctctt aggtctctt aagctctgt 2160
ttctctctt ttctctctt actctctct ctctctctt aagctctct atctctctt 2220
tactctctc tctctctct actctctct ttctctctt aagctctct 2280
atctctctt ttctctctt ctctctctt ttctctctt aagctctct 2340
tactctctt ttctctctt ttctctctt ttctctctt ttctctctt 2400
gtctctctt ttctctctt ttctctctt ttctctctt ttctctctt 2460
ctctctctt ttctctctt ttctctctt ttctctctt ttctctctt 2520
atactctct ctctctctt actctctct ttctctctt ttctctctt 2580

```

<210> 7

<211> 269

<212> PRT

<213> Unknown

<220>

<223> Description of Unknown Organism: Rhodopsin family
transmembrane receptor

<400> 7

Gly Asn Ile Leu Val Ile Trp Val Ile Cys Arg Tyr Arg Arg Met Arg
1 5 10 15

Thr Pro Met Asn Tyr Phe Ile Val Asn Ile Ala Val Ala Asp Leu Leu
20 25 30

Phe Ser Leu Phe Thr Met Pro Phe Trp Met Val Tyr Tyr Val Met Gly
35 40 45

Gly Arg Trp Pro Phe Gly Asp Phe Met Cys Arg Ile Trp Met Tyr Phe
50 55 60

Asp Tyr Met Asn Met Tyr Ala Ser Ile Phe Phe Leu Thr Cys Ile Ser
65 70 75 80

Ile Asp Arg Tyr Leu Trp Ala Ile Cys His Pro Met Arg Tyr Met Arg
85 90 95

Trp Met Thr Pro Arg His Arg Ala Trp Val Met Ile Ile Ile Ile Trp
100 105 110

Val Met Ser Phe Leu Ile Ser Met Pro Pro Phe Leu Met Phe Arg Trp
115 120 125

Ser Thr Tyr Arg Asp Glu Asn Glu Trp Asn Met Thr Trp Cys Met Ile
130 135 140

8

WO 00/23588

PCT/US99/24368

Tyr Asp Trp Pro Gln Trp Met Trp Arg Trp Tyr Val Ile Leu Met Thr
 145 150 155 160
 Ile Ile Met Gly Phe Tyr Ile Pro Met Ile Ile Met Leu Phe Cys Tyr
 165 170 175
 Trp Arg Ile Tyr Arg Ile Ala Arg Leu Trp Met Arg Met Ile Pro Ser
 180 185 190
 Trp Gln Arg Arg Arg Arg Met Ser Met Arg Arg Glu Arg Arg Ile Val
 195 200 205
 Lys Met Leu Ile Ile Ile Met Val Val Phe Ile Ile Cys Trp Leu Pro
 210 215 220
 Tyr Phe Ile Val Met Phe Met Asp Thr Leu Met Met Trp Trp Phe Cys
 225 230 235 240
 Glu Phe Cys Ile Trp Arg Arg Leu Trp Met Tyr Ile Phe Glu Trp Leu
 245 250 255
 Ala Tyr Val Asn Cys Pro Cys Ile Asn Pro Ile Ile Tyr
 260 265

【国際公開パンフレット（コレクトバージョン）】

PCT		WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau	
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)			
(51) International Patent Classification ⁷ : C12N 15/12, C07K 14/705, 16/28, G01N 33/53, C12Q 1/68, A61K 31/70, 38/17, A01K 67/027		A3	(11) International Publication Number: WO 00/23588 (43) International Publication Date: 27 April 2000 (27.04.00)
(21) International Application Number: PCT/US99/24368 (22) International Filing Date: 18 October 1999 (18.10.99) (30) Priority Data: 09/173,899 16 October 1998 (16.10.98) US Not furnished 18 October 1999 (18.10.99) US (71) Applicant (for all designated States except US): MILLEN- NIUM PHARMACEUTICALS, INC. [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): GLUCKSMANN, Maria; Alexandra [AR/US]; 33 Summit Road, Lexington, MA 02173 (US); WEICHL, Nadine, S. [US/US]; 70 Park Street #53, Brookline, MA 02446 (US). (74) Agents: BROWN, Anne et al.; Alston & Bird LLP, P.O. Drawer 34009, Charlotte, NC 28234-4009 (US).		(81) Designated States: AE, AL, AM, AT, AU (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), DM, EE, EE (Utility model), ES, FI, FI (Utility model), G8, GU, GE, GH, GM, HR, HU, ID, IL, IS, IS, JP, KF, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SI), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published With international search report. (88) Date of publication of the international search report: 27 July 2000 (27.07.00)	
(54) Title: G-PROTEIN COUPLED RECEPTORS			
(57) Abstract The present invention relates to newly identified receptors belonging to the superfamily of G-protein-coupled receptors. The invention also relates to polynucleotides encoding the receptors. The invention further relates to methods using the receptor polypeptides and polynucleotides as a target for diagnosis and treatment in receptor-mediated disorders. The invention further relates to drug-screening methods using the receptor polypeptides and polynucleotides to identify agonists and antagonists for diagnosis and treatment. The invention further encompasses agonists and antagonists based on the receptor polypeptides and polynucleotides. The invention further relates to procedures for producing the receptor polypeptides and polynucleotides.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MG	Madagascar	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CJ	Chad d'Issine	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CN	China	KR	Republic of Korea	PL	Poland		
CU	Cuba	KZ	Kazakhstan	PT	Portugal		
CZ	Czech Republic	LC	Saint Lucia	RO	Romania		
DE	Germany	LI	Liechtenstein	RU	Russian Federation		
DK	Denmark	LK	Sri Lanka	SD	Sudan		
EE	Estonia	LT	Lithuania	SE	Sweden		
				SG	Singapore		

【国際公開パンフレット（コレクトバージョン）】

PCT		WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau	
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)			
(51) International Patent Classification 7 : C12N 15/12, C07K 14/705, 16/28, G01N 33/53, C12Q 1/68, A61K 31/70, 38/17, A01K 67/027		A3	(11) International Publication Number: WO 00/23588 (43) International Publication Date: 27 April 2000 (27.04.00)
(21) International Application Number: PCT/US99/24368 (22) International Filing Date: 18 October 1999 (18.10.99) (30) Priority Data: 09/173,869 16 October 1998 (16.10.98) US Not furnished 18 October 1999 (18.10.99) US (71) Applicant (for all designated States except US): MILLEN- NIUM PHARMACEUTICALS, INC. [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US). (72) Inventors; and (73) Inventors/Applicants (for US only): GLUCKSMANN, Maria, Alexandra [AR/US]; 33 Summit Road, Lexington, MA 02173 (US); WEICH, Nadine, S. [US/US]; 70 Park Street #53, Brookline, MA 02446 (US). (74) Agents: BROWN, Anne et al.; Alston & Bird LLP, P.O. Drawer 34009, Charlotte, NC 28234-4009 (US).		(81) Designated States: AE, AI, AM, AT, AU (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), DM, EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SI, TI, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published With international search report. With amended claims. (88) Date of publication of the international search report: 27 July 2000 (27.07.00) Date of publication of the amended claims: 14 September 2000 (14.09.00)	
(54) Title: G-PROTEIN COUPLED RECEPTORS			
(57) Abstract The present invention relates to newly identified receptors belonging to the superfamily of G-protein-coupled receptors. The invention also relates to polynucleotides encoding the receptors. The invention further relates to methods using the receptor polypeptides and polynucleotides as a target for diagnosis and treatment in receptor-mediated disorders. The invention further relates to drug-screening methods using the receptor polypeptides and polynucleotides to identify agonists and antagonists for diagnosis and treatment. The invention further encompasses agonists and antagonists based on the receptor polypeptides and polynucleotides. The invention further relates to procedures for producing the receptor polypeptides and polynucleotides.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	NM	Nauru	UG	Uganda
BY	Belarus	IN	India	NW	Norway	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CG	Congo	JP	Japan	NE	Niger	VN	Viet Nam
CH	Switzerland	KZ	Kazakhstan	NL	Netherlands	YU	Yugoslavia
CI	Côte d'Ivoire	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CM	Cameroun	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CN	China	KR	Republic of Korea	PL	Poland		
CU	Cuba	KZ	Kazakhstan	PT	Portugal		
CZ	Czech Republic	LC	Saint Lucia	RO	Romania		
DE	Germany	LI	Liechtenstein	RU	Russian Federation		
DK	Denmark	LK	Sri Lanka	SD	Sudan		
EE	Estonia	LR	Liberia	SE	Sweden		
				SG	Singapore		

【国際公開パンフレット（コレクトバージョン）】

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

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
27 April 2000 (27.04.2000)

PCT

(10) International Publication Number
WO 00/23588 A3(51) International Patent Classification: C12N 15/12,
C07K 14/705, 16/28, G01N 33/53, C12Q 1/68, A61K
31/70, 39/17, A01K 67/027

(21) International Application Number: PCT/US99/24368

(22) International Filing Date: 18 October 1999 (18.10.1999)

(25) Filing Language: English

(26) Publication Language: English

DK, DK (utility model), DM, EE, EE (utility model), ES,
FI, FI (utility model), GB, GD, GE, GH, GM, HK, HU, ID,
IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT,
LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL,
PT, RO, RU, SD, SE, SG, SI, SK, SK (utility model), SL,
TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW,(84) Designated States (epimod): ARIPPO patent (GH, GM,
KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent
(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,
MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM,
GA, GN, GW, ML, MR, NE, SN, TD, TG).(30) Priority Data:
09/173,869 16 October 1998 (16.10.1998) US
09/420,187 18 October 1999 (18.10.1999) USPublished:
With international search report.
With amended claims.(71) Applicant (for all designated States except US): MIL-
LENNIUM PHARMACEUTICALS, INC. [US/US]; 75
Sidney Street, Cambridge, MA 02139 (US).(88) Date of publication of the international search report:
27 July 2000(72) Inventors; and
(75) Inventors/Applicants (for US only): GLUCKSMANN,
Maria, Alexandra [AR/US]; 33 Summit Road, Lexington,
MA 02173 (US), WEICHI, Nadine, S. [US/US]; 70 Park
Street #53, Brookline, MA 02446 (US).Date of publication of the amended claims:
14 September 2000(74) Agents: BROWN, Anne et al.: Alston & Bird LLP, P.O.
Drawer 34009, Charlotte, NC 28234-4009 (US).(49) Date of publication of this corrected version:
8 March 2001(81) Designated States (national): AE, AL, AM, AT, AT (util-
ity model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN,
CR, CU, CZ, CZ (utility model), DE, DE (utility model),(15) Information about Correction:
see PCT Gazette No. 10/2001 of 8 March 2001, Section II.
For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 00/23588 A3

(54) Title: G-PROTEIN COUPLED RECEPTORS

(57) Abstract: The present invention relates to newly identified receptors belonging to the superfamily of G-protein-coupled recep-
tors. The invention also relates to polynucleotides encoding the receptors. The invention further relates to methods using the receptor
polypeptides and polynucleotides as a target for diagnosis and treatment in receptor-mediated disorders. The invention further relates
to drug-screening methods using the receptor polypeptides and polynucleotides to identify agonists and antagonists for diagnosis and
treatment. The invention further encompasses agonists and antagonists based on the receptor polypeptides and polynucleotides. The
invention further relates to procedures for producing the receptor polypeptides and polynucleotides.

【国際公開パンフレット（コレクトバージョン）】

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

[CORRECTED VERSION]

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
27 April 2000 (27.04.2000)

PCT

(16) International Publication Number
WO 00/23588 A3(51) International Patent Classification: C12N 15/12,
C07K 14/705, 16/28, C01N 33/53, C12Q 1/68, A61K
51/70, 38/17, A61K 67/027

FI, FL (utility model), GB, GD, GE, GH, GM, HR, HU, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RD, RU, SD, SE, SG, SI, SK, SL (utility model), SM, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(21) International Application Number: PCT/US99/24368

(22) International Filing Date: 18 October 1999 (18.10.1999)

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 09/173,869 16 October 1998 (16.10.1998) US
09/420,187 18 October 1999 (18.10.1999) USPublished:
— with international search report
— with amended claims(71) Applicant (for all designated States except US): MIL-
LENNIUM PHARMACEUTICALS, INC. (US/US): 75
Sidney Street, Cambridge, MA 02139 (US).(88) Date of publication of the international search report:
27 July 2000Date of publication of the amended claims:
14 September 2000(72) Inventors; and
(75) Inventors/Applicants (for US only): GLUCKSMANN,
Maria, Alexandra (AR/US); 33 Summit Road, Lexington,
MA 02173 (US); WEICH, Nadine, S. (US/US); 70 Park
Street #53, Brookline, MA 02446 (US).(48) Date of publication of this corrected version:
25 April 2002(74) Agent: SPROLL, Murray, W.; Alston & Bird LLP, P.O.
Drawer 34009, Charlotte, NC 28234-0009 (US).(15) Information about Corrections:
see PCT Gazette No. 17/2002 of 25 April 2002, Section II
Previous Correction:
see PCT Gazette No. 10/2001 of 8 March 2001, Section II

(81) Designated States (national): AE, AL, AM, AT, AU (utility model), AU, AZ, BA, BR, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (utility model), DE, DE (utility model), DK, DK (utility model), DM, EE, EG (utility model), ES,

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

WO 00/23588 A3

(54) Title: G-PROTEIN COUPLED RECEPTORS

(57) Abstract: The present invention relates to newly identified receptors belonging to the superfamily of G-protein-coupled receptors. The invention also relates to polynucleotides encoding the receptors. The invention further relates to methods using the receptor polypeptides and polynucleotides as a target for diagnosis and treatment in receptor-mediated disorders. The invention further relates to drug-screening methods using the receptor polypeptides and polynucleotides to identify agonists and antagonists for diagnosis and treatment. The invention further encompasses agonists and antagonists based on the receptor polypeptides and polynucleotides. The invention further relates to procedures for producing the receptor polypeptides and polynucleotides.

WO 01/23588

PCT/US99/24368

G-PROTEIN COUPLED RECEPTORS

FIELD OF THE INVENTION

The present invention relates to newly identified receptors belonging to the superfamily of G-protein-coupled receptors. The invention also relates to polynucleotides encoding the receptors. The invention further relates to methods using the receptor polypeptides and polynucleotides as a target for diagnosis and treatment in receptor-mediated or -related disorders. The invention further relates to drug-screening methods using the receptor polypeptides and polynucleotides to identify agonists and antagonists for diagnosis and treatment. The invention further encompasses agonists and antagonists based on the receptor polypeptides and polynucleotides. The invention further relates to procedures for producing the receptor polypeptides and polynucleotides.

BACKGROUND OF THE INVENTION

15

G-protein coupled receptors

G-protein coupled receptors (GPCRs) constitute a major class of proteins responsible for transducing a signal within a cell. GPCRs have three structural domains: an amino terminal extracellular domain, a transmembrane domain containing seven transmembrane segments, three extracellular loops, and three intracellular loops, and a carboxy terminal intracellular domain. Upon binding of a ligand to an extracellular portion of a GPCR, a signal is transduced within the cell that results in a change in a biological or physiological property of the cell. GPCRs, along with G-proteins and effectors (intracellular enzymes and channels modulated by G-proteins), are the components of a modular signaling system that connects the state of intracellular second messengers to extracellular inputs.

GPCR genes and gene-products are potential causative agents of disease (Spiegel *et al.* (1993) *J. Clin. Invest.* 92:1119-1125; McKusick *et al.* (1993) *J. Med.*

WO 00/23588

PCT/US99/24368

Genet. 30:1-26) Specific defects in the rhodopsin gene and the V2 vasopressin receptor gene have been shown to cause various forms of retinitis pigmentosa (Nathans *et al.* (1992) *Annu. Rev. Genet.* 26:403-424), and nephrogenic diabetes insipidus (Holzman *et al.* (1993) *Hum. Mol. Genet.* 2:1201-1204). These receptors are of critical importance to both the central nervous system and peripheral physiological processes. Evolutionary analyses suggest that the ancestor of these proteins originally developed in concert with complex body plans and nervous systems.

The GPCR protein superfamily can be divided into five families: Family I, receptors typified by rhodopsin and the β_2 -adrenergic receptor and currently represented by over 200 unique members (Dohlman *et al.* (1991) *Annu. Rev. Biochem.* 60:653-688); Family II, the parathyroid hormone/calcitonin/secretin receptor family (Juppner *et al.* (1991) *Science* 254:1024-1026; Lin *et al.* (1991) *Science* 254:1022-1024); Family III, the metabotropic glutamate receptor family (Nakanishi (1992) *Science* 258:597-603); Family IV, the cAMP receptor family, important in the chemotaxis and development of *D. discoideum* (Klein *et al.* (1988) *Science* 241:1467-1472); and Family V, the fungal mating pheromone receptors such as STE2 (Kurjan (1992) *Annu. Rev. Biochem.* 61:1097-1129).

There are also a small number of other proteins which present seven putative hydrophobic segments and appear to be unrelated to GPCRs; they have not been shown to couple to G-proteins. *Drosophila* expresses a photoreceptor-specific protein, bride of sevenless (boss), a seven-transmembrane-segment protein which has been extensively studied and does not show evidence of being a GPCR (Hart *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:5047-5051). The gene *frizzled* (*fs*) in *Drosophila* is also thought to be a protein with seven transmembrane segments. Like boss, *fs* has not been shown to couple to G-proteins (Vinson *et al.* (1989) *Nature* 338:263-264).

G proteins represent a family of heterotrimeric proteins composed of α , β and γ subunits, that bind guanine nucleotides. These proteins are usually linked to cell surface receptors, e.g., receptors containing seven transmembrane segments. Following ligand binding to the GPCR, a conformational change is transmitted to the G protein, which causes the α -subunit to exchange a bound GDP molecule for a GTP molecule and to dissociate from the $\beta\gamma$ -subunits. The GTP-bound form of the α -subunit typically functions as an effector-modulating moiety, leading to the

WO 00/23588

PCT/US99/24368

production of second messengers, such as cAMP (e.g., by activation of adenylyl cyclase), diacylglycerol or inositol phosphates. Greater than 20 different types of α -subunits are known in humans. These subunits associate with a smaller pool of β and γ subunits. Examples of mammalian G proteins include Gi, Go, Gq, Gs and Gt. G proteins are described extensively in Lodish *et al.* *Molecular Cell Biology*, (Scientific American Books Inc., New York, N.Y., 1995), the contents of which are incorporated herein by reference. GPCRs, G proteins and G protein-linked effector and second messenger systems have been reviewed in *The G-Protein Linked Receptor Fact Book*, Watson *et al.* eds., Academic Press (1994).

Purinoceptors

Purines, and especially adenosine and adenine nucleotides, have a broad range of pharmacological effects mediated through cell-surface receptors. For a general review, see Adenosine and Adenine Nucleotides in *The G-Protein Linked Receptor Facts Book*, Watson *et al.* (Eds.) Academic Press 1994, pp. 19-31.

Some effects of ATP include the regulation of smooth muscle activity, stimulation of the relaxation of intestinal smooth muscle and bladder contraction, stimulation of platelet activation by ADP when released from vascular endothelium, and excitatory effects in the central nervous system. Some effects of adenosine include vasodilation, bronchoconstriction, immunosuppression, inhibition of platelet aggregation, cardiac depression, stimulation of nociceptive afferents, inhibition of neurotransmitter release, pre- and postsynaptic depressant action, reducing motor activity, depressing respiration, inducing sleep, relieving anxiety, and inhibition of release of factors, such as hormones.

Distinct receptors exist for adenosine and adenine nucleotides. Clinical actions of such analogs as methylxanthines, for example, theophylline and caffeine, are thought to achieve their effects by antagonizing adenosine receptors. Adenosine has a low affinity for adenine nucleotide receptors, while adenine nucleotides have a low affinity for adenosine receptors.

There are four accepted subtypes of adenosine receptors, designated A₁, A_{2A}, A_{2B}, and A₃. In addition, an A₄ receptor has been proposed based on labeling by 2-phenylaminoadenosine (Cornfield *et al.* (1992) *Mol. Pharmacol.* 42:552-561).

WO 00/23588

PCT/US99/24368

P_{2X} receptors are ATP-gated cation channels (See *Neuropharmacology* 36 (1977)). The proposed topology for P_{2X} receptors is two transmembrane regions, a large extracellular loop, and intracellular N and C-termini.

Numerous cloned receptors designated P_{2Y} have been proposed to be members of the G-protein coupled family. UDP, UTP, ADP, and ATP have been identified as agonists. To date, P_{2Y1-7} have been characterized although it has been proposed that P_{2Y7} may be a leukotriene B₄ receptor (Yokomizo *et al.* (1997) *Nature* 387:620-624). It is widely accepted, however, that P_{2Y1}, 2, 4, and 6 are members of the G-protein coupled family of P_{2Y} receptors.

At least three P₂ purinoceptors from the hematopoietic cell line HEL have been identified by intracellular calcium mobilization and by photoaffinity labeling (Akbar *et al.* (1996) *J. Biochem.* 271:18363-18367).

The A₁ adenosine receptor was designated in view of its ability to inhibit adenylyl cyclase. The receptors are distributed in many peripheral tissues such as heart, adipose, kidney, stomach and pancreas. They are also found in peripheral nerves, for example intestine and vas deferens. They are present in high levels in the central nervous system, including cerebral cortex, hippocampus, cerebellum, thalamus, and striatum, as well as in several cell lines. Agonists and antagonists can be found on page 22 of *The G-Protein Linked Receptor Facts Book* cited above, herein incorporated by reference. These receptors are reported to inhibit adenylyl cyclase and voltage-dependent calcium channels and to activate potassium channels through a pertussis-toxin-sensitive G-protein suggested to be of the G_i/G_o class. A₁ receptors have also been reported to induce activation of phospholipase C and to potentiate the ability of other receptors to activate this pathway.

The A_{2A} adenosine receptor has been found in brain, such as striatum, olfactory tubercle and nucleus accumbens. In the periphery, A₂ receptors mediate vasodilation, immunosuppression, inhibition of platelet aggregation, and gluconeogenesis. Agonists and antagonists are found in *The G-Protein Linked Receptor Facts Book* cited above on page 25, herein incorporated by reference. This receptor mediates activation of adenylyl cyclase through G_s.

The A_{2B} receptor has been shown to be present in human brain and in rat intestine and urinary bladder. Agonists and antagonists are discussed on page 27 of *The G-Protein Linked Receptor Facts Book* cited above, herein incorporated by reference. This receptor mediates the stimulation of cAMP through G_s.

WO 00/23588

PCT/US99/24368

The A₁ adenosine receptor is expressed in testes, lung, kidney, heart, central nervous system, including cerebral cortex, striatum, and olfactory bulb. A discussion of agonists and antagonists can be found on page 28 of *The G-Protein Linked Receptor Facts Book* cited above, herein incorporated by reference. The receptor
 5 mediates the inhibition of adenylylase through a pertussis-toxin-sensitive G-protein, suggested to be of the G_i/G_o class.

The P_{2Y} purinoceptor shows a similar affinity for ATP and ADP with a lower affinity for AMP. The receptor has been found in smooth muscle, for example, taeni caeci and in vascular tissue where it induces vasodilation through endothelium-
 10 dependent release of nitric oxide. It has also been shown in avian erythrocytes. Agonists and antagonists are discussed on page 30 of *The G-Protein Linked Receptor Facts Book* cited above, herein incorporated by reference. The receptor function through activation of phosphoinositide metabolism through a pertussis-toxin-insensitive G-protein, suggested to be of the G_i/G_o class.

15 Accordingly, GPCRs, and especially purinoceptors, are a major target for drug action and development. Accordingly, it is valuable to the field of pharmaceutical development to identify and characterize previously unknown GPCRs. The present invention advances the state of the art by providing a previously unidentified human GPCR having homology to purinoceptors.

20

SUMMARY OF THE INVENTION

It is an object of the invention to identify novel GPCRs.

25 It is a further object of the invention to provide novel GPCR polypeptides that are useful as reagents or targets in receptor assays applicable to treatment and diagnosis of GPCR-mediated disorders.

It is a further object of the invention to provide polynucleotides corresponding to the novel GPCR receptor polypeptides that are useful as targets and reagents in receptor assays applicable to treatment and diagnosis of GPCR-mediated disorders
 30 and useful for producing novel receptor polypeptides by recombinant methods.

A specific object of the invention is to identify compounds that act as agonists and antagonists and modulate the expression of the novel receptors.

WO 00/23588

PCT/US99/24368

A further specific object of the invention is to provide compounds that modulate expression of the receptors for treatment and diagnosis of GPCR- related disorders.

The invention is thus based on the identification of novel GPCRs, designated the 2838, 14618, and 15334 receptors.

The invention provides isolated 2838 receptor polypeptides including a polypeptide having the amino acid sequence shown in SEQ ID NO.1

The invention provides isolated 14618 receptor polypeptides including a polypeptide having the amino acid sequence shown in SEQ ID NO.3.

The invention provides isolated 15334 receptor polypeptides including a polypeptide having the amino acid sequence shown in SEQ ID NO.5

The invention also provides isolated 2838 receptor nucleic acid molecules having the sequence shown in SEQ ID NO.2

The invention also provides isolated 14618 receptor nucleic acid molecules having the sequence shown in SEQ ID NO.4.

The invention also provides isolated 15334 receptor nucleic acid molecules having the sequence shown in SEQ ID NO.6

The invention also provides variant polypeptides having an amino acid sequence that is substantially homologous to the amino acid sequence shown in SEQ ID NO.1.

The invention also provides variant polypeptides having an amino acid sequence that is substantially homologous to the amino acid sequence shown in SEQ ID NO.3.

The invention also provides variant polypeptides having an amino acid sequence that is substantially homologous to the amino acid sequence shown in SEQ ID NO.5.

The invention also provides variant nucleic acid sequences that are substantially homologous to the nucleotide sequence shown in SEQ ID NO.2.

The invention also provides variant nucleic acid sequences that are substantially homologous to the nucleotide sequence shown in SEQ ID NO.4.

The invention also provides variant nucleic acid sequences that are substantially homologous to the nucleotide sequence shown in SEQ ID NO.6

WO 00/23588

PCT/US99/24368

The invention also provides fragments of the polypeptide shown in SEQ ID NO:1 and nucleotide sequence shown in SEQ ID NO:2, as well as substantially homologous fragments of the polypeptide or nucleic acid.

5 The invention also provides fragments of the polypeptide shown in SEQ ID NO:3 and nucleotide sequence shown in SEQ ID NO:4, as well as substantially homologous fragments of the polypeptide or nucleic acid.

The invention also provides fragments of the polypeptide shown in SEQ ID NO:5 and nucleotide sequence shown in SEQ ID NO:6, as well as substantially homologous fragments of the polypeptide or nucleic acid.

10 The invention further provides nucleic acid constructs comprising the nucleic acid molecules described above. In a preferred embodiment, the nucleic acid molecules of the invention are operatively linked to a regulatory sequence.

The invention also provides vectors and host cells for expressing the receptor nucleic acid molecules and polypeptides and particularly recombinant vectors and
15 host cells.

The invention also provides methods of making the vectors and host cells and methods for using them to produce the receptor nucleic acid molecules and polypeptides.

20 The invention also provides antibodies or antigen-binding fragments thereof that selectively bind the receptor polypeptides and fragments.

The invention also provides methods of screening for compounds that modulate expression or activity of the receptor polypeptides or nucleic acid (RNA or DNA).

The invention also provides a process for modulating receptor polypeptide or nucleic acid expression or activity, especially using the screened compounds.
25 Modulation may be used to treat conditions related to aberrant activity or expression of the receptor polypeptides or nucleic acids.

The invention also provides assays for determining the presence or absence of and level of the receptor polypeptides or nucleic acid molecules in a biological sample, including for disease diagnosis.

30 The invention also provides assays for determining the presence of a mutation in the receptor polypeptides or nucleic acid molecules, including for disease diagnosis.

In still a further embodiment, the invention provides a computer readable means containing the nucleotide and/or amino acid sequences of the nucleic acids and polypeptides of the invention, respectively.

WO 00/23588

PCT/US99/24368

DESCRIPTION OF THE DRAWINGS

Figure 1 shows a comparison of the 2838 receptor against the Prosite data base of protein patterns, specifically showing a high score against the seven transmembrane segment rhodopsin superfamily. The underlined area shows a GPCR signature, and specifically the position of an arginine residue, conserved in GPCRs. The most commonly conserved sequence is an aspartate, arginine, tyrosine (DRY) triplet. DRY is implicated in signal transduction. Arginine is invariant. Aspartate is conservatively placed in several GPCRs. In the present case, the arginine is found in the sequence DRI¹, which matches the position of DRY or invariant arginine in GPCRs of the rhodopsin superfamily of receptors.

Figure 2 shows an analysis of the 2838 amino acid sequence: α turn and coil regions; hydrophilicity; amphipathic regions; flexible regions; antigenic index; and surface probability plot.

Figure 3 shows a 2838 receptor hydrophobicity plot. The amino acids correspond to 1-319 and show the seven transmembrane segments.

Figure 4 shows an analysis of the 2838 open reading frame for amino acids corresponding to specific functional sites. A glycosylation site is found at amino acids 5-8 which corresponds to the amino terminal extracellular domain. A second glycosylation site is found at amino acids 171-174 which corresponds to the second extracellular loop. A protein kinase C phosphorylation site is found at amino acids 134-136 which is in the second intracellular loop. A second protein kinase C phosphorylation site is found at amino acids 178-180 which is in the second extracellular loop. A casein kinase II phosphorylation site is found at amino acids 6-9 which is in the carboxy terminal intracellular domain. A second casein kinase II phosphorylation site is found at amino acids 95-98 which is in the first extracellular loop. An N-myristoylation site is found at amino acids 34-39 which is in the first transmembrane segment. A second N-myristoylation site is found at amino acids 107-112 which is in the third transmembrane segment. A third N-myristoylation site is found at amino acids 140-145 which is in the fourth transmembrane segment. An

WO 00/23588

PCT/US99/24368

amidation site is found at amino acids 209-212 which spans the fifth transmembrane segment and third intracellular loop. In addition, amino acids corresponding in position to the GPCR signature and containing the invariant arginine are found in the sequence DRF at amino acids 118-120.

5

Figure 5 shows a comparison of the 14618 receptor against the Prosite data base of protein patterns, specifically showing a high score against the seven transmembrane segment rhodopsin superfamily. The underlined area shows a GPCR signature, and specifically the position of an arginine residue, conserved in GPCRs.

10 The most commonly conserved sequence is an aspartate, arginine, tyrosine (DRY) triplet. DRY is implicated in signal transduction. Arginine is invariant. Aspartate is conservatively placed in several GPCRs. In the present case, the arginine is found in the sequence DRF, which matches the position of DRY or invariant arginine in GPCRs of the rhodopsin superfamily of receptors.

15

Figure 6 shows an analysis of the 14618 amino acid sequence: α turn and coil regions; hydrophobicity; amphipathic regions; flexible regions; antigenic index; and surface probability plot.

20

Figure 7 shows a 14618 receptor hydrophobicity plot. The amino acids correspond to 1-337 and show the seven transmembrane segments.

Figure 8 shows an analysis of the 14618 open reading frame for amino acids corresponding to specific functional sites. A glycosylation site is found at amino acids 6-9 which corresponds to the amino terminal extracellular domain. A second glycosylation site is found at amino acids 169-172 which corresponds to the second extracellular loop. A third glycosylation site is found at amino acids 180-183 which also corresponds to the second extracellular loop. A fourth glycosylation site found at amino acids 224-227 which corresponds to the third intracellular loop. A fifth glycosylation site is found at amino acids 262-265 which corresponds to the third extracellular loop. A cAMP- and cGMP-dependent protein kinase phosphorylation site is found at amino acids 304-307, 310-313, and 323-326, all in the carboxy terminal intracellular domain. A protein kinase C phosphorylation site is found at amino acids 136-138 which corresponds to the second intracellular loop. Second and

25

30

WO 00/23588

PCT/US99/24368

third protein kinase C phosphorylation sites are found at amino acids 220-222 and 227-229, both corresponding to the third intracellular loop. A fourth protein kinase C phosphorylation site is found at amino acids 308-310, corresponding to the carboxy terminal intracellular domain. Casein kinase II phosphorylation sites are found at amino acids 13-16 and 17-20, both in the amino terminal extracellular domain. A third casein kinase II phosphorylation site is found at amino acids 326-329 corresponding to the carboxy terminal intracellular domain. A microbodies C-terminal targeting signal is found at amino acids 335-338, corresponding to the carboxy terminal intracellular domain. In addition, amino acids corresponding in position to the GPCR signature and containing the invariant arginine are found in the sequence FRC at amino acids 120-122.

Figure 9 shows a comparison of the 15334 receptor against the Prosite data base of protein patterns, specifically showing a high score against the seven transmembrane segment rhodopsin superfamily. The underlined area shows a GPCR signature, and specifically the position of an arginine residue, conserved in GPCRs. The most commonly conserved sequence is an aspartate, arginine, tyrosine (DRY) triplet. DRY is implicated in signal transduction. Arginine is invariant. Aspartate is conservatively placed in several GPCRs. In the present case, the arginine is found in the sequence DRY, which matches the position of DRY or invariant arginine in GPCRs of the rhodopsin superfamily of receptors.

Figure 10 shows an analysis of the 15334 amino acid sequence: α turn and coil regions; hydrophobicity; amphipathic regions; flexible regions; antigenic index; and surface probability plot.

Figure 11 shows a 15334 receptor hydrophobicity plot. The amino acids show the seven transmembrane segments.

Figure 12 shows an analysis of the 15334 open reading frame for amino acids corresponding to specific functional sites. Glycosylation sites are found at amino acids 4-7 and 9-12, which are in the amino terminal extracellular domain. A glycosylation site is also found at amino acids 251-254, which is in the sixth transmembrane segment. A further glycosylation site is found from amino acid 323-

WO 00/23588

PCT/US99/24368

326, which is in the carboxy terminal domain. A cAMP- and cGMP-dependent protein kinase phosphorylation site is found at amino acids 229-232, which is in the third intracellular loop. Protein kinase C phosphorylation sites are found from amino acids 21-23, which corresponds to the amino terminal domain, 211-213, 226-228, and 232-234, which corresponds to the third intracellular loop, and 307-309 and 332-334, which corresponds to the carboxy terminal intracellular domain. A casein kinase II phosphorylation site is found from amino acids 178-181, which is in the second extracellular loop and from 342-345, which is in the carboxy terminal intracellular domain. N-myristoylation sites are found at amino acids 36-41, which is in the amino terminal extracellular domain, 258-263, which spans the sixth transmembrane segment and third extracellular loop, and 324-329, which corresponds to the carboxy terminal intracellular domain. In addition, a GPCR signature, DRY, is found at amino acids 118-120, which is in the second intracellular loop.

15 **Figure 13** shows expression of the 14618 receptor in various normal human tissues.

Figure 14 shows the 14618 receptor expression in various tissues and in cells of the hematopoietic system.

20 **Figure 15** shows expression of the 2838 receptor in various cells and tissues, including cells of the hematopoietic system.

Figure 16 shows expression of the 15334 receptor in normal human tissues.

25 **Figure 17** shows expression of the 15334 receptor in various tissues and in cells of the hematopoietic system.

Figure 18 shows expression of the 15334 receptor in various hematopoietic cells/cell lines, including erythroid and megakaryocyte cells.

30 **Figure 19** shows expression of the 15334 receptor in various hematopoietic cells or cell lines, including erythroid and megakaryocyte cells.

WO 00/23588

PCT/US99/24368

DETAILED DESCRIPTION OF THE INVENTION

Receptor function/signal pathway

The 2838, 14618, and 15334 receptor proteins are GPCRs that participate in signaling pathways. As used herein, a "signaling pathway" refers to the modulation (e.g., stimulation or inhibition) of a cellular function/activity upon the binding of a ligand to the GPCR (2838, 14618, or 15334 protein). Examples of such functions include mobilization of intracellular molecules that participate in a signal transduction pathway, e.g., phosphatidylinositol 4,5-bisphosphate (PIP₂), inositol 1,4,5-triphosphate (IP₃) and adenylate cyclase; polarization of the plasma membrane; production or secretion of molecules; alteration in the structure of a cellular component; cell proliferation, e.g., synthesis of DNA; cell migration; cell differentiation; and cell survival. Since the 2838 receptor protein is expressed in thymus, lymph node, spleen, testes, colon, and peripheral blood lymphocytes including but not limited to activated T helper cells-1, activated T-helper cells-2, CD3(both CD4 and CD8), activated B cells, and granulocytes, cells participating in a 2838 receptor protein signaling pathway include, but are not limited to, these tissues and cells. Since the 14618 receptor protein is expressed in breast, skeletal muscle, thyroid, lymph node, spleen, and peripheral blood lymphocytes including, but not limited to, CD34⁺ cells, resting B cells, and megakaryocytes, cells participating in a 14618 receptor protein signaling pathway include, but are not limited to, cells derived from these tissues and cells. Since the 15334 receptor protein is expressed in colon, pancreas, tonsil, lymph node, spleen, thymus, adrenal gland, heart, and peripheral blood cells, including but not limited to those shown in Figure 17, megakaryocytes, and erythroblasts, cells participating in a 15334 receptor protein signaling pathway include, but are not limited to, cells derived from these tissues and cells.

The response mediated by the receptor protein depends on the type of cell. For example, in some cells, binding of a ligand to the receptor protein may stimulate an activity such as release of compounds, gating of a channel, cellular adhesion, migration, differentiation, etc., through phosphatidylinositol or cyclic AMP metabolism and turnover while in other cells, the binding of the ligand will produce a different result. Regardless of the cellular activity/response modulated by the receptor protein, it is universal that the protein is a GPCR and interacts with G proteins to produce one or more secondary signals, in a variety of intracellular signal

WO 00/23588

PCT/US99/24368

transduction pathways, e.g., through phosphatidylinositol or cyclic AMP metabolism and turnover, in a cell.

As used herein, "phosphatidylinositol turnover and metabolism" refers to the molecules involved in the turnover and metabolism of phosphatidylinositol 4,5-bisphosphate (PIP₂) as well as to the activities of these molecules. PIP₂ is a phospholipid found in the cytosolic leaflet of the plasma membrane. Binding of ligand to the receptor activates, in some cells, the plasma-membrane enzyme phospholipase C that in turn can hydrolyze PIP₂ to produce 1,2-diacylglycerol (DAG) and inositol 1,4,5-triphosphate (IP₃). Once formed IP₃ can diffuse to the endoplasmic reticulum surface where it can bind an IP₃ receptor, e.g., a calcium channel protein containing an IP₃ binding site. IP₃ binding can induce opening of the channel, allowing calcium ions to be released into the cytoplasm. IP₃ can also be phosphorylated by a specific kinase to form inositol 1,3,4,5-tetraphosphate (IP₄), a molecule which can cause calcium entry into the cytoplasm from the extracellular medium. IP₃ and IP₄ can subsequently be hydrolyzed very rapidly to the inactive products inositol 1,4-bisphosphate (IP₂) and inositol 1,3,4-triphosphate, respectively. These inactive products can be recycled by the cell to synthesize PIP₂. The other second messenger produced by the hydrolysis of PIP₂, namely 1,2-diacylglycerol (DAG), remains in the cell membrane where it can serve to activate the enzyme protein kinase C. Protein kinase C is usually found soluble in the cytoplasm of the cell, but upon an increase in the intracellular calcium concentration, this enzyme can move to the plasma membrane where it can be activated by DAG. The activation of protein kinase C in different cells results in various cellular responses such as the phosphorylation of glycogen synthase, or the phosphorylation of various transcription factors, e.g., NF- κ B. The language "phosphatidylinositol activity", as used herein, refers to an activity of PIP₂ or one of its metabolites.

Another signaling pathway in which the receptor may participate is the cAMP turnover pathway. As used herein, "cyclic AMP turnover and metabolism" refers to the molecules involved in the turnover and metabolism of cyclic AMP (cAMP) as well as to the activities of these molecules. Cyclic AMP is a second messenger produced in response to ligand-induced stimulation of certain G protein coupled receptors. In the cAMP signaling pathway, binding of a ligand to a GPCR can lead to the activation of the enzyme adenylyl cyclase, which catalyzes the synthesis of cAMP. The newly synthesized cAMP can in turn activate a cAMP-dependent protein kinase.

WO 00/23588

PCT/US99/24368

This activated kinase can phosphorylate a voltage-gated potassium channel protein, or an associated protein, and lead to the inability of the potassium channel to open during an action potential. The inability of the potassium channel to open results in a decrease in the outward flow of potassium, which normally repolarizes the membrane of a neuron, leading to prolonged membrane depolarization.

Polypeptides

The invention is based on the discovery of novel G-coupled protein receptors. Specifically, an expressed sequence tag (EST) was selected based on homology to G-protein-coupled receptor sequences. This EST was used to design primers based on sequences that it contains and used to identify a 2838 cDNA from a B cell cDNA library, a 14618 cDNA from a liver and spleen cDNA library, and a 15334 cDNA from a spleen cDNA library. Positive clones were sequenced and the overlapping fragments were assembled. Analysis of the assembled sequences revealed that the three cloned cDNA molecules encode G-protein coupled receptors.

The invention thus relates to a novel GPCR having the deduced amino acid sequence shown in SEQ ID NO:1.

The invention also thus relates to a novel GPCR having the deduced amino acid sequence shown in SEQ ID NO:3.

The invention also thus relates to a novel GPCR having the deduced amino acid sequence shown in SEQ ID NO:5.

The Sequence Listing shows the 2838 nucleotide sequence (SEQ ID NO:2) and the deduced 2838 amino acid sequence (SEQ ID NO:1). It is predicted that amino acids 1 to about 24 constitute the amino terminal extracellular domain, amino acids about 25-292 constitute the region spanning the transmembrane domain, and amino acids about 293-319 constitute the carboxy terminal intracellular domain. The transmembrane domain contains seven transmembrane segments, three extracellular loops and three intracellular loops. The transmembrane segments are found from about amino acid 25 to about amino acid 49, from about amino acid 56 to about amino acid 79, from about amino acid 100 to about amino acid 117, from about amino acid 138 to about amino acid 159, from about amino acid 187 to about amino acid 210, from about amino acid 224 to about amino acid 248, and from about amino acid 268 to about amino acid 292. Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular

WO 00/23588

PCT/US99/24368

loops are found from about amino acid 50 to about amino acid 55, from about amino acid 118 to about amino acid 137, and from about amino acid 211 to about amino acid 223. The three extracellular loops are found at from about amino acid 80 to about amino acid 99, from about amino acid 160 to about amino acid 186, and from about amino acid 249 to about amino acid 267.

The Sequence Listing shows the 14618 nucleotide sequence (SEQ ID NO:4) and the deduced 14618 amino acid sequence (SEQ ID NO:3). It is predicted that amino acids 1 to about 28 constitute the amino terminal extracellular domain, amino acids about 29-297 constitute the region spanning the transmembrane domain, and amino acids about 298-337 constitute the carboxy terminal intracellular domain. The transmembrane domain contains seven transmembrane segments, three extracellular loops and three intracellular loops. The transmembrane segments are found from about amino acid 29 to about amino acid 49, from about amino acid 60 to about amino acid 84, from about amino acid 103 to about amino acid 127, from about amino acid 142 to about amino acid 161, from about amino acid 194 to about amino acid 217, from about amino acid 231 to about amino acid 247, and from about amino acid 276 to about amino acid 297. Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular loops are found from about amino acid 50 to about amino acid 59, from about amino acid 128 to about amino acid 141, and from about amino acid 218 to about amino acid 230. The three extracellular loops are found at from about amino acid 85 to about amino acid 102, from about amino acid 162 to about amino acid 193, and from about amino acid 248 to about amino acid 275.

The Sequence Listing shows the 15334 nucleotide sequence (SEQ ID NO:6) and the deduced 15334 amino acid sequence (SEQ ID NO:5). It is predicted that amino acids 1 to about 25 constitute the amino terminal extracellular domain, amino acids about 26-299 constitute the region spanning the transmembrane domain, and amino acids about 300-372 constitute the carboxy terminal intracellular domain. The transmembrane domain contains seven transmembrane segments, three extracellular loops and three intracellular loops. The transmembrane segments are found from about amino acid 26 to about amino acid 48, from about amino acid 56 to about amino acid 77, from about amino acid 99 to about amino acid 115, from about amino acid 140 to about amino acid 157, from about amino acid 188 to about amino acid 209, from about amino acid 235 to about amino acid 259, and from about amino acid

WO 00/23588

PCT/US99/24368

277 to about amino acid 299. Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular loops are found from about amino acid 49 to about amino acid 55, from about amino acid 116 to about amino acid 139, and from about amino acid 210 to about amino acid 234. The three extracellular loops are found at from about amino acid 78 to about amino acid 98, from about amino acid 158 to about amino acid 187, and from about amino acid 260 to about amino acid 276.

The "2838 receptor polypeptide" or "2838 receptor protein" refers to the polypeptide in SEQ ID NO:1. The "14618 receptor polypeptide" or "14618 receptor protein" refers to the polypeptide in SEQ ID NO:3. The "15334 receptor polypeptide" or "15334 receptor protein" refers to the polypeptide in SEQ ID NO:5. The term "receptor protein" or "receptor polypeptide", however, further includes the numerous variants of 2838, 14618, or 15334 polypeptides described herein, as well as fragments derived from the full length 2838, 14618, or 15334 polypeptides and variants.

The present invention thus provides isolated or purified 2838, 14618, and 15334 receptor polypeptides and variants and fragments thereof.

The 2838 polypeptide is a 319 residue protein exhibiting three main structural domains. The amino terminal extracellular domain is identified to be within residues 1 to about 24 in SEQ ID NO:1. The transmembrane domain is identified to be within residues from about 25 to about 292 in SEQ ID NO:1. The carboxy terminal intracellular domain is identified to be within residues from about 293 to 319 in SEQ ID NO:1. The transmembrane domain contains seven segments that span the membrane. The transmembrane segments are found from about amino acid 25 to about amino acid 49, from about amino acid 56 to about amino acid 79, from about amino acid 100 to about amino acid 117, from about amino acid 138 to about amino acid 159, from about amino acid 187 to about amino acid 210, from about amino acid 224 to about amino acid 248, and from about amino acid 268 to about amino acid 292. Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular loops are found from about amino acid 50 to about amino acid 55, from about amino acid 118 to about amino acid 137, and from about amino acid 211 to about amino acid 223. The three extracellular loops are found at from about amino acid 80 to about amino acid 99, from about amino acid 160 to about amino acid 186, and from about amino acid 249 to about amino acid 267.

WO 00/23588

PCT/US99/24368

The transmembrane domain includes a GPCR signal transduction signature, DRF, at residues 118-120. The sequence includes an arginine at residue 119, an invariant amino acid in GPCRs.

Based on a BLAST search, highest homology was shown to purinoceptors.

5 The 14618 polypeptide is a 337 residue protein exhibiting three main structural domains. The amino terminal extracellular domain is identified to be within residues 1 to about 28 in SEQ ID NO:3. The transmembrane domain is identified to be within residues from about 29 to about 297 in SEQ ID NO:3. The carboxy terminal intracellular domain is identified to be within residues from about 298 to 337
10 in SEQ ID NO:3. The transmembrane domain contains seven segments that span the membrane. The transmembrane segments are found from about amino acid 29 to about amino acid 49, from about amino acid 84 to about amino acid 60, from about amino acid 103 to about amino acid 127, from about amino acid 142 to about amino acid 161, from about amino acid 194 to about amino acid 217, from about amino acid
15 231 to about amino acid 247, and from about amino acid 276 to about amino acid 297. Within the region spanning the entire transmembrane domain are three intracellular and three extracellular loops. The three intracellular loops are found from about amino acid 50 to about amino acid 59, from about amino acid 128 to about amino acid 141, and from about amino acid 218 to about amino acid 230. The three extracellular
20 loops are found at from about amino acid 85 to about amino acid 102, from about amino acid 162 to about amino acid 193, and from about amino acid 248 to about amino acid 275.

The transmembrane domain includes a GPCR signal transduction signature, FRC, at residues 121-123. The sequence includes an arginine at residue 122, an
25 invariant amino acid in GPCRs.

Based on a BLAST search, highest homology was shown to purinoceptors.

The 15334 polypeptide is a 372 residue protein exhibiting three main structural domains. The amino terminal extracellular domain is identified to be within residues 1 to about 25 in SEQ ID NO:5. The transmembrane domain is identified to
30 be within residues from about 26 to about 299 in SEQ ID NO:5. The carboxy terminal intracellular domain is identified to be within residues from about 300 to 372 in SEQ ID NO:5. The transmembrane domain contains seven segments that span the membrane. The transmembrane segments are found from about amino acid 26 to about amino acid 48, from about amino acid 56 to about amino acid 77, from about

WO 00/23588

PCT/US99/24368

amino acid 99 to about amino acid 115, from about amino acid 140 to about amino acid 157, from about amino acid 188 to about amino acid 209, from about amino acid 235 to about amino acid 259, and from about amino acid 277 to about amino acid 299. Within the region spanning the entire transmembrane domain are three intracellular
 5 and three extracellular loops. The three intracellular loops are found from about amino acid 49 to about amino acid 55, from about amino acid 116 to about amino acid 139, and from about amino acid 210 to about amino acid 234. The three extracellular loops are found at from about amino acid 78 to about amino acid 98, from about amino acid 158 to about amino acid 187, and from about amino acid 260 to about
 10 amino acid 276.

The transmembrane domain includes a GPCR signal transduction signature, DRY, at residues 118-120. The sequence includes an arginine at residue 119, an invariant amino acid in GPCRs.

Based on a BLAST search, highest homology was shown to purinoceptors.

As used herein, a polypeptide is said to be "isolated" or "purified" when it is
 15 substantially free of cellular material when it is isolated from recombinant and non-recombinant cells, or free of chemical precursors or other chemicals when it is chemically synthesized. A polypeptide, however, can be joined to another polypeptide with which it is not normally associated in a cell and still be considered
 20 "isolated" or "purified."

The receptor polypeptides can be purified to homogeneity. It is understood, however, that preparations in which the polypeptide is not purified to homogeneity are useful and considered to contain an isolated form of the polypeptide. The critical
 25 feature is that the preparation allows for the desired function of the polypeptide, even in the presence of considerable amounts of other components. Thus, the invention encompasses various degrees of purity.

In one embodiment, the language "substantially free of cellular material" includes preparations of the receptor polypeptide having less than about 30% (by dry weight) other proteins (*i.e.*, contaminating protein), less than about 20% other
 30 proteins, less than about 10% other proteins, or less than about 5% other proteins. When the receptor polypeptide is recombinantly produced, it can also be substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, less than about 10%, or less than about 5% of the volume of the protein preparation.

WO 00/23588

PCT/US99/24368

A receptor polypeptide is also considered to be isolated when it is part of a membrane preparation or is purified and then reconstituted with membrane vesicles or liposomes.

The language "substantially free of chemical precursors or other chemicals" includes preparations of the receptor polypeptide in which it is separated from chemical precursors or other chemicals that are involved in its synthesis. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of the polypeptide having less than about 30% (by dry weight) chemical precursors or other chemicals, less than about 20% chemical precursors or other chemicals, less than about 10% chemical precursors or other chemicals, or less than about 5% chemical precursors or other chemicals.

In one embodiment, the receptor polypeptide comprises the amino acid sequence shown in SEQ ID NO:1. However, the invention also encompasses sequence variants. Variants include a substantially homologous protein encoded by the same genetic locus in an organism, *i.e.*, an allelic variant. The receptor maps to chromosome 2, in close proximity to WI-7921. Variants also encompass proteins derived from other genetic loci in an organism, but having substantial homology to the 2838 receptor protein of SEQ ID NO:1. Variants also include proteins substantially homologous to the 2838 receptor protein but derived from another organism, *i.e.*, an ortholog. Variants also include proteins that are substantially homologous to the 2838 receptor protein that are produced by chemical synthesis. Variants also include proteins that are substantially homologous to the 2838 receptor protein that are produced by recombinant methods. It is understood, however, that variants exclude any amino acid sequences disclosed prior to the invention.

In another embodiment, the receptor polypeptide comprises the amino acid sequence shown in SEQ ID NO:3. However, the invention also encompasses sequence variants. Variants include allelic variants. Variants also encompass proteins derived from other genetic loci in an organism, but having substantial homology to the 14618 receptor protein of SEQ ID NO:3. Variants also include substantially homologous orthologs. Variants also include proteins that are substantially homologous to the 14618 receptor protein that are produced by chemical synthesis. Variants also include proteins that are substantially homologous to the 14618 receptor protein that are produced by recombinant methods. It is understood,

WO 00/23588

PCT/US99/24368

however, that variants exclude any amino acid sequences disclosed prior to the invention.

In another embodiment, the receptor polypeptide comprises the amino acid sequence shown in SEQ ID NO:5. However, the invention also encompasses sequence variants. Variants include allelic variants. The receptor maps to chromosome 12 in close proximity to SHGC-30262. Variants also encompass proteins derived from other genetic loci in an organism, but having substantial homology to the 15334 receptor protein of SEQ ID NO:5. Variants also include substantially homologous orthologs. Variants also include proteins that are substantially homologous to the 15334 receptor protein that are produced by chemical synthesis. Variants also include proteins that are substantially homologous to the 15334 receptor protein that are produced by recombinant methods. It is understood, however, that variants exclude any amino acid sequences disclosed prior to the invention.

As used herein, two proteins (or a region of the proteins) are substantially homologous to the 2838 protein when the amino acid sequences are at least about 40-45%, 45-50%, 50-55%, 55-60%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous. A substantially homologous amino acid sequence, according to the present invention, will be encoded by a nucleic acid sequence hybridizing to the nucleic acid sequence, or portion thereof, of the sequence shown in SEQ ID NO:2 under stringent conditions as more fully described below.

As used herein, two proteins (or a region of the proteins) are substantially homologous to the 14618 protein when the amino acid sequences are at least about 35-40%, 40-45%, 45-50%, 55-60%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous. A substantially homologous amino acid sequence, according to the present invention, will be encoded by a nucleic acid sequence hybridizing to the nucleic acid sequence, or portion thereof, of the sequence shown in SEQ ID NO:4 under stringent conditions as more fully described below.

As used herein, two proteins (or a region of the proteins) are substantially homologous to the 15334 protein when the amino acid sequences are at least about 40-45%, 45-50%, 50-55%, 55-60%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous. A

WO 00/23588

PCT/US99/24368

substantially homologous amino acid sequence, according to the present invention, will be encoded by a nucleic acid sequence hybridizing to the nucleic acid sequence, or portion thereof, of the sequence shown in SEQ ID NO.6 under stringent conditions as more fully described below.

5 To determine the percent homology of two amino acid sequences, or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of one protein or nucleic acid for optimal alignment with the other protein or nucleic acid). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When
10 a position in one sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the other sequence, then the molecules are homologous at that position. As used herein, amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity". The percent homology between the two sequences is a function of the number of identical positions shared
15 by the sequences (*i.e.*, per cent homology equals the number of identical positions/total number of positions times 100).

The invention also encompasses polypeptides having a lower degree of identity but having sufficient similarity so as to perform one or more of the same functions performed by the 2838, 14618, or 15334 polypeptides. Similarity is
20 determined by conserved amino acid substitution. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Conservative substitutions are likely to be phenotypically silent. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu, and Ile; interchange of the hydroxyl
25 residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr. Guidance concerning which amino acid changes are likely to be phenotypically silent are found in Bowie *et al.* (1990) *Science* 247:1306-1310.

30

WO 00/23588

PCT/US99/24368

TABLE 1. Conservative Amino Acid Substitutions

Aromatic	Phenylalanine
	Tryptophan
	Tyrosine
Hydrophobic	Leucine
	Isoleucine
	Valine
Polar	Glutamine
	Asparagine
Basic	Arginine
	Lysine
	Histidine
Acidic	Aspartic Acid
	Glutamic Acid
Small	Alanine
	Serine
	Threonine
	Methionine
	Glycine

- Both identity and similarity can be readily calculated (*Computational*
- 5 *Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data. Part 1*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*,
- 10 Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991).

WO 00/23588

PCT/US99/24368

A preferred, non-limiting example of such a mathematical algorithm is described in Karlin *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402. When utilizing

5 BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>. In one embodiment, parameters for sequence comparison can be set at score= 100, wordlength=12, or can be varied (e.g., W=5 or W=20).

In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman *et al.* (1970) (*J. Mol. Biol.* 48:444-453) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a BLOSUM 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between

10 two nucleotide sequences is determined using the GAP program in the GCG software package (Devereux *et al.* (1984) *Nucleic Acids Res.* 12(1):387) (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6.

Another preferred, non-limiting example of a mathematical algorithm utilized

20 for the comparison of sequences is the algorithm of Myers and Miller, CABIOS (1989). Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the CGC sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Additional

25 algorithms for sequence analysis are known in the art and include ADVANCE and ADAM as described in Torellis *et al.* (1994) *Comput. Appl. Biosci.* 10:3-5; and FASTA described in Pearson *et al.* (1988) *PNAS* 85:2444-8.

A variant polypeptide can differ in amino acid sequence by one or more substitutions, deletions, insertions, inversions, fusions, and truncations or a

30 combination of any of these.

Variant polypeptides can be fully functional or can lack function in one or more activities. Thus, in the present case, variations can affect the function, for example, of one or more of the regions corresponding to ligand binding, membrane association, G-protein binding and signal transduction.

WO 00/23588

PCT/US99/24368

Fully functional variants typically contain only conservative variation or variation in non-critical residues or in non-critical regions. Functional variants can also contain substitution of similar amino acids which result in no change or an insignificant change in function. Alternatively, such substitutions may positively or negatively affect function to some degree.

Non-functional variants typically contain one or more non-conservative amino acid substitutions, deletions, insertions, inversions, or truncation or a substitution, insertion, inversion, or deletion in a critical residue or critical region.

As indicated, variants can be naturally-occurring or can be made by recombinant means or chemical synthesis to provide useful and novel characteristics for the receptor polypeptide. This includes preventing immunogenicity from pharmaceutical formulations by preventing protein aggregation.

Useful variations further include alteration of ligand binding characteristics. For example, one embodiment involves a variation at the binding site that results in binding but not release, or slower release, of ligand. A further useful variation at the same sites can result in a higher affinity for ligand. Useful variations also include changes that provide for affinity for another ligand. Another useful variation includes one that allows binding but which prevents activation by the ligand. Another useful variation includes variation in the transmembrane G-protein-binding/signal transduction domain that provides for reduced or increased binding by the appropriate G-protein or for binding by a different G-protein than the one with which the receptor is normally associated. Another useful variation provides a fusion protein in which one or more domains or subregions is operationally fused to one or more domains or subregions from another G-protein coupled receptor.

Amino acids that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham *et al.* (1989) *Science* 244:1081-1085). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as receptor binding or *in vitro*, or *in vitro* proliferative activity. Sites that are critical for ligand-receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al.* (1992) *J. Mol. Biol.* 224:899-904; de Vos *et al.* (1992) *Science* 255:306-312).

WO 00/23588

PCT/US99/24368

Substantial homology can be to the entire nucleic acid or amino acid sequence or to fragments of these sequences.

The invention thus also includes polypeptide fragments of the 2838 receptor protein. Fragments can be derived from the amino acid sequence shown in SEQ ID

5 NO.1. However, the invention also encompasses fragments of the variants of the 2838 receptor protein as described herein.

The invention thus also includes polypeptide fragments of the 14618 receptor protein. Fragments can be derived from the amino acid sequence shown in SEQ ID

10 NO.3. However, the invention also encompasses fragments of the variants of the 14618 receptor protein as described herein.

The invention thus also includes polypeptide fragments of the 15334 receptor protein. Fragments can be derived from the amino acid sequence shown in SEQ ID

NO.5. However, the invention also encompasses fragments of the variants of the 15334 receptor protein as described herein.

15 The fragments to which the invention pertains, however, are not to be construed as encompassing fragments that may be disclosed prior to the present invention.

Fragments can retain one or more of the biological activities of the protein, for example the ability to bind to a G-protein or ligand, as well as fragments that can be used as an immunogen to generate receptor antibodies.

20 Biologically active fragments of 2838 receptor (peptides which are, for example, 8, 12, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) can comprise a domain or motif, e.g., an extracellular or intracellular domain or loop, one or more transmembrane segments, or parts thereof, G-protein binding site, or GPCR signature, glycosylation sites, protein kinase C and casein kinase II

25 phosphorylation sites, N-myristoylation and amidation sites. Such domains or motifs can be identified by means of routine computerized homology searching procedures

Possible fragments include, but are not limited to: 1) soluble peptides comprising the entire amino terminal extracellular domain from amino acid 1 to about

30 amino acid 24 of SEQ ID NO.1, or parts thereof; 2) peptides comprising the entire carboxy terminal intracellular domain from about amino acid 293 to amino acid 319 of SEQ ID NO.1, or parts thereof; 3) peptides comprising the region spanning the entire transmembrane domain from about amino acid 25 to about amino acid 292, or parts thereof; 4) any of the specific transmembrane segments, or parts thereof, from

WO 00/23598

PCT/US99/24368

about amino acid 25 to about amino acid 49, from about amino acid 56 to about amino acid 79, from about amino acid 100 to about amino acid 117, from about amino acid 138 to about amino acid 159, from about amino acid 187 to about amino acid 210, from about amino acid 224 to about amino acid 248, and from about amino acid 268 to about amino acid 292; 5) any of the three intracellular or three extracellular loops, or parts thereof, from about amino acid 50 to about amino acid 55, from about amino acid 118 to about amino acid 137, from about amino acid 211 to about amino acid 223, from about amino acid 80 to about amino acid 99, from about amino acid 160 to about amino acid 186, and from about amino acid 249 to about amino acid 267.

10 Fragments further include combinations of the above fragments, such as an amino terminal domain combined with one or more transmembrane segments and the attendant extra or intracellular loops or one or more transmembrane segments, and the attendant intra or extracellular loops, plus the carboxy terminal domain. Thus, any of the above fragments can be combined. Other fragments include the mature protein

15 from about amino acid 6 to 319. Other fragments contain the various functional sites described herein, such as phosphorylation sites, glycosylation sites, and myristoylation sites and a sequence containing the GPCR signature sequence. Fragments, for example, can extend in one or both directions from the functional site to encompass 5, 10, 15, 20, 30, 40, 50, or up to 100 amino acids. Further, fragments

20 can include sub-fragments of the specific domains mentioned above, which sub-fragments retain the function of the domain from which they are derived. Fragments also include amino acid sequences greater than 7 amino acids from amino acid 1 to about amino acid 264. Fragments also include antigenic fragments and specifically those shown to have a high antigenic index in Figure 2.

25 Accordingly, possible fragments include fragments defining a ligand-binding site, fragments defining a glycosylation site, fragments defining membrane association, fragments defining phosphorylation sites, fragments defining interaction with G proteins and signal transduction, and fragments defining myristoylation sites. By this is intended a discrete fragment that provides the relevant function or allows

30 the relevant function to be identified. In a preferred embodiment, the fragment contains the ligand-binding site.

Biologically active fragments of 14618 receptor (peptides which are, for example, 9, 12, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) can comprise a domain or motif, e.g., an extracellular or intracellular domain

WO 00/23588

PCT/US99/24368

or loop, one or more transmembrane segments, or parts thereof, G-protein binding site, or GPCR signature, glycosylation sites, and cAMP- and cGMP-dependent, protein kinase C, and casein kinase II phosphorylation sites.

Possible fragments include, but are not limited to 1) soluble peptides

- 5 comprising the entire amino terminal extracellular domain about amino acid 1 to about amino acid 28 of SEQ ID NO.3, or parts thereof; 2) peptides comprising the entire carboxy terminal intracellular domain from about amino acid 298 to amino acid 337 of SEQ ID NO.3, or parts thereof; 3) peptides comprising the region spanning the entire transmembrane domain from about amino acid 29 to about amino acid 297, or
- 10 parts thereof; 4) any of the specific transmembrane segments, or parts thereof, from about amino acid 29 to about amino acid 49, from about amino acid 60 to about amino acid 84, from about amino acid 103 to about amino acid 127, from about amino acid 142 to about amino acid 161, from about amino acid 194 to about amino acid 217, from about amino acid 231 to about amino acid 247, and from about amino acid
- 15 276 to about amino acid 297; 5) any of the three intracellular or three extracellular loops, or parts thereof, from about amino acid 50 to about amino acid 59, from about amino acid 128 to about amino acid 141, from about amino acid 218 to about amino acid 230, from about amino acid 85 to about amino acid 102, from about amino acid 162 to about amino acid 193, and from about amino acid 248 to about amino acid 275.
- 20 Fragments further include combinations of the above fragments, such as an amino terminal domain combined with one or more transmembrane segments and the attendant extra or intracellular loops or one or more transmembrane segments, and the attendant intra or extracellular loops, plus the carboxy terminal domain. Thus, any of the above fragments can be combined. Other fragments include the mature protein
- 25 from about amino acid 6 to 337. Other fragments contain the various functional sites described herein, such as phosphorylation sites, glycosylation sites, and a sequence containing the GPCR signature sequence. Fragments, for example, can extend in one or both directions from the functional site to encompass 5, 10, 15, 20, 30, 40, 50, or up to 100 amino acids. Further, fragments can include sub-fragments of the specific
- 30 domains mentioned above, which sub-fragments retain the function of the domain from which they are derived. Fragments also include amino acid sequences greater than 8 amino acids from amino acid 1 to about amino acid 244.

Fragments also include antigenic fragments and specifically those shown to have a high antigenic index in Figure 2.

WO 00/23588

PCT/US99/24368

Accordingly, possible fragments include fragments defining a ligand-binding site, fragments defining a glycosylation site, fragments defining membrane association, fragments defining phosphorylation sites, fragments defining interaction with G proteins and signal transduction, and fragments defining myristoylation sites.

- 5 By this is intended a discrete fragment that provides the relevant function or allows the relevant function to be identified. In a preferred embodiment, the fragment contains the ligand-binding site.

- Biologically active fragments (peptides which are, for example, 8, 12, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) can comprise a
10 domain or motif, e.g., an extracellular or intracellular domain or loop, one or more transmembrane segments, or parts thereof, G-protein binding site, or GPCR signature, glycosylation sites, cAMP, cGMP, protein kinase C, and casein kinase II phosphorylation sites, and N-myristoylation sites.

- Possible fragments include, but are not limited to: 1) soluble peptides
15 comprising the entire amino terminal extracellular domain about amino acid 1 to about amino acid 23 of SEQ ID NO:5, or parts thereof; 2) peptides comprising the entire carboxy terminal intracellular domain from about amino acid 300 to amino acid 372 of SEQ ID NO:5, or parts thereof; 3) peptides comprising the region spanning the entire transmembrane domain from about amino acid 26 to about amino acid 299, or
20 parts thereof; 4) any of the specific transmembrane segments, or parts thereof, from about amino acid 26 to about amino acid 48, from about amino acid 56 to about amino acid 77, from about amino acid 99 to about amino acid 115, from about amino acid 140 to about amino acid 157, from about amino acid 188 to about amino acid 209, from about amino acid 235 to about amino acid 259, and from about amino acid
25 277 to about amino acid 299; 5) any of the three intracellular or three extracellular loops, or parts thereof, from about amino acid 49 to about amino acid 55, from about amino acid 78 to about amino acid 98, from about amino acid 116 to about amino acid 139, from about amino acid 158 to about amino acid 187, from about amino acid 210 to about amino acid 234, and from about amino acid 260 to about amino acid 276.
30 Fragments further include combinations of the above fragments, such as an amino terminal domain combined with one or more transmembrane segments and the attendant extra or intracellular loops or one or more transmembrane segments, and the attendant intra or extracellular loops, plus the carboxy terminal domain. Thus, any of the above fragments can be combined. Other fragments include the mature protein

WO 00/23588

PCT/US99/24368

from about amino acid 6 to 372. Other fragments contain the various functional sites described herein, such as phosphorylation sites, glycosylation sites, and myristoylation sites and a sequence containing the GPCR signature sequence.

Fragments, for example, can extend in one or both directions from the functional site to encompass 5, 10, 15, 20, 30, 40, 50, or up to 100 amino acids. Further, fragments can include sub-fragments of the specific domains mentioned above, which sub-fragments retain the function of the domain from which they are derived. Fragments also include amino acid sequences greater than 7 amino acids.

Fragments also include antigenic fragments and specifically those shown to have a high antigenic index in Figure 2.

These regions can be identified by well-known methods involving computerized homology analysis.

Accordingly, possible fragments include fragments defining a ligand-binding site, fragments defining a glycosylation site, fragments defining membrane association, fragments defining phosphorylation sites, fragments defining interaction with G proteins and signal transduction, and fragments defining myristoylation sites. By this is intended a discrete fragment that provides the relevant function or allows the relevant function to be identified. In a preferred embodiment, the fragment contains the ligand-binding site.

The invention also provides 2838 receptor fragments with immunogenic properties. These contain an epitope-bearing portion of the 2838 receptor protein and variants. These epitope-bearing peptides are useful to raise antibodies that bind specifically to a receptor polypeptide or region or fragment. These peptides can contain at least 8, 12, at least 14, or between at least about 15 to about 30 amino acids.

The invention also provides 14618 receptor fragments with immunogenic properties. These contain an epitope-bearing portion of the 14618 receptor protein and variants. These epitope-bearing peptides are useful to raise antibodies that bind specifically to a receptor polypeptide or region or fragment. These peptides can contain at least 9, 12, at least 14, or between at least about 15 to about 30 amino acids.

The invention also provides 15334 receptor fragments with immunogenic properties. These contain an epitope-bearing portion of the 15334 receptor protein and variants. These epitope-bearing peptides are useful to raise antibodies that bind specifically to a receptor polypeptide or region or fragment. These peptides can contain at least 8, 12, at least 14, or between at least about 15 to about 30 amino acids.

WO 00/23588

PCT/US99/24368

Non-limiting examples of antigenic polypeptides that can be used to generate antibodies include peptides derived from the amino terminal extracellular domain or any of the extracellular loops. Regions having a high antigenicity index are shown in Figures 2, 6, and 10.

5 The epitope-bearing receptor and polypeptides may be produced by any conventional means (Houghten (1985) *Proc. Natl. Acad. Sci. USA* 82:5131-5135). Simultaneous multiple peptide synthesis is described in U.S. Patent No. 4,631,211.

10 Fragments can be discrete (not fused to other amino acids or polypeptides) or can be within a larger polypeptide. Further, several fragments can be comprised within a single larger polypeptide. In one embodiment a fragment designed for expression in a host can have heterologous pre- and pro-polypeptide regions fused to the amino terminus of the receptor fragment and an additional region fused to the carboxyl terminus of the fragment.

15 The invention thus provides chimeric or fusion proteins. These comprise a receptor protein operatively linked to a heterologous protein having an amino acid sequence not substantially homologous to the receptor protein. "Operatively linked" indicates that the receptor protein and the heterologous protein are fused in-frame. The heterologous protein can be fused to the N-terminus or C-terminus of the receptor protein.

20 In one embodiment the fusion protein does not affect receptor function *per se*. For example, the fusion protein can be a GST-fusion protein in which the receptor sequences are fused to the C-terminus of the GST sequences. Other types of fusion proteins include, but are not limited to, enzymatic fusion proteins, for example beta-galactosidase fusions, yeast two-hybrid GAL fusions, poly-His fusions and Ig fusions. 25 Such fusion proteins, particularly poly-His fusions, can facilitate the purification of recombinant receptor protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of a protein can be increased by using a heterologous signal sequence. Therefore, in another embodiment, the fusion protein contains a heterologous signal sequence at its N-terminus.

30 EP-A-O 464 533 discloses fusion proteins comprising various portions of immunoglobulin constant regions. The Fc is useful in therapy and diagnosis and thus results, for example, in improved pharmacokinetic properties (EP-A 0232 262). In drug discovery, for example, human proteins have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists. Bennett *et al*

WO 00/23588

PCT/US99/24368

(1995) *J. Mol. Recog.* 8:52-58 and Johanson *et al.* (1995) *J. Biol. Chem.* 270, 16:9459-9471. Thus, this invention also encompasses soluble fusion proteins containing a receptor polypeptide and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclass (IgG, IgM, IgA, IgE).

5 Preferred as immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. For some uses it is desirable to remove the Fc after the fusion protein has been used for its intended purpose, for example when the fusion protein is to be used as antigen for immunizations. In a particular embodiment, the Fc part can be removed in a simple way by a cleavage sequence which is also incorporated and can be cleaved with factor Xa.

A chimeric or fusion protein can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different protein sequences are ligated together in-frame in accordance with conventional techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see Ausubel *et al.*, (1992) *Current*
20 *Protocols in Molecular Biology*). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST protein). A receptor protein-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the receptor protein.

Another form of fusion protein is one that directly affects receptor functions. Accordingly, a receptor polypeptide is encompassed by the present invention in which one or more of the receptor domains (or parts thereof) has been replaced by homologous domains (or parts thereof) from another G-protein coupled receptor or other type of receptor. Accordingly, various permutations are possible. The amino terminal extracellular domain, or subregion thereof, (for example, ligand-binding) can
30 be replaced with the domain or subregion from another ligand-binding receptor protein. Alternatively, the entire transmembrane domain, or any of the seven segments or loops, or parts thereof, for example, G-protein-binding/signal transduction, can be replaced. Finally, the carboxy terminal intracellular domain or

WO 00/23588

PCT/US99/24368

subregion can be replaced. Thus, chimeric receptors can be formed in which one or more of the native domains or subregions has been replaced.

The isolated 2838 receptor protein can be purified from cells that naturally express it, such as from lymph node, thymus, spleen, testes, colon, and peripheral blood lymphocytes, and from those cells in which it is significantly expressed as shown in Figure 15, such as activated T-helper cells (1 and 2), hypoxic Hep 3B cells, CD3 cells (both CD4 and CD8), activated B cells, Jurkat cells, among others, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods.

The isolated 14618 receptor protein can be purified from cells that naturally express it, such as from breast, skeletal muscle, lymph node, spleen and blood peripheral lymphocytes, as well as those tissues in which the gene is significantly expressed as shown in Figures 13 and 14 including, but not limited to, CD34⁺ cells and megakaryocytes, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods.

The isolated 15334 receptor protein can be purified from cells that naturally express it, such as from colon, placenta, pancreas, tonsil, lymph node, spleen, peripheral blood cells, thymus, adrenal gland and heart, as well as those tissues shown in Figures 16-19, including K562 cells, erythroblasts, and megakaryocytes, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods.

In one embodiment, the protein is produced by recombinant DNA techniques. For example, a nucleic acid molecule encoding the receptor polypeptide is cloned into an expression vector, the expression vector introduced into a host cell and the protein expressed in the host cell. The protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques.

Polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally-occurring amino acids. Further, many amino acids, including the terminal amino acids, may be modified by natural processes, such as processing and other post-translational modifications, or by chemical modification techniques well known in the art. Common modifications that occur naturally in polypeptides are described in basic texts, detailed monographs, and the research literature, and they are well known to those of skill in the art.

WO 00/23588

PCT/US99/24368

Accordingly, the polypeptides also encompass derivatives or analogs in which a substituted amino acid residue is not one encoded by the genetic code, in which a substituent group is included, in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence for purification of the mature polypeptide or a pro-protein sequence.

Known modifications include, but are not limited to, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well-known to those of skill in the art and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as *Proteins - Structure and Molecular Properties*, 2nd Ed., J.E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as by Wold, F., *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York 1-12 (1983); Seifter *et al.* (1990) *Meth. Enzymol.* 182: 626-646 and Rattan *et al.* (1992) *Ann. N.Y. Acad. Sci.* 663:48-62).

As is also well known, polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of post-translation events, including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, branched and branched circular polypeptides may be synthesized by non-translational natural processes and by synthetic methods.

WO 00/23588

PCT/US99/24368

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. Blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally-occurring and synthetic polypeptides. For instance, the amino
5 terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost invariably will be N-formylmethionine.

The modifications can be a function of how the protein is made. For recombinant polypeptides, for example, the modifications will be determined by the host cell posttranslational modification capacity and the modification signals in the
10 polypeptide amino acid sequence. Accordingly, when glycosylation is desired, a polypeptide should be expressed in a glycosylating host, generally a eukaryotic cell. Insect cells often carry out the same posttranslational glycosylations as mammalian cells and, for this reason, insect cell expression systems have been developed to efficiently express mammalian proteins having native patterns of glycosylation.
15 Similar considerations apply to other modifications.

The same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain more than one type of modification.

20 Polypeptide uses

The receptor polypeptides are useful for producing antibodies specific for the 2838, 14618, and 15334 receptor protein, regions, or fragments. Regions having a high antigenicity index score are shown in Figures 2, 6, and 10.

The receptor polypeptides, variants, and fragments (including those which
25 may have been disclosed prior to the present invention) are useful for biological assays related to GPCRs. Such assays involve any of the known GPCR functions or activities or properties useful diagnosis and treatment of GPCR-related conditions.

The receptor polypeptides are also useful in drug screening assays, in cell-based or cell-free systems. Cell-based systems can be native, *i.e.*, cells that normally
30 express the receptor protein, as a biopsy or expanded in cell culture. In one embodiment, however, cell-based assays involve recombinant host cells expressing the receptor protein.

Determining the ability of the test compound to interact with the polypeptide can also comprise determining the ability of the test compound to preferentially bind

WO 08/23588

PCT/US99/24368

to the polypeptide as compared to the ability of the ligand, or a biologically active portion thereof, to bind to the polypeptide

The polypeptides can be used to identify compounds that modulate receptor activity. Such compounds, for example, can increase or decrease affinity or rate of binding to a known ligand, compete with ligand for binding to the receptor, or displace ligand bound to the receptor. 2838, 14618, and 15334 protein and appropriate variants and fragments can be used in high-throughput screens to assay candidate compounds for the ability to bind to the receptor. These compounds can be further screened against a functional receptor to determine the effect of the compound on the receptor activity. Compounds can be identified that activate (agonist) or inactivate (antagonist) the receptor to a desired degree. Modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). Examples include purine analogs such as those discussed above.

The receptor polypeptides can be used to screen a compound for the ability to stimulate or inhibit interaction between the receptor protein and a target molecule that normally interacts with the receptor protein. The target can be ligand or a component of the signal pathway with which the receptor protein normally interacts (for example, a G-protein or other interactor involved in cAMP or phosphatidylinositol turnover and/or adenylate cyclase, or phospholipase C activation). The assay includes the steps of combining the receptor protein with a candidate compound under conditions that allow the receptor protein or fragment to interact with the target molecule, and to detect the formation of a complex between the protein and the target or to detect the biochemical consequence of the interaction with the receptor protein and the target, such as any of the associated effects of signal transduction such as G-protein phosphorylation, cyclic AMP or phosphatidylinositol turnover, and adenylate cyclase or phospholipase C activation.

Determining the ability of the protein to bind to a target molecule can also be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (e.g., BIAcore™). Changes in the optical phenomenon surface

WO 00/23588

PCT/US99/24368

plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including:
 5 biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution: the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide
 10 oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994) *J. Med.*
 15 *Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233. Libraries of compounds may be presented in solution (e.g., Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556),
 20 bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Coll *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390), (Devlin (1990) *Science* 249:404-406), (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 97:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra*).

Candidate compounds include, for example, 1) purine analogs, 2) peptides
 25 such as soluble peptides, including Ig-tailed fusion peptides and members of random peptide libraries (see, e.g., Lam *et al.* (1991) *Nature* 354:82-84; Houghten *et al.* (1991) *Nature* 354:84-86) and combinatorial chemistry-derived molecular libraries made of D- and/or L- configuration amino acids, 3) phosphopeptides (e.g., members
 30 of random and partially degenerate, directed phosphopeptide libraries, see, e.g., Songyang *et al.* (1993) *Cell* 72:767-778), 4) antibodies (e.g., polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, and single chain antibodies as well as Fab, F(ab)₂, Fab expression library fragments, and epitope-binding fragments of

WO 00/23588

PCT/US99/24368

antibodies); and 5) small organic and inorganic molecules (e.g., molecules obtained from combinatorial and natural product libraries).

One candidate compound is a soluble full-length receptor or fragment that competes for ligand binding. Other candidate compounds include mutant receptors or
5 appropriate fragments containing mutations that affect receptor function and thus compete for ligand. Accordingly, a fragment that competes for ligand, for example with a higher affinity, or a fragment that binds ligand but does not allow release, is encompassed by the invention.

The invention provides other end points to identify compounds that modulate
10 (stimulate or inhibit) receptor activity. The assays typically involve an assay of events in the signal transduction pathway that indicate receptor activity. Thus, the expression of genes that are up- or down-regulated in response to the receptor protein dependent signal cascade can be assayed. In one embodiment, the regulatory region of such genes can be operably linked to a marker that is easily detectable, such as
15 luciferase. Alternatively, phosphorylation of the receptor protein, or a receptor protein target, could also be measured.

Binding and/or activating compounds can also be screened by using chimeric receptor proteins in which the amino terminal extracellular domain, or parts thereof, the entire transmembrane domain or subregions, such as any of the seven
20 transmembrane segments or any of the intracellular or extracellular loops and the carboxy terminal intracellular domain, or parts thereof, can be replaced by heterologous domains or subregions. For example, a G-protein-binding region can be used that interacts with a different G-protein than that which is recognized by the native receptor. Accordingly, a different set of signal transduction components is
25 available as an end-point assay for activation. Alternatively, the entire transmembrane portion or subregions (such as transmembrane segments or intracellular or extracellular loops) can be replaced with the entire transmembrane portion or subregions specific to a host cell that is different from the host cell from which the amino terminal extracellular domain and/or the G-protein-binding region
30 are derived. This allows for assays to be performed in other than the specific host cell from which the receptor is derived. Alternatively, the amino terminal extracellular domain (and/or other ligand-binding regions) could be replaced by a domain (and/or other binding region) binding a different ligand, thus, providing an assay for test compounds that interact with the heterologous amino terminal extracellular domain

WO 00/23588

PCT/US99/24368

(or region) but still cause signal transduction. Finally, activation can be detected by a reporter gene containing an easily detectable coding region operably linked to a transcriptional regulatory sequence that is part of the native signal transduction pathway².

5 The receptor polypeptides are also useful in competition binding assays in methods designed to discover compounds that interact with the receptor. Thus, a compound is exposed to a receptor polypeptide under conditions that allow the compound to bind or to otherwise interact with the polypeptide. Soluble receptor polypeptide is also added to the mixture. If the test compound interacts with the
10 soluble receptor polypeptide, it decreases the amount of complex formed or activity from the receptor target. This type of assay is particularly useful in cases in which compounds are sought that interact with specific regions of the receptor. Thus, the soluble polypeptide that competes with the target receptor region is designed to contain peptide sequences corresponding to the region of interest.

15 To perform cell free drug screening assays, it is desirable to immobilize either the receptor protein, or fragment, or its target molecule to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay.

Techniques for immobilizing proteins on matrices can be used in the drug
20 screening assays. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be bound to a matrix. For example, glutathione-S-transferase/2838, 14618, and 15334 fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates (e.g., ³⁵S-labeled) and
25 the candidate compound, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly, or in the supernatant after the complexes are dissociated. Alternatively, the complexes can be dissociated from the
30 matrix, separated by SDS-PAGE, and the level of receptor-binding protein found in the bead fraction quantitated from the gel using standard electrophoretic techniques. For example, either the polypeptide or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin using techniques well known in the art. Alternatively, antibodies reactive with the protein but which do not interfere with

WO 00/23588

PCT/US99/24368

binding of the protein to its target molecule can be derivatized to the wells of the plate, and the protein trapped in the wells by antibody conjugation. Preparations of a receptor-binding protein and a candidate compound are incubated in the receptor protein-presenting wells and the amount of complex trapped in the well can be quantitated. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the receptor protein target molecule, or which are reactive with receptor protein and compete with the target molecule; as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule.

Modulators of receptor protein activity identified according to these drug screening assays can be used to treat a subject with a disorder mediated by the receptor pathway, by treating cells that express the 2838 protein, such as in lymph node, thymus, spleen, testes, colon, and peripheral blood lymphocytes including, but not limited to, T-helper cells (1 and 2), CD3⁺ (CD4 and CD8) cells, B cells and granulocytes. Modulators of receptor protein activity identified according to these drug screening assays can be used to treat a subject with a disorder mediated by the receptor pathway, by treating cells that express the 14618 protein, such as in breast, skeletal muscle, spleen and peripheral blood lymphocytes as well as CD34⁺ cells and megakaryocytes. Modulators of receptor protein activity identified according to these drug screening assays can be used to treat a subject with a disorder mediated by the receptor pathway, by treating cells that express the 15334 protein, such as in lymph node, tonsil, pancreas, colon, spleen, peripheral blood cells, thymus, adrenal gland and heart as well as megakaryocytes and erythroblasts. These methods of treatment include the steps of administering the modulators of protein activity in a pharmaceutical composition as described herein, to a subject in need of such treatment.

Figure 13 shows expression of the 14618 receptor in a variety of normal human tissues. The gene is highly expressed in breast and skeletal muscle. Accordingly, expression of the gene in disorders involving these tissues is particularly relevant. Significant expression also occurs in the thyroid, placenta, fetal kidney, fetal heart, and lymph node. Accordingly, expression of the gene is relevant to disorders involving these tissues as well. Furthermore, as shown in Figure 13, lower levels of

WO 00/23588

PCT/US99/24368

expression are seen in a variety of other tissues. Accordingly, expression of the gene is also relevant in disorders involving these tissues.

The 14618 receptor is also expressed in various hematopoietic cells with and without activation. The gene is highly expressed in CD34⁺ bone marrow cells.

- 5 Accordingly, expression of the gene is relevant in a variety of blood cell progenitors. Expression of this gene is therefore relevant to disorders involving deficiencies in any of the major blood cell types, *i.e.* neutropenia, thrombocytopenia or anemia. The gene is also highly expressed in mobilized peripheral blood cell megakaryocytes (mobilized with G-CSF). Accordingly, expression of this gene is relevant to disorders
- 10 involving platelet function, such as thrombocytopenia. Significant expression is also seen in mobilized peripheral blood cell leukocytes, mobilized bone marrow CD34⁺ cells, and cord blood CD434⁺ cells. Accordingly, expression of the gene is relevant to function of these cells, and therefore relevant to disorders involving immune function or inflammation. Further, expression of the gene occurs in activated peripheral blood
- 15 mononuclear cells. Activated B cells differentially express the gene. Accordingly, expression of the gene is relevant to disorders involving immune function and/or inflammation.

- The 2838 receptor is expressed in the cells and tissues shown in Figure 15. High levels of expression are shown in lymph node and thymus. Accordingly,
- 20 expression of the gene is especially relevant to disorders involving these tissues. Extremely high expression is found in CD8 cells and activated B cells. High expression also occurs in activated T-helper cells (1 and 2), CD4 cells and Jurkat cells (a T-cell line). Expression is differential in activated B cells and activated T-helper cells. Expression increases upon activation in both of these cell types. Accordingly,
- 25 expression of the gene is relevant to disorders involving immune function and inflammation. The gene is also significantly expressed in granulocytes. Accordingly, expression of the gene is relevant to disorders involving these cells.

- Figure 16 shows expression of the 15334 receptor in normal human tissues. The gene is highly expressed in lymph node, tonsil, and pancreas. Accordingly,
- 30 expression of the gene is particularly relevant to disorders involving these tissues. Expression of the gene is also high in colon, testis, placenta, fetal heart, and spleen. Accordingly, expression of the gene is also relevant to disorders involving these tissues. In addition, the figure shows low levels of expression in several other tissues. Accordingly, expression of the gene may be relevant to disorders involving these

WO 00/23588

PCT/U599/24368

tissues. Expression of the 15334 receptor has also been studied in various hematopoietic cells (Figures 17-19). Extremely high expression occurs in primary megakaryocytes and erythroblasts. Accordingly, expression of the gene is relevant to erythrocyte differentiation and megakaryocyte differentiation and thus is relevant to treating anemia and thrombocytopenia. Further, expression of the gene is significantly increased in resting B cells compared to activated B cells. Accordingly, expression of the gene is relevant to B cell immune function. Further, lower levels of expression are found in various other cells of the hematopoietic lineage. See Figure 17. The expression of this gene in hematopoietic cells in a lineage-restricted manner indicates that expression is relevant in regulating the development of the lineage cells, erythrocytes/red blood cells, or megakaryocytes/platelets.

Disorders involving the spleen include, but are not limited to, splenomegaly, including nonspecific acute splenitis, congestive splenomegaly, and splenic infarcts, neoplasms, congenital anomalies, and rupture. Disorders associated with splenomegaly include infections, such as nonspecific splenitis, infectious mononucleosis, tuberculosis, typhoid fever, brucellosis, cytomegalovirus, syphilis, malaria, histoplasmosis, toxoplasmosis, kala-azar, trypanosomiasis, schistosomiasis, leishmaniasis, and echinococcosis; congestive states related to partial hypertension, such as cirrhosis of the liver, portal or splenic vein thrombosis, and cardiac failure; lymphohematogenous disorders, such as Hodgkin disease, non-Hodgkin lymphomas/leukemia, multiple myeloma, myeloproliferative disorders, hemolytic anemias, and thrombocytopenic purpura; immunologic-inflammatory conditions, such as rheumatoid arthritis and systemic lupus erythematosus; storage diseases such as Gaucher disease, Niemann-Pick disease, and mucopolysaccharidoses; and other conditions, such as amyloidosis, primary neoplasms and cysts, and secondary neoplasms.

Disorders involving the lung include, but are not limited to, congenital anomalies; atelectasis; diseases of vascular origin, such as pulmonary congestion and edema, including hemodynamic pulmonary edema and edema caused by microvascular injury, adult respiratory distress syndrome (diffuse alveolar damage), pulmonary embolism, hemorrhage, and infarction, and pulmonary hypertension and vascular sclerosis; chronic obstructive pulmonary disease, such as emphysema, chronic bronchitis, bronchial asthma, and bronchiectasis; diffuse interstitial (infiltrative, restrictive) diseases, such as pneumoconioses, sarcoidosis, idiopathic

WO 00/23588

PCT/US99/24368

pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia (pulmonary infiltration with eosinophilia), *Bronchiolitis obliterans-organizing pneumonia*, diffuse pulmonary hemorrhage syndromes, including Goodpasture syndrome, idiopathic pulmonary hemosiderosis and other hemorrhagic syndromes, pulmonary involvement in collagen vascular disorders, and pulmonary alveolar proteinosis; complications of therapies, such as drug-induced lung disease, radiation-induced lung disease, and lung transplantation, tumors, such as bronchogenic carcinoma, including paraneoplastic syndromes, bronchioloalveolar carcinoma, neuroendocrine tumors, such as bronchial carcinoid, miscellaneous tumors, and metastatic tumors; pathologies of the pleura, including inflammatory pleural effusions, noninflammatory pleural effusions, pneumothorax, and pleural tumors, including solitary fibrous tumors (pleural fibroma) and malignant mesothelioma.

Disorders involving the colon include, but are not limited to, congenital anomalies, such as atresia and stenosis, Meckel diverticulum, congenital aganglionic megacolon-Hirschsprung disease; enterocolitis, such as diarrhea and dysentery, infectious enterocolitis, including viral gastroenteritis, bacterial enterocolitis, necrotizing enterocolitis, antibiotic-associated colitis (pseudomembranous colitis), and collagenous and lymphocytic colitis, miscellaneous intestinal inflammatory disorders, including parasites and protozoa, acquired immunodeficiency syndrome, transplantation, drug-induced intestinal injury, radiation enterocolitis, neutropenic colitis (typhlitis), and diversion colitis; idiopathic inflammatory bowel disease, such as Crohn disease and ulcerative colitis; tumors of the colon, such as non-neoplastic polyps, adenomas, familial syndromes, colorectal carcinogenesis, colorectal carcinoma, and carcinoid tumors.

Disorders involving the liver include, but are not limited to, hepatic injury, jaundice and cholestasis, such as bilirubin and bile formation; hepatic failure and cirrhosis, such as cirrhosis, portal hypertension, including ascites, portosystemic shunts, and splenomegaly; infectious disorders, such as viral hepatitis, including hepatitis A-E infection and infection by other hepatitis viruses, clinicopathologic syndromes, such as the carrier state, asymptomatic infection, acute viral hepatitis, chronic viral hepatitis, and fulminant hepatitis; autoimmune hepatitis; drug- and toxin-induced liver disease, such as alcoholic liver disease; inborn errors of metabolism and pediatric liver disease, such as hemochromatosis, Wilson disease, α_1 -

WO 00/23588

PCT/US99/24368

- antitrypsin deficiency, and neonatal hepatitis; intrahepatic biliary tract disease, such as secondary biliary cirrhosis, primary biliary cirrhosis, primary sclerosing cholangitis, and anomalies of the biliary tree; circulatory disorders, such as impaired blood flow into the liver, including hepatic artery compromise and portal vein obstruction and thrombosis, impaired blood flow through the liver, including passive congestion and centrilobular necrosis and peliosis hepatis, hepatic vein outflow obstruction, including hepatic vein thrombosis (Budd-Chiari syndrome) and veno-occlusive disease; hepatic disease associated with pregnancy, such as preeclampsia and eclampsia, acute fatty liver of pregnancy, and intrahepatic cholestasis of pregnancy; hepatic complications of organ or bone marrow transplantation, such as drug toxicity after bone marrow transplantation, graft-versus-host disease and liver rejection, and nonimmunologic damage to liver allografts; tumors and tumorous conditions, such as nodular hyperplasias, adenomas, and malignant tumors, including primary carcinoma of the liver and metastatic tumors.
- Disorders involving the uterus and endometrium include, but are not limited to, endometrial histology in the menstrual cycle; functional endometrial disorders, such as anovulatory cycle, inadequate luteal phase, oral contraceptives and induced endometrial changes, and menopausal and postmenopausal changes; inflammations, such as chronic endometritis; adenomyosis; endometriosis; endometrial polyps, endometrial hyperplasia; malignant tumors, such as carcinoma of the endometrium; mixed Müllerian and mesenchymal tumors, such as malignant mixed Müllerian tumors; tumors of the myometrium, including leiomyomas, leiomyosarcomas, and endometrial stromal tumors.
- Disorders involving the brain include, but are not limited to, disorders involving neurons, and disorders involving glia, such as astrocytes, oligodendrocytes, ependymal cells, and microglia; cerebral edema, raised intracranial pressure and herniation, and hydrocephalus; malformations and developmental diseases, such as neural tube defects, forebrain anomalies, posterior fossa anomalies, and syringomyelia and hydromyelia; perinatal brain injury; cerebrovascular diseases, such as those related to hypoxia, ischemia, and infarction, including hypotension, hypoperfusion, and low-flow states--global cerebral ischemia and focal cerebral ischemia--infarction from obstruction of local blood supply, intracranial hemorrhage, including intracerebral (intraparenchymal) hemorrhage, subarachnoid hemorrhage and ruptured berry aneurysms, and vascular malformations, hypertensive cerebrovascular disease.

WO 00/23588

PCT/US99/24368

- including lacunar infarcts, slit hemorrhages, and hypertensive encephalopathy;
infections, such as acute meningitis, including acute pyogenic (bacterial) meningitis
and acute aseptic (viral) meningitis, acute focal suppurative infections, including brain
abscess, subdural empyema, and extradural abscess, chronic bacterial
5 meningoencephalitis, including tuberculosis and mycobacterioses, neurosyphilis, and
neuroborreliosis (Lyme disease), viral meningoencephalitis, including arthropod-
borne (Arbo) viral encephalitis, *Herpes simplex* virus Type 1, *Herpes simplex* virus
Type 2, *Varicella-zoster* virus (*Herpes zoster*), cytomegalovirus, poliomyelitis, rabies,
and human immunodeficiency virus 1, including HIV-1 meningoencephalitis
10 (subacute encephalitis), vacuolar myelopathy, AIDS-associated myopathy, peripheral
neuropathy, and AIDS in children, progressive multifocal leukoencephalopathy,
subacute sclerosing panencephalitis, fungal meningoencephalitis, other infectious
diseases of the nervous system, transmissible spongiform encephalopathies (prion
diseases), demyelinating diseases, including multiple sclerosis, multiple sclerosis
15 variants, acute disseminated encephalomyelitis and acute necrotizing hemorrhagic
encephalomyelitis, and other diseases with demyelination, degenerative diseases, such
as degenerative diseases affecting the cerebral cortex, including Alzheimer disease
and Pick disease, degenerative diseases of basal ganglia and brain stem, including
Parkinsonism, idiopathic Parkinson disease (paralysis agitans), progressive
20 supranuclear palsy, corticobasal degeneration, multiple system atrophy, including
striatonigral degeneration, Shy-Drager syndrome, and olivopontocerebellar atrophy,
and Huntington disease; spinocerebellar degenerations, including spinocerebellar
ataxias, including Friedreich ataxia, and ataxia-telangiectasia, degenerative diseases
affecting motor neurons, including amyotrophic lateral sclerosis (motor neuron
25 disease), bulbospinal atrophy (Kennedy syndrome), and spinal muscular atrophy;
inborn errors of metabolism, such as leukodystrophies, including Krabbe disease,
metachromatic leukodystrophy, adrenoleukodystrophy, Pelizaeus-Merzbacher
disease, and Canavan disease, mitochondrial encephalomyopathies, including Leigh
disease and other mitochondrial encephalomyopathies; toxic and acquired metabolic
30 diseases, including vitamin deficiencies such as thiamine (vitamin B₁) deficiency and
vitamin B₁₂ deficiency, neurologic sequelae of metabolic disturbances, including
hypoglycemia, hyperglycemia, and hepatic encephatopathy, toxic disorders, including
carbon monoxide, methanol, ethanol, and radiation, including combined methotrexate
and radiation-induced injury; tumors, such as gliomas, including astrocytoma.

WO 00/23588

PCT/US99/24368

- including fibrillary (diffuse) astrocytoma and glioblastoma multiforme, pilocytic astrocytoma, pleomorphic xanthoastrocytoma, and brain stem glioma, oligodendroglioma, and ependymoma and related paraventricular mass lesions, neuronal tumors, poorly differentiated neoplasms, including medulloblastoma, other
- 5 parenchymal tumors, including primary brain lymphoma, germ cell tumors, and pineal parenchymal tumors, meningiomas, metastatic tumors, paraneoplastic syndromes, peripheral nerve sheath tumors, including schwannoma, neurofibroma, and malignant peripheral nerve sheath tumor (malignant schwannoma), and neurocutaneous syndromes (phakomatoses), including neurofibromatosis, including
- 10 Type 1 neurofibromatosis (NF1) and TYPE 2 neurofibromatosis (NF2), tuberous sclerosis, and Von Hippel-Lindau disease.

- Disorders involving T-cells include, but are not limited to, cell-mediated hypersensitivity, such as delayed type hypersensitivity and T-cell-mediated cytotoxicity, and transplant rejection; autoimmune diseases, such as systemic lupus
- 15 erythematosus, Sjögren syndrome, systemic sclerosis, inflammatory myopathies, mixed connective tissue disease, and polyarteritis nodosa and other vasculitides; immunologic deficiency syndromes, including but not limited to, primary immunodeficiencies, such as thymic hypoplasia, severe combined immunodeficiency diseases, and AIDS; leukopenia; reactive (inflammatory) proliferations of white cells,
- 20 including but not limited to, leukocytosis, acute nonspecific lymphadenitis, and chronic nonspecific lymphadenitis; neoplastic proliferations of white cells, including but not limited to lymphoid neoplasms, such as precursor T-cell neoplasms, such as acute lymphoblastic leukemia/lymphoma, peripheral T-cell and natural killer cell neoplasms that include peripheral T-cell lymphoma, unspecified, adult T-cell
- 25 leukemia/lymphoma, mycosis fungoides and Sézary syndrome, and Hodgkin disease

- Diseases of the skin, include but are not limited to, disorders of pigmentation and melanocytes, including but not limited to, vitiligo, freckle, melasma, lentigo, nevocellular nevus, dysplastic nevi, and malignant melanoma; benign epithelial tumors, including but not limited to, seborrheic keratoses, acanthosis nigricans,
- 30 fibroepithelial polyp, epithelial cyst, keratoacanthoma, and adnexal (appendage) tumors; premalignant and malignant epidermal tumors, including but not limited to, actinic keratosis, squamous cell carcinoma, basal cell carcinoma, and merkel cell carcinoma; tumors of the dermis, including but not limited to, benign fibrous histiocytoma, dermatofibrosarcoma protuberans, xanthomas, and dermal vascular

WO 00/23588

PCT/US99/24368

tumors; tumors of cellular immigrants to the skin, including but not limited to, histiocytosis X, mycosis fungoides (cutaneous T-cell lymphoma), and mastocytosis; disorders of epidermal maturation, including but not limited to, ichthyosis; acute inflammatory dermatoses, including but not limited to, urticaria, acute eczematous dermatitis, and erythema multiforme; chronic inflammatory dermatoses, including but not limited to, psoriasis, lichen planus, and lupus erythematosus; blistering (bullous) diseases, including but not limited to, pemphigus, bullous pemphigoid, dermatitis herpetiformis, and noninflammatory blistering diseases; epidermolysis bullosa and porphyria; disorders of epidermal appendages, including but not limited to, acne vulgaris; panniculitis, including but not limited to, erythema nodosum and erythema induratum; and infection and infestation, such as verrucae, molluscum contagiosum, impetigo, superficial fungal infections, and arthropod bites, stings, and infestations.

In normal bone marrow, the myelocytic series (polymorphonuclear cells) make up approximately 60% of the cellular elements, and the erythrocytic series, 20-30%. Lymphocytes, monocytes, reticular cells, plasma cells and megakaryocytes together constitute 10-20%. Lymphocytes make up 5-15% of normal adult marrow. In the bone marrow, cell types are admixed so that precursors of red blood cells (erythroblasts), macrophages (monoblasts), platelets (megakaryocytes), polymorphonuclear leucocytes (myeloblasts), and lymphocytes (lymphoblasts) can be visible in one microscopic field. In addition, stem cells exist for the different cell lineages, as well as a precursor stem cell for the committed progenitor cells of the different lineages. The various types of cells and stages of each would be known to the person of ordinary skill in the art and are found, for example, on page 42 (Figure 2-8) of *Immunology, Immunopathology and Immunity*, Fifth Edition, Sell *et al.*, Simon and Schuster (1996), incorporated by reference for its teaching of cell types found in the bone marrow. According, the invention is directed to disorders arising from these cells. These disorders include but are not limited to the following: diseases involving hematopoietic stem cells; committed lymphoid progenitor cells; lymphoid cells including B and T-cells; committed myeloid progenitors, including monocytes, granulocytes, and megakaryocytes; and committed erythroid progenitors. These include but are not limited to the leukemias, including B-lymphoid leukemias, T-lymphoid leukemias, undifferentiated leukemias, erythroleukemia, megakaryoblastic leukemia, monocytic, [leukemias are encompassed with and without differentiation]; chronic and acute lymphoblastic leukemia, chronic and acute lymphocytic leukemia,

WO 00/23588

PCT/US99/24368

chronic and acute myelogenous leukemia, lymphoma, myelo dysplastic syndrome, chronic and acute myeloid leukemia, myelomonocytic leukemia; chronic and acute myeloblastic leukemia, chronic and acute myelogenous leukemia, chronic and acute promyelocytic leukemia, chronic and acute myelocytic leukemia, hematologic malignancies of monocyte-macrophage lineage, such as juvenile chronic myelogenous leukemia; secondary AML, antecedent hematological disorder; refractory anemia; aplastic anemia; reactive cutaneous angioendotheliomatosis; fibrosing disorders involving altered expression in dendritic cells, disorders including systemic sclerosis, E-M syndrome, epidemic toxic oil syndrome, eosinophilic fasciitis

5 localized forms of scleroderma, keloid, and fibrosing colonopathy; angiomatoid malignant fibrous histiocytoma; carcinoma, including primary head and neck squamous cell carcinoma; sarcoma, including kaposi's sarcoma; fibroadenoma and phyllodes tumors, including mammary fibroadenoma; stromal tumors; phyllodes tumors, including histiocytoma; erythroblastosis; neurofibromatosis; diseases of the

10 vascular endothelium; demyelinating, particularly in old lesions; gliosis, vasogenic edema, vascular disease, Alzheimer's and Parkinson's disease; T-cell lymphomas; B cell lymphomas.

Disorders involving the heart, include but are not limited to, heart failure, including but not limited to, cardiac hypertrophy, left-sided heart failure, and right-sided heart failure; ischemic heart disease, including but not limited to angina pectoris, myocardial infarction, chronic ischemic heart disease, and sudden cardiac death; hypertensive heart disease, including but not limited to, systemic (left-sided) hypertensive heart disease and pulmonary (right-sided) hypertensive heart disease; valvular heart disease, including but not limited to, valvular degeneration caused by

20 calcification, such as calcific aortic stenosis, calcification of a congenitally bicuspid aortic valve, and mitral annular calcification, and myxomatous degeneration of the mitral valve (mitral valve prolapse), rheumatic fever and rheumatic heart disease, infective endocarditis, and noninfected vegetations, such as nonbacterial thrombotic endocarditis and endocarditis of systemic lupus erythematosus (Libman-Sacks

25 disease), carcinoid heart disease, and complications of artificial valves; myocardial disease, including but not limited to dilated cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy, and myocarditis; pericardial disease, including but not limited to, pericardial effusion and hemopericardium and pericarditis, including acute pericarditis and healed pericarditis, and rheumatoid heart

WO 00/23588

PCT/US99/24368

disease; neoplastic heart disease, including but not limited to, primary cardiac tumors, such as myxoma, lipoma, papillary fibroelastoma, rhabdomyoma, and sarcoma, and cardiac effects of noncardiac neoplasms; congenital heart disease, including but not limited to, left-to-right shunts--late cyanosis, such as atrial septal defect, ventricular septal defect, patent ductus arteriosus, and atrioventricular septal defect, right-to-left shunts--early cyanosis, such as tetralogy of fallot, transposition of great arteries, truncus arteriosus, tricuspid atresia, and total anomalous pulmonary venous connection, obstructive congenital anomalies, such as coarctation of aorta, pulmonary stenosis and atresia, and aortic stenosis and atresia, and disorders involving cardiac transplantation.

Disorders involving blood vessels include, but are not limited to, responses of vascular cell walls to injury, such as endothelial dysfunction and endothelial activation and intimal thickening; vascular diseases including, but not limited to, congenital anomalies, such as arteriovenous fistula, atherosclerosis, and hypertensive vascular disease, such as hypertension; inflammatory disease--the vasculitides, such as giant cell (temporal) arteritis, Takayasu arteritis, polyarteritis nodosa (classic), Kawasaki syndrome (mucocutaneous lymph node syndrome), microscopic polyangiitis (microscopic polyarteritis, hypersensitivity or leukocytoclastic angitis), Wegener granulomatosis, thromboangitis obliterans (Buerger disease), vasculitis associated with other disorders, and infectious arteritis; Raynaud disease; aneurysms and dissection, such as abdominal aortic aneurysms, syphilitic (lentic) aneurysms, and aortic dissection (dissecting hematoma); disorders of veins and lymphatics, such as varicose veins, thrombophlebitis and phlebotrombosis, obstruction of superior vena cava (superior vena cava syndrome), obstruction of inferior vena cava (inferior vena cava syndrome), and lymphangitis and lymphedema; tumors, including benign tumors and tumor-like conditions, such as hemangioma, lymphangioma, glomus tumor (glomangioma), vascular ectasias, and bacillary angiomatosis, and intermediate-grade (borderline low-grade malignant) tumors, such as Kaposi sarcoma and hemangioendothelioma, and malignant tumors, such as angiosarcoma and hemangiopericytoma; and pathology of therapeutic interventions in vascular disease, such as balloon angioplasty and related techniques and vascular replacement, such as coronary artery bypass graft surgery.

Disorders involving red cells include, but are not limited to, anemias, such as hemolytic anemias, including hereditary spherocytosis, hemolytic disease due to

WO 00/23588

PCT/US99/24368

erythrocyte enzyme defects: glucose-6-phosphate dehydrogenase deficiency, sickle cell disease, thalassemia syndromes, paroxysmal nocturnal hemoglobinuria, immunohemolytic anemia, and hemolytic anemia resulting from trauma to red cells; and anemias of diminished erythropoiesis, including megaloblastic anemias, such as

5 anemias of vitamin B12 deficiency: pernicious anemia, and anemia of folate deficiency, iron deficiency anemia, anemia of chronic disease, aplastic anemia, pure red cell aplasia, and other forms of marrow failure.

Disorders involving the thymus include developmental disorders, such as DiGeorge syndrome with thymic hypoplasia or aplasia; thymic cysts; thymic

10 hypoplasia, which involves the appearance of lymphoid follicles within the thymus, creating thymic follicular hyperplasia; and thymomas, including germ cell tumors, lymphomas, Hodgkin disease, and carcinoids. Thymomas can include benign or encapsulated thymoma, and malignant thymoma Type I (invasive thymoma) or Type II, designated thymic carcinoma.

Disorders involving B cells include, but are not limited to precursor B cell neoplasms, such as lymphoblastic leukemia/lymphoma. Peripheral B cell neoplasms

15 include, but are not limited to, chronic lymphocytic leukemia/small lymphocytic lymphoma, follicular lymphoma, diffuse large B cell lymphoma, Burkitt lymphoma, plasma cell neoplasms, multiple myeloma, and related entities, lymphoplasmacytic

20 lymphoma (Waldenström macroglobulinemia), mantle cell lymphoma, marginal zone lymphoma (MALToma), and hairy cell leukemia.

Disorders involving the kidney include, but are not limited to, congenital anomalies including, but not limited to, cystic diseases of the kidney, that include but are not limited to, cystic renal dysplasia, autosomal dominant (adult) polycystic kidney

25 disease, autosomal recessive (childhood) polycystic kidney disease, and cystic diseases of renal medulla, which include, but are not limited to, medullary sponge kidney, and nephronophthisis-uremic medullary cystic disease complex, acquired (dialysis-associated) cystic disease, such as simple cysts; glomerular diseases including pathologies of glomerular injury that include, but are not limited to, in situ immune

30 complex deposition, that includes, but is not limited to, anti-GBM nephritis, Heymann nephritis, and antibodies against planted antigens, circulating immune complex nephritis, antibodies to glomerular cells, cell-mediated immunity in glomerulonephritis, activation of alternative complement pathway, epithelial cell injury, and pathologies involving mediators of glomerular injury including cellular and soluble mediators, acute

WO 00/23588

PCT/US99/24368

glomerulonephritis, such as acute proliferative (poststreptococcal, postinfectious) glomerulonephritis, including but not limited to, poststreptococcal glomerulonephritis and nonstreptococcal acute glomerulonephritis, rapidly progressive (crescentic) glomerulonephritis, nephrotic syndrome, membranous glomerulonephritis (membranous nephropathy), minimal change disease (lipoid nephrosis), focal segmental glomerulosclerosis, membranoproliferative glomerulonephritis, IgA nephropathy (Berger disease), focal proliferative and necrotizing glomerulonephritis (focal glomerulonephritis), hereditary nephritis, including but not limited to, Alport syndrome and thin membrane disease (benign familial hematuria), chronic glomerulonephritis,

glomerular lesions associated with systemic disease, including but not limited to, systemic lupus erythematosus, Henoch-Schönlein purpura, bacterial endocarditis, diabetic glomerulosclerosis, amyloidosis, fibrillary and immunotactoid glomerulonephritis, and other systemic disorders; diseases affecting tubules and interstitium, including acute tubular necrosis and tubulointerstitial nephritis, including but not limited to, pyelonephritis and urinary tract infection, acute pyelonephritis, chronic pyelonephritis and reflux nephropathy, and tubulointerstitial nephritis induced by drugs and toxins, including but not limited to, acute drug-induced interstitial nephritis, analgesic abuse nephropathy, nephropathy associated with nonsteroidal anti-inflammatory drugs, and other tubulointerstitial diseases including, but not limited to, urate nephropathy, hypercalcemia and nephrocalcinosis, and multiple myeloma; diseases of blood vessels including benign nephrosclerosis, malignant hypertension and accelerated nephrosclerosis, renal artery stenosis, and thrombotic microangiopathies including, but not limited to, classic (childhood) hemolytic-uremic syndrome, adult hemolytic-uremic syndrome/thrombotic thrombocytopenic purpura, idiopathic HUS/TTP, and other vascular disorders including, but not limited to, atherosclerotic ischemic renal disease, atheroembolic renal disease, sickle cell disease nephropathy, diffuse cortical necrosis, and renal infarcts; urinary tract obstruction (obstructive uropathy); urolithiasis (renal calculi, stones); and tumors of the kidney including, but not limited to, benign tumors, such as renal papillary adenoma, renal fibroma or hamartoma (renomedullary interstitial cell tumor), angiomyolipoma, and oncocytoma, and malignant tumors, including renal cell carcinoma (hypernephroma, adenocarcinoma of kidney), which includes urothelial carcinomas of renal pelvis.

Disorders of the breast include, but are not limited to, disorders of development; inflammations, including but not limited to, acute mastitis, periductal mastitis, periductal

WO 00/23588

PCT/US99/24368

mastitis (recurrent subareolar abscess, squamous metaplasia of lactiferous ducts), mammary duct ectasia, fat necrosis, granulomatous mastitis, and pathologies associated with silicone breast implants; fibrocystic changes; proliferative breast disease including, but not limited to, epithelial hyperplasia, sclerosing adenosis, and small duct papillomas; 5 tumors including, but not limited to, stromal tumors such as fibroadenoma, phyllodes tumor, and sarcomas, and epithelial tumors such as large duct papilloma; carcinoma of the breast including in situ (noninvasive) carcinoma that includes ductal carcinoma in situ (including Paget's disease) and lobular carcinoma in situ, and invasive (infiltrating) carcinoma including, but not limited to, invasive ductal carcinoma, no special type, 10 invasive lobular carcinoma, medullary carcinoma, colloid (mucinous) carcinoma, tubular carcinoma, and invasive papillary carcinoma, and miscellaneous malignant neoplasms.

Disorders in the male breast include, but are not limited to, gynecomastia and carcinoma.

Disorders involving the testis and epididymis include, but are not limited to, 15 congenital anomalies such as cryptorchidism, regressive changes such as atrophy, inflammations such as nonspecific epididymitis and orchitis, granulomatous (autoimmune) orchitis, and specific inflammations including, but not limited to, gonorrhea, mumps, tuberculosis, and syphilis, vascular disturbances including torsion, testicular tumors including germ cell tumors that include, but are not limited to, 20 seminoma, spermatocytic seminoma, embryonal carcinoma, yolk sac tumor, choriocarcinoma, teratoma, and mixed tumors, tumor of sex cord-gonadal stroma including, but not limited to, leydig (interstitial) cell tumors and sertoli cell tumors (androsthenoma), and testicular lymphoma, and miscellaneous lesions of tunica vaginalis.

Disorders involving the prostate include, but are not limited to, inflammations, 25 benign enlargement, for example, nodular hyperplasia (benign prostatic hypertrophy or hyperplasia), and tumors such as carcinoma.

Disorders involving the thyroid include, but are not limited to, hyperthyroidism; hypothyroidism including, but not limited to, cretinism and myxedema; thyroiditis including, but not limited to, hashimoto thyroiditis, subacute (granulomatous) thyroiditis, 30 and subacute lymphocytic (painless) thyroiditis; Graves disease; diffuse and multinodular goiter including, but not limited to, diffuse nontoxic (simple) goiter and multinodular goiter; neoplasms of the thyroid including, but not limited to, adenomas, other benign tumors, and carcinomas, which include, but are not limited to, papillary

WO 00/23588

PCT/US99/24368

carcinoma, follicular carcinoma, medullary carcinoma, and anaplastic carcinoma; and congenital anomalies.

Disorders involving the skeletal muscle include tumors such as rhabdomyosarcoma

- 5 Disorders involving the pancreas include those of the exocrine pancreas such as congenital anomalies, including but not limited to, ectopic pancreas; pancreatitis, including but not limited to, acute pancreatitis; cysts, including but not limited to, pseudocysts; tumors, including but not limited to, cystic tumors and carcinoma of the pancreas; and disorders of the endocrine pancreas such as, diabetes mellitus, islet cell
10 tumors, including but not limited to, insulinomas, gastrinomas, and other rare islet cell tumors.

Disorders involving the small intestine include the malabsorption syndromes such as, celiac sprue, tropical sprue (postinfectious sprue), whipple disease, disaccharidase (lactase) deficiency, abetalipoproteinemia, and tumors of the small

- 15 intestine including adenomas and adenocarcinoma.

Disorders related to reduced platelet number, thrombocytopenia, include idiopathic thrombocytopenic purpura, including acute idiopathic thrombocytopenic purpura, drug-induced thrombocytopenia, HIV-associated thrombocytopenia, and thrombotic microangiopathies: thrombotic thrombocytopenic purpura and hemolytic-
20 uremic syndrome.

Disorders involving precursor T-cell neoplasms include precursor T lymphoblastic leukemia/lymphoma. Disorders involving peripheral T-cell and natural killer cell neoplasms include T-cell chronic lymphocytic leukemia, large granular lymphocytic leukemia, mycosis fungoides and Sézary syndrome, peripheral T-cell
25 lymphoma, unspecified, angioimmunoblastic T-cell lymphoma, angiocentric lymphoma (NK/T-cell lymphoma^{4a}), intestinal T-cell lymphoma, adult T-cell leukemia/lymphoma, and anaplastic large cell lymphoma.

- Disorders involving the ovary include, for example, polycystic ovarian disease, Stein-leventhal syndrome, Pseudomyxoma peritonei and stromal
30 hyperthecosis; ovarian tumors such as, tumors of coelomic epithelium, serous tumors, mucinous tumors, endometrioid tumors, clear cell adenocarcinoma, cystadenofibroma, brenner tumor, surface epithelial tumors, germ cell tumors such as mature (benign) teratomas, monodermal teratomas, immature malignant teratomas, dysgerminoma, endodermal sinus tumor, choriocarcinoma; sex cord-stromal tumors

WO 00/23588

PCT/US99/24368

such as, granulosa-theca cell tumors, thecoma-fibromas, androblastomas, hili cell tumors, and gonadoblastoma; and metastatic tumors such as Krukenberg tumors.

Bone-forming cells include the osteoprogenitor cells, osteoblasts, and osteocytes. The disorders of the bone are complex because they may have an impact on the skeleton during any of its stages of development. Hence, the disorders may have variable manifestations and may involve one, multiple or all bones of the body. Such disorders include, congenital malformations, achondroplasia and thanatophoric dwarfism, diseases associated with abnormal matrix such as type I collagen disease, osteoporosis, paget disease, rickets, osteomalacia, high-turnover osteodystrophy, low-turnover of aplastic disease, osteonecrosis, pyogenic osteomyelitis, tuberculous osteomyelitis, osteoma, osteoid osteoma, osteoblastoma, osteosarcoma, osteochondroma, chondromas, chondroblastoma, chondromyxoid fibroma, chondrosarcoma, fibrous cortical defects, fibrous dysplasia, fibrosarcoma, malignant fibrous histiocytoma, ewing sarcoma, primitive neuroectodermal tumor, giant cell tumor, and metastatic tumors.

The receptor polypeptides are thus useful for treating a receptor-associated disorder characterized by aberrant expression or activity of a receptor protein, especially as disclosed above. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) expression or activity of the protein. In another embodiment, the method involves administering a protein as therapy to compensate for reduced or aberrant expression or activity of the protein.

Stimulation of protein activity is desirable in situations in which the protein is abnormally downregulated and/or in which increased protein activity is likely to have a beneficial effect. Likewise, inhibition of protein activity is desirable in situations in which the protein is abnormally upregulated and/or in which decreased protein activity is likely to have a beneficial effect. In one example of such a situation, a subject has a disorder characterized by aberrant development or cellular differentiation. In another example of such a situation, the subject has a proliferative disease (e.g., cancer) or a disorder characterized by an aberrant hematopoietic response. In another example of such a situation, it is desirable to achieve tissue regeneration in a subject (e.g., where a subject has undergone brain or spinal cord injury and it is desirable to regenerate neuronal tissue in a regulated manner).

WO 00/23588

PCT/US99/24368

In yet another aspect of the invention, the proteins of the invention can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO 94/10300), to identify other proteins (captured proteins) which bind to or interact with the proteins of the invention and modulate their activity.

The 2838 receptor polypeptides also are useful to provide a target for diagnosing a disease or predisposition to disease mediated by the receptor protein, especially in hematopoietic cells such as T-helper cells (1 and 2), B cells, and granulocytes. The 14618 receptor polypeptides also are useful to provide a target for diagnosing a disease or predisposition to disease mediated by the receptor protein, especially in breast, skeletal muscle and lymphnode, as well as CD34⁺ progenitor cells and cells generated from these progenitors, which include erythroid, lymphoid, and megakaryocytic lineages. The 15334 receptor polypeptides also are useful to provide a target for diagnosing a disease or predisposition to disease mediated by the receptor protein, especially in pancreas, spleen, tonsils and lymphnodes, as well as in cells of the megakaryocyte and erythroblast lineages. Accordingly, methods are provided for detecting the presence, or levels of, the receptor protein in a cell, tissue, or organism. The method involves contacting a biological sample with a compound capable of interacting with the receptor protein such that the interaction can be detected.

One agent for detecting receptor protein is an antibody capable of selectively binding to receptor protein. A biological sample includes tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject.

The receptor protein also provides a target for diagnosing active disease, or predisposition to disease, in a patient having a variant receptor protein. Thus, receptor protein can be isolated from a biological sample, assayed for the presence of a genetic mutation that results in aberrant receptor protein. This includes amino acid substitution, deletion, insertion, rearrangement, (as the result of aberrant splicing events), and inappropriate post-translational modification. Analytic methods include altered electrophoretic mobility, altered tryptic peptide digest, altered receptor activity in cell-based or cell-free assay, alteration in ligand or antibody-binding pattern,

WO 00/23588

PCT/US99/24368

altered isoelectric point, direct amino acid sequencing, and any other of the known assay techniques useful for detecting mutations in a protein.

In vitro techniques for detection of receptor protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. Alternatively, the protein can be detected *in vivo* in a subject by introducing into the subject a labeled anti-receptor antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques. Particularly useful are methods which detect the allelic variant of a receptor protein expressed in a subject and methods which detect fragments of a receptor protein in a sample.

The receptor polypeptides are also useful in pharmacogenomic analysis. Pharmacogenomics deal with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, e.g., Eichelbaum (1997) *Clin. Exp. Pharmacol. Physiol.* 23(10-11):983-985 (1996), and Linder (1997) *Clin. Chem.* 43(2):254-266. The clinical outcomes of these variations result in severe toxicity of therapeutic drugs in certain individuals or therapeutic failure of drugs in certain individuals as a result of individual variation in metabolism. Thus, the genotype of the individual can determine the way a therapeutic compound acts on the body or the way the body metabolizes the compound. Further, the activity of drug metabolizing enzymes effects both the intensity and duration of drug action. Thus, the pharmacogenomics of the individual permit the selection of effective compounds and effective dosages of such compounds for prophylactic or therapeutic treatment based on the individual's genotype. The discovery of genetic polymorphisms in some drug metabolizing enzymes has explained why some patients do not obtain the expected drug effects, show an exaggerated drug effect, or experience serious toxicity from standard drug dosages. Polymorphisms can be expressed in the phenotype of the extensive metabolizer and the phenotype of the poor metabolizer. Accordingly, genetic polymorphism may lead to allelic protein variants of the receptor protein in which one or more of the receptor functions in one population is different from those in another population. The polypeptides thus allow a target to ascertain a genetic predisposition that can affect treatment modality. Thus, in a ligand-based treatment, polymorphism may give rise to amino terminal extracellular domains and/or other ligand-binding regions that are more or less active in ligand binding, and receptor activation. Accordingly, ligand dosage would

WO 00/23588

PCT/US99/24368

necessarily be modified to maximize the therapeutic effect within a given population containing a polymorphism. As an alternative to genotyping, specific polymorphic polypeptides could be identified.

The receptor polypeptides are also useful for monitoring therapeutic effects during clinical trials and other treatment. Thus, the therapeutic effectiveness of an agent that is designed to increase or decrease gene expression, protein levels or receptor activity can be monitored over the course of treatment using the receptor polypeptides as an end-point target. The monitoring can be, for example, as follows: (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression or activity of a specified protein in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the protein in the post-administration samples; (v) comparing the level of expression or activity of the protein in the pre-administration sample with the protein in the post-administration sample or samples; and (vi) increasing or decreasing the administration of the agent to the subject accordingly.

The receptor polypeptides are also useful for treating a receptor-associated disorder. Accordingly, methods for treatment include the use of soluble receptor or fragments of the receptor protein that compete for ligand binding. These receptors or fragments can have a higher affinity for the ligand so as to provide effective competition.

Antibodies

The invention also provides antibodies that selectively bind to the 2838, 14618, and 15334 receptor proteins and variants and fragments. An antibody is considered to selectively bind, even if it also binds to other proteins that are not substantially homologous with the receptor protein. These other proteins share homology with a fragment or domain of the receptor protein. This conservation in specific regions gives rise to antibodies that bind to both proteins by virtue of the homologous sequence. In this case, it would be understood that antibody binding to the receptor protein is still selective.

To generate antibodies, an isolated receptor polypeptide is used as an immunogen to generate antibodies using standard techniques for polyclonal and monoclonal antibody preparation. Either the full-length protein or antigenic peptide

WO 00/23588

PCT/US99/24368

fragment can be used. Regions having a high antigenicity index are shown in Figures 2, 6, and 10.

Antibodies are preferably prepared from these regions or from discrete fragments in these regions. However, antibodies can be prepared from any region of the peptide as described herein. A preferred fragment produces an antibody that diminishes or completely prevents ligand-binding. Antibodies can be developed against the entire receptor or portions of the receptor, for example, the intracellular carboxy terminal domain, the amino terminal extracellular domain, the entire transmembrane domain or specific segments, any of the intra or extracellular loops, or any portions of the above. Antibodies may also be developed against specific functional sites, such as the site of ligand-binding, the site of G protein coupling, or sites that are phosphorylated, glycosylated, or myristoylated.

An antigenic 2838 fragment will typically comprise at least 8 contiguous amino acid residues. The antigenic peptide can comprise, however, a contiguous sequence of at least 12, 14 amino acid residues, at least 15 amino acid residues, at least 20 amino acid residues, or at least 30 amino acid residues. In one embodiment, fragments correspond to regions that are located on the surface of the protein, e.g., hydrophilic regions. These fragments are not to be construed, however, as encompassing any fragments which may be disclosed prior to the invention.

An 14618 antigenic fragment will typically comprise at least 9 contiguous amino acid residues. The antigenic peptide can comprise, however, a contiguous sequence of at least 12, 14 amino acid residues, at least 15 amino acid residues, at least 20 amino acid residues, or at least 30 amino acid residues. In one embodiment, fragments correspond to regions that are located on the surface of the protein, e.g., hydrophilic regions. These fragments are not to be construed, however, as encompassing any fragments which may be disclosed prior to the invention.

An antigenic 15334 fragment will typically comprise at least 8 contiguous amino acid residues. The antigenic peptide can comprise, however, a contiguous sequence of at least 12, 14 amino acid residues, at least 15 amino acid residues, at least 20 amino acid residues, or at least 30 amino acid residues. In one embodiment, fragments correspond to regions that are located on the surface of the protein, e.g., hydrophilic regions. These fragments are not to be construed, however, as encompassing any fragments which may be disclosed prior to the invention.

WO 00/23588

PCT/US99/24368

Antibodies can be polyclonal or monoclonal. An intact antibody, or a fragment thereof (e.g. Fab or F(ab')₂) can be used.

Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ¹²⁵I, ¹³¹I, ³⁵S or ³H.

An appropriate immunogenic preparation can be derived from native, recombinantly expressed, protein or chemically synthesized peptides.

Antibody Uses

The antibodies can be used to isolate a receptor protein by standard techniques, such as affinity chromatography or immunoprecipitation. The antibodies can facilitate the purification of the natural receptor protein from cells and recombinantly produced receptor protein expressed in host cells.

The antibodies are useful to detect the presence of receptor protein in cells or tissues to determine the pattern of expression of the receptor among various tissues in an organism and over the course of normal development.

The antibodies can be used to detect receptor protein *in situ*, *in vitro*, or in a cell lysate or supernatant in order to evaluate the abundance and pattern of expression.

The antibodies can be used to assess abnormal tissue distribution or abnormal expression during development.

Antibody detection of circulating fragments of the full length receptor protein can be used to identify receptor turnover.

Further, the antibodies can be used to assess receptor expression in disease states such as in active stages of the disease or in an individual with a predisposition toward disease related to receptor function. When a disorder is caused by an

WO 00/23588

PCT/US99/24368

inappropriate tissue distribution, developmental expression, or level of expression of the receptor protein, the antibody can be prepared against the normal receptor protein. If a disorder is characterized by a specific mutation in the receptor protein, antibodies specific for this mutant protein can be used to assay for the presence of the specific mutant receptor protein.

5

The antibodies can also be used to assess normal and aberrant subcellular localization of cells in the various tissues in an organism. Antibodies can be developed against the whole receptor or portions of the receptor, for example, portions of the amino terminal extracellular domain or extracellular loops.

10

The diagnostic uses can be applied, not only in genetic testing, but also in monitoring a treatment modality. Accordingly, where treatment is ultimately aimed at correcting receptor expression level or the presence of aberrant receptors and aberrant tissue distribution or developmental expression, antibodies directed against the receptor or relevant fragments can be used to monitor therapeutic efficacy.

15

Antibodies accordingly can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen.

Additionally, antibodies are useful in pharmacogenomic analysis. Thus, antibodies prepared against polymorphic receptor proteins can be used to identify individuals that require modified treatment modalities.

20

The antibodies are also useful as diagnostic tools as an immunological marker for aberrant receptor protein analyzed by electrophoretic mobility, isoelectric point, tryptic peptide digest, and other physical assays known to those in the art.

The antibodies are also useful for tissue typing. Thus, where a specific receptor protein has been correlated with expression in a specific tissue, antibodies that are specific for this receptor protein can be used to identify a tissue type.

25

The antibodies are also useful in forensic identification. Accordingly, where an individual has been correlated with a specific genetic polymorphism resulting in a specific polymorphic protein, an antibody specific for the polymorphic protein can be used as an aid in identification.

30

The antibodies are also useful for inhibiting receptor function, for example, blocking ligand binding.

These uses can also be applied to a therapeutic context in which treatment involves inhibiting receptor function. An antibody can be used, for example, to block

WO 00/23588

PCT/US99/24368

ligand binding. Antibodies can be prepared against specific fragments containing sites required for function or against intact receptor associated with a cell

Completely human antibodies are particularly desirable for therapeutic treatment of human patients. For an overview of this technology for producing human antibodies, see Lonberg and Hulszar (1995, *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, e.g., U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806.

The invention also encompasses kits for using antibodies to detect the presence of a receptor protein in a biological sample. The kit can comprise antibodies such as a labeled or labelable antibody and a compound or agent for detecting receptor protein in a biological sample; means for determining the amount of receptor protein in the sample; and means for comparing the amount of receptor protein in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect receptor protein

Polynucleotides

The nucleotide sequence in SEQ ID NO.2 was obtained by sequencing the human full length cDNA.

The nucleotide sequence in SEQ ID NO.4 was obtained by sequencing the human full length cDNA.

The nucleotide sequence in SEQ ID NO.6 was obtained by sequencing the human full length cDNA.

The specifically disclosed cDNAs comprise the coding region and 5' and 3' untranslated sequences (SEQ ID NO.2, SEQ ID NO.4, and SEQ ID NO.6).

The human 2838 receptor cDNA is approximately 1617 nucleotides in length and encodes a full length protein that is approximately 319 amino acid residues in length. The nucleic acid is expressed in spleen, testes, colon, and peripheral blood lymphocytes as well as those tissues shown in Figure 15. Structural analysis of the amino acid sequence of SEQ ID NO.1 is provided in Figure 3, a hydropathy plot. The figure shows the putative structure of the seven transmembrane segments, the amino terminal extracellular domain and the carboxy terminal intracellular domain.

WO 00/23588

PCT/US99/24368

The human 14618 receptor cDNA is approximately 1358 nucleotides in length and encodes a full length protein that is approximately 337 amino acid residues in length. The nucleic acid is expressed in spleen, peripheral blood lymphocytes, breast, skeletal muscle, and other tissues shown in Figure 13, as well as various hematopoietic cells shown in Figure 14, especially megakaryocytes and CD34⁺ cells. Structural analysis of the amino acid sequence of SEQ ID NO:3 is provided in Figure 7, a hydropathy plot. The figure shows the putative structure of the seven transmembrane segments, the amino terminal extracellular domain and the carboxy terminal intracellular domain.

The human 153340 receptor cDNA is approximately 2559 nucleotides in length and encodes a full length protein that is approximately 372 amino acid residues in length. The nucleic acid is expressed in lymphnode, tonsils, pancreas, colon, spleen, peripheral blood cells, thymus, adrenal gland and heart as well as cells of the megakaryocytic and erythroid lineages as shown in Figures 17-19. Structural analysis of the amino acid sequence of SEQ ID NO:5 is provided in Figure 11, a hydropathy plot. The figure shows the putative structure of the seven transmembrane segments, the amino terminal extracellular domain and the carboxy terminal intracellular domain.

As used herein, the term "transmembrane segment" refers to a structural amino acid motif which includes a hydrophobic helix that spans the plasma membrane. The entire transmembrane domain of 2838 spans from about amino acid 25 to about amino acid 292. The entire transmembrane domain of 14618 spans from about amino acid 29 to about amino acid 297. The entire transmembrane domain of 15334 spans from about amino acid 26 to about amino acid 299. Seven segments span the membrane and there are three intracellular and three extracellular loops in this domain.

The invention provides isolated polynucleotides encoding a 2838 receptor protein. The term "2838 polynucleotide" or "2838 nucleic acid" refers to the sequence shown in SEQ ID NO:2.

The invention provides isolated polynucleotides encoding a 14618 receptor protein. The term "14618 polynucleotide" or "14618 nucleic acid" refers to the sequence shown in SEQ ID NO:4.

The invention provides isolated polynucleotides encoding a 15334 receptor protein. The term "15334 polynucleotide" or "15334 nucleic acid" refers to the sequence shown in SEQ ID NO:6.

WO 00/23588

PCT/US99/24368

The term "receptor polynucleotide" or "receptor nucleic acid" further includes variants and fragments of the 2838, 14618, and 15334 polynucleotides.

An "isolated" receptor nucleic acid is one that is separated from other nucleic acid present in the natural source of the receptor nucleic acid. Preferably, an
 5 "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. However, there can be some flanking nucleotide sequences, for example up to about 5KB. The important point is that the nucleic acid is isolated from flanking sequences such that it can be subjected
 10 to the specific manipulations described herein such as recombinant expression, preparation of probes and primers, and other uses specific to the receptor nucleic acid sequences.

Moreover, an "isolated" nucleic acid molecule, such as a cDNA or RNA molecule, can be substantially free of other cellular material, or culture medium when
 15 produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. However, the nucleic acid molecule can be fused to other coding or regulatory sequences and still be considered isolated.

For example, recombinant DNA molecules contained in a vector are considered isolated. Further examples of isolated DNA molecules include
 20 recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the isolated DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

25 In some instances, the isolated material will form part of a composition (for example, a crude extract containing other substances), buffer system or reagent mix. In other circumstances, the material may be purified to essential homogeneity, for example as determined by PAGE or column chromatography such as HPLC. Preferably, an isolated nucleic acid comprises at least about 50, 80 or 90 % (on a
 30 molar basis) of all macromolecular species present.

The receptor polynucleotides can encode the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor

WO 00/23588

PCT/US99/24368

to a mature form, facilitate protein trafficking, prolong or shorten protein half-life or facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in situ*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

5 The receptor polynucleotides include, but are not limited to, the sequence encoding the mature polypeptide alone, the sequence encoding the mature polypeptide and additional coding sequences, such as a leader or secretory sequence (e.g., a pre-pro or pro-protein sequence), the sequence encoding the mature polypeptide, with or without the additional coding sequences, plus additional non-coding sequences, for
10 example introns and non-coding 5' and 3' sequences such as transcribed but non-translated sequences that play a role in transcription, mRNA processing (including splicing and polyadenylation signals), ribosome binding and stability of mRNA. In addition, the polynucleotide may be fused to a marker sequence encoding, for example, a peptide that facilitates purification.

15 Receptor polynucleotides can be in the form of RNA, such as mRNA, or in the form DNA, including cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The nucleic acid, especially DNA, can be double-stranded or single-stranded. Single-stranded nucleic acid can be the coding strand (sense strand) or the non-coding strand (anti-sense
20 strand).

One receptor nucleic acid comprises the nucleotide sequence shown in SEQ ID NO:2, corresponding to human 2838 cDNA.

One receptor nucleic acid comprises the nucleotide sequence shown in SEQ ID NO:4, corresponding to human 14618 cDNA.

25 One receptor nucleic acid comprises the nucleotide sequence shown in SEQ ID NO:6, corresponding to human 15334 cDNA.

In one embodiment, the receptor nucleic acid comprises only the coding region.

30 The invention further provides variant receptor polynucleotides, and fragments thereof, that differ from the nucleotide sequence shown in SEQ ID NO:2 due to degeneracy of the genetic code and thus encode the same protein as that encoded by the nucleotide sequence shown in SEQ ID NO:2.

The invention further provides variant receptor polynucleotides, and fragments thereof, that differ from the nucleotide sequence shown in SEQ ID NO:4 due to

WO 00/23588

PCT/US99/24368

degeneracy of the genetic code and thus encode the same protein as that encoded by the nucleotide sequence shown in SEQ ID NO: 4

The invention further provides variant receptor polynucleotides, and fragments thereof, that differ from the nucleotide sequence shown in SEQ ID NO: 6 due to degeneracy of the genetic code and thus encode the same protein as that encoded by the nucleotide sequence shown in SEQ ID NO: 6.

The invention also provides receptor nucleic acid molecules encoding the variant polypeptides described herein. Such polynucleotides may be naturally occurring, such as allelic variants (same locus), homologs (different locus), and orthologs (different organism), or may be constructed by recombinant DNA methods or by chemical synthesis. Such non-naturally occurring variants may be made by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. Accordingly, as discussed above, the variants can contain nucleotide substitutions, deletions, inversions and insertions.

Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions.

Typically, variants have a substantial identity with a nucleic acid molecule selected from the group consisting of SEQ ID NOS: 2, 4, and 6 and its complements.

Orthologs, homologs, and allelic variants can be identified using methods well known in the art. 2838 variants comprise a nucleotide sequence encoding a receptor that is 40-45%, 45-50%, 50-55%, 55-60%, at least about 65%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous to the nucleotide sequence shown in SEQ ID NO: 2 or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions, to the nucleotide sequence shown in SEQ ID NO: 2 or a fragment of the sequence.

14618 variants comprise a nucleotide sequence encoding a receptor that is 40-45%, 45-50%, 50-55%, 55-60%, at least about 65%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous to the nucleotide sequence shown in SEQ ID NO: 4 or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions, to the nucleotide sequence shown in SEQ ID NO: 4 or a fragment of the sequence.

WO 00/23588

PCT/US99/24368

15334 variants comprise a nucleotide sequence encoding a receptor that is 45-50%, 50-55%, 55-60%, at least about 65%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more homologous to the nucleotide sequence shown in SEQ ID NO:6 or a fragment of this sequence. Such nucleic acid molecules can readily be identified as being able to hybridize under stringent conditions, to the nucleotide sequence shown in SEQ ID NO:6 or a fragment of the sequence.

It is understood that stringent hybridization does not indicate substantial homology where it is due to general homology, such as poly A sequences, or sequences common to all or most proteins, all GPCRs, or all family I GPCRs or even all purinoceptors. Moreover, it is understood that variants do not include any of the nucleic acid sequences that may have been disclosed prior to the invention.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences encoding a receptor polypeptide at least 50-55%, 55% homologous to each other typically remain hybridized to each other. The conditions can be such that sequences at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 90%, at least about 95% or more identical to each other remain hybridized to one another. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6, incorporated by reference. One example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. In another non-limiting example, nucleic acid molecules are allowed to hybridize in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more low stringency washes in 0.2X SSC/0.1% SDS at room temperature, or by one or more moderate stringency washes in 0.2X SSC/0.1% SDS at 42°C, or washed in 0.2X SSC/0.1% SDS at 65°C for high stringency. In one embodiment, an isolated receptor nucleic acid molecule that hybridizes under stringent conditions to the sequence of SEQ ID NOS: 2, 4, and 6 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

WO 00/23588

PCT/US99/24368

As understood by those of ordinary skill, the exact conditions can be determined empirically and depend on ionic strength, temperature and the concentration of destabilizing agents such as formamide or denaturing agents such as SDS. Other factors considered in determining the desired hybridization conditions include the length of the nucleic acid sequences, base composition, percent mismatch between the hybridizing sequences and the frequency of occurrence of subsets of the sequences within other non-identical sequences. Thus, equivalent conditions can be determined by varying one or more of these parameters while maintaining a similar degree of identity or similarity between the two nucleic acid molecules.

The present invention also provides isolated nucleic acids that contain a single or double stranded fragment or portion that hybridizes under stringent conditions to a nucleotide sequence selected from the group consisting of SEQ ID NOS: 2, 4, and 6 and the complements of SEQ ID NOS: 2, 4, and 6. In one embodiment, the nucleic acid consists of a portion of a nucleotide sequence selected from the group consisting of SEQ ID NOS: 2, 4, and 6 and the complements. The nucleic acid fragments of the invention are at least about 15, preferably at least about 18, 20, 23 or 25 nucleotides, and can be 30, 40, 50, 100, 200 or more nucleotides in length. Longer fragments, for example, 30 or more nucleotides in length, which encode antigenic proteins or polypeptides described herein are useful.

Furthermore, the invention provides polynucleotides that comprise a fragment of the full length receptor polynucleotides. The fragment can be single or double stranded and can comprise DNA or RNA. The fragment can be derived from either the coding or the non-coding sequence.

In one embodiment, an isolated 2838 receptor nucleic acid from nucleotide 1 to around nucleotide 990 is at least 16 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 2. In another embodiment, the nucleic acid from around nucleotide 1487-1617 is at least 20 nucleotides. In other embodiments, the nucleic acid is at least 40, 50, 100, 250 or 500 nucleotides in length or greater.

In another embodiment, an isolated 14168 receptor nucleic acid from around nucleotide 1 to around nucleotide 911 is at least 8 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 4. In other embodiments, the nucleic acid is at least 40, 50, 100, 250, or 500 nucleotides in length or greater.

WO 00/23588

PCT/US99/24368

In another embodiment, an isolated 15334 receptor nucleic acid from nucleotide 1 to around nucleotide 1355 is at least 18 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO.6. In another embodiment, the nucleic acid from
 5 around nucleotide 868 to around 1355 is at least 11 nucleotides. In other embodiments, the nucleic acid is at least 40, 50, 100, 250, or 500 nucleotides in length or greater.

In another embodiment, an isolated 2838 receptor nucleic acid encodes the entire coding region from amino acid 1 to amino acid 319. In another embodiment,
 10 the isolated 2838 receptor nucleic acid encodes a sequence corresponding to the mature protein from about amino acid 6 to about amino acid 319. In another embodiment, an isolated 14618 receptor nucleic acid encodes the entire coding region from amino acid 1 to amino acid 337. In another embodiment, the isolated 14618 receptor nucleic acid encodes a sequence corresponding to the mature protein from
 15 about amino acid 6 to about amino acid 337. In another embodiment, an isolated 15334 receptor nucleic acid encodes the entire coding region from amino acid 1 to amino acid 372. In another embodiment, the isolated 15334 receptor nucleic acid encodes a sequence corresponding to the mature protein from about amino acid 6 to about amino acid 372.

20 Other fragments of all three receptors include nucleotide sequences encoding the amino acid fragments described herein. Further, fragments can include subfragments of the specific domains or sites described herein. Fragments also include nucleic acid sequences corresponding to specific amino acid sequences described above or fragments thereof. Nucleic acid fragments, according to the
 25 present invention, are not to be construed as encompassing those fragments that may have been disclosed prior to the invention.

However, it is understood that a receptor fragment includes any nucleic acid sequence that does not include the entire gene.

2838 receptor nucleic acid fragments further include sequences corresponding
 30 to the domains described herein, subregions also described, and specific functional sites. 2838 receptor nucleic acid fragments include nucleic acid molecules encoding a polypeptide comprising the amino terminal extracellular domain including amino acid residues from 1 to about 24, a polypeptide comprising the region spanning the transmembrane domain (amino acid residues from about 25 to about 292), a

WO 00/23588

PCT/US99/24368

- polypeptide comprising the carboxy terminal intracellular domain (amino acid residues from about 293 to about 319), and a polypeptide encoding the G-protein receptor signature (118-120 or surrounding amino acid residues from about 107 to about 123), nucleic acid molecules encoding any of the seven transmembrane segments, extracellular or intracellular loops, glycosylation sites, phosphorylation sites, myristoylation sites, and amidation site.
- 14618 receptor nucleic acid fragments include nucleic acid molecules encoding a polypeptide comprising the amino terminal extracellular domain including amino acid residues from 1 to about 28, a polypeptide comprising the region spanning the transmembrane domain (amino acid residues from about 29 to about 297), a polypeptide comprising the carboxy terminal intracellular domain (amino acid residues from about 298 to about 337), and a polypeptide encoding the G-protein receptor signature (120-122 or surrounding amino acid residues from about 110 to about 132), nucleic acid molecules encoding any of the seven transmembrane segments, extracellular or intracellular loops, glycosylation sites and phosphorylation sites.
- 15334 receptor nucleic acid fragments include nucleic acid molecules encoding a polypeptide comprising the amino terminal extracellular domain including amino acid residues from 1 to about 25, a polypeptide comprising the region spanning the transmembrane domain (amino acid residues from about 26 to about 299), a polypeptide comprising the carboxy terminal intracellular domain (amino acid residues from about 300 to about 372), and a polypeptide encoding the G-protein receptor signature (118-120 or surrounding amino acid residues from about 110 to about 130), nucleic acid molecules encoding any of the seven transmembrane segments, extracellular or intracellular loops, glycosylation sites, protein kinase C, cAMP, cGMP, and casein kinase II phosphorylation sites, and myristoylation sites.
- Where the location of the domains have been predicted by computer analysis, one of ordinary skill would appreciate that the amino acid residues constituting these domains can vary depending on the criteria used to define the domains.
- Receptor nucleic acid fragments also include combinations of the domains, segments, loops, and other functional sites described above. Thus, for example, a receptor nucleic acid could include sequences corresponding to the amino terminal extracellular domain and one transmembrane segment. A person of ordinary skill in the art would be aware of the many permutations that are possible.

WO 00/23588

PCT/US99/24368

Where the location of the domains or sites have been predicted by computer analysis, one of ordinary skill would appreciate that the amino acid residues constituting these domains can vary depending on the criteria used to define the domains.

The invention also provides receptor nucleic acid fragments that encode epitope bearing regions of the receptor proteins described herein.

The isolated receptor polynucleotide sequences, and especially fragments, are useful as DNA probes and primers.

For example, the coding region of a receptor gene can be isolated using the known nucleotide sequence to synthesize an oligonucleotide probe. A labeled probe can then be used to screen a cDNA library, genomic DNA library, or mRNA to isolate nucleic acid corresponding to the coding region. Further, primers can be used in PCR reactions to clone specific regions of receptor genes.

A probe/primer typically comprises substantially purified oligonucleotide. The 2838 oligonucleotide typically comprises a region of nucleotide sequence from 1-990 and from 1487-1617 that hybridizes under stringent conditions to at least about 16 and 20, respectively, typically about 25, more typically about 40, 50 or 75 consecutive nucleotides of SEQ ID NO.2 sense or anti-sense strand or other receptor polynucleotides. The 14618 oligonucleotide typically comprises a region of nucleotide sequence from 1-911 that hybridizes under stringent conditions to at least about 9, 12, typically about 25, more typically about 40, 50 or 75 consecutive nucleotides of SEQ ID NO.4 sense or anti-sense strand or other receptor polynucleotides. The 15334 oligonucleotide typically comprises a region of nucleotide sequence from 868-1355 that hybridizes under stringent conditions to at least about 11, 12, typically about 25, more typically about 40, 50 or 75 consecutive nucleotides of SEQ ID NO.6 sense or anti-sense strand or other receptor polynucleotides. The 15334 oligonucleotide also typically comprises a region of nucleotide sequence from nucleotide 1-1355 that hybridizes under stringent conditions to at least about 18, typically about 25, more typically about 40, 50, or 75 consecutive nucleotides of SEQ ID NO.6 sense or anti-sense strand or other receptor polynucleotides.

Polynucleotide Uses

The nucleic acid sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example,

WO 00/23588

PCT/US99/24368

identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences

5 homologous to the nucleic acid molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

10 The nucleic acid fragments of the invention provide probes or primers in assays such as those described below. "Probes" are oligonucleotides that hybridize in a base-specific manner to a complementary strand of nucleic acid. Such probes include polypeptide nucleic acids, as described in Nielsen *et al.* (1991) *Science* 254:1497-1500. Typically, a probe comprises a region of nucleotide sequence that
15 hybridizes under highly stringent conditions to at least about 15, typically about 20-25, and more typically about 40, 50 or 75 consecutive nucleotides of a nucleic acid selected from the group consisting of SEQ ID NOS. 2, 4, and 6 and the complements thereof. More typically, the probe further comprises a label, e.g., radioisotope, fluorescent compound, enzyme, or enzyme co-factor.

20 As used herein, the term "primer" refers to a single-stranded oligonucleotide which acts as a point of initiation of template-directed DNA synthesis using well-known methods (e.g., PCR, LCR) including, but not limited to those described herein. The appropriate length of the primer depends on the particular use, but typically ranges from about 15 to 30 nucleotides. The term "primer site" refers to the area of
25 the target DNA to which a primer hybridizes. The term "primer pair" refers to a set of primers including a 5' (upstream) primer that hybridizes with the 5' end of the nucleic acid sequence to be amplified and a 3' (downstream) primer that hybridizes with the complement of the sequence to be amplified.

The receptor polynucleotides are useful for probes, primers, and in biological
30 assays.

Where the polynucleotides are used to assess GPCR properties or functions, such as in the assays described herein, all or less than all of the entire cDNA can be useful. In this case, even fragments that may have been known prior to the invention are encompassed. Thus, for example, assays specifically directed to GPCR functions,

WO 00/23588

PCT/US99/24368

such as assessing agonist or antagonist activity, encompass the use of known fragments. Further, diagnostic methods for assessing receptor function can also be practiced with any fragment, including those fragments that may have been known prior to the invention. Similarly, in methods involving treatment of receptor dysfunction, all fragments are encompassed including those which may have been known in the art.

The 2838 receptor polynucleotides are useful as a hybridization probe for cDNA and genomic DNA to isolate a full-length cDNA and genomic clones encoding the polypeptide described in SEQ ID NO:1 and to isolate cDNA and genomic clones that correspond to variants producing the same polypeptide shown in SEQ ID NO:1 or the other variants described herein. Variants can be isolated from the same tissue and organism from which the polypeptide shown in SEQ ID NO:1 was isolated, different tissues from the same organism, or from different organisms.

The 14618 receptor polynucleotides are useful as a hybridization probe for cDNA and genomic DNA to isolate a full-length cDNA and genomic clones encoding the polypeptide described in SEQ ID NO:3 and to isolate cDNA and genomic clones that correspond to variants producing the same polypeptide shown in SEQ ID NO:3 or the other variants described herein. Variants can be isolated from the same tissue and organism from which the polypeptide shown in SEQ ID NO:3 was isolated, different tissues from the same organism, or from different organisms.

The 15334 receptor polynucleotides are useful as a hybridization probe for cDNA and genomic DNA to isolate a full-length cDNA and genomic clones encoding the polypeptide described in SEQ ID NO:5 and to isolate cDNA and genomic clones that correspond to variants producing the same polypeptide shown in SEQ ID NO:5 or the other variants described herein. Variants can be isolated from the same tissue and organism from which the polypeptide shown in SEQ ID NO:5 was isolated, different tissues from the same organism, or from different organisms.

This method is useful for isolating genes and cDNA that are developmentally-controlled and therefore may be expressed in the same tissue or different tissues at different points in the development of an organism.

The probe can correspond to any sequence along the entire length of the gene encoding the receptor. Accordingly, it could be derived from 5' noncoding regions, the coding region, and 3' noncoding regions. Probes, however, are not to be construed as corresponding to any sequences that may be known prior to the invention.

WO 00/23588

PCT/US99/24368

The 2838 nucleic acid probe can be, for example, the full-length cDNA of SEQ ID NO:1, or a fragment thereof, such as an oligonucleotide of at least 11, 16, 18, 20, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to mRNA or DNA. The 14618 nucleic acid probe can be, for example, the full-length cDNA of SEQ ID NO:3, or a fragment thereof, such as an oligonucleotide of at least 8, 12, 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to mRNA or DNA. The 15334 nucleic acid probe can be, for example, the full-length cDNA of SEQ ID NO:5, or a fragment thereof, such as an oligonucleotide of at least 11, 12, 15, 18, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to mRNA or DNA.

Fragments of the polynucleotides described herein are also useful to synthesize larger fragments or full-length polynucleotides described herein. For example, a fragment can be hybridized to any portion of an mRNA and a larger or full-length cDNA can be produced.

The fragments are also useful to synthesize antisense molecules of desired length and sequence.

Antisense nucleic acids of the invention can be designed using the nucleotide sequences of SEQ ID NOS: 2, 4, and 6 constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N⁶-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N⁶-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, S⁶-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-

WO 00/23588

PCT/US99/24368

N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methyl ester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest).

Additionally, the nucleic acid molecules of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4:5). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptidic backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996), *supra*; Perry-O'Keefe *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:14670. PNAs can be further modified, e.g., to enhance their stability, specificity or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996), *supra*; Finn *et al.* (1996) *Nucleic Acids Res.* 24(17):3357-63; Mag *et al.* (1989) *Nucleic Acids Res.* 17:5973, and Peterser *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5:1119.

The nucleic acid molecules and fragments of the invention can also include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaire *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. WO 88/0918) or the blood brain barrier (see, e.g., PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (see, e.g., Krol *et al.*

WO 00/23588

PCT/US99/24368

(1988) *Bio-Techniques* 6:958-976) or intercalating agents (see, e.g., Zon (1988) *Pharm Res.* 5:539-549).

The receptor polynucleotides are also useful as primers for PCR to amplify any given region of a receptor polynucleotide.

5 The receptor polynucleotides are also useful for constructing recombinant vectors. Such vectors include expression vectors that express a portion of, or all of, the receptor polypeptides. Vectors also include insertion vectors, used to integrate into another polynucleotide sequence, such as into the cellular genome, to alter *in situ* expression of receptor genes and gene products. For example, an endogenous
10 receptor coding sequence can be replaced via homologous recombination with all or part of the coding region containing one or more specifically introduced mutations.

The receptor polynucleotides are also useful for expressing antigenic peptides. Peptide regions having a high antigenicity index are shown in Figures 2, 6, and 10.

The receptor polynucleotides are also useful as probes for determining the
15 chromosomal positions of the receptor polynucleotides by means of *in situ* hybridization methods, such as FISH (for a review of this technique, see Verma *et al.* (1988) *Human Chromosomes: A Manual of Basic Techniques* (Pergamon Press, New York), and PCR mapping of somatic cell hybrids. The mapping of the sequences to chromosomes is an important first step in correlating these sequences with genes
20 associated with disease.

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding
25 sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, *Mendelian Inheritance*
30 *in Man*, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland *et al.* (1987) *Nature* 325:783-787.

WO 00/23588

PCT/US99/24368

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with a specified gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

The receptor polynucleotide probes are also useful to determine patterns of the presence of the gene encoding the receptors and their variants with respect to tissue distribution, for example, whether gene duplication has occurred and whether the duplication occurs in all or only a subset of tissues. The genes can be naturally occurring or can have been introduced into a cell, tissue, or organism exogenously.

The receptor polynucleotides are also useful for designing ribozymes corresponding to all, or a part, of the mRNA produced from genes encoding the polynucleotides described herein.

The receptor polynucleotides are also useful for constructing host cells expressing a part, or all, of the receptor polynucleotides and polypeptides.

The receptor polynucleotides are also useful for constructing transgenic animals expressing all, or a part, of the receptor polynucleotides and polypeptides.

The receptor polynucleotides are also useful for making vectors that express part, or all, of the receptor polypeptides.

The receptor polynucleotides are also useful as hybridization probes for determining the level of receptor nucleic acid expression. Accordingly, the probes can be used to detect the presence of, or to determine levels of, receptor nucleic acid in cells, tissues, and in organisms. The nucleic acid whose level is determined can be DNA or RNA. Accordingly, probes corresponding to the polypeptides described herein can be used to assess gene copy number in a given cell, tissue, or organism. This is particularly relevant in cases in which there has been an amplification of the receptor genes.

Alternatively, the probe can be used in an *in situ* hybridization context to assess the position of extra copies of the receptor genes, as on extrachromosomal

WO 00/23588

PCT/US99/24368

elements or as integrated into chromosomes in which the receptor gene is not normally found, for example as a homogeneously staining region.

These uses are relevant for diagnosis of disorders involving an increase or decrease in receptor expression relative to normal results, such as a proliferative disorder, a differentiative or developmental disorder, or a hematopoietic disorder.

Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant expression or activity of receptor nucleic acid, in which a test sample is obtained from a subject and nucleic acid (e.g., mRNA, genomic DNA) is detected, wherein the presence of the nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant expression or activity of the nucleic acid.

One aspect of the invention relates to diagnostic assays for determining nucleic acid expression as well as activity in the context of a biological sample (e.g., blood, serum, cells, tissue) to determine whether an individual has a disease or disorder, or is at risk of developing a disease or disorder, associated with aberrant nucleic acid expression or activity. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with expression or activity of the nucleic acid molecules.

In vitro techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detecting DNA includes Southern hybridizations and *in situ* hybridization.

Probes can be used as a part of a diagnostic test kit for identifying cells or tissues that express a receptor protein, such as by measuring a level of a receptor-encoding nucleic acid in a sample of cells from a subject e.g., mRNA or genomic DNA, or determining if a receptor gene has been mutated.

Nucleic acid expression assays are useful for drug screening to identify compounds that modulate receptor nucleic acid expression (e.g., antisense, polypeptides, peptidomimetics, small molecules or other drugs). A cell is contacted with a candidate compound and the expression of mRNA determined. The level of expression of receptor mRNA in the presence of the candidate compound is compared to the level of expression of receptor mRNA in the absence of the candidate compound. The candidate compound can then be identified as a modulator of nucleic acid expression based on this comparison and be used, for example to treat a disorder

WO 00/23588

PCT/US99/24368

characterized by aberrant nucleic acid expression. The modulator can bind to the nucleic acid or indirectly modulate expression, such as by interacting with other cellular components that affect nucleic acid expression.

Modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject) in patients or in transgenic animals.

The invention thus provides a method for identifying a compound that can be used to treat a disorder associated with nucleic acid expression of the receptor gene. The method typically includes assaying the ability of the compound to modulate the expression of the receptor nucleic acid and thus identifying a compound that can be used to treat a disorder characterized by undesired receptor nucleic acid expression.

The assays can be performed in cell-based and cell-free systems. Cell-based assays include cells naturally expressing the receptor nucleic acid or recombinant cells genetically engineered to express specific nucleic acid sequences.

Alternatively, candidate compounds can be assayed *in vivo* in patients or in transgenic animals.

The assay for receptor nucleic acid expression can involve direct assay of nucleic acid levels, such as mRNA levels, or on collateral compounds involved in the signal pathway (such as cyclic AMP or phosphatidylinositol turnover). Further, the expression of genes that are up- or down-regulated in response to the receptor protein signal pathway can also be assayed. In this embodiment the regulatory regions of these genes can be operably linked to a reporter gene such as luciferase.

Thus, modulators of receptor gene expression can be identified in a method wherein a cell is contacted with a candidate compound and the expression of mRNA determined. The level of expression of receptor mRNA in the presence of the candidate compound is compared to the level of expression of receptor mRNA in the absence of the candidate compound. The candidate compound can then be identified as a modulator of nucleic acid expression based on this comparison and be used, for example to treat a disorder characterized by aberrant nucleic acid expression. When expression of mRNA is statistically significantly greater in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of nucleic acid expression. When nucleic acid expression is statistically significantly less in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of nucleic acid expression.

WO 00/23588

PCT/US99/24368

Accordingly, the invention provides methods of treatment, with the nucleic acid as a target, using a compound identified through drug screening as a gene modulator to modulate receptor nucleic acid expression. Modulation includes both up-regulation (*i.e.* activation or agonization) or down-regulation (suppression or antagonization) or effects on nucleic acid activity (e.g. when nucleic acid is mutated or improperly modified). Treatment is of disorders characterized by aberrant expression or activity of the nucleic acid.

Alternatively, a modulator for receptor nucleic acid expression can be a small molecule or drug identified using the screening assays described herein as long as the drug or small molecule inhibits the receptor nucleic acid expression.

The receptor polynucleotides are also useful for monitoring the effectiveness of modulating compounds on the expression or activity of the receptor gene in clinical trials or in a treatment regimen. Thus, the gene expression pattern can serve as a barometer for the continuing effectiveness of treatment with the compound, particularly with compounds to which a patient can develop resistance. The gene expression pattern can also serve as a marker indicative of a physiological response of the affected cells to the compound. Accordingly, such monitoring would allow either increased administration of the compound or the administration of alternative compounds to which the patient has not become resistant. Similarly, if the level of nucleic acid expression falls below a desirable level, administration of the compound could be commensurately decreased.

Monitoring can be, for example, as follows: (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a specified mRNA or genomic DNA of the invention in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the mRNA or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the mRNA or genomic DNA in the pre-administration sample with the mRNA or genomic DNA in the post-administration sample or samples; and (vi) increasing or decreasing the administration of the agent to the subject accordingly.

The receptor polynucleotides are also useful in diagnostic assays for qualitative changes in receptor nucleic acid, and particularly in qualitative changes that lead to pathology. The polynucleotides can be used to detect mutations in receptor genes and gene expression products such as mRNA. The polynucleotides

WO 00/23588

PCT/US99/24368

can be used as hybridization probes to detect naturally-occurring genetic mutations in the receptor gene and thereby to determine whether a subject with the mutation is at risk for a disorder caused by the mutation. Mutations include deletion, addition, or substitution of one or more nucleotides in the gene, chromosomal rearrangement, such as inversion or transposition, modification of genomic DNA, such as aberrant methylation patterns or changes in gene copy number, such as amplification. Detection of a mutated form of the receptor gene associated with a dysfunction provides a diagnostic tool for an active disease or susceptibility to disease when the disease results from overexpression, underexpression, or altered expression of a receptor protein.

Mutations in the receptor gene can be detected at the nucleic acid level by a variety of techniques. Genomic DNA can be analyzed directly or can be amplified by using PCR prior to analysis. RNA or cDNA can be used in the same way.

In certain embodiments, detection of the mutation involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.*, (1988) *Science* 241:1077-1080, and Nakazawa *et al.*, (1994) *PNAS* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the gene (see Abrahava *et al.* (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a gene under conditions such that hybridization and amplification of the gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. Deletions and insertions can be detected by a change in size of the amplified product compared to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to normal RNA or antisense DNA sequences.

It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional

WO 00/23588

PCT/US99/24368

amplification system (Kwob *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al.* (1988) *Bio/Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well-known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers

Alternatively, mutations in a receptor gene can be directly identified, for example, by alterations in restriction enzyme digestion patterns determined by gel electrophoresis.

Further, sequence-specific ribozymes (U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

Perfectly matched sequences can be distinguished from mismatched sequences by nuclease cleavage digestion assays or by differences in melting temperature.

Sequence changes at specific locations can also be assessed by nuclease protection assays such as RNase and S1 protection or the chemical cleavage method.

Furthermore, sequence differences between a mutant receptor gene and a wild-type gene can be determined by direct DNA sequencing. A variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, e.g., PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA duplexes (Myers *et al.* (1985) *Science* 230:1242; Cotton *et al.* (1988) *PNAS* 85:4397; Sateeba *et al.* (1992) *Meth. Enzymol.* 217:286-295), electrophoretic mobility of mutant and wild type nucleic acid is compared (Orita *et al.* (1989) *PNAS* 86:2766; Cotton *et al.* (1993) *Mutat. Rev.* 285:125-144; and Hayashi *et al.* (1992) *Genet. Anal. Tech. Appl.* 9:73-79), and movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (Myers *et al.* (1985) *Nature* 313:495). The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex

WO 00/23588

PCT/US99/24368

molecules on the basis of changes in electrophoretic mobility (Koen *et al.* (1991) *Trends Genet.* 7:5). Examples of other techniques for detecting point mutations include, selective oligonucleotide hybridization, selective amplification, and selective primer extension.

5 In other embodiments, genetic mutations can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotide probes (Cronin *et al.* (1996) *Human Mutation* 7:244-255, Kozal *et al.* (1996) *Nature Medicine* 2:753-759). For example, genetic mutations can be identified in two dimensional arrays containing
10 light-generated DNA probes as described in Cronin *et al. supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the
15 characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

The receptor polynucleotides are also useful for testing an individual for a
20 genotype that while not necessarily causing the disease, nevertheless affects the treatment modality. Thus, the polynucleotides can be used to study the relationship between an individual's genotype and the individual's response to a compound used for treatment (pharmacogenomic relationship). In the present case, for example, a mutation in the receptor gene that results in altered affinity for ligand could result in
25 an excessive or decreased drug effect with standard concentrations of ligand that activates the receptor. Accordingly, the receptor polynucleotides described herein can be used to assess the mutation content of the receptor gene in an individual in order to select an appropriate compound or dosage regimen for treatment.

Thus polynucleotides displaying genetic variations that affect treatment
30 provide a diagnostic target that can be used to tailor treatment in an individual. Accordingly, the production of recombinant cells and animals containing these polymorphisms allow effective clinical design of treatment compounds and dosage regimens.

WO 00/23588

PCT/US99/24368

The methods can involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting mRNA, or genomic DNA, such that the presence of mRNA or genomic DNA is detected in the biological sample, and comparing the presence of mRNA or genomic DNA in the control sample with the presence of mRNA or genomic DNA in the test sample.

The receptor polynucleotides are also useful for chromosome identification when the sequence is identified with an individual chromosome and to a particular location on the chromosome. First, the DNA sequence is matched to the chromosome by *in situ* or other chromosome-specific hybridization. Sequences can also be correlated to specific chromosomes by preparing PCR primers that can be used for PCR screening of somatic cell hybrids containing individual chromosomes from the desired species. Only hybrids containing the chromosome containing the gene homologous to the primer will yield an amplified fragment. Sublocalization can be achieved using chromosomal fragments. Other strategies include prescreening with labeled flow-sorted chromosomes and preselection by hybridization to chromosome-specific libraries. Further mapping strategies include fluorescence *in situ* hybridization which allows hybridization with probes shorter than those traditionally used. Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on the chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

The receptor polynucleotides can also be used to identify individuals from small biological samples. This can be done for example using restriction fragment-length polymorphism (RFLP) to identify an individual. Thus, the polynucleotides described herein are useful as DNA markers for RFLP (See U.S. Patent No. 5,272,057).

Furthermore, the receptor sequence can be used to provide an alternative technique which determines the actual DNA sequence of selected fragments in the genome of an individual. Thus, the receptor sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify DNA from an individual for subsequent sequencing.

WO 00/23588

PCT/US99/24368

Panels of corresponding DNA sequences from individuals prepared in this manner can provide unique individual identifications, as each individual will have a unique set of such DNA sequences. It is estimated that allelic variation in humans occurs with a frequency of about once per each 500 bases. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. The receptor sequences can be used to obtain such identification sequences from individuals and from tissue. The sequences represent unique fragments of the human genome. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes.

If a panel of reagents from the sequences is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

The receptor polynucleotides can also be used in forensic identification procedures. PCR technology can be used to amplify DNA sequences taken from very small biological samples, such as a single hair follicle, body fluids (e.g. blood, saliva, or semen). The amplified sequence can then be compared to a standard allowing identification of the origin of the sample.

The receptor polynucleotides can thus be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.*, another DNA sequence that is unique to a particular individual). As described above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to the noncoding region are particularly useful since greater polymorphism occurs in the noncoding regions, making it easier to differentiate individuals using this technique. Fragments are at least 12 bases.

The receptor polynucleotides can further be used to provide polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue. This is useful in cases in

WO 00/23588

PCT/US99/24368

which a forensic pathologist is presented with a tissue of unknown origin. Panels of receptor probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these primers and probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

Alternatively, the receptor polynucleotides can be used directly to block transcription or translation of receptor gene sequences by means of antisense or ribozyme constructs. Thus, in a disorder characterized by abnormally high or undesirable receptor gene expression, nucleic acids can be directly used for treatment

The receptor polynucleotides are thus useful as antisense constructs to control receptor gene expression in cells, tissues, and organisms. A DNA antisense polynucleotide is designed to be complementary to a region of the gene involved in transcription, preventing transcription and hence production of receptor protein. An antisense RNA or DNA polynucleotide would hybridize to the mRNA and thus block translation of mRNA into receptor protein.

Examples of antisense molecules useful to inhibit nucleic acid expression include antisense molecules complementary to a fragment of the 5' untranslated region of SEQ ID NOS:2, 4, or 6 which also includes the start codon and antisense molecules which are complementary to a fragment of the 3' untranslated region of SEQ ID NOS:2, 4, or 6.

Alternatively, a class of antisense molecules can be used to inactivate mRNA in order to decrease expression of receptor nucleic acid. Accordingly, these molecules can treat a disorder characterized by abnormal or undesired receptor nucleic acid expression. This technique involves cleavage by means of ribozymes containing nucleotide sequences complementary to one or more regions in the mRNA that attenuate the ability of the mRNA to be translated. Possible regions include coding regions and particularly coding regions corresponding to the catalytic and other functional activities of the receptor protein, such as ligand binding.

The receptor polynucleotides also provide vectors for gene therapy in patients containing cells that are aberrant in receptor gene expression. Thus, recombinant cells, which include the patient's cells that have been engineered *ex vivo* and returned to the patient, are introduced into an individual where the cells produce the desired receptor protein to treat the individual.

WO 00/23588

PCT/US99/24368

The invention also encompasses kits for detecting the presence of a receptor nucleic acid in a biological sample. For example, the kit can comprise reagents such as a labeled or labelable nucleic acid or agent capable of detecting receptor nucleic acid in a biological sample; means for determining the amount of receptor nucleic acid in the sample; and means for comparing the amount of receptor nucleic acid in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect receptor mRNA or DNA.

10 Computer Readable Means

The nucleotide or amino acid sequences of the invention are also provided in a variety of mediums to facilitate use thereof. As used herein, "provided" refers to a manufacture, other than an isolated nucleic acid or amino acid molecule, which contains a nucleotide or amino acid sequence of the present invention. Such a manufacture provides the nucleotide or amino acid sequences, or a subset thereof (e.g., a subset of open reading frames (ORFs)) in a form which allows a skilled artisan to examine the manufacture using means not directly applicable to examining the nucleotide or amino acid sequences, or a subset thereof, as they exist in nature or in purified form.

20 In one application of this embodiment, a nucleotide or amino acid sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium that can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. The skilled artisan will readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide or amino acid sequence of the present invention.

30 As used herein, "recorded" refers to a process for storing information on computer readable medium. The skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide or amino acid sequence information of the present invention.

WO 00/23588

PCT/US99/24368

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide or amino acid sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. The skilled artisan can readily adapt any number of dataprocessor structuring formats (e.g., text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing the nucleotide or amino acid sequences of the invention in computer readable form, the skilled artisan can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the nucleotide or amino acid sequences of the invention in computer readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

As used herein, a "target sequence" can be any DNA or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences.

WO 00/23588

PCT/US99/24368

Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium for analysis and comparison to other sequences. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, MacPattern (EMBL), BLASTN and BLASTX (NCBI).

For example, software which implements the BLAST (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410) and BLAZE (Brutlag *et al.* (1993) *Comp. Chem.* 17:203-207) search algorithms on a Sybase system can be used to identify open reading frames (ORFs) of the sequences of the invention which contain homology to ORFs or proteins from other libraries. Such ORFs are protein encoding fragments and are useful in producing commercially important proteins such as enzymes used in various reactions and in the production of commercially useful metabolites.

Vectors/host cells

The invention also provides vectors containing the receptor polynucleotides. The term "vector" refers to a vehicle, preferably a nucleic acid molecule, that can transport the receptor polynucleotides. When the vector is a nucleic acid molecule, the receptor polynucleotides are covalently linked to the vector nucleic acid. With this aspect of the invention, the vector includes a plasmid, single or double stranded phage, a single or double stranded RNA or DNA viral vector, or artificial chromosome, such as a BAC, PAC, YAC, OR MAC.

A vector can be maintained in the host cell as an extrachromosomal element where it replicates and produces additional copies of the receptor polynucleotides. Alternatively, the vector may integrate into the host cell genome and produce additional copies of the receptor polynucleotides when the host cell replicates.

The invention provides vectors for the maintenance (cloning vectors) or vectors for expression (expression vectors) of the receptor polynucleotides. The vectors can function in procaryotic or eukaryotic cells or in both (shuttle vectors).

Expression vectors contain cis-acting regulatory regions that are operably linked in the vector to the receptor polynucleotides such that transcription of the

WO 00/23588

PCT/US99/24368

polynucleotides is allowed in a host cell. The polynucleotides can be introduced into the host cell with a separate polynucleotide capable of affecting transcription. Thus, the second polynucleotide may provide a trans-acting factor interacting with the cis-regulatory control region to allow transcription of the receptor polynucleotides from the vector. Alternatively, a trans-acting factor may be supplied by the host cell. Finally, a trans-acting factor can be produced from the vector itself.

It is understood, however, that in some embodiments, transcription and/or translation of the receptor polynucleotides can occur in a cell-free system.

The regulatory sequence to which the polynucleotides described herein can be operably linked include promoters for directing mRNA transcription. These include, but are not limited to, the left promoter from bacteriophage λ , the lac, TRP, and TAC promoters from *E. coli*, the early and late promoters from SV40, the CMV immediate early promoter, the adenovirus early and late promoters, and retrovirus long-terminal repeats.

In addition to control regions that promote transcription, expression vectors may also include regions that modulate transcription, such as repressor binding sites and enhancers. Examples include the SV40 enhancer, the cytomegalovirus immediate early enhancer, polyoma enhancer, adenovirus enhancers, and retrovirus LTR enhancers.

In addition to containing sites for transcription initiation and control, expression vectors can also contain sequences necessary for transcription termination and, in the transcribed region a ribosome binding site for translation. Other regulatory control elements for expression include initiation and termination codons as well as polyadenylation signals. The person of ordinary skill in the art would be aware of the numerous regulatory sequences that are useful in expression vectors. Such regulatory sequences are described, for example, in Sambrook *et al. Molecular Cloning: A Laboratory Manual*, 2nd. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

A variety of expression vectors can be used to express a receptor polynucleotide. Such vectors include chromosomal, episomal, and virus-derived vectors, for example vectors derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, including yeast artificial chromosomes, from viruses such as baculoviruses, papovaviruses such as SV40, Vaccinia viruses, adenoviruses, poxviruses, pseudorabies viruses, and retroviruses.

WO 00/23588

PCT/US99/24368

Vectors may also be derived from combinations of these sources such as those derived from plasmid and bacteriophage genetic elements, e.g. cosmids and phagemids. Appropriate cloning and expression vectors for prokaryotic and eukaryotic hosts are described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1989).

The regulatory sequence may provide constitutive expression in one or more host cells (*i.e.* tissue specific) or may provide for inducible expression in one or more cell types such as by temperature, nutrient additive, or exogenous factor such as a hormone or other ligand. A variety of vectors providing for constitutive and inducible expression in prokaryotic and eukaryotic hosts are well known to those of ordinary skill in the art.

The receptor polynucleotides can be inserted into the vector nucleic acid by well-known methodology. Generally, the DNA sequence that will ultimately be expressed is joined to an expression vector by cleaving the DNA sequence and the expression vector with one or more restriction enzymes and then ligating the fragments together. Procedures for restriction enzyme digestion and ligation are well known to those of ordinary skill in the art.

The vector containing the appropriate polynucleotide can be introduced into an appropriate host cell for propagation or expression using well-known techniques. Bacterial cells include, but are not limited to, *E. coli*, *Streptomyces*, and *Salmonella typhimurium*. Eukaryotic cells include, but are not limited to, yeast, insect cells such as *Drosophila*, animal cells such as COS and CHO cells, and plant cells.

As described herein, it may be desirable to express the polypeptide as a fusion protein. Accordingly, the invention provides fusion vectors that allow for the production of the receptor polypeptides. Fusion vectors can increase the expression of a recombinant protein, increase the solubility of the recombinant protein, and aid in the purification of the protein by acting for example as a ligand for affinity purification. A proteolytic cleavage site may be introduced at the junction of the fusion moiety so that the desired polypeptide can ultimately be separated from the fusion moiety. Proteolytic enzymes include, but are not limited to, factor Xa, thrombin, and enterokinase. Typical fusion expression vectors include pGEX (Smith *et al.* (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST).

WO 00/23588

PCT/US99/24368

maltose B binding protein, or protein A, respectively, to the target recombinant protein. Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Armann *et al.* (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.* (1990) *Gene Expression Technology: Methods in Enzymology* 185:60-89)

- 5 Recombinant protein expression can be maximized in a host bacteria by providing a genetic background wherein the host cell has an impaired capacity to proteolytically cleave the recombinant protein. (Gottesman *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Alternatively, the sequence of the polynucleotide of interest can be
 10 altered to provide preferential codon usage for a specific host cell, for example *E. coli*. (Wada *et al.* (1992) *Nucleic Acids Res.* 20:2111-2118)

- The receptor polynucleotides can also be expressed by expression vectors that are operative in yeast. Examples of vectors for expression in yeast e.g., *S. cerevisiae* include pYepSec1 (Baldari, *et al.* (1987) *EMBO J.* 6:229-234), pMFa (Kurjan *et al.*
 15 (1982) *Cell* 30 933-943), pJRY88 (Schultz *et al.* (1987) *Gene* 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA).

- The receptor polynucleotides can also be expressed in insect cells using, for example, baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf9 cells) include the pAc series
 20 (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow *et al.* (1989) *Virology* 170:31-39).

- In certain embodiments of the invention, the polynucleotides described herein are expressed in mammalian cells using mammalian expression vectors. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and
 25 pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195).

- The expression vectors listed herein are provided by way of example only of the well-known vectors available to those of ordinary skill in the art that would be useful to express the receptor polynucleotides. The person of ordinary skill in the art would be aware of other vectors suitable for maintenance propagation or expression
 30 of the polynucleotides described herein. These are found for example in Sambrook, *et al. T. Molecular Cloning: A Laboratory Manual*, 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

The invention also encompasses vectors in which the nucleic acid sequences described herein are cloned into the vector in reverse orientation, but operably linked

WO 00/23588

PCT/US99/24368

to a regulatory sequence that permits transcription of antisense RNA. Thus, an antisense transcript can be produced to all, or to a portion, of the polynucleotide sequences described herein, including both coding and non-coding regions.

Expression of this antisense RNA is subject to each of the parameters described above

5 in relation to expression of the sense RNA (regulatory sequences, constitutive or inducible expression, tissue-specific expression).

The invention also relates to recombinant host cells containing the vectors described herein. Host cells therefore include prokaryotic cells, lower eukaryotic cells such as yeast, other eukaryotic cells such as insect cells, and higher eukaryotic
10 cells such as mammalian cells.

The recombinant host cells are prepared by introducing the vector constructs described herein into the cells by techniques readily available to the person of ordinary skill in the art. These include, but are not limited to, calcium phosphate transfection, DEAE-dextran-mediated transfection, cationic lipid-mediated
15 transfection, electroporation, transduction, infection, lipofection, and other techniques such as those found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual*, 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Host cells can contain more than one vector. Thus, different nucleotide sequences can be introduced on different vectors of the same cell. Similarly, the
20 receptor polynucleotides can be introduced either alone or with other polynucleotides that are not related to the receptor polynucleotides such as those providing trans-acting factors for expression vectors. When more than one vector is introduced into a cell, the vectors can be introduced independently, co-introduced or joined to the
25 receptor polynucleotide vector.

In the case of bacteriophage and viral vectors, these can be introduced into cells as packaged or encapsulated virus by standard procedures for infection and transduction. Viral vectors can be replication-competent or replication-defective. In
30 the case in which viral replication is defective, replication will occur in host cells providing functions that complement the defects.

Vectors generally include selectable markers that enable the selection of the subpopulation of cells that contain the recombinant vector constructs. The marker can be contained in the same vector that contains the polynucleotides described herein or may be on a separate vector. Markers include tetracycline or ampicillin-resistance

WO 00/23598

PCT/US99/24368

genes for prokaryotic host cells and dihydrofolate reductase or neomycin resistance for eukaryotic host cells. However, any marker that provides selection for a phenotypic trait will be effective.

While the mature proteins can be produced in bacteria, yeast, mammalian cells, and other cells under the control of the appropriate regulatory sequences, cell-free transcription and translation systems can also be used to produce these proteins using RNA derived from the DNA constructs described herein.

Where secretion of the polypeptide is desired, appropriate secretion signals are incorporated into the vector. The signal sequence can be endogenous to the receptor polypeptides or heterologous to these polypeptides.

Where the polypeptide is not secreted into the medium, the protein can be isolated from the host cell by standard disruption procedures, including freeze thaw, sonication, mechanical disruption, use of lysing agents and the like. The polypeptide can then be recovered and purified by well-known purification methods including ammonium sulfate precipitation, acid extraction, anion or cationic exchange chromatography, phosphocellulose chromatography, hydrophobic-interaction chromatography, affinity chromatography, hydroxylapatite chromatography, lectin chromatography, or high performance liquid chromatography.

It is also understood that depending upon the host cell in recombinant production of the polypeptides described herein, the polypeptides can have various glycosylation patterns, depending upon the cell, or maybe non-glycosylated as when produced in bacteria. In addition, the polypeptides may include an initial modified methionine in some cases as a result of a host-mediated process.

25 Uses of vectors and host cells

It is understood that "host cells" and "recombinant host cells" refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

The host cells expressing the polypeptides described herein, and particularly recombinant host cells, have a variety of uses. First, the cells are useful for producing receptor proteins or polypeptides that can be further purified to produce desired

WO 00/23588

PCT/US99/24368

amounts of receptor protein or fragments. Thus, host cells containing expression vectors are useful for polypeptide production

Host cells are also useful for conducting cell-based assays involving the receptor or receptor fragments. Thus, a recombinant host cell expressing a native
5 receptor is useful to assay for compounds that stimulate or inhibit receptor function. This includes ligand binding, gene expression at the level of transcription or translation, G-protein interaction, and components of the signal transduction pathway.

Host cells are also useful for identifying receptor mutants in which these functions are affected. If the mutants naturally occur and give rise to a pathology,
10 host cells containing the mutations are useful to assay compounds that have a desired effect on the mutant receptor (for example, stimulating or inhibiting function) which may not be indicated by their effect on the native receptor.

Recombinant host cells are also useful for expressing the chimeric polypeptides described herein to assess compounds that activate or suppress activation
15 by means of a heterologous amino terminal extracellular domain (or other binding region). Alternatively, a heterologous region spanning the entire transmembrane domain (or parts thereof) can be used to assess the effect of a desired amino terminal extracellular domain (or other binding region) on any given host cell. In this embodiment, a region spanning the entire transmembrane domain (or parts thereof)
20 compatible with the specific host cell is used to make the chimeric vector. Alternatively, a heterologous carboxy terminal intracellular, e.g., signal transduction, domain can be introduced into the host cell.

Further, mutant receptors can be designed in which one or more of the various functions is engineered to be increased or decreased (e.g., ligand binding or G-protein
25 binding) and used to augment or replace receptor proteins in an individual. Thus, host cells can provide a therapeutic benefit by replacing an aberrant receptor or providing an aberrant receptor that provides a therapeutic result. In one embodiment, the cells provide receptors that are abnormally active.

In another embodiment, the cells provide receptors that are abnormally
30 inactive. These receptors can compete with endogenous receptors in the individual.

In another embodiment, cells expressing receptors that cannot be activated, are introduced into an individual in order to compete with endogenous receptors for ligand. For example, in the case in which excessive ligand is part of a treatment modality, it may be necessary to inactivate this ligand at a specific point in treatment.

WO 00/23588

PCT/US99/24368

Providing cells that compete for the ligand, but which cannot be affected by receptor activation would be beneficial

Homologously recombinant host cells can also be produced that allow the *in situ* alteration of endogenous receptor polynucleotide sequences in a host cell genome. The host cell includes, but is not limited to, a stable cell line, cell *in vivo*, or cloned microorganism. This technology is more fully described in WO 93/09222, WO 91/12650, WO 91/06667, U.S. 5,272,071, and U.S. 5,641,670. Briefly, specific polynucleotide sequences corresponding to the receptor polynucleotides or sequences proximal or distal to a receptor gene are allowed to integrate into a host cell genome by homologous recombination where expression of the gene can be affected. In one embodiment, regulatory sequences are introduced that either increase or decrease expression of an endogenous sequence. Accordingly, a receptor protein can be produced in a cell not normally producing it. Alternatively, increased expression of receptor protein can be effected in a cell normally producing the protein at a specific level. Further, expression can be decreased or eliminated by introducing a specific regulatory sequence. The regulatory sequence can be heterologous to the receptor protein sequence or can be a homologous sequence with a desired mutation that affects expression. Alternatively, the entire gene can be deleted. The regulatory sequence can be specific to the host cell or capable of functioning in more than one cell type. Still further, specific mutations can be introduced into any desired region of the gene to produce mutant receptor proteins. Such mutations could be introduced, for example, into the specific functional regions such as the ligand-binding site.

In one embodiment, the host cell can be a fertilized oocyte or embryonic stem cell that can be used to produce a transgenic animal containing the altered receptor gene. Alternatively, the host cell can be a stem cell or other early tissue precursor that gives rise to a specific subset of cells and can be used to produce transgenic tissues in an animal. See also Thomas *et al.* (1987) *Cell* 51:503 for a description of homologous recombination vectors. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced gene has homologously recombined with the endogenous receptor gene is selected (see e.g., Li *et al.* (1992) *Cell* 69:915). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras (see e.g., Bradley in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into

WO 00/23588

PCT/US99/24368

a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for
5 constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos. WO 90/11354; WO 91/011140; and WO 93/04169.

The genetically engineered host cells can be used to produce non-human
10 transgenic animals. A transgenic animal is preferably a mammal, for example a rodent, such as a rat or mouse, in which one or more of the cells of the animal include a transgene. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal in one or more cell types or tissues of the transgenic animal. These
15 animals are useful for studying the function of a receptor protein and identifying and evaluating modulators of receptor protein activity.

Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, and amphibians.

In one embodiment, a host cell is a fertilized oocyte or an embryonic stem cell
20 into which receptor polynucleotide sequences have been introduced.

A transgenic animal can be produced by introducing nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Any of the receptor nucleotide sequences can be introduced as a transgene into the genome of a
25 non-human animal, such as a mouse.

Any of the regulatory or other sequences useful in expression vectors can form part of the transgenic sequence. This includes intronic sequences and polyadenylation signals, if not already included. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the receptor protein to
30 particular cells.

Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.* U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan

WO 00/23588

PCT/US99/24368

Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the transgene in its genome and/or expression of transgenic mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene can further be bred to other transgenic animals carrying other transgenes. A transgenic animal also includes animals in which the entire animal or tissues in the animal have been produced using the homologously recombinant host cells described herein.

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso *et al.* (1992) *PNAS* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *S. cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355). If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein is required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyst and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, e.g., the somatic cell, is isolated.

WO 00/23588

PCT/US99/24368

Transgenic animals containing recombinant cells that express the polypeptides described herein are useful to conduct the assays described herein in an *in vivo* context. Accordingly, the various physiological factors that are present *in vivo* and that could effect ligand binding, receptor activation, and signal transduction, may not be evident from *in vitro* cell-free or cell-based assays. Accordingly, it is useful to provide non-human transgenic animals to assay *in vivo* receptor function, including ligand interaction, the effect of specific mutant receptors on receptor function and ligand interaction, and the effect of chimeric receptors. It is also possible to assess the effect of null mutations, that is mutations that substantially or completely eliminate one or more receptor functions.

In general, methods for producing transgenic animals include introducing a nucleic acid sequence according to the present invention, the nucleic acid sequence capable of expressing the receptor protein in a transgenic animal, into a cell in culture or *in vivo*. When introduced *in vivo*, the nucleic acid is introduced into an intact organism such that one or more cell types and, accordingly, one or more tissue types, express the nucleic acid encoding the receptor protein. Alternatively, the nucleic acid can be introduced into virtually all cells in an organism by transfecting a cell in culture, such as an embryonic stem cell, as described herein for the production of transgenic animals, and this cell can be used to produce an entire transgenic organism. As described, in a further embodiment, the host cell can be a fertilized oocyte. Such cells are then allowed to develop in a female foster animal to produce the transgenic organism.

Pharmaceutical compositions

The receptor nucleic acid molecules, protein (particularly fragments such as the amino terminal extracellular domain), modulators of the protein, and antibodies (also referred to herein as "active compounds") can be incorporated into pharmaceutical compositions suitable for administration to a subject, e.g., a human. Such compositions typically comprise the nucleic acid molecule, protein, modulator, or antibody and a pharmaceutically acceptable carrier.

As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for

WO 00/23588

PCT/US99/24368

pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, such media can be used in the compositions of the invention. Supplementary active compounds can also be incorporated into the compositions. A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as

WO 00/23588

PCT/US99/24368

mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

5 Sterile injectable solutions can be prepared by incorporating the active compound (e.g., a receptor protein or anti-receptor antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above.

10 In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They

15 can be enclosed in gelatin capsules or compressed into tablets. For oral administration, the agent can be contained in enteric forms to survive the stomach or further coated or mixed to be released in a particular region of the GI tract by known methods. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules.

20 Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or

25 compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or

30 orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

WO 00/23588

PCT/US99/24368

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated, each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

WO 00/23588

PCT/US99/24368

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (U.S. 5,328,470) or by stereotactic injection (see e.g., Chen *et al.* (1994) *PNAS* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g. retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

As defined herein, a therapeutically effective amount of protein or polypeptide (i.e., an effective dosage) ranges from about 0.001 to 50 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight.

The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments. In a preferred example, a subject is treated with antibody, protein, or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody, protein, or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids,

WO 00/23588

PCT/US99/24368

amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds.

It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention. Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (e.g., about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (e.g., a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

This invention may be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments

WO 00/23598

PCT/US99/24368

are provided so that this disclosure will fully convey the invention to those skilled in the art. Many modifications and other **embodiments** of the invention will come to mind in one skilled in the art to which this invention pertains having the benefit of the teachings presented in the foregoing description. Although specific terms are
5 employed, they are used as in the art unless otherwise indicated.

WO 00/23588

PCT/US99/24368

THAT WHICH IS CLAIMED

1. An isolated polypeptide having an amino acid sequence selected from the group consisting of:
 - 5 (a) The amino acid sequence shown in SEQ ID NO 1, the amino acid sequence shown in SEQ ID NO 3, and the amino acid sequence shown in SEQ ID NO 5;
 - (b) The amino acid sequence of an allelic variant of the amino acid sequence shown in SEQ ID NO 1, the amino acid sequence of an allelic variant of the amino acid sequence shown in SEQ ID NO 3, and the amino acid sequence of an allelic variant of the amino acid sequence shown in SEQ ID NO 5;
 - 10 (c) The amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 1, wherein the sequence variant is encoded by a nucleic acid molecule hybridizing to the nucleic acid molecule shown in SEQ ID NO 2 under stringent conditions, the amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 3, wherein the sequence variant is encoded by a nucleic acid molecule hybridizing to the nucleic acid molecule shown in SEQ ID NO 4 under stringent conditions, and the amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 5, wherein the sequence variant is encoded by a nucleic acid molecule hybridizing to the nucleic acid molecule shown in SEQ ID NO 6 under stringent conditions;
 - 20 (d) A fragment of the amino acid sequence shown in SEQ ID NO 1, wherein the fragment comprises at least 8 contiguous amino acids from 1-264, a fragment of the amino acid sequence shown in SEQ ID NO 3, wherein the fragment comprises at least 9 contiguous amino acids from 1-244, and a fragment of the amino acid sequence shown in SEQ ID NO 5, wherein the fragment comprises at least 8 contiguous amino acids;
 - 25 (e) The amino acid sequence of the mature receptor polypeptide from about amino acid 6 to about amino acid 319, shown in SEQ ID NO 1, the amino acid sequence of the mature receptor polypeptide from about amino acid 6 to about amino acid 337, shown in SEQ ID NO 3, and the amino acid sequence of the mature receptor polypeptide from about amino acid 6 to about amino acid 372, shown in SEQ ID NO 5;
 - 30 (f) The amino acid sequence of the polypeptide shown in SEQ ID NO 1, from about amino acid 1 to about amino acid 24, the amino acid sequence of the polypeptide shown in SEQ ID NO 3, from about amino acid 1 to about amino acid 28,

WO 00/23588

PCT/US99/24368

and the amino acid sequence of the polypeptide shown in SEQ ID NO 5, from about amino acid 1 to about amino acid 25,

(g) The amino acid sequence of an epitope bearing region of any one of the polypeptides of (a)-(f).

5

2. An isolated antibody that selectively binds to a polypeptide of claim 1, (a)-(g).

3. An isolated nucleic acid molecule having a nucleotide sequence selected from the group consisting of

(a) The nucleotide sequence shown in SEQ ID NO 2, the nucleotide sequence shown in SEQ ID NO 4, and the nucleotide sequence shown in SEQ ID NO 6;

(b) A nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO 1, a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO 3, and a nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO 5;

(c) A nucleotide sequence complementary to any of the nucleotide sequences in (a) or (b).

4. An isolated nucleic acid molecule having a nucleotide sequence selected from the group consisting of:

(a) A nucleotide sequence encoding an amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 1 that hybridizes to the nucleotide sequence shown in SEQ ID NO 2 under stringent conditions, a nucleotide sequence encoding an amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 3 that hybridizes to the nucleotide sequence shown in SEQ ID NO 4 under stringent conditions, and a nucleotide sequence encoding an amino acid sequence of a sequence variant of the amino acid sequence shown in SEQ ID NO 5 that hybridizes to the nucleotide sequence shown in SEQ ID NO 6 under stringent conditions;

(b) A nucleotide sequence complementary to the nucleotide sequences in (a).

WO 00/23588

PCT/US99/24368

5. An isolated nucleic acid molecule a polynucleotide having a nucleotide sequence selected from the group consisting of:
- (a) A nucleotide sequence encoding a fragment of the amino acid sequence shown in SEQ ID NO 1, wherein the fragment comprises at least 8 contiguous amino acids from 1-264, a nucleotide sequence encoding a fragment of the amino acid sequence shown in SEQ ID NO 3, wherein the fragment comprises at least 9 contiguous amino acids from 1-244, and a nucleotide sequence encoding a fragment of the amino acid sequence shown in SEQ ID NO 5, wherein the fragment comprises at least 8 contiguous amino acids;
- (b) A nucleotide sequence complementary to the nucleotide sequences in (a).
6. A nucleic acid vector comprising the nucleic acid sequences in any of claims 3-5.
7. A host cell containing the vector of claim 6.
8. A method for producing any of the polypeptides in claim 1 comprising introducing a nucleotide sequence encoding any of the polypeptide sequences in (a)-(f) into a host cell, and culturing the host cell under conditions in which the proteins are expressed from the nucleic acid.
9. A method for detecting the presence of any of the polypeptides in claim 1 in a sample, said method comprising contacting said sample with an agent that specifically allows detection of the presence of the polypeptide in the sample and then detecting the presence of the polypeptide.
10. The method of claim 9, wherein said agent is capable of selective physical association with said polypeptide.
11. The method of claim 10, wherein said agent binds to said polypeptide.
12. The method of claim 11, wherein said agent is an antibody.

WO 00/23588

PCT/US99/24368

13. The method of claim 11, wherein said agent is a ligand.
14. A kit comprising reagents used for the method of claim 9, wherein the reagents comprise an agent that specifically binds to said polypeptide.
- 5 15. A method for detecting the presence of any of the nucleic acid sequences in any of claims 3-5 in a sample, the method comprising contacting the sample with an oligonucleotide that hybridizes to the nucleic acid sequences under stringent conditions and determining whether the oligonucleotide binds to the nucleic acid sequence in the sample.
- 10 16. The method of claim 15, wherein the nucleic acid, whose presence is detected, is mRNA.
- 15 17. A kit comprising reagents used for the method of claim 15, wherein the reagents comprise a compound that hybridizes under stringent conditions to any of the nucleic acid molecules.
18. A method for identifying an agent that interacts with any of the polypeptides in claim 1, said method comprising contacting said agent with a cell capable of allowing an interaction between said polypeptide and said agent such that said polypeptide can interact with said agent and measuring the interaction.
- 20 19. A method of screening a cell to identify an agent that interacts with any of the polypeptides in claim 1, said method comprising contacting said agent with a cell capable of allowing an interaction between said polypeptide and said agent such that said polypeptide can interact with said agent and measuring the interaction.
- 25 20. A method for identifying an agent that modulates the level or activity of any of the polypeptides of claim 1 in a cell, said method comprising contacting said agent with a cell capable of expressing said polypeptide such that said polypeptide level or activity can be modulated in said cell by said agent and measuring said polypeptide level or activity.
- 30

WO 00/23588

PCT/US99/24368

21. A method of screening a cell to identify an agent that modulates the level or activity of any of the polypeptides of claim 1 in a cell, said method comprising contacting said agent with a cell capable of expressing said polypeptide such that said polypeptide level or activity can be modulated in said cell by said agent
5 and measuring said polypeptide level or activity.

22. The method of either of claims 19 or 20 wherein said interaction is binding.

10 23. The method of claim 22, wherein a fragment of the polypeptide is contacted.

24. A method for modulating the activity of any of the polypeptides in claim 1, the method comprising contacting any of the polypeptides of claim 1 with an agent under conditions that allow the agent to modulate the activity of the polypeptide.
15

25. A method for identifying an agent that interacts with any of the nucleic acid molecules of claims 3-5, said method comprising contacting said agent with a cell capable of allowing an interaction between said nucleic acid molecule and said agent such that said nucleic acid molecule can interact with said agent and measuring the interaction.
20

26. A method of screening a cell to identify an agent that interacts with any of the nucleic acid molecules of claims 3-5, said method comprising contacting said agent with a cell capable of allowing an interaction between said nucleic acid molecule and said agent such that said nucleic acid molecule can interact with said agent and measuring the interaction.
25

30 27. A method for identifying an agent that modulates the level or activity of any of the nucleic acid molecules of claims 3-5 in a cell, said method comprising contacting said agent with a cell capable of expressing said nucleic acid molecule such that said nucleic acid molecule level or activity can be modulated in said cell by said agent and measuring said nucleic acid molecule level or activity.

WO 00/23588

PCT/US99/24368

28. A method of screening a cell to identify an agent that modulates the level or activity of any of the nucleic acid molecules of claims 3-5 in a cell, said method comprising contacting said agent with a cell capable of expressing said nucleic acid molecule such that said nucleic acid molecule level or activity can be modulated in said cell by said agent and measuring said nucleic acid molecule level or activity.

29. The method of either of claims 25 or 26 wherein said interaction is binding.

30. A method for modulating the activity of any of the nucleic acid molecules in claim 3-5, the method comprising contacting any of the nucleic acid molecules of claim 1 with an agent under conditions that allow the agent to modulate the activity of the nucleic acid molecule.

31. A pharmaceutical composition containing any of the polypeptides in claim 1 in a pharmaceutically acceptable carrier.

32. A pharmaceutical composition containing any of the nucleic acid molecules of claims 3-5 in a pharmaceutically acceptable carrier.

33. A non-human transgenic animal wherein one or more of the cells of said animal contains any of the nucleic acid sequences of claims 3-5.

34. A non-human transgenic animal wherein one or more of cells of said animal contains any of the nucleic acid sequences of claims 3-5 wherein said cell expresses any of the polypeptides of claim 1.

35. A method for producing a transgenic animal according to claims 32 and 33, said method comprising introducing any of the nucleic acid sequences of claim 1 into a cell, wherein said cell is present in said animal or gives rise to said animal.

WO 00/23588

PCT/US99/24368

- 36 An agent identified using any of the methods of claims 19-22 or 25-28.
- 37 A method of treating a disorder in a subject comprising administering the pharmaceutical compositions of either of claims 31 or 32.
- 5 38 A method of treating a disorder in a subject by modulating the level or expression of any of the polypeptides of claim 1 in said subject.
- 39 A method of treating a disorder in a subject by modulating the level or activity of any of the nucleic acid molecules of claims 3-5 in said subject.
- 10 40 The method of either of claims 9 or 15 wherein said detecting is in a cell derived from a subject having a disorder involving said cell.
- 15 41 The method of either of claims 24 or 30 wherein said modulation is in a subject having a disorder involving said cell.
- 42 The method of claim 39 wherein said disorder is anemia, neutropenia or thrombocytopenia.
- 20 43 The method of claim 41 wherein said disorder is anemia, neutropenia or thrombocytopenia.

WO 00/23588

PCT/US99/24368

1/19

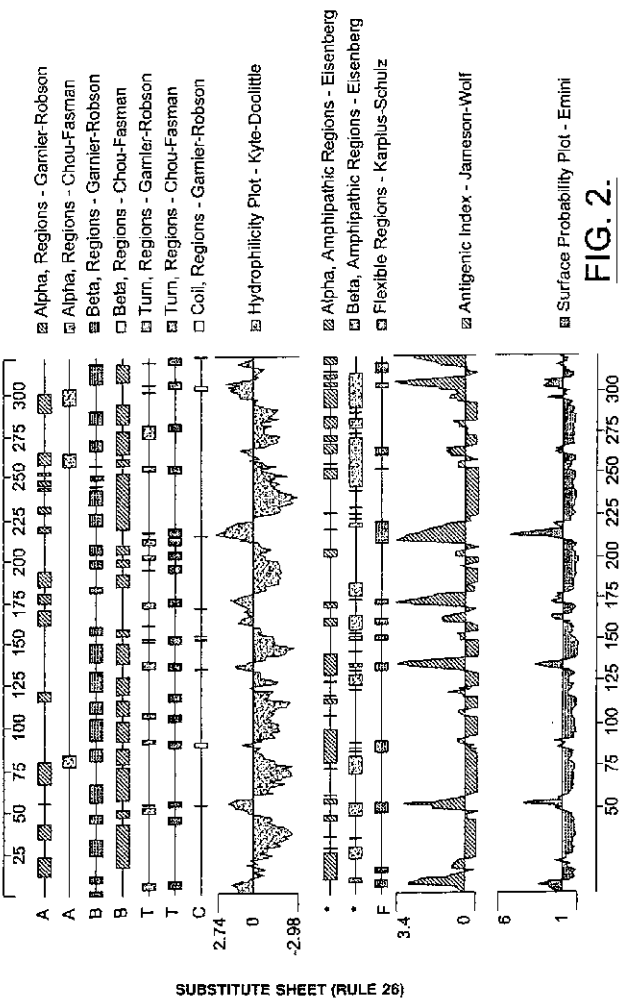
Protein Family/Domain HMM Matches for flh2838

>PF00001/7tm_1 7 transmembrane receptor (rhodopsin family)

Score: 154.54 Seq: 37 288 Model: 1 269
 *GN:LVIVicRyRRMR...TPMNYFIVNLAVADLLFsfMPFWVvYy
 N+L I ++ + R +++++NLAV DLL+ +++PF MV
 flh2838 37 LNLLAIHGFSTFLKNRVDPDYAATSIYHINLAVFDLL-VLSLFFKMYLS 84
 vMqgRWpFGdFMCrIWvYFDYMMMYASIFFLtCISIDRYLWAICHPMvYn
 PF++ +C ++ +++++MY S+F+++IS+DR+L AI +P+
 flh2838 85 QV-Q-SPPPS-LCTLVECLYFVSMYGSVFTICFISMREL-AIRYPLLVS 130
 RWMTPRHRAWvMIiIWvNSF!ISMPPFLMFrVstyrDenevNnTWOnly
 ++PR ++ +++ IWV+ S+P+ + F + E + M C+
 flh2838 131 HLRSPR-KIFGICCTIWVLVWTGSIPI-YSF-HGKV--EKY--M--CFH- 170
 DWPeWwWvYvILmtlingFYIPMITMIFCYwRIYRIoRIWMRNipswQr
 + ++W + V + + ++GF++PM IM FC +R I ++ + ++ Q
 flh2838 171 NMSDDTWSAKVFFPLEVFGLLPMGIMGFCSSRS--I-HILLGRRDHTQD 217
 RRnSnRrERRivKMliIWvVFICWLPYFivnFMDTLM.MwvFCefC.
 + + ++ + + + VF++++LP ++ +F++ L + + E C
 flh2838 171 WVQQ-----KACIYSTAASLAVFVVSFLPVHLGFLQFLYRNSFIVE-CR 261
 IwrrIWn.Y.IfeWLoYvNCPGiNPliY*
 + +++ + + + + + VNC C++ + Y
 flh2838 262 AKQSTISFFLQLSMCFSNVNC-CLDVFCY 288

FIG. 1.

SUBSTITUTE SHEET (RULE 26)



WO 00/23588

PCT/US99/24368

3/19

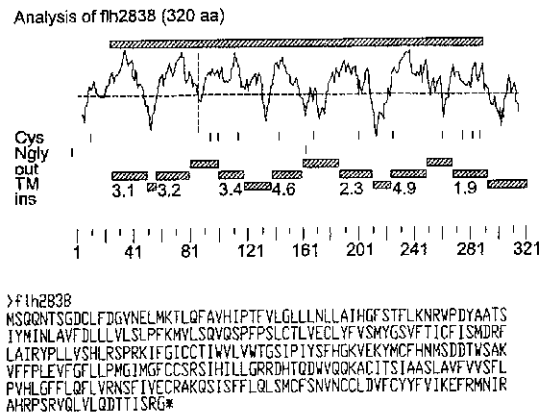


FIG. 3.

WO 00/23588

PCT/US99/24368

4/19

Prosite Pattern Matches for flh2838

>PS00001/PDOC00001/ASN_GLYCOSYLATION N-glycosylation site.

Query: 5 NTSG 8

Query: 171 NTSE 174

>PS00005/PDOC00005/PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 134 SPR 136

Query: 178 SAK 180

>PS00006/PDOC00006/CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query: 6 TSGD 9

Query: 95 TLVE 98

>PS00008/PDOC00008/MYRISTYL N-myristoylation site.

Query: 34 GLLLNL 39

Query: 107 GSVFTI 112

Query: 140 GICCTI 145

>PS00009/PDOC00009/AMIDATION Amidation site.

Query: 209 LGRR 212

>PS00237/PDOC00210/G_PROTEIN_RECEPTOR G-protein coupled receptors signature.

Query: 107 GSVFTICFISMDRFLAI 123

>PS00339/PDOC00363/AA_TRNA_L1(GASE_II_2 Aminoacyl-transfer RNA synthetases class - II sign

Query: 114 FISMDRFLAI 123

FIG. 4.

WO 00/23588

PCT/US99/24368

5/19

Protein Family/Domain HMM Maatches for flh14618

>PF00001/7tm_1 7 transmembrane receptor (rhodopsin family)

Score: 230.98 Seq: 40 295 Model: 1 269

```

*GNILVIVVLCRYRRMTPMNYFIVNLAVADLLFsIfMPFVWvYvMqg
GN++V++V++++ + ++ + +++NLAVADLL + T+P+ VYYY+ G
flh14618 40 GNGFVLYVLKTYHKSAFQVYMINLAVADLLC-VCTLPLRVVYVYHKG 87
RVpFGdFMdriWnYFDYNNMYASIFLTcISIDRYLWATChPMrYmRWMT
V+FGDF+CR+++Y Y+N+Y SIFF+T++S+ R +Al+ P++ ++ T
flh14618 88 IWLFGFLCRLSTYALYNLYCSIFFMTAMSPERC-IAIVFPVGNINLYT 136
pRHRAWvMIi iWvNSFIISMPPFLMFrVstYrDeneWnTWcnlyDWPe
+ A+++++iW++ tL S P F ++ +DE + T C++ ++
flh14618 137 OK-KARFVCGIWFIFVILTSSP-F-LW-AKPKKDEKN--NTKDFE-PPQD 179
...wMwVYvILmtiinGFIYIPM;IMIFCYwRIYRJaRIWMRMipswGrR
+++ +++ +++++ GF+P+ J+++CY++I I L++++ S+++
flh14618 180 NQTKNHVLVLYVSLVGGFIIPFVILVCYTM1--I--LTLK-KSMKK- 223
RsnSmRrERRivKH1i iJhVVF LICW2PYFIvMFMDILMwvF.....
+++++++ M++++ +f+++++PY+I + +++++
flh14618 224 ----NLSSHKKAIGMIMVVTAAFLVSFMPYHIQRT-----IHLHFLHNETK 265
.CeFCIwrrjWvY. IfeWiaYvNCpC:NP1IY*
C+ +++R++ I++ LA +NC C++P++Y
flh14618 266 PCD-SVLRHQKSVVITLSLAASNC-CFDPLLY 295

```

FIG. 5.

SUBSTITUTE SHEET (RULE 26)

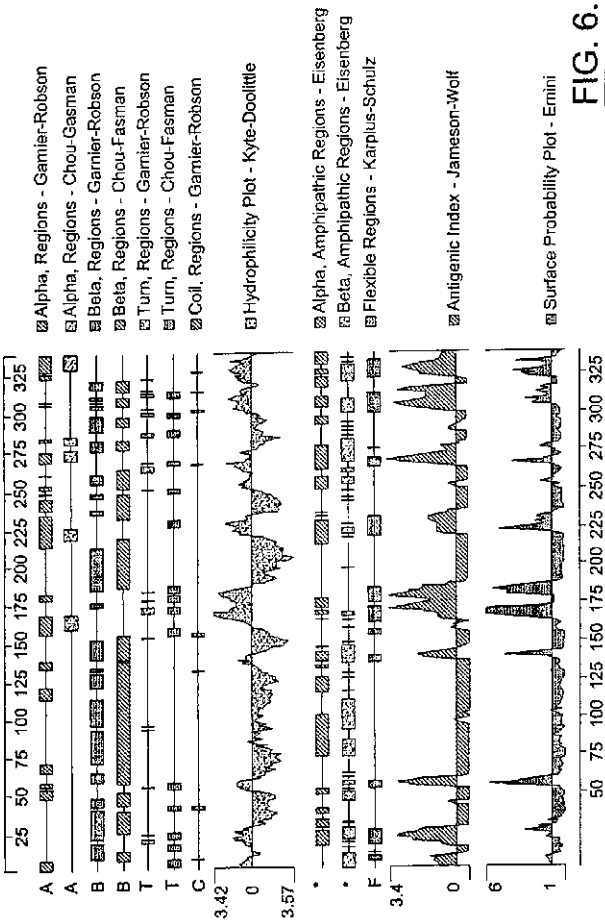
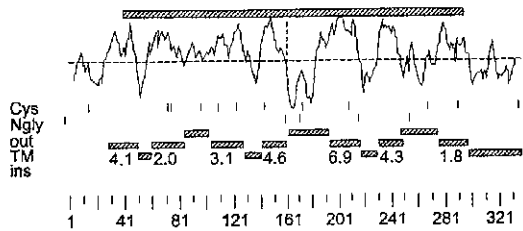


FIG. 6.

Analysis of flh1461B (338 aa)



>flh1461B
MDETGNLTSSATCHDTIDDFRHOVYSTLYSMISVVGFFQNGFVLYVLIKTYHKKSARQV
YMINLAVADLLCVCTPLRVVYVYHKGIVLFGDFLCRLSTYALYVNLVCSIFFMTAMSF
RCIATVFPVGNILVTGKKARFVCGVWIFVILTSSPFLMAKPKIDEKNNTKCFEPPQBN
DTKNHYLVLYHYSLVGGFIIPFVITIVCYTMIILTLKKSHKKNLSSHKKAIIGMIMVYTA
AFLVSFMPYHTQRTIHLHFLHNETKPCDSVLRMOKSVVITLSLAASNCCEPPLL YFFSGG
NFRKRLSTFRKHSLSSTVYYPKKASLPEKGEETCKV*

FIG. 7.

WO 00/23588

PCT/US99/24368

8/19

Prosite Pattern Matches for flh14618

>PS00001/PDOC0001/ASN_GLYCOSYLATION N-glycosylation site.

Query: 6 NLTV 9
Query: 169 NLTV 172
Query: 180 NLTV 183
Query: 224 NLTV 227
Query: 262 NLTV 265

>PS00004/PDOC0004/CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation

Query: 304 KRLS 307
Query: 310 RKHS 313
Query: 323 KKAS 326

>PS00005/PDOC0005/PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 136 TQK 138
Query: 220 SMK 222
Query: 227 SHK 229
Query: 308 TFR 310

>PS00006/PDOC0006/CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query: 13 TCHD 16
Query: 17 TIDD 20
Query: 326 SLPE 329

>PS00342/PDOC00299/MICROBODIES_CTER Microbodies C-terminal targeting signal.

Query: 335 CKV* 338

FIG. 8.

WO 00/23588

PCT/US99/24368

9/19

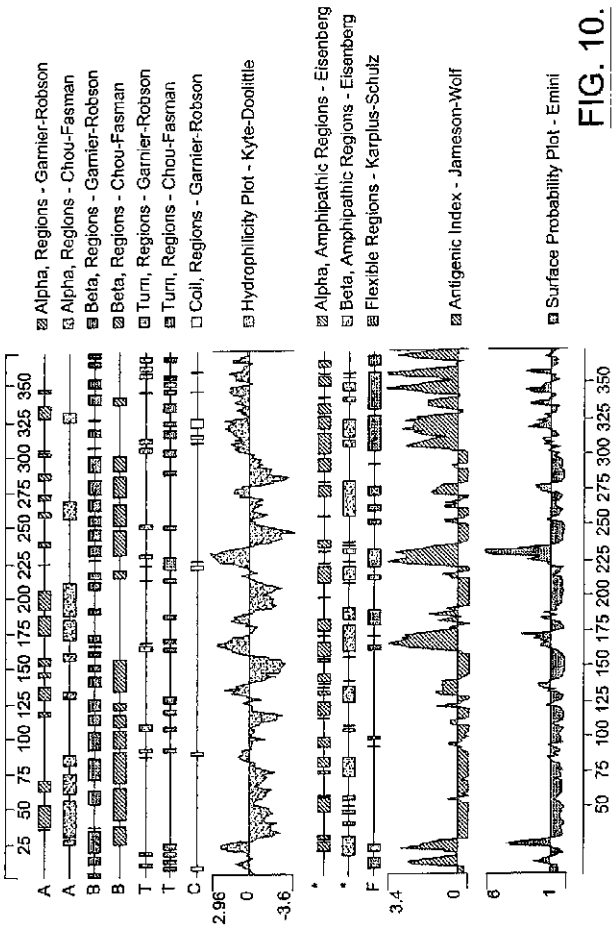
```

>PF00001/PF00001.7 transmembrane receptor (rhodopsin family)
  Score: 217.21 Seq: 39 297 Model: 1 269
  *GNILVIVVichYRRWRTPMNYFYVNLAVADLLFsIFMPFWNvYvMqg
    N L +VV++R R +++ +++++NL A +DLLF +++++P+ + YY++ +
15334orfaa 39 LNALALVFLRALRVHSVSVYMCNLAASDLLF-TLSLPVRESYAL-H 85
  RVppGdFNCrIWhYFDYMMNYASIFLTcISIDRYLVAICHPMYvRWMT
  +VPP+D++C ++++++ MHHY S .FL++I +DRY AI+HP+R + +++
15334orfaa 86 HWPPDILLQDTTGAIPQMNHYGSCIFLHLINVRV-AAIVHPLRLRLRR 134
  pRHRAWvMiiIiVvMSFIISMPPFLMFr. WstynDeNeVWnTWcNiYDWP
  PR +A+++++V++ +++++P + ++R ++R+++ + + C++++
15334orfaa 135 PR-VARLLCLGVWALILVFAPVPAARVHRPS-RCRYRDLEVRL-CFESFSD 181
  e...mWvYVYVlNtitingFYIPMiIMFCYwRIYRLoRIWHRMipswDr
  E + +VHL +GF++P+ +++ +R+++++ + + +
15334orfaa 182 ELWKGRLLPLVLLAEALGFLPLAAVYVYSSGRVFWLARPDAQ---SQ 227
  RRrnSnRxERRivKMiIiIMvVFIDWIPVFIvFMDTL...M.MwWFC
  R+R++V+ L++ +V+F++C+ PY+ ++ ++ L + ++
15334orfaa 228 -----RRRTVRLLLANLVIFLLCFVPYNSTLAVYGLLRSLVAASVP 270
  eFCIwrrIWhYIfvLaYvNCpCINPIIY*
  + R + M + ++LA +HC +++P+Y
15334orfaa 271 ARDRYRGVLM-VWVLLAGANC-VLDPLVY 297

>PF00010/PF00010 Helix-loop-helix DNA-binding domain
  Score: 0.33 Seq: 157 177 Model: 1 21
  *RRRHMMREGRRRNvNEaFE*
  +R+ +R R R +Y++ FE
15334orfaa 157 ARVHRPSRCRYRDLEVRLCFE 177

```

FIG. 9.



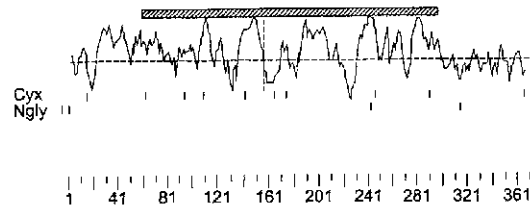
SUBSTITUTE SHEET (RULE 26)

WO 00/23589

PCT/US99/24368

11/19

Analysis of sequence28390 (372 aa)



```

>sequence28390
MLANSSSTNGSYLPCPDYRPTHRLHLVYYSVLAAGLPLNALAVVFLRALRVHSVVSYY
MDNLAASDLFLTLSPVRLSYALHWPFPDILLCTTGATFQMMYGCIFLMLINVDYR
AATVHPLRLRHLRPPVARELLCLGWALTLVFAVPAARVHRPSRCRYRDLVRLCFESF
DELWGRLLPLVLLACALGFLPLAAVYSSGRVFWTLARPDATGSRRRKTVRLLLANL
VIFLLCFVPYNSTLAVYQLRSKLYAASVPARDVVRQVLMVWVLLAGANCVLDPVYVFS
AEGFRNTLRGLGTPHARTSATNGTRALAGSERSAVTTDATRPBAASQGLLRPSDSHSL
SFTGCPQDSAL

```

FIG. 10.

WO 00/23588

PCT/US99/24368

12/19

Prosite Pattern Matches for sequence 28390
Prosite version: Release 12.2 of February 1995

>PS00001/PDOC0001/ASN_GLYCOSYLATION N-glycosylation site.

Query: 4 NSSS 7
Query: 9 NSSV 12
Query: 251 NSTL 254
Query: 323 NGTR 326

>PS00004/PDOC0004/CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation

Query: 229 RRKT 232

>PS00005/PDOC0005/PKC_PHOSPHO_SITE Protein kinase C phosphorylation site.

Query: 21 THR 23
Query: 211 SGR 213
Query: 226 SGR 228
Query: 232 TVR 234
Query: 307 TLR 309
Query: 332 SER 334

>PS00006/PDOC0006/CK2_PHOSPHO_SITE Casein kinase II phosphorylation site.

Query: 178 SFSD 181
Query: 342 TRPD 345

>PS00008/PDOC0008/MYRISTYL N-myristoylation site.

Query: 36 GLPLNA 41
Query: 258 GLLRSK 263
Query: 324 GTRAAL 329

>PS00237/PDOC00210/G_PROTEIN_RECEPTOR G-protein coupled receptors signature.

Query: 107 GSCIFLMLINVIYAAI 123

FIG. 12.

SUBSTITUTE SHEET (RULE 26)

WO 00/23588

PCT/US99/24368

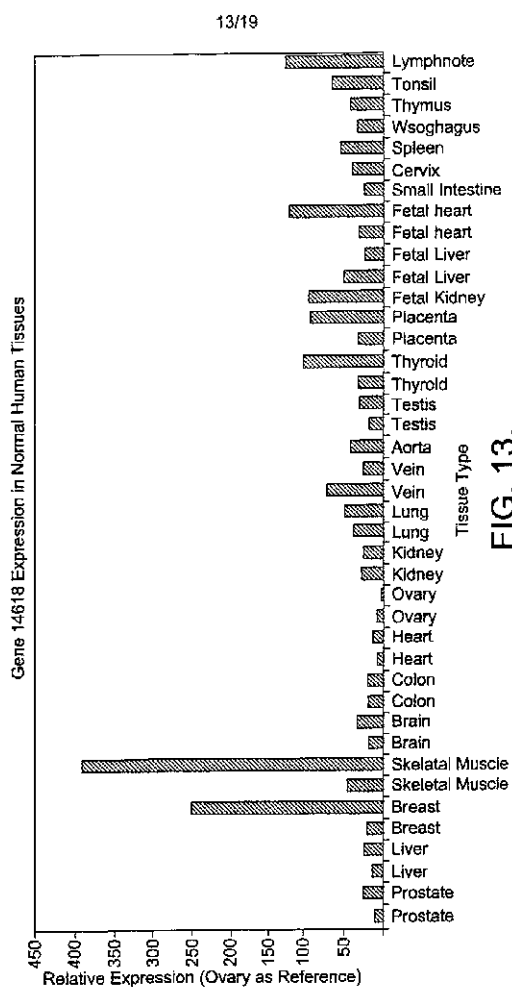
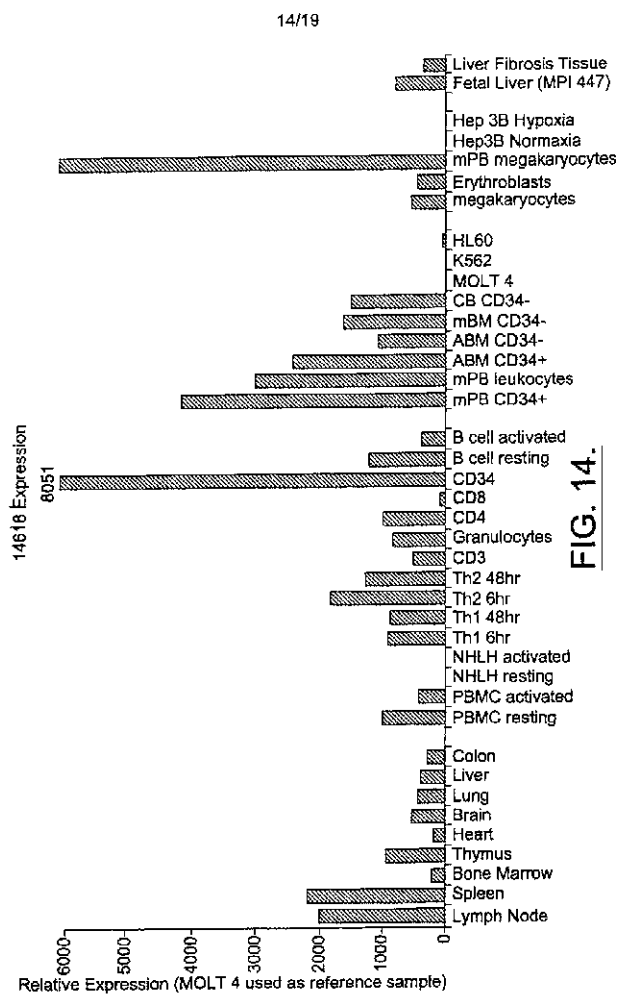
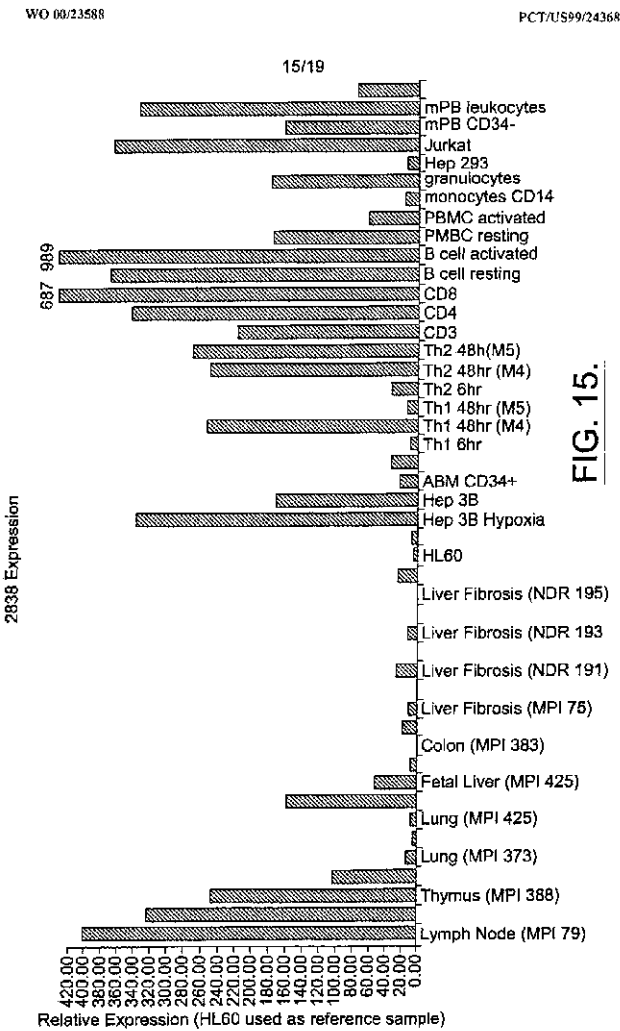


FIG. 13.

WO 00/23588

PCT/US99/24368

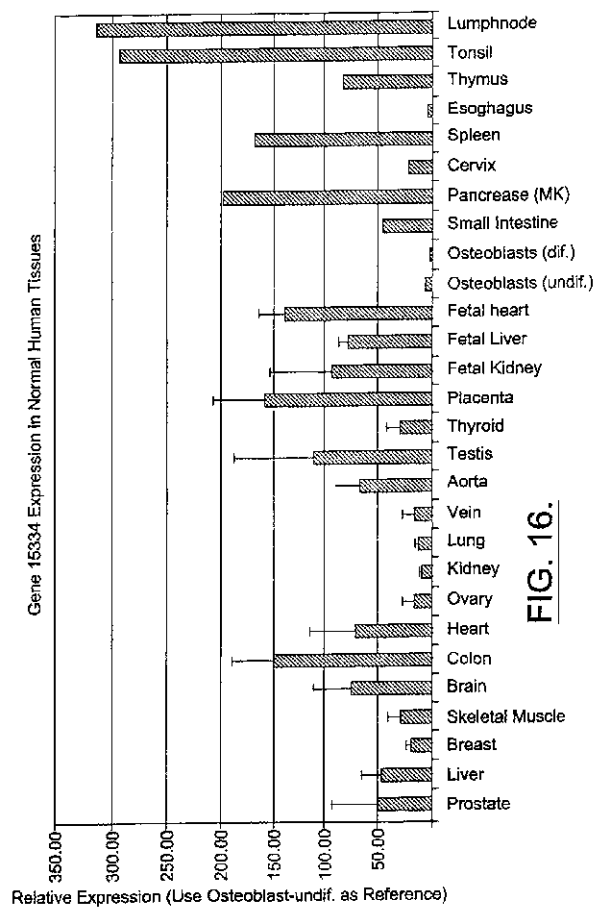




WO 06/23588

PCT/US99/24368

16/19



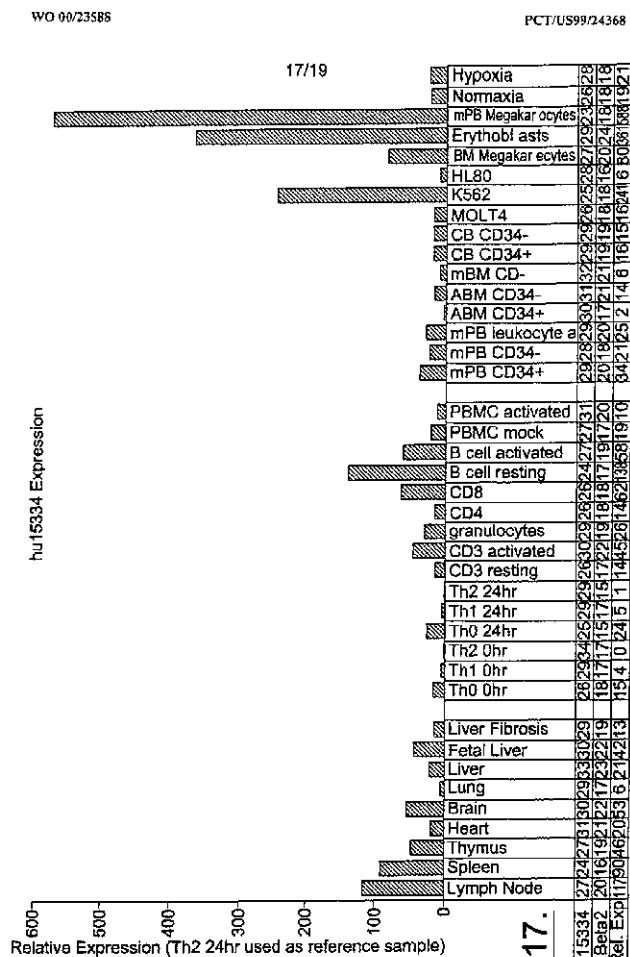
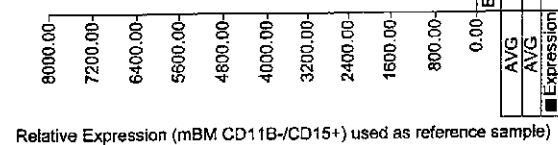
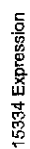
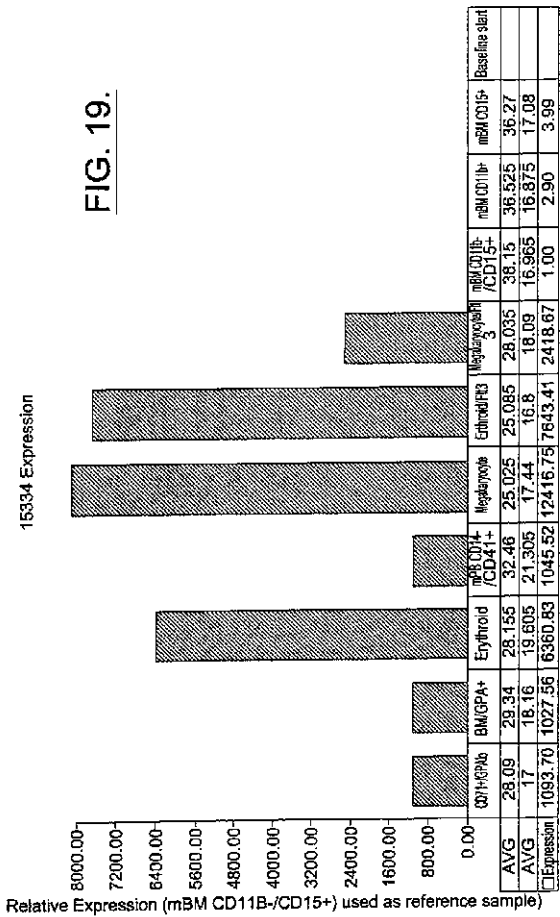


FIG. 17.

FIG. 18.



SUBSTITUTE SHEET (RULE 26)



WO 00/23588

PCT/US99/24368

SEQUENCE LISTING:

<110> Glucksmann, Maria A.

<120> Novel G-Protein Coupled Receptors

<130> SR26-11-1

<140>

<141>

<160> 7

<170> PatentIn Ver. 2.0

<210> 1

<211> 319

<212> PRT

<213> Homo sapiens

<400> 1

Met Ser Gln Gln Asn Thr Ser Gly Asp Cys Leu Phe Asp Gly Val Asn
 1 5 10 15

Glu Leu Met Lys Thr Leu Gln Phe Ala Val His Ile Pro Thr Phe Val
 20 25 30

Leu Gly Leu Leu Leu Asn Leu Leu Ala Thr His Gly Phe Ser Thr Phe
 35 40 45

Leu Lys Asn Arg Trp Pro Asp Tyr Ala Ala Thr Ser Ile Tyr Met Ile
 50 55 60

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

Met Val Leu Ser Gln Val Gln Ser Pro Phe His Ser Leu Cys Thr Leu
 85 90 95

Val Glu Cys Leu Tyr Phe Val Ser Met Tyr Gly Ser Val Phe Thr Trp
 100 105 110

Cys Phe Ile Ser Met Asp Arg Phe Leu Ala Ile Arg Tyr Pro Leu Leu
 115 120 125

Val Ser His Leu Arg Ser Pro Arg Lys Ile Phe Gly Ile Cys Cys Thr
 130 135 140

WO 00/23588

PCT/US99/24368

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

<210> 2

<211> 1617

<212> DNA

<213> Homo sapiens

<400> 2

aggtacagcc ttgtgacatt aaagaatlea ggcagggaac tccacacatga ctttactctt 60
 gtggactgca gtgtgtattc aatgggaatg aatgtgtgtg tgaactgtgg aaggtgtttg 120
 tgattcttga caaagtcatt tgaatccalc acttcaagag agtgaagga gcccggtctg 180
 atctgttggg gtgtgtaaaa gaacacatgag tcaatcaaac accagtgggg actgtctggt 240
 tgacgggtgt aacgagctga tgaataacatt acagtttgaq gtccacatcc caactctt 300
 cctgggcttg cttctcaaac tgcctggccat ccatggcttt agaaccttcc ttaagaaacag 360
 gtggcccgat tatgtgtgca cctccatata catgatcaaz ctggaagtc ttagccttct 420
 gctggtgtct tttctccat taaagatggt cctgttccag gtacagltcc ccttcccgta 480

WO 00/23588

PCT/US99/24368

ccctgtgacac ctggggaggt gccttctctt ccttcggcgt tacgaaagcg tcttcacacat 540
 ctgttccatc agcatgggac ggcttcttggg ccttccttacc cggctacgtg tgagccacat 600
 ccggcttctt aggaagatct ttgqgacatcg cggccacatc ttgqgacatcg ttgqgacatcg 660
 aagcatctct atctacagtt tccatgggga agtgggaaaa tacatgttgt taccacacat 720
 gctctgatgat acctggagcg aaaggttctt ccttcgtgtg gacgtgtttg gcttctctct 780
 tccatggggt atcatggggt tctgtgtctc caggagcacc cactctctgc tgggctgctg 840
 agaccacacc caggaatggg tccagcagaa ggcctgcctc taccagcctg caggcagcct 900
 ggtgtctctc gtgtctctct tcttcacagt ccaacgtggg ttcttctctg agttctctgt 960
 gagaaacggt tctatctgtg agtgcagagc caacacagag atnagcttct ccttgcacat 1020
 gctaatgtgt tcttcacagc tcaactgtgt cctgtgtgtt tctgtctctc accttctcat 1080
 caaagacatc cgtctgagaa tcaagggcga cggcctctct aggtctctgc tggctctgca 1140
 ggacacccatg atcttcctcg gctaacggaa ggcacatctg ttcaggggaa gaaagcctg 1200
 ggcctgaatt ctgttaacgt attctgcctt cagggtgtt gctgtgtgtg gatgatccgc 1260
 accatctctt ctgatgtgt tctcttctgt ggcctctgag tgcacgtctt gctcattata 1320
 ccccaagac ctttctctca cggcctgac agcttatcag accaggtgtt cagggaatct 1380
 tgaagacccc acagccaggt tgaattctc atttctaacg ccaaaacata lagagccgaa 1440
 gaattcgaaa aaggaacgag accatgtcaa caaggtctgt ccccaactct ccatcttctc 1500
 tgtctgactg ggaggttctg gaaagaaaaa gagagagaga aagaggttaa cggagggagc 1560
 caagagagtc agttattggg gagggtgtct tgggcagagg tgggtgtgta gaaatga 1617

<210> 3

<211> 337

<212> PRT

<213> Homo sapiens

<400> 3

Met Asp Glu Thr Gly Asn Leu Thr Val Ser Ser Ala Thr Cys His Asp

1 5 10 15

Thr Ile Asp Asp Phe Arg Asn Gln Val Tyr Ser Thr Leu Tyr Ser Met

20 25 30

Ile Ser Val Val Gly Phe Phe Gly Asn Gly Phe Val Leu Tyr Val Leu

35 40 45

Ile Lys Thr Tyr His Lys Lys Ser Ala Phe Gln Val Tyr Met Ile Asn

50 55 60

Leu Ala Val Ala Asp Leu Leu Cys Val Cys Thr Leu Pro Leu Arg Val

65 70 75 80

Val Tyr Tyr Val His Lys Gly Ile Trp Leu Phe Gly Asp Phe Leu Cys

85 90 95

Arg Leu Ser Thr Tyr Ala Leu Tyr Val Asn Leu Tyr Cys Ser Ile Phe

100 105 110

Phe Met Thr Ala Met Ser Phe Phe Arg Cys Ile Ala Ile Val Phe Pro

WO 00/23588

PCT/US99/24368

```

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

```

Val

<210> 4

<211> 1358

<212> DNA

WO 00/23588
 <213> Homo sapiens

PCT/US99/24368

<400> 4
 accttcaggc cagaattcgg caggaqgctg gcagatcgaa tttatlgag actlggagct 60
 tgcctctcgg agcaaaagca aaaggacgt aaactgtgga ccttgaggt agnaggctgg 120
 gcttcctcta atattacac gtaaaagcca tggatcaca taagaaggaa cacttgctga 180
 ggtactccag tgcagaaag aggcacaaa cagacattcg taqaaacaa tggatgaac 240
 aggaatcta aaagtatct ntgcacacg ccctggcact attgatgact tggcaatca 300
 aatgtattcc acctcgtact ctatgatct tcttclaggg tttcttgcca atggcttctg 360
 gctctatgta cctctaaaa cctalcacaa gaactcagcc ttcacagta acatgallca 420
 tctagcagta gacgatctac tctgtgtgtg cccactgccc ctccgtgtgg cctattacgt 480
 tcacaaggc atttggtct tctgtgactt cttgtgagc ctacacaccl atgcttcta 540
 tgtcaacatc tattgtagca tctctttat gacagccag agcttttcc ggtgcattgg 600
 aattgctttt ccaatcaga acattaattt ggttccacag aaaaaagcca ggttctgtgt 660
 tctaggctat tggattttt tgattttgat cagctctcca ttttcaatg ccaaaccaaa 720
 aaaaagtgag aaaaaataa ccaagtgcct tgcgcctcca caagacatc aaactaaaa 780
 tcatgtttt gtttgcatt atgtgtcatt ggttctggg tttatctcc ctcttcttat 840
 tataattgta tgttccaaa tgaatcattt gaccttacta ccaaaatcaa tgaaaaaaa 900
 cctgtcagat cctaaaaagg cctatggat gatctatgt ctcacagctg cctttttagt 960
 aagttctatg cctalcata ttcacagta ccttcacat cctttttat acaatgaac 1020
 taacacctg gattctgtcc ttgaatgca gaactcgtg gctataact tgcctctggc 1080
 tgaatccat ttttctttg acctcttctt atattctctt tctgggggtt actttagga 1140
 aaggtctgtt acattagaa agcattcttt gtcacagctg acttatgtac ccagaagaa 1200
 ggcctcttta ccaaaaaag gacagaaat atgtacagta taqtttaaa ccatlccag 1260
 tcaaaaccaa tgaatatgt ttcacaaa atattttgt ccaatcattt acaaaaaaa 1320
 taataatttt acttaaaaa aaaaaaaa aaggaaaa 1380

<210> 5
 <211> 372
 <212> PRT
 <213> Homo sapiens

<400> 5
 Met Leu Ala Asn Ser Ser Thr Asn Ser Ser Val Leu Pro Cys Pro
 1 5 10 15
 Asp Tyr Arg Pro Thr His Arg Leu His Leu Val Val Tyr Ser Leu Val
 20 25 30
 Leu Ala Ala Gly Leu Pro Leu Asn Ala Leu Ala Leu Trp Val Phe Leu
 35 40 45
 Arg Ala Leu Arg Val His Ser Val Val Ser Val Tyr Met Cys Asn Leu
 50 55 60
 Ala Ala Ser Asp Leu Leu Phe Thr Leu Ser Leu Pro Val Arg Leu Ser
 65 70 75 80

WO 00/23588

PCT/US99/24368

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
 Met Leu Ile Asp Val Asp Arg Tyr Ala Gly Ile Val His Pro Leu Arg
 115 120 125
 Leu Arg His Leu Arg Arg Ala 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
 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 Ile Pro Leu Val
 180 185 190
 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
 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
 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
 Ala Thr Asn Gly Thr Arg Ala Ala Leu Ala Gln Ser Glu Arg Ser Ala
 325 330 335

WO 00/23588 PCT/US99/24368

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 Phe Gln
 355 360 365

Asp Ser Ala Leu
 370

<210> 6
 <211> 2339
 <212> DNA
 <213> Homo sapiens

<400> 6
 ccattgaaccc tctctgcttg ttttgggaac atgltctgac agcctctagg ccccgagccc 60
 ggaggtgaat gccaatgccat gattcttggtg tgcctaatgg caccaccagc ctatgctccca 120
 atcccaacttt ggcacgatgt tagucaacag ctacccaacc aacagttctg ttctcccggt 180
 tcttgactac cgaactaccc aacgcttgcg cttaqctggt cacagcttgg tgetgggtac 240
 cgggtctccac ctcaacgggc taaccccttg ggtcttcttg cgcgcgctgc gctgacactc 300
 ggtgggtgagc gtagacatgt gtaacctggc ggcacagcag ctgctcttcc cctctctggt 360
 gcccggtctgt cctctctact aagcaatgca ccaactggcc ctcccgagac tcttggtcca 420
 gacgaggggc gccatcttcc agatgaacat gtacgggagc tgcctcttcc tgaatctcat 480
 caacgtggag cgtacggcg gcatcgtgca cccgctggga ctgcgccacc cgcggcgggc 540
 ccgctggcg cggctgctct gcttggtggt gggggcgctc atcttggtgt ttgcgttacc 600
 cgcgcgcgcg gtagaacagg cctccgcttg ccgtacagcg gacctcaggg tgcgctatg 660
 cttcagagagc ttccagcagc agctgtggaa aggcaggctg ctgcacctcg tctgcttgg 720
 cagggcgctg ggtctctgga tgcctctggc ggcgttgctc tctctgctgg gctgagttct 780
 ctggagcctg ggcgcgcgcg acgcacagca gtagcagcgg cggcggaaga cgtgagct 840
 ctgctgggtt aacctctgca tcttctgtgt glecttctg cctcacaca gccagctggc 900
 ggtctacggc cgtctgggga ccaagcttgt ggcggcagc gtccctgccc gcatcgcgr 960
 gccggggctg ctgctggtga tggctgtgct ggcggcgccc aactgctgc tggatcact -1020
 ggtgtactac ttagcgccg agggctctcg caaaccttg cgcggcttgg gcatccgca 1080
 ccgggcccag aacctggcca ccaacgggac ggcggcgccg ctgcgcgcat cggaaaggcc 1140
 cgcgctcacu accagagcca ccaggccggc tgcgcgcagt cagggcgctg tccgacctc 1200
 cgaactccac tctctgttct ccttcacaca gtgtcccaag gattccgcc tctgaacaca 1260
 catgcatctg cgtctctcgt gcccgactcc caacgctctc cgttctggga gcttcaagg 1320
 gtgtacacac aagaaggctg gctgggcaat tggaccttgg ggtggcaatt ccagcttacc 1380
 aacgcagagc agtacaaagt gtggaaagcca gggcccaagg aagtcagtcg tctgggaaat 1440
 ggccttctta aactgtgagc accgagagca cccctctctc agcggtagga agtgatccag 1500
 aaagcccaac cgtgcagagg gcaagaaggc acgaaatgcc ttgggtggg cagggtatta 1560
 aactgtaaa agctggctag atggacccga aaatgggat tccggatttt aaccgcacac 1620
 aggggcttga gagttgaaga gccccaggtt tgggtggcaa agctactgag atgctgttcc 1680
 atctgtgac ttctgtctag gccatggac gccacccctt ttcatttgg cttagggttc 1740
 cctgtctcac caatgagggc taatcacagc gtctctatgg acagaaetac attctttctc 1800
 gcatagtgac ttgtgaacat tttagcttgg catccagctt gggataatg gggcaaggca 1860
 aaactaactt agagtttccc ccttaacacg atccagcttc aaacctttt taggttctcc 1920

WO 00/23588

PCT/US99/24368

```

ttttttccat aatattccct ttctcagccc tcttcacatt taggtcccca atattctctta 1980
ttttttctct tcttcagccc ctctctctct tcttcctctc tcttcctctc ctctctctct 2040
ctctctccct ctctctccct gcttcaggtc aggtctctct ctctctctc aggtctctct 2100
ctctctctct ttgtctctc gcttcaggtc ttatctctct aggtctctc aggtctctct 2160
ttatctctc ttctctctc ctctctctc ctctctctc aggtctctc atttttctct 2220
taattctctc tcttcaggtc aggtctctc aggtctctc aggtctctc aggtctctc 2280
atgagctctc aggtctctc tcttcaggtc ttgtctctc aggtctctc aggtctctc 2340
taattctctc gcttcaggtc ttatctctc gcttcaggtc gcttcaggtc gcttcaggtc 2400
ggtctctctc tcttcaggtc aggtctctc gcttcaggtc gcttcaggtc gcttcaggtc 2460
ctctctctct gcttcaggtc tcttcaggtc ttatctctc gcttcaggtc gcttcaggtc 2520
ataattctct ctctctctct aggtctctc aggtctctc aggtctctc 2580

```

<210> 7
 <211> 269
 <212> PRT
 <213> Unknown

<220>
 <221> Description of Unknown Organism: Rhodopsin family
 transmembrane receptor

<400> 7
 Gly Asn Ile Leu Val Ile Trp Val Ile Cys Arg Tyr Arg Arg Met Arg
 1 5 10 15
 Thr Pro Met Asn Tyr Phe Ile Val Asn Leu Ala Val Ala Asp Leu Leu
 20 25 30
 Phe Ser Leu Phe Thr Met Pro Phe Trp Met Val Tyr Tyr Val Met Gly
 35 40 45
 Gly Arg Trp Pro Phe Gly Asp Phe Met Cys Arg Ile Trp Met Tyr Phe
 50 55 60
 Asp Tyr Met Asn Met Tyr Ala Ser Ile Phe Phe Leu Thr Cys Ile Ser
 65 70 75 80
 Ile Asp Arg Tyr Leu Trp Ala Ile Cys His Pro Met Arg Tyr Met Arg
 85 90 95
 Trp Met Thr Pro Arg His Arg Ala Trp Val Met Ile Ile Ile Ile Trp
 100 105 110
 Val Met Ser Phe Leu Ile Ser Met Pro Pro Phe Leu Met Phe Arg Trp
 115 120 125
 Ser Thr Tyr Arg Asp Glu Asn Glu Trp Asn Met Thr Trp Cys Met Ile
 130 135 140

WO 00/23588

PCT/US99/24368

Tyr Asp Trp Pro Glu Trp Met Trp Arg Trp Tyr Val Ile Met Met Thr
 145 150 155 160
 Ile Ile Met Gly Phe Tyr Ile Pro Met Ile Ile Met Leu Phe Cys Tyr
 165 170 175
 Trp Arg Ile Tyr Arg Ile Ala Arg Leu Trp Met Arg Met Ile Pro Ser
 180 185 190
 Trp Gln Arg Arg Arg Arg Met Ser Met Arg Arg Glu Arg Arg Ile Val
 195 200 205
 Lys Met Leu Ile Ile Ile Met Val Val Phe Ile Ile Cys Trp Leu Pro
 210 215 220
 Tyr Phe Ile Val Met Phe Met Asp Thr Leu Met Met Trp Trp Phe Cys
 225 230 235 240
 Glu Phe Cys Ile Trp Arg Arg Leu Trp Met Tyr Ile Phe Glu Trp Leu
 245 250 255
 Ala Tyr Val Asn Cys Pro Cys Ile Asn Pro Ile Ile Tyr
 260 265

【手続補正書】

【提出日】平成12年7月7日(2000.7.7)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

- (a) 配列番号5に示されるアミノ酸配列と、
- (b) A T C C 寄託番号____に含まれるc D N Aによりコードされるアミノ酸配列と、
- (c) 配列番号5に示したアミノ酸配列の対立遺伝子変異体のアミノ酸配列と、
- (d) A T C C 寄託番号____に含まれるc D N Aによりコードされるアミノ酸配列の対立遺伝子変異体のアミノ酸配列と、
- (e) 配列番号5に示したアミノ酸配列の配列変異体であって、ストリンジェントな条件下で、配列番号6に示した核酸分子にハイブリッド形成する核酸分子によりコードされる配列変異体のアミノ酸配列と、
- (f) A T C C 寄託番号____に含まれるc D N Aクローンによりコードされるアミノ酸配列の配列変異体であって、A T C C 寄託番号____に含まれるc D N Aに、ストリンジェントな条件下でハイブリッド形成する核酸分子によりコードされる配列変異体のアミノ酸配列と、
- (g) 少なくとも8つの連続したアミノ酸を含む、配列番号5に示したアミノ酸配列の断片と、
- (h) A T C C 寄託番号____に含まれるc D N Aによりコードされる、少なくとも8つの連続したアミノ酸を含むアミノ酸配列の断片と、
- (i) 配列番号5に示したアミノ酸6～372番目の成熟受容体ポリペプチドのアミノ酸配列と、
- (j) A T C C 寄託番号____に含まれるc D N Aクローンによりコードされるアミノ酸6～372番目の成熟ポリペプチドのアミノ酸配列と、
- (k) アミノ酸1～25番目の、配列番号5に示したポリペプチドのアミノ酸配列と、
- (l) A T C C 寄託番号____に含まれるc D N Aによりコードされるポリペプチドのアミノ酸1～25番目のアミノ酸配列と、
- (m) (a)から(l)までのポリペプチドのいずれか1つのエピトープ含有領域のアミノ酸配列と

からなる群から選択されるアミノ酸配列を有する単離ポリペプチド。

【請求項2】

請求項1の(a)～(m)に選択的に結合する単離抗体。

【請求項3】

- (a) 配列番号6に示したヌクレオチド配列と、
- (b) A T C C 寄託番号____に含まれるc D N Aのヌクレオチド配列と、
- (c) 配列番号5に示したアミノ酸配列をコードするヌクレオチド配列と、
- (d) A T C C 寄託番号____に含まれるc D N Aによりコードされるアミノ酸配列をコードするヌクレオチド配列と、
- (e) (a)、(b)、(c)または(d)のヌクレオチド配列のいずれかに相補的なヌクレオチド配列と

からなる群から選択されるヌクレオチド配列を有する単離核酸分子。

【請求項4】

- (a) 配列番号6に示したヌクレオチド配列にストリンジェントな条件下でハイブリッド形成する、配列番号5に示したアミノ酸配列の配列変異体のアミノ酸配列をコードするヌクレオチド配列と、

(b) A T C C 寄託番号 _____ に含まれる c D N A によりコードされるアミノ酸配列の配列変異体のアミノ酸配列をコードするヌクレオチド配列と、ここで、前記配列変異体の核酸配列はストリンジェントな条件下で A T C C 寄託番号 _____ に含まれる c D N A にハイブリッド形成し、

(c) (a) または (b) のヌクレオチド配列のいずれかに相補的なヌクレオチド配列とからなる群から選択される、ヌクレオチド配列を有する単離核酸分子。

【請求項 5】

(a) 少なくとも 8 つの連続したアミノ酸を含む配列番号 5 に示したアミノ酸配列の断片をコードするヌクレオチド配列と、

(b) 少なくとも 8 つの連続したアミノ酸を含む、A T C C 寄託番号 _____ に含まれる c D N A によりコードされるアミノ酸配列の断片をコードするヌクレオチド配列と、

(c) (a) または (b) のヌクレオチド配列のいずれかに相補的なヌクレオチド配列とからなる群から選択されるヌクレオチド配列を有する単離核酸分子。

【請求項 6】

請求項 3 ～ 5 のいずれかに記載の核酸配列を含む核酸ベクター。

【請求項 7】

請求項 6 のベクターを含む宿主細胞。

【請求項 8】

(a) ～ (m) のポリペプチド配列のいずれかをコードするヌクレオチド配列を宿主細胞に導入するステップと、タンパク質が核酸から発現される条件下で前記宿主細胞を培養するステップとを含む請求項 1 のポリペプチドのいずれかの製造方法。

【請求項 9】

サンプル中の請求項 1 のポリペプチドのいずれかの存在を検出する方法であって、サンプル中のポリペプチドの存在の検出を特異的に可能とする物質と前記サンプルとを接触させるステップと、前記ポリペプチドの存在を検出するステップとを含む方法。

【請求項 10】

前記物質は、前記ポリペプチドとの選択的な物理的会合が可能である請求項 9 の方法。

【請求項 11】

前記物質は前記ポリペプチドに結合する請求項 10 の方法。

【請求項 12】

前記物質は抗体である請求項 11 の方法。

【請求項 13】

前記物質はリガンドである請求項 11 の方法。

【請求項 14】

請求項 9 の方法に使用するための試薬を含むキットであって、前記試薬は、前記ポリペプチドに特異的に結合する物質を含むことを特徴とするキット。

【請求項 15】

サンプル中の請求項 3 ～ 5 のいずれかの核酸配列のいずれかの存在の検出方法であって、核酸配列にストリンジェントな条件下でハイブリッド形成するオリゴヌクレオチドとサンプルとを接触させるステップと、オリゴヌクレオチドがサンプル中の核酸配列と結合するかを決定するステップとを含む方法。

【請求項 16】

存在を検出する核酸は、m R N A である請求項 15 の方法。

【請求項 17】

請求項 15 の方法に使用する試薬を含むキットであって、試薬が、ストリンジェントな条件下で核酸分子のいずれかにハイブリッド形成する化合物を含むことを特徴とするキット。

【請求項 18】

請求項 1 のポリペプチドのいずれかに結合する物質を同定する方法であって、ポリペプチドに結合する物質とポリペプチドとを接触させるステップと、ポリペプチドに結合した物

質により形成された複合体をアッセイするステップとを含む方法。

【請求項 19】

ポリペプチドの断片を接触させる請求項 18 の方法。

【請求項 20】

請求項 1 のポリペプチドのいずれかの活性を調節する方法であって、物質がポリペプチドの活性を調節できる条件下で、前記物質と請求項 1 のポリペプチドのいずれかとを接触させるステップを含む方法。

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 99/24368
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C07K14/705 C07K16/28 G01N33/53 C12Q1/68 A61K31/70 A61K38/17 A01K67/027		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K G01N C12Q A61K A01K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data bases consulted during the international search (name of data base and, where practical, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL - EMBEST28 [Online] Entry HSA33905, Acc.no. AA33905, 26 August 1996 (1996-08-26) HILLIER, L ET AL.: "zi05b04.r1 Soares fetal liver spleen 1MFLS S1 Homo sapeins cDNA clone 429871 5'." XP002129413 the whole document --- -/--	3-7
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubt on priority claims or which is cited to establish the publication date of another claim or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
25 April 2000		10.05.00
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patenthaus 2 NL - 2280 HV Rijswijk Tel: (+31-70) 340-5240, Tx: 31 851 epo nl Fax: (+31-70) 340-3018		Authorized officer Smalt, R

Form PCT/IS/H/220 (European sheet) 1, as of 1999

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 99/24368

C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL - EMEST5 [Online] Entry/Acc.no. AA993247, 5 June 1998 (1998-06-05) STRAUSBERG, R.: "ot60c09.s1 Soares testis NHT Homo sapiens cDNA clone IMAGE:1621168 3' similar to TR:Q60485 Q60485 ACROSOMAL PROTEIN SP32 PRECURSOR. ;" XP002136321 the whole document	3,4,6,7
X	DATABASE EMBL - EMEST3 [Online] Entry/Acc.no. AA769338, 31 January 1998 (1998-01-31) STRAUSBERG, R. ET AL.: "nz30e08.s1 NCI CGAP GCB1 Homo sapiens cDNA clone IMAGE:1289318 3'," XP002136322 the whole document	4,6,7
A	O'DOWD B F ET AL: "Cloning and chromosomal mapping of four putative novel human G-protein-coupled receptor genes" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES, 88, ELSEVIER SCIENCE PUBLISHERS, BARKING, vol. 187, no. 1, 10 March 1997 (1997-03-10), pages 75-81, XP004093242 ISSN: 0378-1119 the whole document	
A	WO 98 34948 A (CORNELL RES FOUNDATION INC) 13 August 1998 (1998-08-13) the whole document	
P,X	WO 99 42582 A (ELSHOURBAGY NABIL A ;SMITHKLINE BEECHAM CORP (US); TSUI PING (US);) 26 August 1999 (1999-08-26) the whole document	1-32, 37-43
P,X	DATABASE EMBL - EMHTG1 [Online] Entry/Acc.no. AC006087, 7 December 1998 (1998-12-07) MUNZY, D. ET AL.: "Homo sapiens chromosome 12p13.3 clone RPC111-433J6, *** SEQUENCING IN PROGRESS ***, 100 unordered pieces." XP002136323 From nt.131859-134408. *****	4-7

INTERNATIONAL SEARCH REPORT		national application No. PCT/US 99/24368
Box I: Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)		
This international Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:		
1	☒	Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: Although claims 37-39, 42, and 24, 30, 41 in as far as they relate to in vivo use, are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2	☒	Claims Nos.: 36 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210
3	☐	Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box II: Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)		
The international Searching Authority found multiple inventions in this international application, as follows:		
see additional sheet		
1	☐	As all required additional search fees were timely paid by the applicant, this international Search Report covers all searchable claims.
2	☐	As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3	☒	As only some of the required additional search fees were timely paid by the applicant, this international Search Report covers only those claims for which fees were paid, specifically claims Nos.: 1-43, all partially (inventions 1 and 3)
4	☐	No required additional search fees were timely paid by the applicant. Consequently, this international Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest. ☒ No protest accompanied the payment of additional search fees.		

International Application No. PCT/US 99/24368

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 36

Claim 36 refers to a compound identifiable with a method, without giving a true technical characterization of the compound. Moreover, no such compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT). No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the results to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

International Application No. PCT/US 99/24368

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-43, all partially

A G-protein coupled receptor protein, designate REC2838, represented by seq.ID.2, nucleic acids encoding it as represented by seq.ID.1, fragments thereof, antisense thereto, antibody directed against the protein, vector comprising said nucleic acid, host transformed with the vector, and method for producing the protein using the host. Methods for detecting the protein or the nucleic acid in a sample, method for identifying compounds that interact with said protein or said nucleic acid, method for identifying compounds that modulate the level or activity of the protein or the nucleic acid, pharmaceutical compositions of the protein, the nucleic acid, or the modulating compound, transgenic animal expressing said protein and method for producing such an animal.

2. Claims: 1-43, all partially

A G-protein coupled receptor protein, designate REC14618, represented by seq.ID.3, nucleic acids encoding it as represented by seq.ID.4, fragments thereof, antisense thereto, antibody directed against the protein, vector comprising said nucleic acid, host transformed with the vector, and method for producing the protein using the host. Methods for detecting the protein or the nucleic acid in a sample, method for identifying compounds that interact with said protein or said nucleic acid, method for identifying compounds that modulate the level or activity of the protein or the nucleic acid, pharmaceutical compositions of the protein, the nucleic acid, or the modulating compound, transgenic animal expressing said protein and method for producing such an animal.

3. Claims: 1-43, all partially

A G-protein coupled receptor protein, designate REC15334, represented by seq.ID.5, nucleic acids encoding it as represented by seq.ID.6, fragments thereof, antisense thereto, antibody directed against the protein, vector comprising said nucleic acid, host transformed with the vector, and method for producing the protein using the host. Methods for detecting the protein or the nucleic acid in a sample, method for identifying compounds that interact with said protein or said nucleic acid, method for identifying compounds that modulate the level or activity of the protein or the nucleic acid, pharmaceutical compositions of the protein, the nucleic acid, or the modulating compound, transgenic animal expressing said protein and method for producing such an animal.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/us 99/24368

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9834948	A	11-08-1998	AU 6149898 A	26-08-1998
WO 9942582	A	26-08-1999	NONE	

Form PCT/ISA/E1 (patent family member) (July 1999)

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード (参考)
A 6 1 P 7/06	C 0 7 K 14/705	4 C 0 8 4
C 0 7 K 14/705	C 0 7 K 16/28	4 C 0 8 6
C 0 7 K 16/28	C 1 2 N 1/15	4 H 0 4 5
C 1 2 N 1/15	C 1 2 N 1/19	
C 1 2 N 1/19	C 1 2 N 1/21	
C 1 2 N 1/21	C 1 2 P 21/02	C
C 1 2 N 5/10	C 1 2 Q 1/02	
C 1 2 P 21/02	C 1 2 Q 1/68	A
C 1 2 Q 1/02	G 0 1 N 33/15	Z
C 1 2 Q 1/68	G 0 1 N 33/50	Z
G 0 1 N 33/15	G 0 1 N 33/53	D
G 0 1 N 33/50	G 0 1 N 33/53	M
G 0 1 N 33/53	G 0 1 N 33/566	
G 0 1 N 33/566	C 1 2 N 5/00	A
	A 6 1 K 37/02	

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

(特許庁注：以下のものは登録商標)

フロッピー

(72)発明者 グラックスマン, マリア・アレクサンドラ

アメリカ合衆国マサチューセッツ州02173, レキシントン, サミット・ロード 33

(72)発明者 ウェイチ, ナディン・エス

アメリカ合衆国マサチューセッツ州02446, ブルックリン, パーク・ストリート 70, #53

Fターム(参考) 2G045 DA12 DA13 DA36

4B024 AA01 AA11 BA63 CA02 DA02 EA04 GA11 HA12

4B063 QA01 QA05 QA18 QQ08 QQ13 QQ53 QQ79 QQ91 QR32 QR56

QR77 QR80 QS25 QS34

4B064 AG20 CA01 CC24 DA01

4B065 AA26X AA46X AA50X AA80X AA90X AA90Y AB01 AC14 BA02 CA24

CA44 CA46

4C084 AA02 AA07 BA01 BA22 NA14 ZA51 ZA55 ZB21

4C086 AA01 AA03 EA16 MA01 MA04 NA14 ZA51 ZA55 ZB21

4H045 AA10 AA11 AA20 AA30 BA10 BA15 BA18 CA40 DA50 DA75

EA24 EA50