

(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2004-528812

(P2004-528812A)

(43) 公表日 平成16年9月24日(2004.9.24)

(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 6 1 K 45/00	A 6 1 K 45/00	2 G O 4 5
A 6 1 P 3/14	A 6 1 P 3/14	4 B O 2 4
A 6 1 P 7/02	A 6 1 P 7/02	4 B O 5 0
A 6 1 P 9/10	A 6 1 P 9/10	4 B O 6 3
		4 B O 6 5
	審査請求 未請求 予備審査請求 有	(全 157 頁) 最終頁に続く

(21) 出願番号	特願2002-534503 (P2002-534503)	(71) 出願人	500297535
(86) (22) 出願日	平成13年10月12日 (2001.10.12)		フェリング ベスローテン フェンノート
(85) 翻訳文提出日	平成15年4月14日 (2003.4.14)		シャップ
(86) 国際出願番号	PCT/US2001/031874		オランダ エヌエル-2132 ジェイエ
(87) 国際公開番号	W02002/031134		ックス ホーフドルブ ポラリス アベニ
(87) 国際公開日	平成14年4月18日 (2002.4.18)		ュー 144
(31) 優先権主張番号	60/240,117	(74) 代理人	100077481
(32) 優先日	平成12年10月12日 (2000.10.12)		弁理士 谷 義一
(33) 優先権主張国	米国 (US)	(74) 代理人	100088915
			弁理士 阿部 和夫
		(72) 発明者	スティーブ キー
			アメリカ合衆国 92130 カリフォル
			ニア州 サン ディエゴ ウェスト オー
			シャン エア ドライブ 11034 ア
			パートメント 24ビー
			最終頁に続く

(54) 【発明の名称】 DPP IVに関連する新規セリンプロテアーゼ遺伝子

(57) 【要約】

D P P I V に対し有意な配列相同性を有する新規タンパク質またはポリペプチド、それらをコードする核酸、それらのタンパク質を発現するようにそのような核酸で改変された細胞、それらのタンパク質に対する抗体、それらのタンパク質または関連タンパク質の活性の阻害剤である新たな治療薬を発見するためのスクリーニング方法、およびそのようなスクリーニング方法によって発見された治療薬、ならびに新たな治療的措置をすべて提供する。

【特許請求の範囲】

【請求項 1】

(a) 配列番号 1、3 および 5 のうち 1 つのアミノ酸配列を含むポリペプチド、または (b) それらと少なくとも約 70 % 類似するアミノ酸配列を有するポリペプチドをコードし、しかも同じ生物学的機能を示すか、

または、配列番号 2、4 および 6 のうち 1 つの選択的スプライスバリエントであり、または

(a) もしくは (b) をコードする前記核酸からの少なくとも 14 個の隣接ヌクレオチドを含むプローブであるか、または

前記のいずれか 1 つに相補的である

ことを特徴とする単離核酸。

10

【請求項 2】

DNA または RNA であることを特徴とする請求項 1 に記載の単離核酸。

【請求項 3】

配列番号 2、4 および 6 のいずれか 1 つの全長を含む DNA 転写物であるか、または配列番号 2、4 および 6 のうち 1 つの全コード領域に相補的であることを特徴とする請求項 1 に記載の単離核酸。

【請求項 4】

請求項 3 に記載の DNA を対象とすることを特徴とするアンチセンスオリゴヌクレオチド。

20

【請求項 5】

配列番号 2、4 および 6 のいずれか 1 つの全長を含む RNA 転写物であることを特徴とする請求項 1 に記載の単離核酸。

【請求項 6】

配列番号 2、4 および 6 のうち 1 つの選択的スプライスバリエントであることを特徴とする請求項 1 に記載の単離核酸。

【請求項 7】

請求項 6 に記載の核酸によってコードされることを特徴とするポリペプチド。

【請求項 8】

配列番号 1、3 および 5 のうち 1 つと少なくとも約 90 % 類似しているアミノ酸を有するポリペプチドをコードすることを特徴とする請求項 1 に記載の単離核酸。

30

【請求項 9】

配列番号 1、3 および 5 のうち 1 つと少なくとも約 95 % 類似しているアミノ酸を有するポリペプチドをコードすることを特徴とする請求項 1 に記載の単離核酸。

【請求項 10】

配列番号 1、3 および 5 のうち 1 つと少なくとも約 90 % の同一性を有するポリペプチドをコードすることを特徴とする請求項 1 に記載の単離核酸。

【請求項 11】

配列番号 2、4 および 6 のうち 1 つからの少なくとも 14 個の隣接ヌクレオチドを含むことを特徴とする請求項 1 による核酸プローブ。

40

【請求項 12】

請求項 1 に記載の核酸に加えてそれらの発現を操作するために前記核酸と実施可能に連結された発現制御要素を含むことを特徴とする単離組換えポリヌクレオチド分子。

【請求項 13】

動作可能にプロモーターと連結された配列番号 1、3 および 5 のいずれか 1 つで示される全アミノ酸配列を有するポリペプチドをコードする請求項 1 に記載の核酸を含み、前記発現ベクターは適合する宿主細胞中に存在することを特徴とする発現ベクター。

【請求項 14】

配列番号 1、3 または 5 のアミノ酸配列の少なくとも成熟タンパク質部分をコードする請求項 1 に記載の核酸の挿入により遺伝子組み換えされたことを特徴とする哺乳類、昆虫ま

50

たは細菌宿主細胞。

【請求項 15】

配列番号 1、3 および 5 のうち 1 つの成熟タンパク質部分を含むポリペプチドの産生方法であって、その方法は、前記ポリペプチドの産生に十分な条件下で請求項 11 に記載の宿主細胞を培養することを含むことを特徴とする方法。

【請求項 16】

前記ポリペプチドは前記細胞の表面において発現され、培養液からポリペプチドまたはその断片を回収するステップをさらに含むことを特徴とする請求項 15 に記載の方法。

【請求項 17】

任意選択的に、グリコシル化されていてもよく、かつ

10

(a) 配列番号 1、3 および 5 のいずれか 1 つで示される成熟タンパク質のアミノ酸配列を有し；(b) (a) の成熟タンパク質のうち 1 つと少なくとも約 70% の類似性を有する成熟タンパク質のアミノ酸配列を有し、かつ同じ生物学的機能を示し；(c) 配列番号 1、3 および 5 のいずれかの成熟タンパク質と少なくとも約 90% の同一性を有する成熟タンパク質のアミノ酸配列を有し；または (d) (a) の免疫学的に反応性の断片であることを特徴とするポリペプチド。

【請求項 18】

(a) の成熟タンパク質と少なくとも約 95% の類似性を有する成熟タンパク質であることを特徴とする請求項 14 に記載のポリペプチド。

【請求項 19】

20

(a) の成熟タンパク質と少なくとも約 95% の類似性を有する成熟タンパク質であることを特徴とする請求項 14 に記載のポリペプチド。

【請求項 20】

配列番号 1、3 および 5 のうち 1 つの成熟タンパク質のアミノ酸配列を有し、またはそれぞれの成熟タンパク質と同じ生物学的機能を示すそれらの断片であることを特徴とする請求項 14 に記載のポリペプチド。

【請求項 21】

請求項 17、18 および 19 に記載の前記成熟タンパク質のうち 1 つの生物学的機能を阻害することを特徴とする DPPR アнтаゴニスト。

【請求項 22】

30

請求項 17 に記載のポリペプチドまたは断片を認識することを特徴とする抗体。

【請求項 23】

配列番号 1 または 3 または 5 のアミノ酸配列を有するポリペプチドを認識することを特徴とする請求項 22 に記載の抗体。

【請求項 24】

請求項 17 に記載の少なくとも 1 つの成熟タンパク質の酵素活性を阻害することができる化合物をスクリーニングする方法であって、1 個または複数の試験化合物またはその塩の存在下で前記成熟タンパク質および前記成熟タンパク質に適した基質をインキュベートすること、前記成熟タンパク質の酵素活性を測定すること、前記活性を試験化合物の不存在下で測定された同様の活性と比較すること、および酵素活性を低下させる 1 種または複数の試験化合物を選択することを含むことを特徴とする方法。

40

【請求項 25】

請求項 20 に記載の成熟タンパク質の少なくとも 1 つの酵素活性を阻害せず、DPPIV の酵素活性を阻害することができる化合物をスクリーニングする方法であって、1 個または複数の DPPIV の阻害剤またはその塩の存在下で前記成熟タンパク質および前記成熟タンパク質に適した基質をインキュベートすること、前記成熟タンパク質の酵素活性を測定すること、前記活性を DPPIV 阻害剤の不存在下で測定された同様の活性と比較すること、および前記成熟タンパク質の酵素活性を低下させない化合物を選択することを含むことを特徴とする方法。

【発明の詳細な説明】

50

【0001】

(発明の分野)

本発明は、ジペプチジルペプチダーゼIV(DPPIV)に関連する新規セリンプロテアーゼ、およびこれらのプロテアーゼをコードする単離核酸に関し、これらの単離核酸はすべて、新たな治療薬の発見、プロテアーゼ活性の測定、およびこれらのプロテアーゼに対する化合物の阻害活性の測定に有用である。

【0002】

(発明の背景)

プロテアーゼおよびペプチダーゼは、ペプチドアミド結合の加水分解を触媒する酵素である。プロテアーゼは、細菌からウイルスから哺乳類までのほぼすべての生命体において生物学的プロセスの調節に重要な役割を果たす。プロテアーゼは、例えば、消化、血液凝固、アポトーシス、免疫応答の活性化、チモーゲン活性化、ウイルス成熟、タンパク質分泌およびタンパク質輸送において重要な機能を成す。作用部位、基質優先度、および機序などの多くの基準に従ってプロテアーゼを分類することができる。したがって、例えばアミノペプチダーゼは、ペプチドのN末端残基において優先的に作用するが、カルボキシペプチダーゼは、C末端において優先的に作用し、エンドペプチダーゼは、2つの末端から離れた部位において作用する。カルボキシペプチダーゼおよびアミノペプチダーゼの中で、ペプチジルペプチダーゼは、基質から単一アミノ酸残基を切断し、ジペプチジルペプチダーゼは、基質からジペプチド単位(2個のアミノ酸)を切断し、トリペプチダーゼは、基質から3個のアミノ酸を切断する。基質優先度は、切断部位に直接接するN末端のアミノ酸残基に関して発現されることが多い。例えば、トリプシン様ペプチダーゼは、塩基性アミノ酸(アルギニンまたはリジン)の次にあるペプチドを優先的に切断し、すなわち加水分解される結合は、Arg/Lys-Xaa結合である。別の例として、ペプチダーゼのキモトリプシン様ファミリーは、芳香族残基に隣接するペプチドを優先的に加水分解する。機構的に、ペプチダーゼは、セリン依存性、システイン依存性、アスパラギン酸依存性または亜鉛依存性であると分類される。

10

20

【0003】

ペプチダーゼおよびプロテアーゼは多くの生理学的プロセスの調節に関与しているため、治療薬の開発にとって魅力的な標的である。例えば、プロテアーゼおよびペプチダーゼ阻害剤は、高血圧症、凝固障害、およびウイルス感染の治療に用いられる。

30

【0004】

触媒活性にセリンを利用するタンパク質分解酵素は広範に分布しており、ウイルス、細菌および真核生物に見いだされる。セリンプロテアーゼの20を超えるファミリー(S1~S27で表される)が同定されており、それらは、構造類似性およびその他の機能的証拠に基づき、6種類のクラン(c1an)(SA、SB、SC、SE、SFおよびSG)に分類されている。4種類のクラン(SA、SB、SCおよびSE)に関しては構造が知られているが、それらは全体として無関係のようであり、セリンペプチダーゼについて少なくとも4種類、恐らくそれ以上の進化上の起源を示唆している、RawlingsおよびBarrett、Meth.Enzymol.244:19~61(1994)。

40

【0005】

プロリルオリゴペプチダーゼファミリーは、多くの進化的に関連するペプチダーゼからなり、それらの触媒活性は、セリンプロテアーゼのトリプシンファミリーと類似した電荷リレー系によってもたらされるようであるが、独立した収束進化により進化したものである。触媒機序には保存されたセリン残基が必要であることが実験的に(大腸菌(E.coli)プロテアーゼIIならびにブタおよび細菌PEにおいて)明らかとなっている。このセリンは、触媒三つ組残基(catalytic triad)(Ser、His、Asp)の一部であり、通常はこれらの酵素(約700から800個のアミノ酸を含むすべてのタンパク質である)のC末端から約150残基離れて位置している。

【0006】

最も集中して研究されているプロリルオリゴペプチダーゼの1つは、II型糖タンパク質

50

であるジペプチジルペプチダーゼ I V (D P P I V、E C 3 . 4 1 4 . 5) であり、これは、原形質膜の外側にあることが知られている特徴が十分に明らかにされた唯一のジペプチジルアミノペプチダーゼである。前述のように、ジペプチジルアミノペプチダーゼの特徴は、様々な小さなペプチドから N 末端ジペプチドを切断するそれらの能力である。ジペプチジルアミノペプチダーゼは、様々な基質特異性および細胞局在化を示し、このことはペプチドプロセッシングにおける各活性の様々な機能を示唆している。D P P I V の特徴は、最後から 2 番目の残基としてプロリンまたはアラニンを含む N 末端ジペプチドを切断する能力である。D P P I V 遺伝子は、約 7 0 k b の長さがあり、サイズが 4 5 b p から 1 . 4 k b までの 2 6 個のエキソンを含む。c D N A のヌクレオチド配列 (3 , 4 6 5 b p) は、7 6 6 個のアミノ酸を含むポリペプチドをコードするオープンリーディングフレームを含む。活性部位配列 (G - W - S - Y - G) をコードするヌクレオチドは、2 個のエキソン間に分割されている。このことが、プロリルオリゴペプチダーゼファミリーのゲノム組織化を古典的セリンプロテアーゼファミリーのゲノム組織化とを明確に区別している。

10

20

30

40

50

【 0 0 0 7 】

D P P I V は哺乳類の組織に広く分布し、腎臓、腸管上皮および胎盤に極めて多量に見いだされる (Y a r o n , A . および N a i d e r , F .、B i o c h e m . M o l . B i o l . 中の C r i t i c a l R e v i e w s、1 9 9 3 [1]、3 1)。ヒト免疫系では、この酵素は、C D 4 ⁺ 型の活性化 T リンパ球によってほぼ排他的に発現され、この酵素は細胞表面抗原 C D 2 6 と同義であることが明らかにされている。ヒト生理学における D P - I V の正確な役割はいまだ完全には理解されていないが、最近の研究から、ヒトの生理学および病態生理学において、この酵素は明らかに重要な役割を有していることが明らかとなった。

【 0 0 0 8 】

ヒト T 細胞上では、D P P I V 発現は胸腺分化の後期に現れ、C D 4 ⁺ ヘルパー / メモリー集団に優先的に限定され、C D 2 6 は、強力な共刺激 T 細胞活性化シグナルを送ることができる。したがって、T 細胞活性化抗原 C D 2 6 としても知られる D P P I V は、C D 4 5 チロシンホスファターゼとの会合を介する免疫応答で重要な役割を果たし、アデノシンデアミナーゼ (A D A) を T 細胞表面に結合させる能力により、T 細胞をアデノシン媒介性増殖阻害から守る。さらに、C D 2 6 / D P P I V によるケモキネシスの機能の調節は、リンパ球輸送および H I V 株の感染力にとって不可欠のようである。D P P I V は、T 細胞活性化、細胞接着、腎臓および腸管におけるプロリン含有ペプチドの消化、H I V 感染およびアポトーシス、ならびにある種の黒色腫細胞における腫瘍形成能の調節への関与を含む多くの機能に関与している、P e t h i y a g o d a 他、C l i n . E x p . M e t a s t a s i s 2 0 0 0 ; 1 8 (5) : 3 9 1 ~ 4 0 0。また、D P P I V は、内分泌調節および代謝生理にも関与している。より具体的には、D P P I V は、G L P - 1 のアミノ末端 H i s - A l a ジペプチドを切断し、G L P - 1 受容体アンタゴニストを生成し、それによって G L P - 1 に対する生理学的応答を短縮する。グルカゴン様ペプチド - 1 (G L P - 1) は、グルコース依存性インスリン分泌を引き起こすインクレチンであり、D P P I V によって速やかに分解され、D P P I V 切断の半減期が循環から G L P - 1 が除去される半減期よりかなり短いために、D P P - I V 阻害による G L P - 1 生物活性の著しい上昇 (5 から 1 0 倍) が予想される。D P P I V の阻害剤は、現在、2 型糖尿病および耐糖能異常のための強力な治療薬として臨床的に研究されている。

【 0 0 0 9 】

1 9 9 3 年には D P P I V の様々な異なる阻害剤が知られていた。それらのうちの 1 つが自殺阻害剤の N - A l a - P r o - O - (ニトロベンゾイル -) ヒドロキシルアミンである。もう 1 つは競合的阻害剤の e - (4 - ニトロ) ベンゾキシカルボニル - L y s - P r o であり、もう 1 つはポリクローナルウサギ抗ブタ腎臓 D P P I V 免疫グロブリンである。それ以後も他の阻害剤が開発され、米国特許第 5 , 9 3 9 , 5 6 0 号、第 6 , 1 1 0 , 9 4 9 m 号、第 6 , 0 1 1 , 1 5 5 号および第 5 , 4 6 2 , 9 2 8 号に詳細に記載されて

いる。

【0010】

そのセリン型触媒活性に加え、しかしそれに依存せず、DPPIVは、可溶性細胞外酵素アデノシンデアミナーゼ(ADA)としっかり結合し、受容体として作用し、シグナル伝達を媒介すると考えられている。DPPIV構造の特徴は、2個の細胞外ドメイン、ノフォールド加水分解酵素ドメインおよび約50個のアミノ酸の反復ベータシートからなる7枚羽根の(7-blade)ベータ-プロペラドメインである。最近、大きさによって基質を選択するほかに、12個の高度に保存されたシステイン残基のうち10個を含むベータ-プロペラドメインがペプチダーゼドメインの触媒反応に寄与していることが明らかにされた。さらに、システインが豊富なドメインは、コラーゲンIおよび細胞外ADAに対するDPPIVの結合を担っている。また、DPPIVは、細胞の細胞外マトリックスとのフィブロネクチン媒介性相互作用において役割を果たしていると報告されている。最近の研究は、DPPIVのプロテアーゼ活性がその抗侵襲活性に必要なのではないのは、細胞外セリンプロテアーゼ活性を欠くDPPIVの変異体がそのような活性を保持しているためであることを示している。

10

【0011】

DPPIVと類似性を共有する多くのタンパク質が文献に報告されている。DPP-I、DPP-II、DPP-III、DPP-Xおよび線維芽細胞活性化タンパク質(FAP)を含むこれらのタンパク質のいくつかがクローニングされている。それらは同定され、分子クローニングおよび発現タンパク質の機能試験により、または組織抽出物における生化学的活性として特徴付けられている。DPPIVとの機能的類似性を備えるDPPIV-ベータおよびその他の新規ペプチダーゼは未だにクローニングされていない。プロリルオリゴペプチダーゼのファミリーのさらなるメンバーの同定、特徴付けおよび/または適切な分類、それらの生理学的(および特に病態生理学的)役割の解明、ならびに新たな治療薬の開発への知識の応用は有意義な挑戦である。

20

【0012】

(発明の概要)

本発明は、DPPIVに関連するタンパク質のファミリーのうち3種類の新規メンバーを構成するプロリルオリゴペプチダーゼ(プロリン後切断)活性を持つタンパク質であって、完全長タンパク質、選択的スプライスフォーム、サブユニット、および変異体を含むタンパク質、ならびにそれらをコードするヌクレオチド配列を提供する。また、本発明は、上記タンパク質の基質、相互作用タンパク質、アゴニスト、アンタゴニストまたは阻害剤をスクリーニングする方法を提供し、さらにタンパク質および/または変異体、それらの誘導体および/または類縁体および/またはそれらに対するリガンドを含む薬剤組成物を提供する。

30

【0013】

DPPIVと有意な配列相同性を有するこれらの新規タンパク質をジペプチジルペプチダーゼIV関連タンパク質1、2および3(DPRP-1、DPRP-2およびDPRP-3)と呼ぶ。DPRP-1、DPRP-2およびDPRP-3のアミノ酸配列を、それぞれ、配列番号(SEQ.ID NOS):1、3および5に示す。これらのタンパク質をコードする核酸配列をさらに開示する(配列番号2、4および6)。表1は、新規タンパク質DPRP-1、DPRP-2およびDPRP-3およびその他の知られているセリンプロテアーゼ間の相同性(すなわち類似性)を表している。

40

【0014】

【表1】

表1—これら3種類の新規タンパク質とDPPIVおよびその他のクランSC、ファミリーS9メンバーならびにサブファミリーBメンバーとの配列の比較

プロテアーゼファミリー	プロテアーゼ名	アミノ酸数	DPPIVとの相同性	TM領域	Ser-Asp-His三つ組残基	遺伝子座決定	最適pH
クランCA、ファミリーC1	DPPI	463	N	N	N	11q14.1~q14.3	-
クランSC、ファミリーS28	DPPII	500	N	Y	N	-	4.5~6.0
	QPP	492	N	N	N	-	4.5~7.5
	PCP	496	N	N	N	-	-
割当未決定	DPPIII	737	N	N	N	-	-
クランSC、ファミリーS9、サブファミリーB	DPPIV	766	100	Y	Y	2q24.3	7.5~8.0
	DPPVI	865	52	Y	突然変異	7	-
	FAP	760	70	Y	Y	2q23	7.5~8.0
	DPRP-1	882	41	N	Y	15q22.1~15q22.2	7.5~8.0
	DPRP-2	864	39	N	Y	19p13.3	7.5~8.0
	DPRP-3	796	54	Y	突然変異	2q12.3~2q14.1	-

10

20

40

50

【0015】

DPRP-1、DPRP-2とDPPIV間の最大の相同性はC末端配列に見られる。DPPIVとの配列相同性に基づき(図1を参照)、これらのDPRPタンパク質は酵素としての役割を含むが、それらに限定されない機能を有することが予想されるであろう。クローニング、発現、生化学的および分子化学的特徴付けはこの仮説を裏付けた。

30

【0016】

DPRPの発現パターンならびに特殊化された上皮細胞および形質細胞(ライディッヒ細胞、前立腺上皮細胞、リンパ球、B細胞)への局在化は、分化、増殖および炎症における役割と一致している。ホルモン感受性癌(乳癌、前立腺癌、精巣癌)、テストステロンによって調節される組織におけるDPRP-1遺伝子の局在化、および分化が不十分な癌における大量の発現は、DPRPを活性化または阻害する分子が、無秩序な成長、分化およびステロイドまたはポリペプチドホルモン合成および分解を特徴とする障害の治療において多くの治療用途を有していることを示している。本明細書に開示されるデータは、DPRP-1およびDPRP-2が、当業者によく知られている前立腺癌および精巣癌の*in vitro*モデルの増殖の調節に関与している仮説を裏付けている。

40

【0017】

本明細書に記載のDPRP-1およびDPRP-2活性ならびにそれらの発現パターンは、ペプチドおよびケモキネスのような生化学的メディエーターの酵素修飾を介し、免疫系および神経内分泌系の生理学的レギュレーターとしての機能的役割を有することと一致する。阻害剤の使用法に基づきDPPIVについて前述した多くの機能は、その作用およびDPRPのような類似したタンパク質の作用に部分的に起因していると思われる。したがって、DPPIV、DPRPおよびFAPのようなその他の関連プロテアーゼの選択的かつ強力な阻害剤を発見することは、それらの阻害剤および新たに同定されるセリンプロテ

アーゼ阻害剤、ならびにそのようなタンパク質の機能を改変するその他の活性化化合物の有効かつ安全な薬剤使用法を実現する中核を成すと考えられる。

【0018】

したがって、本発明は、新規タンパク質またはポリペプチド、それらをコードする核酸、これらのタンパク質を発現するようにその核酸で改変された細胞、これらのタンパク質に対する抗体、これらのタンパク質の活性の阻害剤である（またはD P P I Vの阻害剤であるがタンパク質ではない）新たな治療薬を発見するためのスクリーニング方法、およびそのようなスクリーニング方法によって発見される治療薬を提供する。新規タンパク質およびそれらをコードする核酸を用い、例えば生殖障害、炎症性障害および代謝性障害などのある種の疾患を治療するための新たな治療薬を発見することができ、治療上または診断的価値を有する抗体の調製にも使用することができる。

10

【0019】

本発明の一態様によれば、主にヒト起源の新規な成熟した生物活性タンパク質を提供する。このようなタンパク質は、標準的技法により適当な動物（ヒトを含む）組織または体液から少量単離してもよい。しかしながら、そのタンパク質を発現するように遺伝子が組み換えられた細胞の培養液中でより大量に調製することがより好都合である。

【0020】

本発明のもう一つの態様によれば、本発明のポリペプチドをコードするmRNA、DNA、cDNA、それらのゲノムDNAを含む単離核酸分子を提供する。

【0021】

本発明の別の態様によれば、本発明の核酸配列と特異的にハイブリッド形成するのに十分な長さの核酸分子を含む核酸プローブを提供する。

20

【0022】

本発明のさらに別の態様によれば、*in vitro*における科学研究、例えばDNAの合成およびDNAベクターの製造に有用なこのようなポリペプチドを製造するために組み換え技法を利用するプロセスを提供する。このようなポリペプチドを製造するためのプロセスには、このようなポリペプチドおよび/または成熟タンパク質をコードする核酸配列を含むDNAベクターが形質移入された組み換え原核生物および/または真核生物の宿主細胞を、このようなタンパク質の発現を促進する条件下で培養することと、続く発現産物のこのようなタンパク質または断片の回収が含まれる。

30

【0023】

さらにもう一つの態様によれば、本発明は、D P R Pポリペプチドおよびポリヌクレオチドを使用する方法を提供し、その使用方法には、細菌、真菌、原虫およびウイルス感染などの感染、特にH I V - 1およびH I V - 2による感染、疼痛、糖尿病、性早熟症、不妊症、肥満症、食欲不振、過食症、パーキンソン病、急性心不全、低血圧症、高血圧症、尿閉、骨粗鬆症、狭心症、心筋梗塞、脳卒中、潰瘍、喘息、アレルギー、良性前立腺肥大、ホルモン感受性およびアンドロゲン非依存性癌を含む癌、片頭痛、嘔吐、不安を含む精神障害および神経障害、精神分裂病、躁鬱病、うつ病、痴呆、および重症精神遅滞、ならびに運動障害（以降はまとめて「疾患」と呼ぶ）の治療が含まれる。

【0024】

本発明のさらにもう一つの態様によれば、このようなポリペプチド、またはこのようなポリペプチドをコードするポリヌクレオチドを、それらの成熟タンパク質の生物活性を、例えばN末端ジペプチドを切断することにより阻害する化合物の発見に利用するプロセスを提供し、そのような阻害剤も提供する。

40

【0025】

より具体的な態様では、本発明は、(a)配列番号1、3および5のうち1つのアミノ酸配列が含まれるポリペプチド、(b)それらと少なくとも約70%類似し、同一の生物学的機能を示すアミノ酸配列を有するポリペプチドをコードする、または配列番号2、4および6のうち1つの選択的スプライズバリエーションである、または(a)もしくは(b)をコードする前記核酸からの少なくとも14個の隣接ヌクレオチドを含むプローブである、

50

または前記のいずれか1つに相補的である単離核酸を提供する。

【0026】

もう1つの具体的な態様では、本発明は、場合によりグリコシル化されていてもよく、(a)配列番号1、3および5のいずれか1つで示される成熟タンパク質のアミノ酸配列を有し；(b)(a)の成熟タンパク質のうち1つと少なくとも約70%の類似性を有する成熟タンパク質のアミノ酸配列を有し、同一の生物学的機能を示し；(c)配列番号1、3および5のいずれかの成熟タンパク質と少なくとも約90%の同一性を有する成熟タンパク質のアミノ酸配列を有し；または(d)(a)の免疫学的に反応性の断片であるポリペプチドを提供する。

【0027】

さらにもう1つの具体的な態様では、本発明は、本発明の少なくとも1つの成熟タンパク質の酵素活性を阻害することができる化合物をスクリーニングする方法であって、前記成熟タンパク質および前記成熟タンパク質に適した基質を1個または複数の試験化合物またはその塩の存在下でインキュベートすること、前記成熟タンパク質の酵素活性を測定すること、試験化合物の存在下で測定された同様の活性を前記活性と比較すること、および酵素活性を低下させる1種または複数の試験化合物を選択することを含む方法を提供し、また、本発明は、1個または複数のDPPIVの阻害剤またはその塩の存在下で少なくとも1個の成熟タンパク質および適当な基質の酵素活性を阻害しないDPPIVの酵素活性を阻害することができる化合物をスクリーニングすること、前記成熟タンパク質の酵素活性を測定すること、DPPIV阻害剤の存在下で測定された同様の活性を前記活性と比較すること、および前記成熟タンパク質の酵素活性を低下させない化合物を選択することに関する方法を提供する。

【0028】

本発明のこれらの態様またはその他の態様は以下の詳細な説明により当業者には明らかになるはずである。

【0029】

(好ましい実施形態の詳細な説明)

本発明の一態様によれば、単離核酸配列(ポリヌクレオチド)を提供し、これらは3種類のDPRP(配列番号1、3および5)の推定アミノ酸配列を有する成熟ポリペプチドをコードする。

【0030】

本発明のポリヌクレオチドは、ヒト精巢cDNAライブラリー(DPRP-1)、ヒト大腸ライブラリー(DPRP-2)およびヒト視床下部cDNAライブラリー(DPRP-3)を用いて発見された。DPRP-1の単離核酸は、長さが約882個のアミノ酸のタンパク質をコードするオープンリーディングフレームを含み、ヒトDPPIVと構造的に関連し、ヒトDPPIVタンパク質配列全体にわたり26%の同一性および41%の類似性を示す。DPRP-2の単離核酸は、約864個のアミノ酸のタンパク質をコードするオープンリーディングフレームを含み、全DPPIVアミノ酸配列と39%類似している。疎水性プロットを用いるDPRP-1およびDPRP-2主要アミノ酸配列の分析から、これら2種類のタンパク質は膜貫通ドメインを有していないことが予測される。この事実にもかかわらず、細胞活性化によってこれらの細胞内セリンプロテアーゼが分泌されることが可能である。休止(quiet)細胞プロリンジペプチダーゼ(QPP)は、リソソームとは異なる細胞内小胞を標的とするセリンプロテアーゼである(Chiravuri M、他、J. Immunol. 2000年11月15日；165(10)：5695~702)。この仮説は、ケモカイン、サイトカイン、ペプチドおよびポリペプチドの翻訳後調節に関する機序へのDPRP-1およびDPRP-2の関与が起こりうる部位および範囲を拡大する。完全長DPRP-3配列は、796個のアミノ酸、1から48番までのシグナルペプチド、および34から56番の間の膜貫通ドメインを含む。成熟タンパク質はII型膜タンパク質であると予想され、切断されて可溶体を生成することができる可能性がある。アミノ酸配列を配列番号5に示すが、これは配列番号6から推定さ

10

20

30

40

50

れ、DPPIVと54%の類似性を有する。

【0031】

プロリルオリゴペプチダーゼ酵素サブファミリーS9Bのメンバーとこれらのポリペプチドとのアミノ酸配列アラインメントは、3種類のDPRPタンパク質すべてが、DPPIVおよびFAPと全体的に配列および構造的相同性を有していることを示している。DPRPは、Ser、Asp、Hisという順序の触媒残基および活性部位配列(G-W-S-Y-G)を持つ酵素クランSC(セリン求核性)のメンバーであることが予想される。

【0032】

【表2】

表2. DPRP-1、DPRP-2、DPRP-3およびプロリルオリゴペプチダーゼファミリーS9B酵素のメンバー間の相同性(すなわち類似性)

DPPIV					
41	DPRP-1				
39	74	DPRP-2			
54	39	40	DPRP-3		
70	41	39	52	FAP	
52	40	42	68	54	DPPVI

10

20

【0033】

DPRP-1、DPRP-2、およびDPRP-3は、古典的セリンプロテアーゼファミリーのメンバー、キモトリプシンおよびサブチリシンのいずれとも配列類似性を示さない。触媒的三つ組残基の順序は、3種類の主な関連SCクランファミリーで異なり、キモトリプシンではHis-Asp-Serであり、サブチリシンではAsp-His-Serであり、プロリルオリゴペプチダーゼではSer-Asp-Hisである。

【0034】

表2に示すように、DPRP-3は、DPPVIと最も高い相同性を有する(68%の相同性および51%の同一性)。Wada他は、DPPVI、DPPIV関連タンパク質についてcDNAクローンをウシ、ラット(Wada他、Proc.Nat.Acad.Sci.89:197~201(1992))およびヒト(Yokotani他、Hum.Molec.Genet.2:1037~1039(1993))脳ライブラリーから単離した。Wada他は、DPPIVとは違って、DPPVIにおける触媒三つ組残基が最初のセリン残基を有しないことを明らかにした。DPRP-3では、セリンプロテアーゼファミリーに特徴的な触媒三つ組残基中のアミノ酸のうち2個が保存されている。しかしながら、セリン残基自体はグリシンで置き換えられている。セリン残基がないことは、この部位におけるプロテアーゼ活性を妨げそうであるが、タンパク質の他の機能的ドメインによって媒介される複数のその他の機能は無傷のままである可能性がある。

30

【0035】

前述のように、DPPIVは、そのペプチダーゼとしての触媒活性の他に、発現細胞および組織に応じて重要な機能を発揮する多機能分子である。また、DPRP-3およびDPPVIは、完全な触媒三つ組残基がないにもかかわらず複数の機能を維持している可能性が高い。例えば、DPPVIは、ニューロンの可塑性の調節と関連づけられてきた。DPPVIは、海馬、視床、視床下部および被殻(stiatum)で高度に発現される。さらに、rump white Rw/Rw胚の発育停止および胚死亡率は、DPPIV遺伝子の破壊に起因すると考えられている。Rw変異は、マウス第5染色体の近位部の30cMに及ぶ染色体反転と関係している。Rw染色体上のDPPVI遺伝子のゲノム解析は、C末端領域の有意な分画の喪失をもたらすコード領域における反転ブレイクポイントを位置づけている、Hough R.B.他、Proc.Nat.Acad.Sci.、9

40

50

5、13800～13805(1998)。

【0036】

ヒトDPRP-1遺伝子は、長さが32668bpであると予想されており、少なくとも22個のエキソンおよび8個の転写物を有する。この遺伝子は、第15染色体(NT_010265)の15q21.1～15q22.1位に位置づけている。予想される選択的スプライスパリアント転写物の長さは、602bpから4523bp間と異なる(配列番号7～22を参照)。これは、ノーザンブロット分析により観察される複数の転写物と一致している(実施例2を参照)。転写物を表すESTは、老化線維芽細胞、Tリンパ球、胚中心B細胞、生殖細胞精上皮腫、精巣、メラニン細胞、子宮、卵巣乳房、多発性硬化症病変、膵臓および胎盤を含む多数の組織中に見いだされた。

10

【0037】

ヒトDPRP-2は、少なくとも27個のエキソンおよび9個のスプライスパリアントを持つ遺伝子に属する(配列番号23～40を参照)。1種類のSNPが3'UTRで観察された。(88%(37)C対12%(5)T)。DPRP-2遺伝子は、第19染色体の領域19p13.3に位置づけている。この位置は、多くの疾患マーカーに対するホストであり、低カルシウム尿性高カルシウム血症、II型小脳性運動失調症、筋ジストロフィー、痙攣、アテローム性動脈硬化症感受性、乾癬、外胚葉性形成異常、および急性骨髄性白血病を含む様々な障害と関係している。ノーザンブロット分析により観察されるmRNAの遍在性分布に一致して(実施例2を参照)、DPRP-2は、ESTの範囲の調査の際、多種多様な組織中で発現された(例えば、肝臓、膵臓、筋肉、メラニン細胞、心臓、肺、胎盤、皮膚、膵臓、胃、脳副甲状腺で64個を超えるESTが発現された)。

20

【0038】

ヒトDPRP-3は、少なくとも23個のエキソンおよび2個のスプライスパリアントを持つ遺伝子に属する(配列番号41～44を参照)。この遺伝子は、第2染色体(NT_005445)2q12.3～2q14.1位に位置づけている。DPRP-3の転写物は、DPRP-1およびDPRP-2のような広い分布は示さない。実施例2におけるノーザンブロットによって示されるように、DPRP-3発現は脳および膵臓に限定されている。DPRP-3mRNAを表すESTは、多発性硬化症病変、視床下部、全脳および神経に由来する組織中に多くあり、少数の転写物が子宮および大腸に見いだされる。

【0039】

DPRP-1、DPRP-2およびDPRP-3を含むクランSCにおけるヒトプロテアーゼと齧歯類プロテアーゼとの関係を近隣結合(Neighbor Joining)法(NJ)を用いて分析した、SaitouおよびNei、Mol. Biol. Evol.、4、406～525(1987)を参照。系統発生解析は、S9プロテアーゼの中で、DPRP-1とDPRP-2は共に膜貫通ドメインを欠き、DPPIVおよびFAPのような密接に関係のあるタンパク質とは区別されることを示している。しかしながら、DPPIVとFAPの間およびDPRP-3とDPPVIの間には類似性が見られ、これらはすべてII型膜タンパク質である。

30

【0040】

さらにDPRP関連遺伝子についてデータベース検索することにより、DPRP-1に関連するマウス配列の存在が明らかとなった。このマウス配列と新規なヒトプロテアーゼのアラインメントは、mDPRP-1がそのヒト対応物とかなりの相同性を示すことを示している(図2)。当業者が容易に認めるように、新規マウスプロテアーゼ遺伝子は、本明細書に開示されている配列情報を用いて単離でき、当技術分野でよく知られ、ルーチン的に用いられる発現構築物の1つへ容易に組み入れることができる。トランスジェニックマウスモデルを作製するための当業者によるこの開示配列の使用法は、例えば、マウス胚性幹細胞における相同的組換えをもたらす遺伝子標的ベクターの開発を用いている。DPRP遺伝子の機能をさらに分析する際にノックアウトマウスを用いることは、価値ある手段である。

40

【0041】

50

本発明のポリヌクレオチドは、RNAの形態またはDNAの形態であってもよく、DNAには、cDNA、ゲノムDNA、および合成DNAが含まれると理解すべきである。DNAは、二本鎖または一本鎖であってもよく、一本鎖であれば、コード鎖または非コード（アンチセンス）鎖であってもよい。成熟ポリペプチドをコードするコード配列は、それぞれ配列番号2、4および6で示されるコード配列と同一であるか、遺伝子コードの重複性もしくは同義性または一塩基多型の結果として、同じ成熟ポリペプチドをコードする異なるコード配列であってもよい。例えば、配列番号2、4および6のいずれか1つの全長を含むRNA転写物であってもよい。

【0042】

配列番号1、3、5の成熟タンパク質をそれぞれコードするポリヌクレオチドには、成熟タンパク質単独のコード配列；成熟ポリペプチドのコード配列に加えてリーダーもしくは分泌配列またはプロタンパク質配列などの追加のコード配列；および成熟タンパク質のコード配列（および任意選択的に追加のコード配列）に加えてイントロンもしくは成熟タンパク質のコード配列の非コード配列5'および/または3'などの非コード配列が含まれるが、それらに限定されるものではない。

10

【0043】

したがって、用語「ポリペプチドをコードするポリヌクレオチド」または用語「ポリペプチドをコードする核酸」は、成熟タンパク質のコード配列のみが含まれるポリヌクレオチドまたは核酸ならびに追加のコード配列および/または非コード配列が含まれるポリヌクレオチドまたは核酸を包含すると理解されるべきである。用語ポリヌクレオチドおよび核酸は互換的に使用される。

20

【0044】

また、本発明には、成熟タンパク質のコード配列が、同一のリーディングフレーム内で宿主細胞からのポリペプチドの発現および分泌を助けるポリヌクレオチド配列と融合していてもよいポリヌクレオチドが含まれる。例えば、細胞からのポリペプチドの輸送を制御するために分泌配列として機能するリーダー配列が融合していてもよい。このようなリーダー配列を有するポリペプチドはプレタンパク質またはプロタンパク質と呼ばれ、宿主細胞によって切断され成熟型のタンパク質を生成するリーダー配列を有することができる。これらのポリヌクレオチドは、プロタンパク質をコードするように5'拡張領域を有し、このプロタンパク質は、成熟タンパク質に加えてN末端における追加のアミノ酸残基である。このようなプロ配列を有する発現産物はプロタンパク質と呼ばれ、これは不活性型の成熟タンパク質であるが、ひとたびプロ配列が切断されると活性な成熟タンパク質が残る。したがって、例えば、本発明のポリヌクレオチドは、成熟タンパク質、またはプロ配列を有するタンパク質、またはプロ配列とプレ配列（リーダー配列）を共に有するタンパク質をコードすることができる。

30

【0045】

また、本発明のポリヌクレオチドは、本発明のポリペプチドの精製を可能にするマーカ配列とフレーム内で融合したコード配列を有することができる。マーカ配列は、ポリヒスチジンタグ、赤血球凝集素（HA）タグ、c-mycタグまたは哺乳類宿主、例えばCOS-1細胞を使用する場合のV5タグであってもよい。HAタグは、インフルエンザ赤血球凝集素タンパク質に由来するエピトープに相当すると思われ（Wilson, I. 他、Cell, 37: 767 (1984)）、およびc-mycタグは、ヒトMycタンパク質に由来するエピトープであってもよい（Evans, G. I. 他、Mol. Cell Biol., 5: 3610~3616 (1985)）。

40

【0046】

用語「遺伝子」は、ポリペプチド鎖の産生に關与するDNAのセグメントを意味し、遺伝子にはコード領域の前および後にくる領域（リーダーおよびトレイラー）ならびに個々のコードセグメント（エキソン）間の介在配列（イントロン）が含まれる。用語「有意な配列相同性」は、アミノ酸残基の少なくとも25%、好ましくは少なくとも40%が保存され、非保存残基のうち、少なくとも40%が保存的置換であることを示すことを意図して

50

いる。

【0047】

本発明の完全長遺伝子の断片は、完全長cDNAを単離するため、ならびに遺伝子に対して有意な配列相同性を有し、類似の生物活性または機能を有するタンパク質またはポリペプチドをコードする他のcDNAを単離するために、cDNAライブラリーのためのハイブリダイゼーションプローブとして使用することができる。この用途のためには、類似の生物活性または機能とは、最後から2番目の残基としてAlaもしくはProまたは他のアミノ酸を有するN末端ジペプチドを切断する能力を意味する。このタイプのこのようなプローブは、少なくとも14個の塩基（配列番号2、4または6のうち1つに由来する少なくとも14個の隣接ヌクレオチド）、好ましくは少なくとも30個の塩基を有し、この

10

【0048】

本発明は、配列間に少なくとも70%、好ましくは少なくとも90%、より好ましくは少なくとも95%の同一性または類似性があり、類似の生物活性を有するタンパク質をコードする、前述の配列とハイブリッド形成するポリヌクレオチドをさらに提供すると考えられる。さらに、当技術分野で知られているように、アミノ酸配列が配列中の各々の個別残基に対して同一または保存アミノ酸置換を含む場合には、2つのポリペプチド間には「類似性」が存在する。同一性および類似性は、配列解析ソフトウェア（例えば、Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705）を用いて測定することができる。

20

30

特に、本発明は、厳密な条件の下で前述のポリヌクレオチドとハイブリッド形成するようなポリヌクレオチドを提供する。本明細書で使用する用語「厳密な条件」は、ポリヌクレオチド配列と配列番号2、4および6のポリヌクレオチド配列との間のハイブリダイゼーションを可能にし、少なくとも約70%の同一性のある条件を意味する。好適な厳密な条件は、例えばプレハイブリダイゼーションおよびハイブリダイゼーション溶液中の塩もしくはホルムアミドの濃度により、またはハイブリダイゼーション温度によって規定することができる。当技術分野でよく知られている。特に、厳密性は、塩の濃度を下げることにより、ホルムアミドの濃度を上げることにより、および/またはハイブリダイゼーション温度を上げることにより増すことができる。

【0049】

例えば、高厳密性条件下のハイブリダイゼーションは、約37 から42 において約50%のホルムアミドを用いることができ、低厳密性条件下のハイブリダイゼーションは、約30 から35 において約35%から25%のホルムアミドを用いる。高厳密性条件下でハイブリッド形成するための特定の条件設定の1つは、42 、50%ホルムアミド、5x.SSPE、0.3%SDS、および200μg/ml剪断および変性サケ精子DNAを用いる。低厳密性下のハイブリダイゼーションについては、前述と類似の条件を35 の低温において35%ホルムアミド中で用いることができる。特定レベルの厳密性に対応する温度範囲は、当該核酸のプリンとピリミジンの比を計算すること、およびそれに従って温度を調整することによってさらに狭めることができる。上記範囲および条件に対する変更は当技術分野でよく知られている。ハイブリダイゼーションは、配列間に少なく

40

50

とも95%、より好ましくは少なくとも97%の同一性のある場合にのみ行うことが好ましい。好ましい実施形態において前述のポリヌクレオチドとハイブリッド形成するポリヌクレオチドは、配列番号2、4および6のcDNAのうち1つによってコードされる成熟タンパク質と実質的に同一の生物学的機能または活性を示すポリペプチドをコードする。

【0050】

前述のように、好適なポリヌクレオチドプローブは、少なくとも14個の塩基、好ましくは30個の塩基、より好ましくは少なくとも50個の塩基を有し、前述のようにそれらとの同一性を有し、活性を維持していても維持していなくともよい本発明のポリヌクレオチドとハイブリッド形成するはずである。例えば、このようなポリヌクレオチドは、それぞれ配列番号2、4および6のポリヌクレオチドとハイブリッドを形成させ、例えば、この
10
ようなポリヌクレオチドを回収するためのプローブとして、または診断用プローブ、またはPCRプライマーとして用いることができる。したがって、本発明には、それぞれ配列番号1、3および5のポリペプチドをコードするポリヌクレオチド、ならびにそのようなポリヌクレオチドによってコードされるポリペプチドに対し、好ましくは少なくとも30個の塩基、より好ましくは少なくとも50個の塩基を有するそれらの断片をコードするポリヌクレオチドに対し、少なくとも70%の同一性、好ましくは少なくとも90%の同一性、より好ましくは少なくとも95%の同一性を有するポリヌクレオチドが含まれる。

【0051】

当技術分野でよく知られているように、遺伝子コードは、2種類以上のヌクレオチドトリプレット(コドン)によって特定のアミノ酸がコードされているという点で重複性であり
20
、本発明には、本明細書の配列中で具体的に例示されているコドンとは異なるコドンを用いて同一のアミノ酸をコードするこれらのポリヌクレオチド配列が含まれる。このようなポリヌクレオチド配列は、本明細書で「等価」ポリヌクレオチド配列と呼ぶ。さらに、本発明には、成熟タンパク質の一部またはすべて、それぞれ配列番号1、3および5の予想アミノ酸配列を有するポリペプチドのうち1つの類縁体および誘導体などの断片をコードする前述のポリヌクレオチドの変異体が含まれる。変異体形のポリヌクレオチドは、天然に存在するポリヌクレオチドの対立遺伝子変異体、または天然に存在しないポリヌクレオチドの変異体であってもよい。例えば、核酸中の変異体は、単に遺伝子コードの縮重によるアミノ酸のコドン配列の差であってもよく、または欠失変異体、置換変異体および付加または挿入変異体であってもよい。当技術分野で知られているように、対立遺伝子変異体
30
は、コードされたポリペプチドの生物学的機能を実質的には変化させない1個または複数のヌクレオチドの置換、欠失または付加を有することができる代替形態のポリヌクレオチド配列である。

【0052】

さらに、本発明には、配列番号1、3および5の予想アミノ酸配列を有するポリペプチド、ならびにそのようなポリペプチドの断片、類縁体および誘導体が含まれる。配列番号1、3および5のポリペプチドに言及する場合の用語「断片」、「誘導体」および「類縁体」は、そのようなポリペプチドと同一の生物学的機能または生物活性を本質的に維持しているポリペプチドを意味する。例えば、類縁体には、プロタンパク質部分の切断によって活性化され活性な成熟タンパク質を生成することができるプロタンパク質が含まれる。本
40
発明のポリペプチドは、組換えポリペプチド、天然ポリペプチドまたは合成ポリペプチドであってもよいが、グリコシル化、または非グリコシル化組換えポリペプチドであることが好ましい。

【0053】

それぞれ配列番号1、3および5のポリペプチドの断片、誘導体または類縁体は、(i) 1個または複数のアミノ酸残基が、保存または非保存アミノ酸残基(好ましくは保存アミノ酸残基)で置換され、このような置換アミノ酸残基が、遺伝子コードによってコードされている、またはされていない断片、誘導体または類縁体、(ii) 1個または複数のアミノ酸残基に置換基が含まれる断片、誘導体または類縁体、または(iii) リーダーもしく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に用いられる配
50

列などの成熟タンパク質に追加のアミノ酸が融合された断片、誘導体または類縁体であってもよい。このような断片、誘導体および類縁体は、当業者の範囲内にあり、本明細書の教示に基づいて提供されると見なされる。

【0054】

本発明のポリペプチドおよびポリヌクレオチドは、単離形態でなければならず、実質的な均質性または純度まで精製されていることが好ましい。実質的な均質性とは、少なくとも約85%の純度を意味する。

【0055】

用語「単離」は、物質がその元の環境（例えば、天然に存在する場合には自然環境）から取り出されたことを意味するために使用される。例えば、生きている動物中に存在する天然に存在するポリヌクレオチドまたはポリペプチドは、単離されているとは見なされないが、同一のポリヌクレオチドまたはポリペプチドが自然系に共存する物質のほぼすべてから分離されている場合には単離されていると見なされる。DNAの場合、この用語には、例えば、ベクター中に、自己複製プラスミドまたはウイルス中に、または原核生物または真核生物のゲノムDNA中に組み入れられる組換えDNA；または他の配列とは関係なく別の分子（例えば、cDNAもしくはゲノムDNAまたはポリメラーゼ連鎖反応（PCR）または制限エンドヌクレアーゼ消化によって生成したcDNA断片）として存在する組換えDNAが含まれる。また、この用語には、追加のポリペプチド配列、例えば融合タンパク質をコードするハイブリッド遺伝子の一部である組換えDNAも含まれる。さらに、配列番号2、4または6の1つで示され、DPRPの選択的スプライズバリエーションをコードするヌクレオチドの一部が含まれる組換えDNAが含まれる。様々な選択的スプライズバリエーションを、配列番号8、10、12、14、16、18、20、22、24、26、28、30、32、34、36、38、40、42、44および46に例示する。

【0056】

本発明のポリペプチドには、配列番号1、3および5のポリペプチド（特に成熟タンパク質における）、ならびに配列番号1、3および5のポリペプチドのうち1つと少なくとも70%の類似性（例えば、好ましくは少なくとも60%、より好ましくは少なくとも70%の同一性）、より好ましくは配列番号1、3および5のポリペプチドのうち1つと少なくとも90%の類似性（例えば、好ましくは少なくとも90%の同一性）、最も好ましくは配列番号1、3および5のポリペプチドのうち1つと少なくとも95%の類似性（例えば、好ましくは少なくとも95%の同一性）を有するポリペプチドのいずれか1つが含まれる。さらに、少なくとも30個のアミノ酸、より好ましくは少なくとも50個のアミノ酸の配列を含むようなポリペプチドの正確な部分が含まれることが好ましい。

【0057】

本発明のポリペプチドの断片または部分は、ペプチド合成によって対応する完全長ポリペプチドを産生するための中間体として用いてもよい。本発明のポリヌクレオチドの断片または部分を用い、本発明の完全長ポリヌクレオチドを合成することもできる。

【0058】

また、本発明には、このようなポリヌクレオチドが含まれるベクター、そのようなベクターを用いて遺伝子組み換えされた宿主細胞、および前記を用いる組換え技法によるポリペプチドの産生が含まれる。宿主細胞は、例えばクローニングベクターまたは発現ベクターであってもよい。このようなベクターにより一般的に遺伝子組み換え（形質導入または形質転換または形質移入）される。このベクターは、例えばプラスミド、ウイルス粒子、ファージなどの形態であってもよい。組換え宿主細胞は、プロモーターを活性化し、形質転換体を選択しまたは本発明の遺伝子を増幅するため必要に応じて改良された従来の栄養培地中で培養することができる。温度、pHなどの培養条件は、当業者によく知られているように、発現のために選択された宿主細胞について一般的に用いられる条件である。

【0059】

本発明のポリヌクレオチドは、組換え技法によりポリペプチドを産生するために用いてもよい。したがって、例えば、ポリヌクレオチドは、ポリペプチドを発現するための様々な

10

20

30

40

50

発現ベクターのいずれかが1つに含まれる。このようなベクターには、染色体、非染色体および合成DNA配列、例えばSV40の誘導體；細菌プラスミド；ファージDNA；バキュロウイルス；酵母プラスミド；プラスミドおよびファージDNA、ワクシニア、アデノウイルス、家禽ポックスウイルス、および仮性狂犬病ウイルスなどのウイルスDNAの組み合わせに由来するベクターが含まれる。しかしながら、宿主中で複製可能かつ生存可能である限り、その他のベクターも使用することができる。

【0060】

様々な手順のいずれかにより、適切なDNA配列をベクター中に挿入することができる。一般に、当技術分野でよく知られている手順によりDNA配列を適切な制限エンドヌクレアーゼ部位に挿入するが、この手順は当業者の範囲内にあると見なされる。

10

【0061】

発現ベクター中のDNA配列は、適切な発現制御配列（プロモーター）と動作可能に連結され、mRNA合成を指揮する。このようなプロモーターの代表例として、LTRまたはSV40プロモーター、大腸菌lacまたはtrp、ファージラムダP_{sub}.Lプロモーターおよび原核生物もしくは真核生物細胞またはそれらのウイルスにおいて遺伝子の発現を制御することが知られているその他のプロモーターを挙げることができる。また、発現ベクターは、翻訳開始および転写終結のためのリボソーム結合部位を含まなければならない。また、ベクターには、発現を増幅するのに適した配列も含まれる。さらに、発現ベクターは、真核生物細胞培養の場合にはジヒドロ葉酸還元酵素もしくはネオマイシン耐性など、または大腸菌においてはテトラサイクリンもしくはアンピシリン耐性などの、形質転換宿主細胞を選択するための表現型形質を提供するための1個または複数の選択可能なマーカー遺伝子を含むことが好ましい。

20

【0062】

前述の適切なDNA配列、ならびに適切なプロモーターまたは制御配列を含むベクターを用いて適切な宿主を形質転換し、宿主がタンパク質を発現することを可能にすることができる。適切な宿主の代表例として、大腸菌、ストレプトマイセス属（Streptomyces）、ネズミチフス菌（Salmonella typhimurium）などの細菌細胞；酵母などの真菌細胞；ショウジョウバエ（Drosophila）S2およびSpodoptera Sf9などの昆虫細胞；CHO、COSまたはBowes黒色腫などの動物細胞；アデノウイルス；植物細胞などを挙げることができる。適切な宿主の選択は、本明細書の教示から当業者の範囲内にあると見なされる。

30

【0063】

CLONTECH 95/96 Catalogue、215～216ページ、CLONTECH、1020 East Meadow Circle、Palo Alto、California、94303から明らかなように、核酸配列の合成的製造は当技術分野でよく知られている。したがって、本発明には、本発明のタンパク質を製造するのに有用な発現ベクターも含まれる。

【0064】

さらに、本発明には、広く前述される1個または複数の配列を含む組換え構築物が含まれる。この構築物は、プラスミドまたはウイルスベクターなどのベクターを含むことができ、これらには本発明の配列が前方または後方に挿入されている。この実施形態の好ましい態様では、構築物は、例えば、配列に動作可能に連結されたプロモーターを含むレギュレーター配列を含む。多数の好適なベクターおよびプロモーターが当業者に知られており、市販されている。例えば、以下のベクターを示す。細菌性：pQE70、pQE60、pQE-9（Qiagen）、pBS、pD10、ファージスクリプト、psiX174、pブラスクリプトSK、pbsks、pNH8A、pNH16a、pNH18A、pNH46A（Stratagene）、ptrc99a、pKK223-3、pKK233-3、pDR540およびpRIT5（Pharmacia）；および真核生物：pWLEO、pSV2CAT、pOG44、pXT1、pSG（Stratagene）pSVK3、pBPV、pMSG、およびpSVL（Pharmacia）。しかしながら、

40

50

宿主中で複製可能かつ生存可能である限り、その他の好適なプラスミドまたはベクターを使用することができる。

【0065】

プロモーター領域は、選択可能なマーカーと共にCAT（クロラムフェニコールアセチル転移酵素）ベクターまたは他のベクターを用い、望ましいどの遺伝子からも選択することができる。2種類の適切なベクターはpKK232-8およびpCM7である。特に有名な細菌プロモーターには、lacI、lacZ、T3、T7、gpt、ラムダP.sub.R、P.sub.Lおよびtrpが含まれる。真核生物プロモーターには、極初期CMV、HSVチミジンキナーゼ、初期および後期SV40、レトロウイルスからのLTR、およびマウスメタロチオネイン-Iが含まれる。適切なベクターおよびプロモーターの選択は、当業者のレベルの範囲にある。

10

【0066】

発現ベクターの構成要素には、一般的に1) 選択マーカーとしてのネオマイシンホスホトランスフェラーゼ(G418)、またはハイグロマイシンBホスホトランスフェラーゼ(hyg)遺伝子、2) 大腸菌複製起点、3) T7およびSP6ファージプロモーター配列、4) lacオペレーター配列、5) ラクトースオペロン抑制遺伝子(lacIq)および6) マルチプルクローニング部位リンカー領域が含まれる。このような複製起点(oriC)は、pUC19(LTI、Gaithersburg、Md.)から誘導することができる。

【0067】

適切な制限酵素認識部位を有し、ポリペプチド配列番号2、4および6のうち1つをコードするヌクレオチド配列は、例えば、以下の実施例1に記載のPCRプロトコルに従い、DPRP-1の場合にはKpnI(5'プライマーとして)およびNotIまたはSacI(3'プライマーとして)の制限酵素認識部位、またはDPRP-2の場合にはHindIII(5'プライマーとして)およびNotIまたはBamHI(3'プライマーとして)の部位を用いるPCRプライマーを用いて作製する。PCR挿入断片をゲル精製し、適合する制限酵素で消化する。挿入断片およびベクターを標準プロトコルに従ってライゲーションする。

20

【0068】

他の実施形態では、本発明は、前述の構築物を含む宿主細胞を提供する。宿主細胞は、哺乳類細胞などの高等な真核生物細胞、酵母細胞などの下等な真核生物細胞であってもよく、または宿主細胞は、細菌細胞などの原核生物細胞であってもよい。宿主細胞中への構築物の導入は、リン酸カルシウム形質移入、DEAE-Dextran媒介性形質移入、リポフェクションまたは電気穿孔法によって行うことができる(Davis, L., Digner, M., Battley, I., Basic Methods in Molecular Biology, (1986))。

30

【0069】

宿主細胞中のこのような構築物は、組換え配列によってコードされる遺伝子産物を産生するために従来の方法で用いられることが好ましい。あるいは、従来のペプチド合成装置により、またはそのようにして調製した好適な断片の化学的ライゲーションにより、本発明のポリペプチドを合成的に製造することができる。

40

【0070】

成熟タンパク質は、適切なプロモーターの制御の下で哺乳類細胞、酵母、細菌、または他の細胞中で発現させることができる。また、無細胞翻訳系を用い、本発明のDNA構築物に由来するRNAを用いてこのようなタンパク質を産生することができる。原核生物および真核生物宿主と共に用いるのに適したクローニングおよび発現ベクターは、Sambrook他、Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989)によって記載されている。

【0071】

50

高等な真核生物による本発明のポリペプチドをコードするDNAの転写は、ベクター中にエンハンサー配列を挿入することによって増加する。エンハンサーには、プロモーター上で作用してその転写を増加させるDNA（通常は約10から300bp）のシス作用性因子が含まれる。その例には、転写起点bp100から270の後側上のSV40エンハンサー、サイトメガロウイルス初期プロモーターエンハンサー、転写起点の後側上のポリオマエンハンサー、およびアデノウイルスエンハンサーが含まれる。

【0072】

一般的に、組換え発現ベクターには、複製起点および宿主細胞の形質転換を可能にする選択可能なマーカー、例えば大腸菌のアンピシリン耐性遺伝子およびサッカロマイセス・セラヴィシエ (*S. cerevisiae*) TRP1 遺伝子、および下流の構造配列の転写を指揮するために高度に発現した遺伝子から誘導されるプロモーターが含まれる。このようなプロモーターは、3-ホスホグリセリン酸キナーゼ (PGK)、アルファ-因子、酸性ホスファターゼ、または熱ショックタンパク質などの解糖酵素をコードするオペロンから誘導することができる。異種構造配列は、適切な段階で、翻訳開始および終結配列、好ましくは翻訳タンパク質の細胞膜周辺腔または細胞外培地への分泌を方向付けることができるリーダー配列と一緒に組み立てる。場合により、異種配列は、望ましい特徴、例えば発現組換え産物の安定化または簡単な精製を付与するN末端識別ペプチドを含む融合タンパク質をコードすることができる。

10

【0073】

細菌使用に有用な発現ベクターは、望ましいタンパク質をコードする構造DNA配列を、機能性プロモーターにより動作可能なリーディング段階で好適な翻訳開始および終結シグナルと一緒に挿入することにより構築する。このベクターは、1個または複数の表現型選択可能なマーカーならびにベクターの維持を保証し、望ましい場合には宿主内の複製を提供するための複製起点を含むものと思われる。形質転換に適した原核生物宿主には、大腸菌、枯草菌、ネズミチフス菌およびシュードモナス属の様々な菌種、ストレプトマイセス属、ならびにブドウ球菌属が含まれるが、その他の宿主を選択して用いることができる。

20

【0074】

代表的であるが非限定的な例として、細菌使用に有用な発現ベクターは、選択可能なマーカーおよびよく知られているクローニングベクター pBR322 (ATCC 37017) の遺伝子を含む市販のプラスミドに由来する細菌の複製起点を含むことができる。このような市販ベクターには、例えば pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) および GEM1 (Promega Biotech, Madison, Wis., U.S.A.) が含まれる。これらの pBR322 「主鎖」 切片を、適切なプロモーターおよび発現される構造配列と結合させる。

30

【0075】

好適な宿主株の形質転換および適切な細胞密度までの宿主株の培養後、適切な手段（例えば、温度変化または化学誘導）により選択プロモーターを誘導し、さらに細胞を培養する。通常は遠心分離により細胞を収集し、次いで物理的または化学的手段により破壊し、得られる粗製抽出物をさらに精製するために保持する。タンパク質の発現で用いられる微生物細胞は、凍結融解循環、超音波処理、機械的破壊および細胞溶解剤の使用を含む好都合ないかなる方法によっても破壊でき、このような方法は当業者によく知られている。

40

【0076】

また、様々な哺乳類細胞培養系を用い、組換えタンパク質を発現させることができる。哺乳類発現系の例には、Gluzman, Cell. 23: 175 (1981) によって記載されたサル腎臓線維芽細胞のCOS-7系が含まれる。適合するベクターを発現することができるその他の細胞系には、例えばC127、3T3、CHO、HeLaおよびBHK細胞系が含まれる。哺乳類発現ベクターは、一般に複製起点、好適なプロモーターおよびエンハンサー、および必要なりボソーム結合部位、ポリアデニル化部位、スプライスドナーおよびスプライスアクセプター部位、転写終結配列、および5'隣接非転写配列を含むものと思われる。SV40スプライスに由来するDNA配列、およびポリアデニル化部

50

位を用い、必要な非転写遺伝因子を提供することができる。

【0077】

組換え細胞培養物から、硫酸アンモニウムまたはエタノール沈殿、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、ヒドロキシルアパタイトクロマトグラフィーおよびレクチンクロマトグラフィーを含む方法によりポリペプチドを回収しかつ精製することができる。細胞の表面でポリペプチドが発現されている場合には容易に回収することができるが、そのようなことは必要条件ではない。また、より長い形態のポリペプチドを発現させた後で切断産物を回収することも望ましい。必要に応じ、当技術分野で知られているタンパク質リフォールディングステップを用いて成熟タンパク質の立体配置を完成することができる。最終精製ステップには高速液体クロマトグラフィー（HPLC）を用いることができる。

10

【0078】

本発明のポリペプチドは、精製天然生成物であってもよく、または原核生物または真核生物宿主（例えば、培養物中の細菌、酵母、高等植物、昆虫または哺乳類細胞による）からの組換え技法により産生させてもよい。組換え産生手順に用いる宿主に応じて、本発明のポリペプチドはグリコシル化されてもよく、またはグリコシル化されなくてもよい。また、本発明のポリペプチドには、最初のメチオニンアミノ酸残基が含まれる。

【0079】

好ましい実施形態では、本発明のタンパク質は、他のタンパク質の混入が実質的にないように単離されかつ精製される。例えば、本発明のタンパク質は、サンプル中に存在する全タンパク質の少なくとも80重量%、より好ましくは少なくとも90重量%、さらにより好ましくは少なくとも95重量%、最も好ましくは全タンパク質の少なくとも98重量%を構成していなければならない。

20

【0080】

これらのタンパク質は、水、ジメチルスルホキシド（DMSO）またはエタノールなどの他の好適な溶媒、または好適な溶媒の混合物中で溶液の形態であってもよい。溶媒の混合物の例には、10（重量）%エタノール水溶液および水に溶けた2（重量）%DMSO水溶液が含まれる。さらに、溶液は、塩、緩衝剤、カオトロピック剤、界面活性剤、保存剤などを含むことができる。あるいは、タンパク質は凍結乾燥粉末または結晶性固体などの固体の形態であってもよく、これらは残留溶媒、塩などを含んでいてもよい。

30

【0081】

本明細書で使用する用語「抗体」には、ポリクローナル抗体、アフィニティー精製ポリクローナル抗体、モノクローナル抗体、ならびにF(ab')₂およびFab'タンパク質分解性断片などの抗原結合性断片が含まれる。キメラ抗体、Fv断片、単鎖抗体などの遺伝子組み換えされたインタクトな抗体または断片、ならびに合成抗原結合性ペプチドおよびポリペプチドも含まれる。ヒト以外の抗体は、ヒトフレームワークおよび定常領域上にヒト以外のCDRを移植することにより、またはヒト以外の可変領域全体を組み入れることにより（場合により露出残基の置換によりヒト様表面で抗体を「覆う（cloaking）」（その結果、生成するものは「化粧被覆（veneered）」抗体である））ヒト化することができる。ある場合には、ヒト化抗体は、ヒト可変領域フレームワークドメイン内にヒト以外の残基を保有し、適切な結合特性を高めることができる。抗体をヒト化することにより、生物学的半減期を増加させ、ヒトへの投与の際の有害な免疫反応の可能性を軽減しなければならない。

40

【0082】

本明細書で有用な抗体を作製または選択するための代替技法には、ヒトプロホルモンDPRPタンパク質またはそれに由来するペプチドへのin vitroにおけるリンパ球の曝露、およびファージまたは類似のベクターにおける抗体ディスプレイライブラリーの選択（例えば、固定化または標識ヒトDPRPタンパク質またはペプチドの使用による）が含まれる。潜在的なヒトDPRPポリペプチド結合ドメインを有するポリペプチドをコー

50

ドする遺伝子は、ファージ（ファージディスプレイ）上または大腸菌などの細菌上に示されるランダムペプチドライブラリーをスクリーニングすることによって得ることができる。このようなポリペプチドをコードするヌクレオチド配列は、当技術分野でよく知られている多くの方法で得ることができる。

【0083】

当業者には明らかなように、ポリクローナル抗体は、ウマ、ウシ、ヤギ、ヒツジ、イヌ、ニワトリ、ウサギ、マウスまたはラットなどの様々な温血動物にヒトDPRPポリペプチドまたはその断片を接種することにより作製することができる。ヒトプロホルモンDPRPポリペプチドの免疫原性は、ミョウバン（水酸化アルミニウム）またはフロイント完全もしくは不完全アジュバントなどのアジュバント、リソレシチン、プルロニックポリオール、ポリアニオン、ペプチド、油乳濁液、KLHまたはジニトロフェノールなどの界面活性物質の使用により増加させることができる。ヒトで用いるアジュバントの中で、BCG（カルメットゲラン菌）およびコリネバクテリウムパルブム（*Corynebacterium parvum*）が特に好ましい。また、免疫化に有用なポリペプチドには、免疫グロブリンポリペプチドまたはマルトース結合タンパク質とDPRPもしくはその一部の融合物などの融合ポリペプチドが含まれる。ポリペプチド免疫原は、完全長分子またはその一部であってもよい。ポリペプチド部分が「ハプテン様」である場合には、そのような部分をキーホールリンペットヘモシニアン（KLH）、ウシ血清アルブミン（BSA）または破傷風トキソイドなどの巨大分子キャリアと免疫化のために好都合に結合または連結することができる。また、DPRPに対する抗体は、当技術分野でよく知られている方法を用いて作製することができる。このような抗体には、ポリクローナル、モノクローナル、キメラ、および単鎖抗体、Fab断片、およびFab発現ライブラリーにより製造される断片が含まれるが、これらだけに限定されるものではない。治療上の使用には中和抗体（すなわち、活性部位において相互作用を遮断または修正する抗体）が特に好ましい。

10

20

【0084】

抗体を産生するために、結合タンパク質、またはDPRPと特異的に結合するペプチド、単鎖抗体のライブラリー、Fab断片、その他の抗体断片、非抗体タンパク質ドメイン、またはペプチドをスクリーニングすることができる。これらのライブラリーは、ファージディスプレイ、その他の組換えDNA法、またはペプチド合成を用いて作製することができる（Vaughan, T. J. 他、*Nature Biotechnology* 14 : 309 ~ 314 (1966)）。このようなライブラリーは、当技術分野でよく知られている方法を用いて普通にスクリーニングし、DPRPに対して特異的結合を示す配列を同定することができる。

30

【0085】

DPRPに対する抗体を誘導するために用いられるオリゴペプチド、ペプチド、または断片は、少なくとも約5個のアミノ酸、より好ましくは少なくとも約10個のアミノ酸からなるアミノ酸配列を有することが好ましい。また、これらのオリゴペプチド、ペプチド、または断片は、天然タンパク質のアミノ酸配列の一部と一致していることが好ましい。また、DPRPアミノ酸の短いストレッチをKLHなどの別のタンパク質の短いストレッチと融合させることができ、キメラ分子に対する抗体を産生することができる。

40

【0086】

DPRPに対するモノクローナル抗体は、培養における連続継代細胞系による抗体分子の産生を提供するよく知られているどの技法を用いても調製することができる。これらの技法には、ハイブリドーマ技法、ヒトB細胞ハイブリドーマ技法、およびEBVハイブリドーマ技法が含まれ、これらに限定されるものではないが、ハイブリドーマ細胞により産生されるモノクローナル抗体が好ましい。

【0087】

さらに、適切な抗原特異性および生物活性のある分子を得るためのヒト抗体遺伝子に対するマウス抗体遺伝子のスプライシングなどの「キメラ抗体」の産生のために開発された技法を用いることができる、Neuberger, M. S. 他、*Nature* 312 : 6

50

04~608(1984)を参照。あるいは、単鎖抗体を産生するために記載された技法を当技術分野で知られている方法を用いて適合させ、DPRP特異的単鎖抗体を産生することができる。関連する特異性を持つが異なるイディオタイプの組成の抗体は、ランダムコンビナトリアル免疫グロブリンライブラリーからのチェーンシャフリングによって作製することができる(Burton D. R., Proc. Natl. Acad. Sci. 88:11120~11123(1991))。

【0088】

また、リンパ球集団において*in vivo*における産生を誘導することにより、または文献に記載の極めて特異的な結合試薬の免疫グロブリンライブラリーまたはパネルをスクリーニングすることにより抗体を産生することができる。(Orlandi, R. 他、Proc. Natl. Acad. Sci. 86:3833~3837(1989))。 10

【0089】

また、DPRPに対する特異的結合部位を含む抗体断片も作製することができる。例えば、このような断片には、抗体分子のペプシン消化によって生成されるF(ab')₂断片およびF(ab')₂断片のジスルフィド架橋の還元によって作製されるFab断片が含まれるが、これらに限定されるものではない。あるいは、Fab発現ライブラリーを構築し、望ましい特異性を持つモノクローナルFab断片の迅速かつ容易な同定を可能にすることができる。(Huse, W. D. 他、Science 254:1275~1281(1989))。

【0090】

様々な免疫学的検査法を用い、望ましい特異性を有する抗体を同定することができる。確立された特異性を持つポリクローナルまたはモノクローナル抗体のどちらかを用いる競合的結合アッセイまたは免疫放射線アッセイのための多くのプロトコルが当技術分野でよく知られている。このような免疫学的検査法は、通常、DPRPとその特異的抗体の間の複合体形成を測定するものである。2つの非干渉DPRPエピトープと反応するモノクローナル抗体を利用する2部位のモノクローナルベースのイムノアッセイが好ましいが、競合的結合アッセイを用いることもできる。 20

【0091】

前述のように、DPRPを疾患の治療に用いることができる。本発明のこの態様で用いるのに適した薬剤組成物には、有効量の活性成分が含まれ、疾患の1つに関して意図した目的を達成する組成物が含まれる。治療上有効な投与量の決定は、当業者の能力で十分に行える範囲であり、最初は、細胞培養アッセイ、例えば新生細胞のアッセイ、または動物モデル、通常はマウス、ラット、ウサギ、イヌ、もしくはブタのどちらでも推定することができる。また、動物モデルを用い、適切な濃度範囲および投与の経路を決定し、次いで一般的にその情報を用いてヒトにおける投与に有用な投与量および経路を決定することができる。 30

【0092】

治療上有効な投与量は、疾患の特定の症状または状態を寛解させる活性成分、例えばDPRPもしくはその断片、DPRPの抗体、またはDPRPのアゴニスト、アンタゴニストもしくは阻害剤の量を指す。例えば、投与される量は、望ましい標的基質を基質と接触して切断するのに有効であってもよい。同様に、治療上の有効性および毒性は、ED50(集団の50%において治療上有効な投与量)またはLD50(集団の50%に対して致死的な投与量)統計値を計算することによるなどの細胞培養液中または実験動物を用いる標準的医薬手順により決定することができる。毒性効果と治療効果の投与量比は治療係数であり、LD50/ED50比として表すことができる。大きな治療係数を示す薬剤組成物が好ましい。細胞培養アッセイおよび動物試験から得られるデータは、ヒト使用のための用量範囲を策定する際に使用される。このような組成物に含まれる用量は、ED50が含まれ、毒性がほとんどまたはまったくない循環濃度の範囲内にあることが好ましい。用量は、用いられる剤形、患者の感受性、および投与の経路に応じてこの範囲内で変化する。 40

【0093】

正確な用量は、通常治療を必要とする対象に関連する要素に鑑みて医師によって決定され、用量および投与は、十分なレベルの活性部分を提供するため、または望ましい効果を維持するために調整される。考慮すべき要素には、疾患状態の重症度、対象の一般健康状態、対象の年齢、体重、および性別、食事、投与の時間および頻度、薬物の併用、反応感受性、および治療法に対する耐性/応答が含まれる。特定の製剤の半減期およびクリアランス速度に応じ、長時間作用性の薬剤組成物を3から4日ごとに、毎週、または2週ごとに1回投与することができる。

【0094】

本発明のさらに別の態様は、DPRP1、DPRP2およびDPRP-3ポリヌクレオチドのmRNA転写物に対してアンチセンスである配列を有するポリヌクレオチド分子を提供する。アンチセンスポリヌクレオチド分子の投与は、DPRP-1、DPRP2またはDPRP-3によってコードされるタンパク質の産生を遮断することができる。アンチセンスポリヌクレオチド分子を調製し、そのような分子を投与するための技法は、当技術分野で知られている。例えば、アンチセンスポリヌクレオチド分子を細胞との融合のためにリポソーム中に封入することができる。

10

【0095】

特に、特殊化した上皮細胞、免疫細胞（リンパ球およびB細胞）、アストロサイトの腫瘍、および様々なホルモン感受性癌におけるDPRP-1、DPRP-2およびDPRP-3の発現は、癌、異形成および転移の病態生理学における潜在的役割の証拠を提供している。したがって、他の態様では、本発明は、不適当なDPRP活性または発現レベルに係る疾患を検出するための診断アッセイに関する。DPRPと特異的に結合する抗体を、DPRPの発現を特徴とする障害を診断するため、またはDPRPもしくはDPRPのアゴニストもしくはアンタゴニスト（阻害剤）で治療される患者をモニターするためのアッセイで使用することができる。診断目的に有用な抗体は、治療薬について前述した方法と同様の方法で調製することができる。DPRPの診断アッセイには、抗体およびヒト体液中または細胞もしくは組織の抽出物中のDPRPを検出するための標識を利用する方法が含まれる。抗体は改変の有無にかかわらず使用でき、リポーター分子との共有結合性または非共有結合性連結によって標識することができる。多種多様のリポーター分子が当技術分野では知られている。触媒的に不活性となるように改変された組換えDPRPタンパク質を優性ネガティブ阻害因子として用いることもできる。このような改変には、例えば

20

30

【0096】

ELISA、RIAおよびFACSを含む、DPRPを測定するための様々なプロトコルが当技術分野で知られており、DPRP発現の変化レベルまたは異常レベルを診断するための基礎を提供する。DPRP発現の正常値または標準値は、正常な哺乳類対象、好ましくはヒトから得られる体液または細胞抽出物を、複合体形成に適した条件の下でDPRPに対する抗体と合わせることによって確立される。生物学的サンプル中のDPRPを検出する方法は、a)生物学的サンプルを提供するステップと、b)DPRPと抗体の間で複合体形成を起こさせるのに適した条件の下で生物学的サンプルと抗DPRP抗体を合わせるステップと、c)DPRPと抗体の間の複合体形成を検出することにより、生物学的

40

【0097】

本発明の別の実施形態では、DPRPをコードするポリヌクレオチドが診断目的に用いられ、ポリヌクレオチドにはオリゴヌクレオチド配列、相補的RNAおよびDNA分子、ならびにPNAが含まれる。これらのポリヌクレオチドを用い、DPRPの発現が疾患の1つと相関している可能性がある生検組織中の遺伝子発現を検出しかつ定量することができる。診断アッセイを用い、DPRPの不存在、存在、および過剰発現を区別し、治療的介

50

在中のDPRPレベルの調節をモニターすることができる。さらに、DPRP遺伝子に関する薬理ゲノム学的な一塩基多型(SNP)の解析は、疾患になりやすい素因または薬物に対する応答変化を示す変異体をスクリーニングするための方法として用いることができる。

【0098】

DPRPポリヌクレオチドおよびポリペプチド配列、それらの断片、DPRPの抗体、ならびにDPRPのアゴニスト、アンタゴニストもしくは阻害剤を発見ツールとして用い、分子認識事象ならびにDPRPタンパク質と相互作用するタンパク質、ポリペプチドおよびペプチドを同定することができる。具体例はファージディスプレイペプチドライブラリーであり、1回のパンニングで108個を超えるペプチド配列をスクリーニングすることができる。このような方法ならびにその他の方法は、当技術分野で知られており、それらを利用してDPRP-1、DPRP-2またはDPRP-3活性を阻害または高める化合物を同定することができる。カップリングした結合は、複合体または経路などの機能的相互作用を表し、DPRPと相互作用するタンパク質は、酵母の2つのハイブリッド系、プロテオミクス(示差2Dゲル分析および質量分析法)およびゲノミクス(マイクロアレイによる示差的遺伝子発現または遺伝子発現SAGEの逐次解析)により同定することができる。機能的にDPRPと連結すると同定されたタンパク質および相互作用のプロセスは、これらのDPRP-タンパク質相互作用の阻害剤、アゴニストおよびアンタゴニストならびにモジュレーターをスクリーニングする方法の基礎を形成する。

10

【0099】

本明細書で使用する用語「アンタゴニスト」は、DPRPと結合した場合に、DPRPの生物活性または免疫活性効果の量または期間を減少させる、例えば、N末端ジペプチドを切断するペプチダーゼの酵素活性を低下させる阻害剤分子を指す。アンタゴニストには、タンパク質、核酸、炭水化物、抗体、DPRPの効果を減少させるその他の分子が含まれ、例えば、競合型または非競合型機序によりDPRPと結合しかつDPRPを不活性化する小さな分子および有機化合物が含まれる。DPRPテトラペプチドペプチド酵素活性阻害剤の具体例を実施例6および7に記載する。阻害剤は、例えばDPRPプロテアーゼ活性の阻害剤か、またはそれらが相互作用するタンパク質に対するDPRPの結合活性の阻害剤であってもよい。このような阻害剤の具体例には、例えば抗DPRP抗体、ペプチド、タンパク質断片、または小さなペプチジルプロテアーゼ阻害剤、または望ましい細胞タイプ中への導入を可能にする媒体中で調合された小さな非ペプチド有機分子阻害剤が含まれる。あるいは、このような阻害剤を、細胞媒介性エンドサイトーシスおよび他の受容体媒介性事象による導入のための標的リガンドに取り付けることができる。このような方法をさらに以下に説明するが、本明細書に記載のDPRPヌクレオチドおよびアミノ酸配列が与えられれば、当業者によって実施することができる。

20

30

【0100】

DPRPのその他の使用法は、治療薬として使用するため、例えばDPRPとの結合を阻害するための潜在的アンタゴニストのスクリーニング、ならびにアゴニストのスクリーニングのためである。DPRP、その免疫原生断片、またはそのオリゴペプチドは、様々な薬物スクリーニング技法のいずれかにおける予測されるアゴニストまたはアンタゴニストである化合物のライブラリーをスクリーニングするために用いることができる。このようなスクリーニングに用いられる断片は、溶液中に遊離していても、固体支持体に固定されていても、細胞表面上に保持されていても、または細胞内に位置していてもよい。次いで、DPRPと被験試剤との間の結合複合体の形成を測定する。DPRPを阻害するアンタゴニストを発見するためのその他のアッセイは、DPPIVの阻害剤について記載する米国特許第6,011,155号、第6,107,317号、第6,110,949号、第6,124,305号および第6,166,063号の開示から明らかである。これらのDPRPの価値ある別の使用法は、1個または複数のDPRPを阻害することにより、望ましくない副作用を有していないことを示すためのDPPIVの阻害剤のスクリーニングである。

40

50

【0101】

DPRPと結合する分子を同定するために小さな分子のライブラリーをスクリーニングするために提供される方法は、一般にa)小さな分子のライブラリーを提供すること、b)複合体形成に適した条件の下で、小さな分子のライブラリーを、配列番号1、3もしくは5のいずれかのポリペプチドと、またはそれらの断片と合わせることを、およびc)複合体形成を検出すること(ここで、このような複合体の存在は、DPRPと結合する小さな分子を同定する)を含む。

【0102】

アンタゴニストを同定するための一方法は、通常は切断が起きると考えられる条件の下で、DPRPを発現するベクターにより形質転換された細胞からの抽出物中に、DPRPと結合する小さな分子を色素生産性基質(例えばAla-Pro-AFCまたはAla-Pro-AMC)と一緒に送達し、次いで分光光度法によって蛍光、または紫外光吸収の変化をモニターすることにより酵素による切断の阻害についてアッセイし、切断を阻害する分子を同定することを含む。分子の存在下で反応速度または蛍光もしくは紫外光吸収の総量が減少することは、小さな分子がDPRPの触媒/酵素活性を減少させるアンタゴニストであることを証明している。このような分子を同定したら、それらを投与し、DPRPによる切断を軽減または阻害することができる。

10

【0103】

本明細書で使用する用語「アゴニスト」は、DPRPと結合した場合に、DPRPの効果の期間を増加または延長させる分子を指す。アゴニストには、タンパク質、核酸、炭水化物、またはDPRPと結合しかつDPRPの効果調節するその他の分子が含まれる。小さな分子が有効なDPRPアゴニストであると分かることはあまりないが、アゴニストとしてDPRPと結合するこのような小さな分子を同定する方法は、DPRPを発現するベクターにより形質転換された細胞中に、DPRPと結合する色素生産性形態の小さな分子を送達すること、分光光度法により蛍光、または紫外光吸収変化をアッセイすることを含む。紫外吸収または蛍光の量の増加は、小さな分子がDPRP活性を増加させるアゴニストであることを証明するであろう。

20

【0104】

使用することができる薬物スクリーニングのための別の技法は、公開PCT出願WO84/03564に記載のように、当該タンパク質に対して好適な結合親和性を有する化合物のハイスループットスクリーニングを提供する。この方法では、プラスチックピンなどの固体基体上、または他の表面上で多数の異なる小さな試験化合物が合成される。試験化合物をDPRP、またはその断片と反応させ、次いで洗浄する。次いで、当技術分野でよく知られている方法により、結合したDPRPを検出する。また、前述の薬物スクリーニング技法で使用するために、精製DPRPをプレート上へ直接コーティングすることができる。あるいは、非中和抗体を用い、ペプチドを捕捉し、それを固体支持体上に固定化することができる。

30

【0105】

別の実施形態では、DPRPと結合することができる中和抗体をDPRPとの結合についての試験化合物と特異的に競合させる競合的薬物スクリーニングアッセイを用いることができる。この方法では、抗体を用い、DPRPと1個または複数の抗原決定基を共有するいずれのペプチドの存在も検出することができる。

40

【0106】

前述のように、結合部位を検討することにより、例えば、天然リガンドよりもDPRPとの相互作用の大きなリガンドを設計することができる。このようなアンタゴニストリガンドは、より大きな親和性でDPRPと結合し、その結果競合的リガンドとして機能する。あるいは、天然DPRPのリガンド結合部位と相同または類似した合成または組換えタンパク質を、DPRPに対して大きな親和性を有する他の分子と同様に設計することができる。また、このような分子は、DPRPと置き換わることができ、防御効果を提供することができるべきである。

50

【0107】

前述のように、DPRPの構造についての知識は、合成結合部位相同体および類縁体の設計を可能にする。このような分子は、潜在的治療薬を標的とするための結合特性の使用法を極めて容易にするはずであり、それらを用いて潜在的治療薬をスクリーニングすることができる。さらに、それらの分子をモノクローナル抗体の産生における免疫原として用い、抗体自体は、前述の診断および/または治療に用いることができる。

【0108】

プロリルオリゴペプチダーゼS9Bファミリーのいくつかのメンバーについての遍在性の発現を考えると、DPPIV、DPRP-1、DPRP-2、DPRP-3、FAPおよびDPPVIの標的遺伝子の破壊が証明されているヌル表現型の細胞系は、選択的かつ効力のある化合物のスクリーニングを支援する価値が大きいはずである。したがって、本発明は、体細胞遺伝子標的ベクターを構築するためのLox-Neo IRES tkカセットおよびGFP-IRES-Neo Knock-in/outカセットDNAエレメントで組み換えられた細胞系を提供する。

10

【0109】

実施例1

哺乳類発現系を用いるDPRP遺伝子のクローニングおよび発現

遺伝子の5'および3'配列に対応するPCRオリゴヌクレオチドプライマー、すなわち配列番号45および46を用い、完全長ポリペプチドDPRP-1をコードするDNA断片を増幅させた。さらに、その遺伝子の5'および3'配列に対応するPCRオリゴヌクレオチドプライマー、すなわち配列番号50および51を用い、完全長ポリペプチドDPRP-2をコードするDNA断片を増幅させた。さらに、その遺伝子の5'および3'配列に対応するPCRオリゴヌクレオチドプライマー、すなわち配列番号55および56を用い、完全長ポリペプチドDPRP-3をコードするDNA断片を増幅させた。

20

【0110】

この3つの増幅配列は、市販のキット(GFX PCR DNAおよびGel Band Purification Kit、Amersham Pharmacia Biotech Inc.、Piscataway NJ、USA)を用いて0.7%アガロースゲルからそれぞれ単離した。次いで、断片をクローニングベクター、pGEM-7Zf(-)(Promega Corporation、Madison WI、USA)中にライゲーションし、配列決定した。対応するクローニング構築物は、それぞれpGEM7-DPRP1、pGEM7-DPRP2およびpGEM7-DPRP3と命名した。テンプレートとしてのpGEM7-DPRP1またはpGEM7-DPRP2またはpGEM7-DPRP3およびPCRオリゴヌクレオチドプライマーを用い、切断型DPRP-1またはDPRP-2またはDPRP-3をコードするDNA配列を増幅させた。DPRP-1の場合は配列番号45および47を用い、DPRP-2の場合は配列番号50および52を用い、DPRP-3の場合は配列番号57および58を用いた。増幅配列は、さらにまた、同じ精製キットを用いて0.7%アガロースゲルから単離し、pGEM-7Zf(-)中にサブクローニングした。得られた構築物は、pGEM7-DPRP1f、pGEM7-DPRP2fおよびpGEM7-DPRP3fと命名した。

30

40

【0111】

DPRP-1哺乳類発現構築物を作製するため、pGEM7-DPRP1を制限酵素KpnIおよびNotIで消化し、完全長DPRP-1遺伝子を遊離させた。DPRP-1遺伝子を運ぶDNA断片は、上記キットを用いてゲルバンド精製し、次いで発現ベクターpcDNA3(Invitrogen、Carlsbad CA、USA)中に挿入してネイティブなDPRP-1発現構築物を作製し、これをpcDNA-DPRP1と命名した。pGEM7-DPRP1fを制限酵素XbaIおよびHindIIIで消化し、切断型DPRP-1f遺伝子を遊離させた。DPRP-1f遺伝子を運ぶDNA断片は、上記キットを用いてゲルバンド精製し、次いで発現ベクターpcDNA3.1(-)/myc-HisA(Invitrogen、Carlsbad CA、USA)中に挿入し、タ

50

グ付きDPRP - 1発現構築物pcDNA - MycHis - DPRP 1を作製した。

【0112】

DPRP - 2哺乳類発現構築物を作製するため、pGEM7 - DPRP 2を制限酵素HindIIIおよびBamHIで消化し、完全長DPRP - 2遺伝子を遊離させた。DPRP - 2遺伝子を運ぶDNA断片は、上記キットを用いてゲルバンド精製し、次いで発現ベクターpcDNA3 (Invitrogen、Carlsbad CA、USA)中に挿入してネイティブなDPRP - 2発現構築物を作製し、これをpcDNA - DPRP 2と命名した。pGEM7 - DPRP 2fを制限酵素EcoRIおよびBamHIで消化し、切断型DPRP - 2f遺伝子を遊離させた。DPRP - 2f遺伝子を運ぶDNA断片は、上記キットを用いてゲルバンド精製し、次いで発現ベクターpcDNA3.1(-)/myc-His B (Invitrogen、Carlsbad CA、USA)中に挿入し、pcDNA - MycHis - DPRP 2と命名したタグ付きDPRP - 2発現構築物を作製した。

10

【0113】

DPRP - 3哺乳類発現構築物を作製するため、pGEM7 - DPRP 3を制限酵素EcoRIおよびXhoIで消化し、完全長DPRP - 3遺伝子を遊離させた。DPRP - 3遺伝子を運ぶDNA断片は、上記キットを用いてゲルバンド精製し、次いで発現ベクターpcDNA3 (Invitrogen、Carlsbad CA、USA)中に挿入し、pcDNA - DPRP 3と命名されたネイティブなDPRP - 3発現構築物を作製した。pGEM7 - DPRP 3fを制限酵素NheIおよびApaIで消化し、切断型DPRP - 3f遺伝子を遊離させた。DPRP - 3f遺伝子を運ぶDNA断片は、上記キットを用いてゲルバンド精製し、次いで発現ベクターpcDNA3.1(-)/myc-His B (Invitrogen、Carlsbad CA、USA)中に挿入し、タグ付きDPRP - 3発現構築物pcDNA - MycHis - DPRP 3を作製した。

20

【0114】

実施例2

ヒト組織におけるDPRP遺伝子の発現パターン

定量的PCR分析を行い、ヒト組織における本発明のポリペプチドについてのmRNAの発現レベルを調べた。前立腺癌細胞(LNCaP、PC3、DU145)、MLTC-1系(マウス精巢)、およびMDA-MB231細胞(乳癌)を含む多数のヒト細胞系に対してはRT-PCRも行った。DPRP - 1、DPRP - 2およびDPPIVについては予想サイズのバンドは様々な癌細胞系ですべて発現され、FAPも極めて低レベルで発現された。

30

【0115】

ノーザンプロット分析

ノーザンプロット分析は、DPRPプローブを用い、8種類の異なる組織から単離された2 μ gのポリ(A)⁺RNAについて行った。具体的には、³²PdCTP(A.P.Feinberg他、Anal.Biochem.、132、6(1983))の存在下で、ランダムプライミングによって放射能標識された1kbのN末端断片についてヒトのマルチプルティッシュノーザン(MTN)プロット(Clontech、Palo Alto、Calif.)を探索した。ハイブリダイゼーションは、ExpressHyb(商標)ハイブリダイゼーション溶液(Clontech、Palo Alto、Calif.)中で68 $^{\circ}$ Cにおいて一夜行った。まず、プロットを2倍SSCおよび0.05%SDS中で室温において洗浄し、次いで0.1倍SSCおよび0.1%SDS中で60 $^{\circ}$ C(DPRP - 1およびDPRP - 2)および50 $^{\circ}$ C(DPRP - 3)において洗浄した。

40

【0116】

ノーザン分析は、いくつかの組織におけるDPRP - 1の発現を示し、精巢、前立腺、筋肉および脳で最も大きなシグナルであった。精巢は、長さが約7.5、4.5および2.5kbの3個の転写物を示した。より短いmRNA種は精巢に極めて多かったが、試験した他の組織ではごくわずかであった。DPRP - 2は、すべての組織で遍在的に発現され

50

、肝臓および筋肉で最も高いレベルであり、主に5 kbの転写物であった。DPRP-3発現は脳および膵臓に限定された。具体的な脳領域(小脳、皮質、髄質、脊髄、後頭葉、前頭葉、側頭葉および被殻)では、3種類のプロテアーゼについてさらに分析を行った。DPRP-1はすべての領域で発現され、脊髄では低レベルであり、一方DPRP-2は試験したすべての脳領域で発現された。

【0117】

DPRP-1の定量的PCRにはオリゴヌクレオチドプライマー配列番号48および49を用い、DPRP-2の定量的PCRにはオリゴヌクレオチドプライマー配列番号53および54を用いた。ヒトマルチプルテリッシュcDNA(MTC(商標))Panel IおよびPanel II(Clontech, Palo Alto CA, USA)を正規化cDNAテンプレートとして用いた。各cDNA0.5 ngを25 μlのPCR反応で用い、各プライマーの最終濃度は300 nMとした。SYBR Green PCR Core Reagents Kit(Applied Biosystems, Foster City CA, USA)を用いてPCR反応を行い、Applied Biosystems GeneAmp 5700配列検出システムで検出した。製造者の推奨熱サイクルパラメータ、例えば50 2分間、95 10分間、続いて95 15秒間と60 1分間の40サイクルを用いた。得られたデータは、膵臓、卵巣および精巣における比較的高率なDPRP-1とDPRP-2の発現、および肝臓におけるDPRP-2の特に高率な発現を示している。

10

【0118】

実施例3-DPRPポリクローナル抗体の産生およびウエスタンブロッティング
DPRP-1をコードするcDNAから推定されるアミノ酸配列をDNASTARソフトウェア(DNASTAR, Inc.)を用いて分析して高い免疫原性の領域を決定し、対応するオリゴペプチドを合成して抗DPRP-1抗体を産生させるために用いた。DPRP-2およびDPRP-3についてこの手順を繰り返した。適切なペプチド配列の選択および抗体産生の技法は、当業者によく知られている方法である。C末端近辺または親水性領域中のエピトープなどの適切なエピトープの選択は、当技術分野でよく知られている。

20

【0119】

通常は、長さが約15から20個の残基であるオリゴペプチド、例えばDPRP-1の場合には配列番号59、DPRP-2の場合には配列番号60、DPRP-3の場合には配列番号61を、Applied Biosystems Peptide Synthesizer Model 431Aを用いて合成した。Fmoc化学反応を用い、N-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステル(MBS)との反応により、19または15残基のペプチドを、それぞれキーホールリンペットヘモシアニン(KLH, Sigma, St. Louis, Mo.)とカップリングさせた。完全フロイントアジュバントにおいて溶かしたオリゴペプチド-KLH複合体でウサギを免疫化した。得られた抗血清は、例えばペプチドをプラスチックに結合させ、1%BSAでブロックし、ウサギ抗血清と反応させ、洗浄し、放射性ヨウ素標識ヤギ抗ウサギIgGと反応させることにより、抗ペプチド活性について試験した。

30

【0120】

Clontechから入手した正常ヒトタンパク質サンプル(Protein Medley)(総タンパク質約36 μg)を用いてウエスタンブロッティングを行った。10%SDS-ポリアクリルアミドゲルによりタンパク質を分画し、0.45 mmのニトロセルロース膜に移した。Tris緩衝塩水(TBS)中で0.05%Tween 20および1%BSAにより膜をブロックした。抗DPRP-1またはDPRP-2特異抗体を一次抗体として用い、0.05%Tween 20(TBST)を含むTris緩衝塩水で1:5,000に希釈し、アルカリホスファターゼ(AP)結合ヤギ抗ウサギIgG(Promega)を使用前に同じ緩衝液で1:5,000に希釈した。AP用のWestern Blue Stabilized Substrate(Promega)中で膜をインキュベートすることにより、当該バンドが望ましい強度に達するまで陽性反応を可視化し

40

50

た。DPRP-1およびDPRP-2タンパク質は、脳、筋肉、腎臓、前立腺、精巣および卵巣組織で検出された。DPRP-1およびDPRP-2をそれぞれ約101kDaおよび100kDa体として合成し、これらは、表3に示す一次構造から推定される分子量とよく一致している。

【0121】

【表3】

表3. HoppおよびWoods、Proc. Nat. Acad. Sci. 78:3824~3828 (1981) によって開発された方法を用いる配列分析に基づく、DPRP-1、DPRP-2およびDPRP-3の予測分子量、潜在的N結合グリコシル化部位(Asn残基)数および予測pI値。

	分子量 (ダルトン)	Asn数	pI
DPRP1	101422	26	5.39
DPRP2	98263	27	6.01
DPRP3	90914	33	6.11

10

20

【0122】

類似する分子量の追加のバンドがいくつか観察された。これらのバンドは、タンパク質の翻訳後グリコシル化の存在に起因すると考えられる。また、表3は、DPRPタンパク質についての潜在的Nグリコシル化部位数を示す。タンパク質のグリコシル化および非グリコシル化体の存在は、オリゴ糖合成の阻害剤であるツニカマイシンを用いて評価した。より小さな形が非グリコシル化体であることは明らかである。DPRP-1についてmRNA(ノーザン分析)とタンパク質含量(ウエスタン分析)の相関関係を表4に示す。

【0123】

【表4】

表4. ヒト組織におけるDPRP-1のmRNAとタンパク質発現の相関関係

	心臓	脳	胎盤	筋肉	腎臓	前立腺	精巣	卵巣
ノーザン	++	+++	+	+++	++	+++	++++	+
ウエスタン	-	++++	-	+	++	+	+++	+++

30

【0124】

実施例4

ヒト組織におけるDPRPタンパク質の免疫組織化学的局在化

多くの様々なホルマリン固定、パラフィン包埋ヒト組織から4ミクロンの切片を調製した。キシレン中5分間と、続く段階的なアルコール系から蒸留水への4回の浸漬により組織切片からパラフィンを除いた。2種類の異なる濃度において、酵素消化組織を含む、および含まないいくつかの異なるSHIER溶液とともに蒸気熱誘導エピトープ回収(steam heat induced epitope recovery)(SHIER)を用いた(Ladner他、Cancer Res.; 60, p3493~3503, 2000)。処理および用いた抗体希釈液について以下に概要を説明する。

40

【0125】

1. ブロッキング試薬15分間(正常ヤギ血清)

50

2. 一次抗体 25、60 分間または一夜インキュベーション
3. 二次抗体 25 分間 (ビオチン化ヤギ抗ウサギ IgG)
4. 内因性ペルオキシダーゼブロッキング 3 × 1.5 分間
5. ABC (アビジン - ビオチン複合体) / 西洋ワサビペルオキシダーゼ 25 分間
6. DAB 色素原 3 × 5 分間 (ブラウン反応生成物)
7. Light Hematoxylin Counter 染色 1 分

【0126】

検出化学反応を保証するために陽性対照を実施し、抗原前処理を適切に行った。ウサギ IgG を陰性対照として用いた。DPRP - 1 抗体の検出には、アビジン - ビオチンをベースとする組織染色系を用いた。色素原としての DAB と共に、西洋ワサビペルオキシドをリポーター酵素として用いた。染色後、アルコール系から無水エタノールと、続くキシレン洗浄によりスライドを脱水した。スライドをガラス製カバースリップおよびパーマウント (permount) で恒久的に覆った。陽性染色が暗褐色の色素原 (DAB - HRP 反応生成物) によって示されている代表的な染色のデジタル画像を、Olympus 製のビデオカメラを用いて捕えた。ヘマトキシリン対比染色により、細胞および組織形態を評価するための青い核染色が得られた。

10

【0127】

DPRP - 1 ウサギポリクローナル抗体は、正常な精巣、前立腺、子宮内膜腺、扁桃腺および膵臓を含むホルマリン固定、パラフィン包埋ヒト組織を標識する。この抗体は、正常な卵巣、膀胱および腎臓の上皮細胞にも存在した。染色は、線維芽細胞などの上皮および一部の間質細胞、内皮細胞ならびにリンパ球の細胞質に局在化した。興味深いことに、DPRP - 1 抗体について試験した正常な精巣では、ライディッヒ細胞に特有な発現があり、精細管を取り囲む空間である間質組織には多核化マクロファージが見いだされた。扁桃腺 B 細胞は DPRP - 1 抗体で染色された。

20

【0128】

実施例 5

DPRP タンパク質の哺乳類および昆虫細胞発現および精製
製造者によって推奨されている LipofectAmine (Life Technologies, Gaithersburg MD, USA) 法を用い、pcDNA - DPRP 1、pcDNA - MycHis - DPRP 1、pcDNA - DPRP - 2 または pcDNA - MycHis - DPRP 2 のプラスミド DNA を PEAK (Edge Biosystems, Gaithersburg MD, USA) または COS - 1 (ATCC CRL - 1650) 中に形質移入した。形質移入された細胞を、5% FBS を含む DMEM 中で 48 時間、5% CO₂ と共に 37 °C において維持した。次いで、細胞を集め、組換えタンパク質抽出に用いた。形質移入 48 時間後に細胞を収集し、ホモジナイズし、次いで 18,000 × g で 40 分間回転させた。上清をサイトゾル分画として集めた。この分画を TALON スピнкаラム (Clontech) 上にロードし、His タグ付きタンパク質を 50 mM PBS、150 mM イミダゾール、pH 7 で溶離した。次いで、組換えタンパク質を、抗 myc 抗体によるウエスタンブロッティングにより検出し、Protoblot II AP システム (Promega) を用いて可視化した。DPRP - 1 および DPRP - 2 の組換え親和性精製融合物をウエスタンブロットにより検出し、DPRP - 1 および DPRP - 2 を、予測される 112 kDa および 109 kDa 体として合成した。

30

40

【0129】

天然に存在する DPRP タンパク質または組換え DPRP タンパク質は、DPRP - 1、DPRP - 2 または DPRP - 3 に特異的な抗体を用いる免疫親和性クロマトグラフィーにより実質的に精製した。免疫親和性カラムは、CNBr 活性化セファロース (Pharmacia & Upjohn) などの活性化クロマトグラフィー用樹脂に DPRP 抗体を共有結合でカップリングさせることにより作成した。カップリング後、製造者の使用説明書に従って樹脂をブロックし洗浄した。

50

【0130】

DPRPタンパク質を含有する培地または細胞抽出物を免疫親和性カラムに通し、DPRPの優先的吸収を可能にする条件（例えば、界面活性剤存在下における高イオン強度の緩衝液）の下でカラムを洗浄した。抗体/DPRP結合を破壊する条件（例えば、pH 2～3の緩衝液または尿素もしくはチオシアネートイオンなどの高濃度のカオトロープ）の下でカラムを溶離し、精製DPRPを収集した。

【0131】

実施例6

DPRPタンパク質の酵素活性および阻害剤のスクリーニング方法

組換えDPRP-1およびDPRP-2の反応速度特性を連続蛍光定量アッセイで測定した。緩衝液、pHおよび温度依存性最適化の結果、以下のアッセイ条件となった。酵素アッセイは50mM PBS、pH 7.4中において行い、精製酵素50 μ l（50 μ g/ml）を様々な濃度のAla-Pro-AMC（Enzyme Systems）1 μ lと混合した。次いで、プレートを37において30分間インキュベートし、Wallac 1420 Fluorimeterを用いex 40355およびem 535で蛍光を検出した。DPRP-1およびDPRP-2の K_m 値は類似していた（それぞれ208および161 μ M）。

【0132】

さらなる生化学的特徴付けは、DPRP-1およびDPRP-2がDPPIVに類似したプロフィールを有していることを示している。2種類の精製プロテアーゼおよびDPPIVを室温において30分間、阻害剤と共にプレインキュベートした。次いで、基質のAla-Pro-AMC（100 μ M）を加え、蛍光強度を60分間のあいだに60個の読み取り値として記録した。不可逆的セリンプロテアーゼ阻害剤のAEBSFは、3種類すべての酵素に強い阻害を示す唯一の被験阻害剤であった（表5）。このことは、これらの酵素がセリンプロテアーゼスーパーファミリーに属するという構造的およびドメイン分析予測を支持している。

【0133】

【表5】

表5. プロテアーゼ阻害剤によるDPRP-1およびDPRP-2の阻害

阻害剤	阻害剤特性	濃度	残留活性 (対照に対する%)		
			DPRP-1	DPRP-2	DPPIV
AEBSF	セリン、不可逆	5mM	29.6	23.9	21.1
アプロチニン	セリン、可逆	5 μ g/ml	77.5	63.2	80.2
ペプスタチン	アスパラギン酸、 可逆	2 μ g/ml	97.3	95.0	93.5
DTT	システイン	2mM	100.1	94.8	98.3
B-メルカプトエタノール	システイン	100mM	93.2	84.0	98.0
EDTA	メタロ、可逆	2mM	91.5	86.0	93.5
ロイペプチン	セリン、可逆	50 μ g/ml	91.1	90.4	90.7

【0134】

Ala-Pro-AMCの他に、試験した別の基質もDPRP-1およびDPRP-2がジペプチジルペプチダーゼであることを支持した。データは、各基質（125 μ M）を酵

素と共に30分間インキュベートした後の蛍光変化を、Ala-Pro-AMCで測定した蛍光のパーセンテージとして決定することにより導かれ、試験した基質の中でGly-Pro-AMCが唯一の良好な基質であった。

【0135】

【表6】

表6. DPRP-1およびDPRP-2はジペプチジルペプチダーゼである。

基質	30分における蛍光の%変化		
	DPRP-1	DPRP-2	DPPIV
Ala-Pro-AMC	239.0	127.5	379.0
Gly-Pro-AMC	341.5	205.0	444.0
Ala-Pro-pNA	45.5	44.0	29.5
Pro-pNA	-1	-2.5	0.0
Gly-Arg-pNA	-4.5	-0.5	0.0
Lys-Ala-pNA	2.5	0.5	0.5
Ala-Phe-Pro-pNA	-4	-0.5	2.0

10

20

【0136】

DPPIVをインキュベートした場合に活性の低下を示すであろう各DPRPタンパク質のそれぞれを試験するための最適な基質を見つけるため、さらに天然および非天然アミノ酸ジ、トリおよびテトラペプチドを試験した。

【0137】

本明細書に記載の酵素アッセイ法は、DPRP酵素のペプチド阻害剤および非ペプチド阻害剤をスクリーニングするために利用することができる多くの方法の一つである。酵素活性の阻害剤を発見するために、テトラペプチド阻害剤のライブラリーを試験した。候補阻害剤は、10~20mMのDMSO原液として調製し、-20℃で保存した。アッセイ緩衝液で希釈液を調製した。阻害された酵素の蛍光変化を対照(溶媒)酵素の蛍光変化と比較することにより阻害を測定した。100-(サンプルのf1単位/対照のf1単位×100)が阻害パーセント値を示す。阻害パーセントおよび酵素が50%阻害される阻害剤濃度(IC₅₀)は、阻害パーセントを対数スケール上の阻害剤濃度に対してプロットすることによって確定された。図3に示すように、いくつかのテトラペプチドアミドが酵素活性を阻害し、データは溶媒(0.02%DMSO)のみが存在する下での活性に対する%として表す。化合物は1mMを加えた。最も興味深いのは、DPPIVに比べ、DPRP-1およびDPRP-2に対する一部のテトラペプチドの明らかな差次的活性であった。3種類すべての酵素がペプチド-1によって阻害されたが、DPRP-1およびDPRP-2のみがペプチド-4およびペプチド-5によって有意に阻害された。このことは、精製酵素の選択的阻害が可能であることを証明している。

30

40

【0138】

また、本実施例に記載のアッセイを用い、DPRP活性を阻害または増強する薬剤として、巨大分子を含む追加の合成化合物または天然に存在する化合物のライブラリーをスクリーニングすることができる。このアッセイで用いられるDPRP-1およびDPRP-2ポリペプチドは、例えば*in vitro*における翻訳、組換え発現(実施例5を参照)または生化学的手順によって入手することができる。また、本明細書に記載した以外の方法を用い、DPRP-1、DPRP-2またはDPRP-3を阻害する化合物をスクリーニングしかつ同定することができ、それらの方法には、例えばELISAおよびRIAなどの結合アッセイが含まれる。

【0139】

50

実施例 7

*In vitro*におけるヒト癌細胞の増殖に対するDPRP阻害剤の効果

DPRP-1およびDPRP-2活性のいくつかの阻害剤がヒト癌細胞の増殖に対して有していると思われる効果を評価する試みにおいて、LNCap、PC3およびDu145、マウス精巣系MLTC-1およびMDA-MB231乳癌細胞を96ウエルの組織培養プレート中で平板培養し(ウエル当たり 10^4 個)、CO₂インキュベータ中37において24時間培養して付着させた。次いで、24時間から96時間までの様々なインキュベーション時間に様々な希釈(最終希釈:0.1nM~10μM)で化合物をウエルに加え、毎日新しい化合物を置き換えた。希釈DMSOのみの添加を対照とした。これらの化合物と3回ずつ繰り返しインキュベートした後、XTT細胞増殖アッセイ(Roché 1-465-015)を用いて細胞の増殖を測定した。XTTミックスを加えて5時間後に490および650nmにおいてプレートを読み取った。3種類の阻害剤について0.1、1、10および100×IC₅₀に等しい濃度において細胞増殖の増加を観察したが、PC3細胞についての結果を図4A、4Bおよび4Cに示す。

10

【0140】

全体的に見て、mRNA増幅、ウエスタンブロットイングおよび免疫組織化学によって明らかにされたように、DPRPは多種多様な組織中で発現される。ノーザンブロットおよびウエスタンブロットによれば、DPRP-1は精巣中に最も多かった。また、DPRP-1と相同の精巣cDNA源からの多数の発現配列タグ(EST)も精巣における多量のDPRP-1の発現を支持している。実施例4は、特異的DPRP-1抗体を用いるヒト精巣におけるDPRP-1タンパク質の免疫組織化学的局在化について記載している。DPRP-1は類上皮(epithelioid)ライディッヒ細胞において強く発現され、ライディッヒ細胞は、哺乳類の雄における精巣アンドロゲン(雄のステロイドホルモン)の主要な源である。精巣の間質では、ライディッヒ細胞およびマクロファージは、マクロファージ表面上に延びるライディッヒ細胞突起の「指状組織(digitation)」と密接に結合している。ライディッヒ細胞に極めて近い多核細胞もDPRP-1抗体で染色され、このプロテアーゼがマクロファージにおいても発現され、精巣中のマクロファージがライディッヒ細胞の傍分泌調節において重要な役割を果たしていることを示唆している。精巣のマクロファージによって分泌されるサイトカインは、ライディッヒ細胞に対して分裂促進的であり、間葉性前駆細胞が成熟ライディッヒ細胞に分化する際に重要な役割を果たす。精巣の成熟に関与するタンパク質および経路をより明確に理解することは、性早熟症の新たな治療の発見に重要である。さらに、ライディッヒ細胞は、性ステロイド産生(主にテストステロン)を介して性索間質腫瘍などの腫瘍を引き起こす。テストステロンは、乳癌および子宮癌、卵巣癌および男性ホルモン性脱毛症(ヘアロス)などのいくつかの腫瘍形成および疾患と関係している。テストステロンおよびその他の男性ホルモンを産生する体内のその他の腺(例えば、副腎)におけるDPRPタンパク質の局在化についてのさらなる調査を現在検討中である。DPRP-1とステロイドおよびポリペプチドホルモン生合成経路機能との可能な関係を検討しており、実施例7は、*in vitro*の細胞モデルにおける前立腺、精巣および乳房におけるDPRPタンパク質の役割を理解することと関連している。

20

30

40

【0141】

免疫組織化学的分析は、子宮内の子宮内膜腺(実施例4を参照)、膵臓腺房、腎臓の糸球体、膀胱内の形質細胞、扁桃腺内のB細胞の一部、前立腺の円柱上皮細胞および分化が不十分な前立腺扁平上皮化生、Gleasonグレード4の前立腺癌、および良性前立腺肥大における過形成腺にDPRP-1が局在している。乳癌、ならびに精上皮腫および前立腺扁平上皮化生における陽性染色は、ホルモン感受性組織、特に分化が不十分な細胞とDPRP-1との一般的関係を示唆している。特殊化した上皮細胞および炎症性形質細胞(リンパ球)におけるDPRP-1の存在にも興味がある。炎症性乳癌では多数の浸潤リンパ球が見られ、全般的予後も不良である。DPRP-1およびその他のDPRPタンパク質は、通常、腫瘍の周辺に一定して浸潤しているリンパ形質細胞成分を有する髄様癌に出

50

現し、このことは新生物に対する宿主組織の反応を表すと考えられている。大部分のリンパ球はT細胞であり、大部分の形質細胞は、IgG産生型である。いくつかの抗原が、B細胞、乳癌細胞のサブグループ、およびその他の上皮癌細胞上に多くあり、これらの抗原は、新しい種類の治療用モノクローナル抗体にとっての標的であり、B細胞特異的抗原CD20に対するヒト化モノクローナル抗体でいくつかの注目すべき成功が実現している。したがって、DPRPタンパク質に対するモノクローナル抗体は、癌を含む、DPRPタンパク質が関与する疾患を診断しかつ治療するのに有用であると思われる。

【0142】

多くの組織の特殊化された上皮細胞におけるDPRP-1の発現は、DPRP-1およびその他のDPRPタンパク質が細胞の成長およびそれらの分化に関与していることを示唆している。前立腺癌および精巣癌の*in vitro*モデルにおいて実施例6に記載の阻害剤を用いる試験は(実施例7)、図4A~4Cに示すように、DPRP-1/DPRP-2阻害剤がnM濃度でPC3細胞の増殖に50~60%の増加を引き起こしたことを示した。

10

【0143】

本発明をその好ましい実施形態にしたがって説明してきたが、これらの実施形態は発明者らが現在知りうる最良の方法を構成するもので、当業者にとって明らかのように、本明細書に添付する特許請求の範囲に示される範囲を逸脱することなく変更および修正を行えることは理解されるであろう。例えば、本開示は、ある場合にはDPRP-1およびDPRP-2を中心に扱っているが、DPRP-3およびその断片もそれらをコードする核酸であることから、同様に有用であると見なされる。本発明の具体的特徴は、特許請求の範囲で強調されている。

20

【図面の簡単な説明】

【図1A】

DPRP-1、DPRP-2、DPRP-3およびDPPIVを一緒に線状に配置した整列(alignment)を示し、陰影は、特定の位置における同一(黒色)または類似(灰色)のアミノ酸残基を示すために付けた図である。

【図1B】

DPRP-1、DPRP-2、DPRP-3およびDPPIVを一緒に線状に配置した整列(alignment)を示し、陰影は、特定の位置における同一(黒色)または類似(灰色)のアミノ酸残基を示すために付けた図である。

30

【図2】

図1と同じようにヒトおよびマウスDPRP-2を一緒に線状に配置した整列を示す図である。

【図3】

様々なテトラペプチドアミド阻害剤のジペプチジルペプチダーゼ酵素活性に対する効果を示すグラフである。

【図4A】

3種類の阻害剤化合物の様々な用量におけるPC3前立腺癌細胞系の増殖に対する効果を示すグラフである。

40

【図4B】

3種類の阻害剤化合物の様々な用量におけるPC3前立腺癌細胞系の増殖に対する効果を示すグラフである。

【図4C】

3種類の阻害剤化合物の様々な用量におけるPC3前立腺癌細胞系の増殖に対する効果を示すグラフである。

【 図 1 A 】

```

DPP4 1 .....
DPRP1 1 .....
DPRP2 1 .....
DPRP3 1 .....

DPP4 20 .....
DPRP1 61 .....
DPRP2 50 .....
DPRP3 48 .....

DPP4 75 .....
DPRP1 121 .....
DPRP2 110 .....
DPRP3 101 .....

DPP4 130 .....
DPRP1 179 .....
DPRP2 168 .....
DPRP3 155 .....

DPP4 187 .....
DPRP1 235 .....
DPRP2 224 .....
DPRP3 215 .....

DPP4 235 .....
DPRP1 294 .....
DPRP2 284 .....
DPRP3 263 .....

DPP4 293 .....
DPRP1 354 .....
DPRP2 344 .....
DPRP3 314 .....

DPP4 322 .....
DPRP1 414 .....
DPRP2 404 .....
DPRP3 343 .....

```

【 図 1 B 】

```

DPP4 379 E.....
DPRP1 473 .....
DPRP2 464 .....
DPRP3 394 .....

DPP4 424 .....
DPRP1 529 .....
DPRP2 520 .....
DPRP3 445 .....

DPP4 484 .....
DPRP1 588 .....
DPRP2 579 .....
DPRP3 505 .....

DPP4 542 .....
DPRP1 648 .....
DPRP2 639 .....
DPRP3 563 .....

DPP4 600 .....
DPRP1 708 .....
DPRP2 699 .....
DPRP3 621 .....

DPP4 659 .....
DPRP1 768 .....
DPRP2 759 .....
DPRP3 680 .....

DPP4 719 .....
DPRP1 828 .....
DPRP2 819 .....
DPRP3 738 .....

```

【 図 2 】

```

hDPRP1 1 MAAMETEQGLVETETACBENLDSODRPKLEPFTVVERYSNOLKLLADTRKYHGV
mDPRP1 1 MAAMETEQGLVETETACBENLDSODRPKLEPFTVVERYSNOLKLLADTRKYHGV

hDPRP1 60 KAKRPIIDDFVYKIDDFEHSDDYHLSMSENRRENTLYSEIERTINRAAVLHLSKPL
mDPRP1 61 KAKRPHDFVAKIDDFEHSDDYHLSMSENRRENTLYSEIERTINRAAVLHLSKPL

hDPRP1 120 LQLPQATLDYGYRREELLRRERRIGTVGTPADYHNSGTFLDQAGSGIYHNDGGPO
mDPRP1 121 LQLPQATLDYGYRREELLRRERRIGTVGTPADYHNSGTFLDQAGSGIYHNDGGPR

hDPRP1 180 SFTQPLRPNLVTSCPNIRMDPKLCPADPDWIAFHSNDWISNRYTREERRVYVHNE
mDPRP1 181 SFTQPLRPNLVTSCPNIRMDPKLCPADPDWIAFHSNDWISNRYTREERRVYVHNE

hDPRP1 240 LANMEDIARSAGVATVLEQEEEDRYSGYVWCFAEFTPSGGKILRILVENDDESEVETIH
mDPRP1 241 LANMEDIARSAGVATVLEQEEEDRYSGYVWCFAEFTPSGGKILRILVENDDESEVETIH

hDPRP1 300 VTSPLLETTRRADSPYKPTGTANPKVTFKMSERLDAEAGTLDVLDKELDPPEFLLEGV
mDPRP1 301 VTSPLLETTRRADSPYKPTGTANPKVTFKMSERLDAEAGTLDVLDKELDPPEFLLEGV

hDPRP1 360 SYTARAGWTPGCRHWSILLDSQSTLOIVLISPELETVEDEAARORLIESVPSDVTB
mDPRP1 361 SYTARAGWTPGCRHWSILLDSQSTLOIVLISPELETVEDEAARORLIESVPSDVTB

hDPRP1 420 LIYGETTQIWIINDIHFVFEKHEHLEIFEFASECKTGRHLRYKITSILAESNYKRS
mDPRP1 421 LIYGETTQIWIINDIHFVFEKHEHLEIFEFASECKTGRHLRYKITSILAESNYKRS

hDPRP1 480 SGLPAPSDFKCPKLEIHTSGENFVLCRHSMLVDEYVGLVYFTGTRDSPLEHHLVYV
mDPRP1 481 SGLPAPSDFKCPKLEIHTSGENFVLCRHSMLVDEYVGLVYFTGTRDSPLEHHLVYV

hDPRP1 540 SYVHGEVRLTDRGYSNCSQSHCDFRISKYSNQRNHCYSIYATSFEDDFCTRE
mDPRP1 541 SYVHGEVRLTDRGYSNCSQSHCDFRISKYSNQRNHCYSIYATSFEDDFCTRE

hDPRP1 600 EWATLLDSAGLPDYTPPEIISPESTTGFYGLYKPHDLQPKRYPTVLIYGGPQVQ
mDPRP1 601 EWATLLDSAGLPDYTPPEIISPESTTGFYGLYKPHDLQPKRYPTVLIYGGPQVQ

hDPRP1 660 LWNRRFKGVYFRLNPLASLGVVVVIDNRRSGRCHLRFCAKFKYMGQIFIDQVBECL
mDPRP1 661 LWNRRFKGVYFRLNPLASLGVVVVIDNRRSGRCHLRFCAKFKYMGQIFIDQVBECL

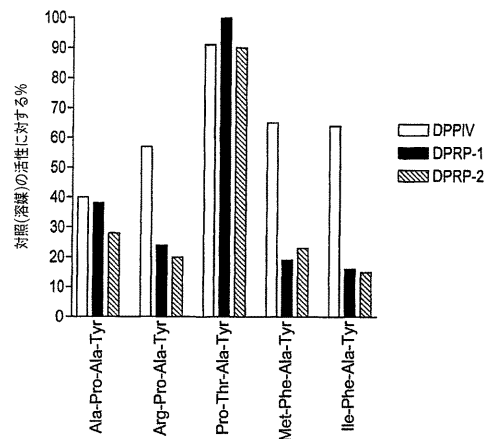
hDPRP1 720 VLASRDFDLDNRVGHGRSYKWTSLMAMQKSDIHWALACAPVTHIYDTGYTERY
mDPRP1 721 VLASRDFDLDNRVGHGRSYKWTSLMAMQKSDIHWALACAPVTHIYDTGYTERY

hDPRP1 780 MHPDQNEQGYLYGSVAQAKRFPSEPRLLHIGFLDENVHFAHTSILLSFLVRAKPY
mDPRP1 781 MHPDQNEQGYLYGSVAQAKRFPSEPRLLHIGFLDENVHFAHTSILLSFLVRAKPY

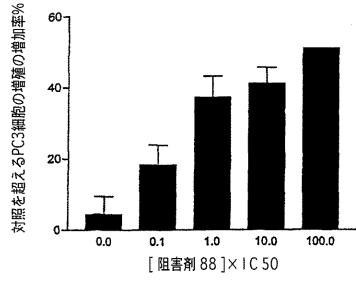
hDPRP1 840 DLQIYPOBRHSIRVPSGPHYLHLLHYLQENLGSRIAALKVI
mDPRP1 841 DLQIYPOBRHSIRVPSGPHYLHLLHYLQENLGSRIAALKVI

```

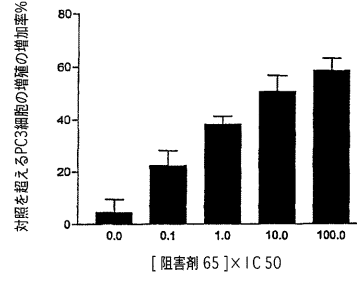
【 図 3 】



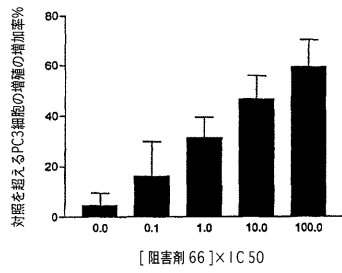
【 図 4 A 】



【 図 4 B 】



【 図 4 C 】



【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
18 April 2002 (18.04.2002)

PCT

(10) International Publication Number
WO 02/31134 A2

- (51) International Patent Classification: C12N 9/64 (74) Agents: SAMPLES, Kenneth, H. et al.; Fitch, Even, Tabin & Flannery, 120 South LaSalle Street, Suite 1600, Chicago, IL 60603 (US).
- (21) International Application Number: PCT/US01/31874
- (22) International Filing Date: 12 October 2001 (12.10.2001) (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/240,117 12 October 2000 (12.10.2000) US (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, 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, TR), OAPI patent (BF, BI, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant (for all designated States except US): FERRING BV [US/US]; Polaris Avenue 144, NL-2132 JX Hoofddorp (NL).
- (72) Inventors; and (75) Inventors/Applicants (for US only): QI, Steve [GB/US]; 11034 West Ocean Air Drive, Apt. 24B, San Diego, CA 92130 (US); AKINSANYA, Karen, O. [GB/US]; 7675 Palmsilla Drive, Apt. 642B, San Diego, CA 92122-5094 (US); RIVIERE, Pierre, J.-M. [FR/US]; 3993 Via Cangrejo, San Diego, CA 92130 (US); JUNIEN, Jean-Louis [FR/FR]; Ferring SAS, F-75507 Paris (FR).
- Published: — without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 02/31134 A2

(54) Title: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

(57) Abstract: Novel proteins or polypeptides having significant sequence homology to DPPIV, nucleic acids coding therefor, cells which have been modified with such nucleic acid so as to express these proteins, antibodies to these proteins, screening methods for the discovery of new therapeutic agents which are inhibitors of the activity of these proteins or of related proteins, and therapeutic agents discovered by such screening methods, as well as new therapeutic treatments, are all provided.

WO 02/31134

PCT/US01/31874

Novel Serine Protease Genes Related to DPPiV**Field of the Invention**

The present invention relates to novel serine proteases related to dipeptidyl peptidase IV (DPPiV), and to isolated nucleic acids coding for these proteases, all of which are useful for the discovery of new therapeutic agents, for measuring protease activity, and for determining the inhibitory activity of compounds against these proteases.

Background of the Invention

Proteases and peptidases are enzymes that catalyse the hydrolysis of peptidic amide bonds. Proteases play an important role in the regulation of biological processes in almost every life-form from bacteria to virus to mammals. They perform critical functions in, for example, digestion, blood clotting, apoptosis, activation of immune responses, zymogen activation, viral maturation, protein secretion and protein trafficking. They can be classified according to a number of criteria, such as site of action, substrate preference, and mechanism. So, for example, aminopeptidases act preferentially at the N-terminal residues of a peptide, while carboxypeptidases act preferentially at the C-terminus and endopeptidases act at sites removed from the two termini. Among the carboxy- and aminopeptidases, peptidyl peptidases cleave a single amino acid residue from the substrate, dipeptidyl peptidases cleave a dipeptide unit (two amino acids) from the substrate, and tripeptidases cleave three amino acids from the substrate. Substrate preference is frequently expressed in terms of the amino acid residue immediately N-terminal to the cleavage site. For example, trypsin-like peptidases will preferentially cleave a peptide next to a basic amino acid (arginine or lysine), i.e. where the bond hydrolysed is the Arg/Lys-Xaa bond. As another example, the chymotrypsin-like family of peptidases preferentially hydrolyse peptides adjacent to an aromatic residue. Mechanistically, peptidases are classified as being serine-dependent, cysteine-dependent, aspartic acid-dependent or zinc-dependent.

Because peptidases and proteases are involved in the regulation of many physiological processes, they are attractive targets for the development of therapeutic agents. Protease and peptidase inhibitors are, for example, used in the treatment of hypertension, coagulation disorders, and viral infection.

Proteolytic enzymes that exploit serine in their catalytic activity are ubiquitous, being found in viruses, bacteria and eukaryotes. Over 20 families (denoted S1 - S27) of serine protease have been identified; these are grouped into 6 clans (SA, SB, SC, SE, SF

WO 02/31134

PCT/US01/31874

and SG) on the basis of structural similarity and other functional evidence. Structures are known for four of the clans (SA, SB, SC and SE); these appear to be totally unrelated, suggesting at least four evolutionary origins of serine peptidases and possibly many more, Rawlings and Barrett, *Meth. Enzymol.*, 244: 19-61 (1994).

5 The prolyl oligopeptidase family consists of a number of evolutionarily related peptidases whose catalytic activity seems to be provided by a charge relay system similar to that of the trypsin family of serine proteases, but which evolved by independent convergent evolution. A conserved serine residue has been shown experimentally (in *E. coli* protease II as well as in pig and bacterial PE) to be necessary for the catalytic
10 mechanism. This serine, which is part of the catalytic triad (Ser, His, Asp), is generally located about 150 residues away from the C-terminal extremity of these enzymes (which are all proteins that contains about 700 to 800 amino acids).

One of the most intensively studied prolyl oligopeptidases is dipeptidyl peptidase IV (DPPIV, EC 3.4.14.5), a type II glycoprotein, which is the only well characterised
15 dipeptidyl aminopeptidase known to be located on the outer side of plasma membranes. As indicated above, dipeptidyl aminopeptidases are characterised by their ability to cleave N-terminal dipeptides from a variety of small peptides. Dipeptidyl aminopeptidases show different substrate specificities and cellular localisation, suggesting different functions of each activity in peptide processing. DPPIV is characterised by its capacity to cleave N-
20 terminal dipeptides containing proline or alanine as the penultimate residue. The DPPIV gene spans approximately 70 kb and contains 26 exons, ranging in size from 45 bp to 1.4 kb. The nucleotide sequence (3,465 bp) of the cDNA contains an open reading frame encoding a polypeptide comprising 766 amino acids. The nucleotides that encode the active site sequence (G-W-S-Y-G) are split between 2 exons. This clearly distinguishes
25 the genomic organisation of the prolyl oligopeptidase family from that of the classic serine protease family.

DPPIV is widely distributed in mammalian tissues and is found in great abundance in the kidney, intestinal epithelium and placenta (Yaron, A. and Naider, F., Critical Reviews in *Biochem. Mol. Biol.*, 1993 [1], 31). In the human immune system,
30 the enzyme is expressed almost exclusively by activated T-lymphocytes of the CD4⁺ type where the enzyme has been shown to be synonymous with the cell-surface antigen CD26. Although the exact role of DP-IV in human physiology is still not completely understood, recent research has shown that the enzyme clearly has a major role in human physiology and pathophysiology.

WO 02/31134

PCT/US01/31874

On human T cells, DPPIV expression appears late in thymic differentiation and is preferentially restricted to the CD4⁺ helper/memory population, and CD26 can deliver a potent co-stimulatory T-cell activation signal. DPPIV, also known as T-cell activation antigen CD26, therefore plays an important role in the immune response via association with CD45 tyrosine phosphatase and, through its ability to bind adenosine deaminase (ADA) to the T-cell surface, protects the T-cell from adenosine-mediated inhibition of proliferation. Furthermore, the regulation of the function of chemokines by CD26/DPPIV appears to be essential for lymphocyte trafficking and infectivity of HIV strains. DPPIV has been associated with numerous functions including involvement in T-cell activation, cell adhesion, digestion of proline containing peptides in the kidney and intestines, HIV infection and apoptosis, and regulation of tumorigenicity in certain melanoma cells, Pethiyagoda et al., *Clin. Exp. Metastasis* 2000;18(5):391-400. DPPIV is also implicated in the endocrine regulation and metabolic physiology. More particularly, DPPIV cleaves the amino-terminal His-Ala dipeptide of GLP-1, generating a GLP-1 receptor antagonist, and thereby shortens the physiological response to GLP-1. Glucagon-like peptide-1 (GLP-1), an incretin that induces glucose-dependent insulin secretion, is rapidly degraded by DPPIV, and since the half-life for DPPIV cleavage is much shorter than the half-life for removal of GLP-1 from circulation, a significant increase in GLP-1 bioactivity (5- to 10- fold) is anticipated from DPP-IV inhibition. Inhibitors of DPPIV are currently being studied in the clinic as potential therapeutic agents for type 2 diabetes and impaired glucose tolerance.

Various different inhibitors of DPPIV were known in 1993. One of these is a suicide inhibitor N-Ala-Pro-O-(nitrobenzoyl-) hydroxylamine. Another is a competitive inhibitor: e-(4-nitro) benzoxycarbonyl-Lys-Pro, and another is a polyclonal rabbit anti-porcine kidney DPPIV immunoglobulin. Others have since been developed and are described in detail in U.S. Patents Nos. 5,939,560, 6,110,949m 6,011,155 and 5,462,928.

In addition to, but independent of, its serine type catalytic activity, DPPIV binds closely to the soluble extracellular enzyme adenosine deaminase (ADA), acting as a receptor and is thought to mediate signal transduction. DPPIV structure is characterized by two extracellular domains, an α/β fold hydrolase domain and a 7-blade beta-propeller domain consisting of repeated beta sheets of about 50 amino acids. Recently it has been shown that, besides selecting substrates by size, the beta-propeller domain, containing 10 of the 12 highly conserved cysteine residues, contributes to catalysis of the peptidase domain. In addition, the cysteine-rich domain is responsible for DPPIV-binding to collagen I and to extracellular ADA. DPPIV is also reported to play a role in fibronectin-

WO 02/31134

PCT/US01/31874

mediated interactions of cells with extracellular matrix. Recent studies show that the protease activity of DPPIV is not required for its anti-invasive activity because mutants of DPPIV that lack the extracellular serine protease activity maintain such activity.

5 A number of proteins that share similarities with DPPIV have been reported in the literature. Several of these proteins have been cloned including DPP-I, DPP-II, DPP-III, DPP-X and fibroblast activation protein (FAP). These have been identified and characterised either by molecular cloning and functional studies of expressed proteins or as biochemical activities in tissue extracts. DPPIV-beta and other novel peptidases with functional similarities to DPPIV are not yet cloned. The identification, characterization
10 and/or appropriate classification of further members of the family of prolyl oligopeptidases, the elucidation of their physiological (and particularly pathophysiological) role, and the application of that knowledge to the development of new therapeutic agents are significant challenges.

Summary of the Invention

15 The present invention provides proteins with prolyl oligopeptidase (post-proline cleaving) activities that constitute three novel members of a family of proteins related to DPPIV, including the full-length proteins, alternative splice forms, subunits, and mutants, as well as nucleotide sequences encoding the same. The present invention also provides methods of screening for substrates, interacting proteins, agonists, antagonists
20 or inhibitors of the above proteins, and furthermore to pharmaceutical compositions comprising the proteins and/or mutants, derivatives and/or analogues thereof and/or ligands thereto.

These novel proteins having significant sequence homology to DPPIV are termed dipeptidyl peptidase IV-related protein-1, 2 & 3 (DPRP-1, DPRP-2 and DPRP-3). The
25 amino acid sequences of DPRP-1, DPRP-2 and DPRP-3 are given in SEQ. ID NOS:1, 3 and 5 respectively. Further disclosed are nucleic acid sequences coding for these proteins (SEQ. ID NOS:2, 4 and 6). Table 1 illustrates the homology (i.e. similarity) between the novel proteins DPRP-1, DPRP-2 and DPRP-3 and other known serine proteases.

WO 02/31134

PCT/US01/31874

Table 1 – Comparison of the sequences of these three novel proteins with DPPIV and other Clan SC, Family S9 members and Subfamily B members

	Protease Family	Protease name	No. of a.a.	Homology with DPPIV	TM region	Ser-Asp-His Triad	Gene location	Optimal pH
5	Clan CA, Family C1	DPPI	463	N	N	N	11q14.1-q14.3	-
		DPPH	500	N	Y	N	-	4.5-6.0
	Clan SC, Family S28	QPP	492	N	N	N	-	4.5-7.5
		PCP	496	N	N	N	-	-
	Unassigned	DPPHII	737	N	N	N	-	-
10	Clan SC, Family S9, Subfamily B	DPPIV	766	100	Y	Y	2q24.3	7.5-8.0
		DPPVI	865	52	Y	Mutation	7	-
		FAP	760	70	Y	Y	2q23	7.5-8.0
		DPRP-1	882	41	N	Y	15q22.1-15q22.2	7.5-8.0
		DPRP-2	864	39	N	Y	19p13.3	7.5-8.0
		DPRP-3	798	54	Y	Mutation	2q12.3-2q14.1	-

The greatest homology between DPRP-1, DPRP-2 and DPPIV is seen in the C-terminal sequences. On the basis of sequence homology with DPPIV (see Figure 1), one might predict that these DPRP proteins would have functions that include, but are not limited to, roles as enzymes. Cloning, expression, biochemical and molecular characterization have confirmed this hypothesis.

The expression pattern of DPRPs and the localization to specialized epithelial cells and plasma cells (Leydig cells, prostate epithelial cells, lymphocytes, B cells) is consistent with a role in differentiation, proliferation and inflammation. The localization of the DPRP-1 gene in hormone sensitive cancers (breast, prostate, testicular), tissues regulated by testosterone and the abundant expression in poorly differentiated cancers, demonstrate that DPRP-activating or inhibiting molecules will have numerous therapeutic applications in the treatment of disorders characterized by dysregulated growth, differentiation and steroid or polypeptide hormone synthesis and degradation. Data disclosed herein supports the hypothesis that DPRP-1 and DPRP-2 are involved in the regulation of proliferation of *in vitro* models of prostate and testis cancer well known to those skilled in the art.

DPRP-1 and DPRP-2 activities described herein and their expression patterns are compatible with their having functional roles as physiological regulators of the immune and neuroendocrine systems through the enzymatic modification of biochemical mediators like peptides and chemokines. The numerous functions previously described

WO 02/31134

PCT/US01/31874

for DPPIV based upon the use of inhibitors may be due in part to its action and that of similar proteins, like the DPRPs. Therefore, the discovery of selective and potent inhibitors of DPPIV, of the DPRPs and of other related proteases like FAP is considered central to achieving effective and safe pharmaceutical use of these and any newly identified serine protease inhibitors, as well as other active compounds that modify the function(s) of such proteins.

The invention thus provides novel proteins or polypeptides, the nucleic acids coding therefor, cells which have been modified with the nucleic acid so as to express these proteins, antibodies to these proteins, a screening method for the discovery of new therapeutic agents which are inhibitors of the activity of these proteins (or which are inhibitors of DPPIV and not of the proteins), and therapeutic agents discovered by such screening methods. The novel proteins and the nucleic acids coding therefor can be used to discover new therapeutic agents for the treatment of certain diseases, such as for example, reproductive, inflammatory and metabolic disorders and also in the preparation of antibodies with therapeutic or diagnostic value.

In accordance with one aspect of the present invention, there are provided novel, mature, biologically active proteins, principally of human origin. Such proteins may be isolated in small quantities from suitable animal (including human) tissue or biological fluids by standard techniques; however, larger quantities are more conveniently prepared in cultures of cells genetically modified so as to express the protein.

In accordance with another aspect of the present invention, there are provided isolated nucleic acid molecules encoding polypeptides of the present invention including mRNAs, DNAs, cDNAs, genomic DNAs thereof.

In accordance with a further aspect of the present invention, nucleic acid probes are also provided comprising nucleic acid molecules of sufficient length to specifically hybridize to a nucleic acid sequence of the present invention.

In accordance with a still further aspect of the present invention, processes utilizing recombinant techniques are provided for producing such polypeptides useful for *in vitro* scientific research, for example, synthesis of DNA and manufacture of DNA vectors. Processes for producing such polypeptides include culturing recombinant prokaryotic and/or eukaryotic host cells that have been transfected with DNA vectors containing a nucleic acid sequence encoding such a polypeptide and/or the mature protein under conditions promoting expression of such protein and subsequent recovery of such protein or a fragment of the expressed product.

WO 02/31134

PCT/US01/31874

In accordance with still another aspect, the invention provides methods for using DPRP polypeptides and polynucleotides, including the treatment of infections, such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2, pain, diabetes, precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers, asthma, allergies, benign prostatic hypertrophy, cancers including hormone-sensitive and androgen-independent cancers, migraines, vomiting, psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, depression, dementia, and severe mental retardation, and dyskinesias, hereinafter collectively referred to as "the Diseases".

In accordance with yet another aspect of the present invention, there is provided a process for utilizing such polypeptides, or polynucleotides encoding such polypeptides, for the discovery of compounds that inhibit the biological activity of the mature proteins thereof, e.g. by cleaving an N-terminal dipeptide, and such inhibitors are thus also provided.

In accordance with a more specific aspect, the invention provides isolated nucleic acid which encodes (a) a polypeptide which includes the amino acid sequence of one of SEQ ID NOS:1, 3 and 5, or (b) a polypeptide having an amino acid sequence that is at least about 70% similar thereto and exhibits the same biological function, or which is an alternative splice variant of one of SEQ ID NOS:2, 4 and 6, or which is a probe comprising at least 14 contiguous nucleotides from said nucleic acid encoding (a) or (b), or which is complementary to any one of the foregoing.

In accordance with another specific aspect, the invention provides a polypeptide which may be optionally glycosylated, and which (a) has the amino acid sequence of a mature protein set forth in any one of SEQ ID NOS:1, 3 and 5; (b) has the amino acid sequence of a mature protein having at least about 70% similarity to one of the mature proteins of (a) and which exhibits the same biological function; (c) has the amino acid sequence of a mature protein having at least about 90% identity with a mature protein of any of SEQ ID NOS:1, 3 and 5; or (d) is an immunologically reactive fragment of (a).

In accordance with still another specific aspect, the invention provides a method for the screening for a compound capable of inhibiting the enzymatic activity of at least one mature protein of the invention, which method comprises incubating said mature protein and a suitable substrate for said mature protein in the presence of one or more test compounds or salts thereof, measuring the enzymatic activity of said mature protein, comparing said activity with comparable activity determined in the absence of a test

WO 02/31134

PCT/US01/31874

compound, and selecting the test compound or compounds that reduce the enzymatic activity, and it also provides a method for screening for a compound capable of inhibiting the enzymatic activity of DPPIV that does not inhibit the enzymatic activity of at least one mature protein and a suitable substrate in the presence of one or more inhibitors of DPPIV or salts thereof, measuring the enzymatic activity of said mature protein, comparing said activity with comparable activity determined in the absence of the DPPIV inhibitor, and selecting a compound that does not reduce the enzymatic activity of said mature protein.

These and other aspects of the present invention should be apparent to those skilled in the art from the detailed description which follows.

Brief Description of the Drawings

FIGS. 1A and 1B show the co-linear alignment of DPRP-1, DPRP-2, DPRP-3 and DPPIV, with shading being supplied to indicate the same (black) or similar (gray) amino acid residues at a particular location.

FIG. 2 is similar to FIG. 1 and shows co-linear alignment of human and mouse DPRP-2.

FIG. 3 is a graph which shows the effects of various tetrapeptide amide inhibitors on dipeptidyl peptidase enzyme activity.

FIGS. 4A-4C show the effects of three inhibitor compounds on the proliferation of PC3 prostate cancer cell lines at various doses.

Detailed Description of the Preferred Embodiments

In accordance with an aspect of the present invention, there are provided isolated nucleic acid sequences (polynucleotides), which encode the mature polypeptides having the deduced amino acid sequences of the three DPRP's (SEQ ID NOS:1, 3 and 5).

The polynucleotides of this invention were discovered using a human testis cDNA library (DPRP-1), a human colon library (DPRP-2) and a human hypothalamus cDNA library (DPRP-3). Isolated nucleic acid for DPRP-1 contains an open reading frame encoding a protein of approximately 882 amino acids in length which is structurally related to human DPPIV, showing 26% identity, and 41% similarity over the entire human DPPIV protein sequence. Isolated nucleic acid for DPRP-2 contains an open reading frame encoding for a protein of approximately 864 amino acids, which is 39% similar to the entire DPPIV amino acid sequence. Analysis of DPRP-1 and DPRP-2 primary amino acid sequence using hydrophobicity plots predicts that these two proteins do not have a transmembrane domain. Despite this fact, it is possible that these intracellular serine proteases are secreted upon cellular activation. Quiescent cell proline

WO 02/31134

PCT/US01/31874

dipeptidase (QPP) is a serine protease that is targeted to intracellular vesicles that are distinct from lysosomes (Chiravuri M, et al., J. Immunol. 2000 Nov 15;165(10):5695-702). This hypothesis expands the potential site(s) and scope of DPRP-1 and DPRP-2 involvement in mechanisms for post-translational regulation of chemokines, cytokines, peptides and polypeptides. The full length DPRP-3 sequence contains 796 amino acids, a signal peptide from 1 to 48, and a transmembrane domain between 34 and 56. The mature protein is predicted to be a type II membrane protein and may be cleaved to produce a soluble form. The amino acid sequence is set forth in SEQ ID NO:5, which was deduced from SEQ ID NO:6 and has 54% similarity with DPPIV.

10 Amino acid sequence alignments of these polypeptides with members of the prolyl oligopeptidase enzyme subfamily S9B show that all three DPRP proteins have overall sequence and structural homology to DPPIV and FAP. DPRPs are predicted to be members of the enzyme Clan SC (Serine nucleophile) with catalytic residues in the order Ser, Asp, His and the active site sequence (G-W-S-Y-G).

15 **Table 2.** Homology (i.e. similarity) between DPRP-1, DPRP-2, DPRP-3 and members of the prolyl oligopeptidase family S9B enzymes.

DPPIV					
41	DPRP-1				
39	74	DPRP-2			
54	39	40	DPRP-3		
70	41	39	52	FAP	
52	40	42	68	54	DPPVI

DPRP-1, DPRP-2 and DPRP-3 do not exhibit sequence similarity with any members of the classical serine protease families, chymotrypsin and subtilisin. The order of the catalytic triad residues is different in the three main related SC clan families: His-Asp-Ser in chymotrypsin, Asp-His-Ser in subtilisin and Ser-Asp-His in the prolyl oligopeptidases.

As shown in Table 2, DPRP-3 has the highest homology with DPPVI (68% homology and 51% identity). Wada et al isolated cDNA clones for DPPVI, a DPPIV-related protein, from bovine, rat (Wada et al., *Proc. Nat. Acad. Sci.* 89: 197-201. (1992)) and human (Yokotani et al., *Hum. Molec. Genet.* 2:1037-1039 (1993)) brain libraries. They demonstrated that, unlike DPPIV, the catalytic triad in DPPVI does not have the first serine residue. In DPRP-3 two of the amino acids in the catalytic triad

WO 02/31134

PCT/US01/31874

characteristic of the serine protease family are conserved. However, the serine residue itself is replaced by glycine. While the absence of the serine residue is likely to prevent protease activity at this site, it is possible that multiple other functions mediated by other functional domains of the protein remain intact.

5 As briefly described above, DPPIV is a multifunctional molecule that exerts important functions depending on the expressed cells and tissues, in addition to its catalytic activity as a peptidase. DPRP-3 and DPPVI are also likely to maintain multiple functions despite the absence of an intact catalytic triad. For example, DPPVI has been implicated in the regulation of neuronal plasticity. DPPVI is highly expressed in the
10 hippocampus, thalamus, hypothalamus and striatum. In addition, developmental arrest and embryonic lethality of *rump white Rw/Rw* embryos is thought to be due to disruption of the DPPIV gene. *Rw* mutation is associated with a chromosomal inversion spanning 30 cM of the proximal portion of mouse chromosome 5. Genomic analysis of the DPPVI gene on the *Rw* chromosome places the inversion breakpoint in the coding region
15 resulting in loss of a significant fraction of the C-terminal region, Hough R.B. et al., Proc. Nat. Acad. Sci., *95*, 13800-13805 (1998).

The human DPRP-1 gene, predicted to be 32668bp in length, has at least 22 exons and eight transcripts. It maps to chromosome 15 (NT_010265) at position 15q21.1 – 15q22.1. The lengths of predicted alternative splice variant transcripts vary
20 between 602bp and 4523bp (see SEQ ID NOS: 7-22). This is in agreement with the multiple transcripts observed by Northern blot analysis (See Example 2). ESTs representing the transcripts were found in numerous tissues including senescent fibroblasts, T-lymphocytes, germinal center B-cells, germ cell seminoma, testis, melanocytes, uterus, ovary breast, multiple sclerosis lesions, pancreas and placenta.

25 Human DPRP-2 belongs to a gene with at least 27 exons and nine splice variants (see SEQ ID NOS:23-40). One SNP was observed in the 3' UTR. (88% (37) C vs. 12% (5) T). The DPRP-2 gene maps to region 19p13.3 of chromosome 19. This location is host to a number of disease markers and is associated with various disorders including hypocalcemic hypercalcemia, type II cerebellar ataxia, muscular dystrophy, convulsions,
30 susceptibility to atherosclerosis, psoriasis, ectodermal dysplasia, and acute myeloid leukemia. In agreement with the ubiquitous distribution of the mRNA observed by Northern blot analysis (see Example 2), DPRP-2 was expressed in a wide variety of tissues upon examination of EST's coverage (e.g. over 64 EST's expressed in liver, spleen, muscle, melanocytes, heart, lung, placenta, skin, pancreas, stomach, brain
35 parathyroid gland).

WO 02/31134

PCT/US01/31874

Human DPRP-3 belongs to a gene with at least 23 exons and two splice variants (see SEQ ID NOS:41-44). The gene maps to chromosome 2 (NT_005445) at position 2q12.3-2q14.1. Transcripts for DPRP-3 did not show as wide a distribution as DPRP-1 and DPRP-2. As shown by Northern blot in Example 2, DPRP-3 expression is restricted
5 to brain and pancreas. ESTs representing the DPRP-3 mRNA were abundant in tissue derived from multiple sclerosis lesions, hypothalamus, whole brain and nerves, with a few transcripts being found in uterus and colon.

The relationships among human and rodent proteases in clan SC, including DPRP-1 DPRP-2 and DPRP-3, were analyzed using Neighbor Joining method (NJ), see
10 Saitou and Nei, *Mol. Biol. Evol.*, 4, 406-525 (1987). Phylogenetic analysis shows that among the S9 proteases, DPRP-1 and DPRP-2, both lacking a transmembrane domain, are distinguished from DPPIV and its closely related proteins like FAP. Similarity is shown however between DPPIV and FAP and between DPRP-3 and DPPVI, which are all type II membrane proteins.

A database search for additional DPRP-related genes revealed the presence of a
15 murine sequence related to DPRP-1. Alignment of this mouse sequence with the novel human proteases shows that the mDPRP-1 displays considerable homology with its human counterpart (FIG. 2). One skilled in the art will readily recognize that the novel mouse protease gene can be isolated using the sequence information disclosed herein and can be readily incorporated into one of the routinely used expression constructs which
20 are well known in the art. Use of this disclosed sequence by those skilled in the art to generate a transgenic mouse model will employ development of gene-targeting vectors, for example, that result in homologous recombination in mouse embryonic stem cells. The use of knockout mice in further analysis of the function of DPRP genes is a valuable
25 tool.

The polynucleotides of the present invention may be in the form of RNA or in the form of DNA; DNA should be understood to include cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded and, if single-stranded, may be the coding strand or non-coding (antisense) strand. The coding
30 sequence which encodes the mature polypeptide may be identical to the coding sequence shown in SEQ ID NOS:2, 4 and 6 respectively, or it may be a different coding sequence encoding the same mature polypeptide, as a result of the redundancy or degeneracy of the genetic code or a single nucleotide polymorphism. For example, it may also be an RNA transcript which includes the entire length of any one of SEQ ID NOS:2, 4 and 6.

WO 02/31134

PCT/US01/31874

The polynucleotides which encode the mature proteins of SEQ ID NOS:1, 3, 5, respectively, may include but are not limited to the coding sequence for the mature protein alone; the coding sequence for the mature polypeptide plus additional coding sequence, such as a leader or secretory sequence or a proprotein sequence; and the
5 coding sequence for the mature protein (and optionally additional coding sequence) plus non-coding sequence, such as introns or a non-coding sequence 5' and/or 3' of the coding sequence for the mature protein.

Thus, the term "polynucleotide encoding a polypeptide" or the term "nucleic acid encoding a polypeptide" should be understood to encompass a polynucleotide or nucleic
10 acid which includes only coding sequence for the mature protein as well as one which includes additional coding and/or non-coding sequence. The terms polynucleotides and nucleic acid are used interchangeably.

The present invention also includes polynucleotides where the coding sequence for the mature protein may be fused in the same reading frame to a polynucleotide
15 sequence which aids in expression and secretion of a polypeptide from a host cell; for example, a leader sequence which functions as a secretory sequence for controlling transport of a polypeptide from the cell may be so fused. The polypeptide having such a leader sequence is termed a preprotein or a proprotein and may have the leader sequence cleaved, by the host cell to form the mature form of the protein. These
20 polynucleotides may have a 5' extended region so that it encodes a proprotein, which is the mature protein plus additional amino acid residues at the N-terminus. The expression product having such a prosequence is termed a proprotein, which is an inactive form of the mature protein; however, once the prosequence is cleaved an active mature protein remains. Thus, for example, the polynucleotides of the present invention
25 may encode mature proteins, or proteins having a prosequence, or proteins having both a prosequence and a prosequence (leader sequence).

The polynucleotides of the present invention may also have the coding sequence fused in frame to a marker sequence which allows for purification of the polypeptides of the present invention. The marker sequence may be a polyhistidine tag, a hemagglutinin
30 (HA) tag, a c-myc tag or a V5 tag when a mammalian host, e.g. COS-1 cells, is used. The HA tag would correspond to an epitope derived from the influenza hemagglutinin protein (Wilson, I., et al., *Cell*, 37:767 (1984)), and the c-myc tag may be an epitope from human Myc protein (Evans, G.I. et al., *Mol. Cell. Biol.* 5: 3610-3616 (1985)).

The term "gene" means the segment of DNA involved in producing a polypeptide
35 chain; it includes regions preceding and following the coding region (leader and trailer)

WO 02/31134

PCT/US01/31874

as well as intervening sequences (introns) between individual coding segments (exons). The term "significant sequence homology" is intended to denote that at least 25%, preferably at least 40%, of the amino acid residues are conserved, and that, of the non-conserved residues, at least 40% are conservative substitutions.

5 Fragments of the full-length genes of the present invention may be used as a hybridization probe for a cDNA library to isolate full-length cDNA as well as to isolate other cDNAs which have significant sequence homology to the gene and will encode proteins or polypeptides having similar biological activity or function. By similar biological activity or function, for purposes of this application, is meant the ability to
10 cleave an N-terminal dipeptide having Ala or Pro as the penultimate residue or other amino acids. Such a probe of this type has at least 14 bases (at least 14 contiguous nucleotides from one of SEQ ID NOS:2, 4 or 6), preferably at least 30 bases, and such may contain, for example, 50 or more bases. Such probe may also be used to identify a
15 cDNA clone corresponding to a full-length transcript and/or a genomic clone or clones that contains the complete gene, including regulatory and promoter regions, exons, and introns. Labelled oligonucleotides having a sequence complementary to that of the gene of the present invention are useful to screen a library of human cDNA, genomic DNA or mRNA to locate members of the library to which the probe hybridizes. As an example, a known DNA sequence may be used to synthesize an oligonucleotide probe which is then
20 used in screening a library to isolate the coding region of a gene of interest.

The present invention is considered to further provide polynucleotides which hybridize to the hereinabove-described sequences wherein there is at least 70%, preferably at least 90%, and more preferably at least 95% identity or similarity between the sequences, and thus encode proteins having similar biological activity. Moreover, as
25 known in the art, there is "similarity" between two polypeptides when the amino acid sequences contain the same or conserved amino acid substitutes for each individual residue in the sequence. Identity and similarity may be measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue,
30 Madison, WI 53705). The present invention particularly provides such polynucleotides which hybridize under stringent conditions to the hereinabove-described polynucleotides. As herein used, the term "stringent conditions" means conditions which permit hybridization between polynucleotides sequences and the polynucleotide sequences of SEQ ID NOS:2, 4 and 6 where there is at least about 70% identity.
35 Suitably stringent conditions can be defined by, e.g., the concentrations of salt or

WO 02/31134

PCT/US01/31874

formamide in the prehybridization and hybridization solutions, or by the hybridization temperature, and are well known in the art. In particular, stringency can be increased by reducing the concentration of salt, by increasing the concentration of formamide, and/or by raising the hybridization temperature.

5 For example, hybridization under high stringency conditions may employ about 50% formamide at about 37°C to 42°C, whereas hybridization under reduced stringency conditions might employ about 35% to 25% formamide at about 30°C to 35°C. One particular set of conditions for hybridization under high stringency conditions employs 42°C, 50% formamide, 5x SSPE, 0.3% SDS, and 200 µg/ml sheared and denatured
10 salmon sperm DNA. For hybridization under reduced stringency, similar conditions as described above may be used in 35% formamide at a reduced temperature of 35°C. The temperature range corresponding to a particular level of stringency can be further narrowed by calculating the purine to pyrimidine ratio of the nucleic acid of interest and adjusting the temperature accordingly. Variations on the above ranges and conditions
15 are well known in the art. Preferably, hybridization should occur only if there is at least 95%, and more preferably at least 97%, identity between the sequences. The polynucleotides which hybridize to the hereinabove described polynucleotides in a preferred embodiment encode polypeptides which exhibit substantially the same biological function or activity as the mature protein encoded by one of the cDNAs of
20 SEQ ID NOS:2, 4 and 6.

As mentioned, a suitable polynucleotide probe may have at least 14 bases, preferably 30 bases, and more preferably at least 50 bases, and will hybridize to a polynucleotide of the present invention which has an identity thereto, as hereinabove described, and which may or may not retain activity. For example, such polynucleotides
25 may be employed as a probe for hybridizing to the polynucleotides of SEQ ID NOS:2, 4 and 6 respectively, for example, for recovery of such a polynucleotide, or as a diagnostic probe, or as a PCR primer. Thus, the present invention includes polynucleotides having at least a 70% identity, preferably at least a 90% identity, and more preferably at least a 95% identity to a polynucleotide which encodes the polypeptides of SEQ ID NOS:1, 3
30 and 5 respectively, as well as fragments thereof, which fragments preferably have at least 30 bases and more preferably at least 50 bases, and to polypeptides encoded by such polynucleotides.

As is well known in the art, the genetic code is redundant in that certain amino acids are coded for by more than one nucleotide triplet (codon), and the invention
35 includes those polynucleotide sequences which encode the same amino acids using a

WO 02/31134

PCT/US01/31874

different codon from that specifically exemplified in the sequences herein. Such a polynucleotide sequence is referred to herein as an "equivalent" polynucleotide sequence. The present invention further includes variants of the hereinabove described polynucleotides which encode for fragments, such as part or all of the mature protein, analogs and derivatives of one of the polypeptides having the deduced amino acid sequence of SEQ ID NOS:1, 3 and 5 respectively. The variant forms of the polynucleotides may be a naturally occurring allelic variant of the polynucleotides or a non-naturally occurring variant of the polynucleotides. For example, the variant in the nucleic acid may simply be a difference in codon sequence for the amino acid resulting from the degeneracy of the genetic code, or there may be deletion variants, substitution variants and addition or insertion variants. As known in the art, an allelic variant is an alternative form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides that does not substantially alter the biological function of the encoded polypeptide.

15 The present invention further includes polypeptides which have the deduced amino acid sequence of SEQ ID NOS:1, 3 and 5, as well as fragments, analogs and derivatives of such polypeptides. The terms "fragment," "derivative" and "analog", when referring to the polypeptides of SEQ ID NOS:1, 3 and 5, means polypeptides that retain essentially the same biological function or activity as such polypeptides. An analog might, for example, include a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature protein. The polypeptides of the present invention may be recombinant polypeptides, natural polypeptides or synthetic polypeptide; however, they are preferably recombinant polypeptides, glycosylated or unglycosylated.

25 The fragment, derivative or analog of a polypeptide of SEQ ID NOS:1, 3 and 5 respectively, may be (i) one in which one or more of the amino acid residues is substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which additional amino acids are fused to the mature protein, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art to provide upon the basis of the teachings herein.

WO 02/31134

PCT/US01/31874

The polypeptides and polynucleotides of the present invention should be in an isolated form, and preferably they are purified to substantial homogeneity or purity. By substantial homogeneity is meant a purity of at least about 85%.

5 The term "isolated" is used to mean that the material has been removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally occurring polynucleotide or polypeptide present in a living animal is not considered to be isolated, but the same polynucleotide or polypeptide, when separated from substantially all of the coexisting materials in the natural system, is considered isolated. For DNA, the term includes, for example, a recombinant DNA
10 which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (e.g., a cDNA or a genomic or cDNA fragment produced by polymerase chain reaction (PCR) or restriction endonuclease digestion) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional
15 polypeptide sequence, e.g., a fusion protein. Further included is recombinant DNA which includes a portion of the nucleotides shown in one of SEQ ID NO:2,4 or 6 which encodes an alternative splice variant of the DPRP. Various alternative splice variants are exemplified in SEQ ID NOS:8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 and 46.

20 The polypeptides of the present invention include any one of the polypeptide of SEQ ID NOS:1, 3 and 5 (in particular the mature protein), as well as polypeptides which have at least 70% similarity (e.g. preferably at least 60% and more preferably at least 70% identity) to one of the polypeptides of SEQ ID NOS:1, 3 and 5, more preferably at least 90% similarity (e.g. preferably at least 90% identity) to one of the polypeptides of
25 SEQ ID NOS:1, 3 and 5, and most preferably at least 95% similarity (e.g. preferably at least 95% identity) to one of the polypeptides of SEQ ID NOS:1, 3 and 5. Moreover, they should preferably include exact portions of such polypeptides containing a sequence of at least 30 amino acids, and more preferably at least 50 amino acids.

Fragments or portions of the polypeptides of the present invention may be
30 employed as intermediates for producing the corresponding full-length polypeptides by peptide synthesis. Fragments or portions of the polynucleotides of the present invention may also be used to synthesize full-length polynucleotides of the present invention.

The present invention also includes vectors which include such polynucleotides, host cells which are genetically engineered with such vectors and the production of
35 polypeptides by recombinant techniques using the foregoing. Host cells are genetically

WO 02/31134

PCT/US01/31874

engineered (transduced or transformed or transfected) with such vectors which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those commonly used with the host cell selected for expression, as well known to the ordinarily skilled artisan.

The polynucleotides of the present invention may be employed for producing polypeptides by recombinant techniques. Thus, for example, the polynucleotides may be included in any one of a variety of expression vectors for expressing polypeptides. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by any of a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures well known in the art, which procedures are deemed to be within the scope of those skilled in this art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the *E. coli* lac or trp, the phage lambda P.sub.L promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector should also contain a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression. In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells, such as dihydrofolate reductase or neomycin-resistance for eukaryotic cell culture, or such as tetracycline- or ampicillin-resistance in *E. coli*.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein. As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as *E. coli*, Streptomyces, Salmonella typhimurium; fungal cells, such as yeast; insect cells, such as

WO 02/31134

PCT/US01/31874

Drosophila S2 and *Spodoptera Sf9*; animal cells, such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

Synthetic production of nucleic acid sequences is well known in the art as is apparent from CLONTECH 95/96 Catalogue, pages 215-216, CLONTECH, 1020 East Meadow Circle, Palo Alto, Calif. 94303. Thus, the present invention also includes expression vectors useful for the production of the proteins of the present invention

The present invention further includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs may comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example: Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBS, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene), ptrc99a, pKK223-3, pKK233-3, pDR540 and pRIT5 (Pharmacia); and Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, and pSVL (Pharmacia). However, any other suitable plasmid or vector may be used as long as it is replicable and viable in the host.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol acetyl transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P.sub.R, P.sub.L and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

Components of the expression vector may generally include: 1) a neomycin phosphotransferase (G418), or hygromycin B phosphotransferase (hyg) gene as a selection marker, 2) an E. coli origin of replication, 3) a T7 and SP6 phage promoter sequence, 4) lac operator sequences, 5) the lactose operon repressor gene (lacIq) and 6) a multiple cloning site linker region. Such an origin of replication (oriC) may be derived from pUC19 (LTI, Gaithersburg, Md.).

A nucleotide sequence encoding one of the polypeptides SEQ ID NOS:2,4 and 6 having the appropriate restriction sites is generated, for example, according to the PCR

WO 02/31134

PCT/US01/31874

protocol described in Example 1 hereinafter, using PCR primers having restriction sites for KpnI (as the 5' primer) and NotI or SacI (as the 3' primer) for DPRP-1, or sites for HindIII (as the 5' primer) and NotI or BamHI (as the 3' primer) for DPRP-2. The PCR inserts are gel-purified and digested with compatible restriction enzymes. The insert and
5 vector are ligated according to standard protocols.

In a further embodiment, the present invention provides host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell
10 can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, lipofection or electroporation (Davis, L., Dibner, M., Battey, J., Basic Methods in Molecular Biology, (1986)).

Such constructs in host cells are preferably used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the
15 polypeptides of the invention can be synthetically produced by conventional peptide synthesizers or by chemical ligation of suitable fragments thus prepared.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of
20 the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989).

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector.
25 Enhancers include cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin-resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes, such as 3-
35 phosphoglycerate kinase (PGK), alpha-factor, acid phosphatase, or heat shock proteins,

WO 02/31134

PCT/US01/31874

among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion
5 protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter.
10 The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desired, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of
15 choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example,
20 pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, Wis., U.S.A.). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g.,
25 temperature shift or chemical induction), and cells are cultured for an additional period. Cells are typically harvested by centrifugation and then disrupted by physical or chemical means, with the resulting crude extract being retained for further purification. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption and use of
30 cell-lysing agents; such methods are well known to those skilled in the art.

Various mammalian cell culture systems can also be employed to express a recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell*, 23:175 (1981). Other cell lines capable of expressing a compatible vector include, for example, the C127, 3T3,
35 CHO, HeLa and BHK cell lines. Mammalian expression vectors will generally comprise

WO 02/31134

PCT/US01/31874

an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide required nontranscribed genetic elements.

The polypeptides can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Recovery can be facilitated if the polypeptide is expressed at the surface of the cells, but such is not a prerequisite. Recovery may also be desirable of cleavage products that are cleaved following expression of a longer form of the polypeptide. Protein refolding steps as known in this art can be used, as necessary, to complete configuration of the mature protein. High performance liquid chromatography (HPLC) can be employed for final purification steps.

The polypeptides of the present invention may be purified natural products, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect or mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. Polypeptides of the invention may also include an initial methionine amino acid residue.

In a preferred embodiment, the proteins of the invention are isolated and purified so as to be substantially free of contamination from other proteins. For example, the proteins of the invention should constitute at least 80% by weight of the total protein present in a sample, more preferably at least 90%, even more preferably at least 95%, and most preferably at least 98% by weight of the total protein.

These proteins may be in the form of a solution in water, another suitable solvent, such as dimethyl sulphoxide (DMSO) or ethanol, or a mixture of suitable solvents. Examples of mixtures of solvents include 10% (by weight) ethanol in water and 2% (by weight) DMSO in water. A solution may further comprise salts, buffering agents, chaotropic agents, detergents, preservatives and the like. Alternatively, the proteins may be in the form of a solid, such as a lyophilised powder or a crystalline solid, which may also comprise a residual solvent, a salt or the like.

As used herein, the term "antibodies" includes polyclonal antibodies, affinity-purified polyclonal antibodies, monoclonal antibodies, and antigen-binding

WO 02/31134

PCT/US01/31874

fragments, such as F(ab')₂ and Fab' proteolytic fragments. Genetically engineered intact antibodies or fragments, such as chimeric antibodies, Fv fragments, single chain antibodies and the like, as well as synthetic antigen-binding peptides and polypeptides, are also included. Non-human antibodies may be humanized by grafting non-human CDRs onto human framework and constant regions, or by incorporating the entire non-human variable domains (optionally "cloaking" them with a human-like surface by replacement of exposed residues, wherein the result is a "veneered" antibody). In some instances, humanized antibodies may retain non-human residues within the human variable region framework domains to enhance proper binding characteristics. Through humanizing antibodies, biological half-life may be increased, and the potential for adverse immune reactions upon administration to humans should be reduced.

Alternative techniques for generating or selecting antibodies useful herein include *in vitro* exposure of lymphocytes to human prohormone DPRP protein or a peptide therefrom, and selection of antibody display libraries in phage or similar vectors (for instance, through use of immobilized or labeled human DPRP protein or peptide). Genes encoding polypeptides having potential human DPRP polypeptide binding domains can be obtained by screening random peptide libraries displayed on phage (phage display) or on bacteria, such as *E. coli*. Nucleotide sequences encoding such polypeptides can be obtained in a number of ways well known in this art.

As would be evident to one of ordinary skill in the art, polyclonal antibodies can be generated from inoculating a variety of warm-blooded animals, such as horses, cows, goats, sheep, dogs, chickens, rabbits, mice and rats, with a human DPRP polypeptide or a fragment thereof. The immunogenicity of a human prohormone DPRP polypeptide may be increased through the use of an adjuvant, such as alum (aluminum hydroxide) or Freund's complete or incomplete adjuvant, or surface active substances, such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH or dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable. Polypeptides useful for immunization also include fusion polypeptides, such as fusions of DPRP or a portion thereof with an immunoglobulin polypeptide or with maltose binding protein. The polypeptide immunogen may be a full-length molecule or a portion thereof. If the polypeptide portion is "hapten-like", such portion may be advantageously joined or linked to a macromolecular carrier, such as keyhole limpet hemocyanin (KLH), bovine serum albumin (BSA) or tetanus toxoid, for immunization. Antibodies to DPRP may also be generated using methods that are well known in the art. Such antibodies may

WO 02/31134

PCT/US01/31874

include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which block or modify interactions at the active sites) are especially preferred for therapeutic use.

5 For the production of antibodies, binding proteins, or peptides which bind specifically to DPRP, libraries of single chain antibodies, Fab fragments, other antibody fragments, non-antibody protein domains, or peptides may be screened. The libraries could be generated using phage display, other recombinant DNA methods, or peptide synthesis (Vaughan, T. J. et al. Nature Biotechnology 14: 309-314 (1996)). Such
10 libraries would commonly be screened using methods which are well known in the art to identify sequences which demonstrate specific binding to DPRP.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to DPRP have an amino acid sequence consisting of at least about 5 amino acids and, more preferably, of at least about 10 amino acids. It is also preferable that
15 these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of DPRP amino acids may also be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to DPRP may be prepared using any well known
20 technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique, although monoclonal antibodies produced by hybridoma cells may be preferred.

In addition, techniques developed for the production of "chimeric antibodies",
25 such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used, see Neuberger, M.S. et al. Nature 312: 604-608 (1984). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce DPRP-specific single chain antibodies. Antibodies with related
30 specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (Burton D. R. Proc. Natl. Acad. Sci. 88: 11120-11123 (1991)).

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly

WO 02/31134

PCT/US01/31874

specific binding reagents as disclosed in the literature. (Orlandi, R. et al. Proc. Natl. Acad. Sci. **86**: 3833-3837 (1989)).

Antibody fragments which contain specific binding sites for DPRP may also be generated. For example, such fragments include, but are not limited to, F(ab')₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments
5 fragments generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (Huse, W. D. et al. Science **254**: 1275-1281 (1989)).

10 Various immunoassays may be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between DPRP and its specific antibody. A two-site, monoclonal-based
15 immunoassay utilizing monoclonal antibodies reactive to two non-interfering DPRP epitopes is preferred, but a competitive binding assay may also be employed.

As earlier mentioned, the DPRPs can be used in treatment of the Diseases. Pharmaceutical compositions suitable for use in this aspect of the invention include compositions wherein the active ingredients are contained in an effective amount to
20 achieve the intended purpose relating to one of the Diseases. The determination of a therapeutically effective dose is well within the capability of those skilled in the art and can be estimated initially either in cell culture assays, e.g. of neoplastic cells, or in animal models, usually mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration, which
25 information is then commonly used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, e.g. a DPRP or fragment thereof, antibodies of DPRP, or an agonist, antagonist or inhibitor of DPRP, which ameliorates particular symptoms or conditions of the Disease. For
30 example, the amount to be administered may be effective to cleave a desired target substrate upon contact therewith. Therapeutic efficacy and toxicity may likewise be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose
35 ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the

WO 02/31134

PCT/US01/31874

LD50/ED50 ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

An exact dosage will normally be determined by the medical practitioner in light of factors related to the subject requiring treatment, with dosage and administration being adjusted to provide a sufficient level of the active moiety or to maintain a desired effect. Factors to be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or even once every two weeks, depending on the half-life and clearance rate of the particular formulation.

Yet another aspect of the invention provides polynucleotide molecules having sequences that are antisense to mRNA transcripts of DPRP1, DPRP2 and DPRP-3 polynucleotides. Administration of an antisense polynucleotide molecule can block the production of the protein encoded by DPRP-1, DPRP2 or DPRP-3. The techniques for preparing antisense polynucleotide molecules and administering such molecules are known in the art. For example, antisense polynucleotide molecules can be encapsulated into liposomes for fusion with cells.

In particular, the expression of DPRP-1, DPRP-2 and DPRP-3 in specialized epithelial cells, immune cells (lymphocytes and B cells), astrocytic tumors, and in various hormone sensitive cancers provides evidence of a potential role in the pathophysiology of cancer, metaplasia and metastasis. Therefore in a further aspect, the invention relates to diagnostic assays for detecting diseases associated with inappropriate DPRP activity or expression levels. Antibodies that specifically bind DPRP may be used for the diagnosis of disorders characterized by expression of DPRP, or in assays to monitor patients being treated with DPRP or with agonists or antagonists (inhibitors) of DPRP. Antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for DPRP include methods that utilize the antibody and a label to detect DPRP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and they may be labeled by covalent or non-covalent joining with a reporter molecule. A wide

WO 02/31134

PCT/US01/31874

variety of reporter molecules are known in the art. Recombinant DPRP proteins that have been modified so as to be catalytically inactive can also be used as dominant negative inhibitors. Such modifications include, for example, mutation of the active site.

A variety of protocols for measuring DPRP, including ELISAs, RIAs and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of DPRP expression. Normal or standard values for DPRP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to DPRP under conditions suitable for complex formation. The method for detecting DPRP in a biological sample would comprise the steps of: a) providing a biological sample; b) combining the biological sample and an anti-DPRP antibody under conditions which are suitable for complex formation to occur between DPRP and the antibody; and c) detecting complex formation between DPRP and the antibody, thereby establishing the presence of DPRP in the biological sample. The amount of complex formation then may be quantified by various methods, preferably by photometric means. Quantities of DPRP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding DPRP are used for diagnostic purposes, which polynucleotides may include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. These polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of DPRP may be correlated with one of the Diseases. The diagnostic assay may be used to distinguish between absence, presence, and excess expression of DPRP and to monitor regulation of DPRP levels during therapeutic intervention. Moreover, pharmacogenomic, single nucleotide polymorphisms (SNP) analysis of the DPRP genes can be used as a method to screen for mutations that indicate predisposition to disease or modified response to drugs.

DPRP polynucleotide and polypeptide sequences, fragments thereof, antibodies of DPRPs, and agonists, antagonists or inhibitors of DPRPs can be used to as discovery tools to identify molecular recognition events and therefore proteins, polypeptides and peptides that interact with DPRP proteins. A specific example is phage display peptide libraries where greater than 10⁸ peptide sequences can be screened in a single round of panning. Such methods as well as others are known within the art and can be utilized to identify compounds that inhibit or enhance DPRP-1, DPRP-2 or DPRP-3 activity.

Coupled links represent functional interactions such as complexes or pathways, and

WO 02/31134

PCT/US01/31874

proteins that interact with DPRPs can be identified by a yeast two-hybrid system, proteomics (differential 2D gel analysis and mass spectrometry) and genomics (differential gene expression by microarray or serial analysis of gene expression SAGE). Proteins identified as functionally linked to DPRPs and the process of interaction form
5 the basis of methods of screening for inhibitors, agonists and antagonists and modulators of these DPRP-protein interactions.

The term "antagonist," as it is used herein, refers to an inhibitor molecule which, when bound to DPRP, decreases the amount or the duration of the effect of the biological or immunological activity of DPRP, e.g. decreasing the enzymatic activity of
10 the peptidase to cleave the N-terminal dipeptide. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of DPRP; for example, they may include small molecules and organic compounds that bind to and inactivate DPRPs by a competitive or non-competitive type mechanism. Specific examples of DPRP tetrapeptide peptidic enzyme activity inhibitors are
15 described in Example 6 and 7. Inhibitors can be, for example, inhibitors of the DPRP protease activity, or alternatively inhibitors of the binding activity of the DPRP to proteins with which they interact. Specific examples of such inhibitors can include, for example, anti-DPRP antibodies, peptides, protein fragments, or small peptidyl protease inhibitors, or small non-peptide, organic molecule inhibitors which are formulated in a
20 medium that allows introduction into the desired cell type. Alternatively, such inhibitors can be attached to targeting ligands for introduction by cell-mediated endocytosis and other receptor mediated events. Such methods are described further below and can be practiced by those skilled in the art given the DPRP nucleotide and amino acid sequences described herein.

25 A further use for DPRPs is for the screening of potential antagonists for use as therapeutic agents, for example, for inhibiting binding to DPRP, as well as for screening for agonists. DPRP, its immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds which are prospective agonists or antagonists in any of a variety of drug screening techniques. The fragment employed in such screening
30 may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between DPRP and the agent being tested is then measured. Other assays to discover antagonists that will inhibit DPRP are apparent from the disclosures of U.S. Patents Nos. 6,011,155, 6,107,317, 6,110,949, 6,124,305 and 6,166,063, which describe inhibitors of DPPIV. Another worthwhile use

WO 02/31134

PCT/US01/31874

of these DPRPs is the screening of inhibitors of DPPIV to show that they will not have undesired side effects by also inhibiting one or more of the DPRPs.

A method provided for screening a library of small molecules to identify a molecule which binds DPRP generally comprises: a) providing a library of small
5 molecules; b) combining the library of small molecules with the polypeptide of either SEQ ID NOS:1, 3 or 5, or with a fragment thereof, under conditions which are suitable for complex formation; and c) detecting complex formation, wherein the presence of such a complex identifies a small molecule which binds DPRP.

One method for identifying an antagonist comprises delivering a small molecule
10 which binds DPRP into extracts from cells transformed with a vector expressing DPRP along with a chromogenic substrate (e.g. Ala-Pro-AFC or Ala-Pro-AMC) under conditions where cleavage would normally occur, and then assaying for inhibition of cleavage by the enzyme by monitoring changes in fluorescence, or UV light absorption, by spectrophotometry to identify molecules that inhibit cleavage. A reduced rate of
15 reaction or total amount of fluorescence or UV light absorption, in the presence of the molecule, establishes that the small molecule is an antagonist which reduces DPRP catalytic/enzymatic activity. Once such molecules are identified, they may be administered to reduce or inhibit cleaving by a DPRP.

The term "agonist," as used herein, refers to a molecule which, when bound to
20 DPRP, increases or prolongs the duration of the effect of DPRP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules that bind to and modulate the effect of DPRP. Although it is less likely that small molecules will prove to be effective DPRP agonists, a method for identifying such a small molecule, which binds DPRP as an agonist, comprises delivering a chromogenic form of a small molecule that
25 binds DPRP into cells transformed with a vector expressing DPRP and assaying for fluorescence or UV light absorption changes by spectrophotometry. An increased amount of UV absorption or fluorescence would establish that the small molecule is an agonist that increases DPRP activity.

Another technique for drug screening which may be used provides for high
30 throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with DPRP, or with fragments thereof, and then washed. Bound DPRP is then detected by methods well
35 known in the art. Purified DPRP can also be coated directly onto plates for use in the

WO 02/31134

PCT/US01/31874

aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding DPRP specifically compete with a test
5 compound for binding DPRP. In this manner, antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants with DPRP.

As indicated above, by investigating the binding sites, ligands may be designed that, for example, have more interactions with DPRP than do its natural ligands. Such antagonist ligands will bind to DPRP with higher affinity and so function as competitive
10 ligands. Alternatively, synthetic or recombinant proteins homologous or analogous to the ligand binding site of native DPRP may be designed, as may other molecules having high affinity for DPRP. Such molecules should also be capable of displacing DPRP and provide a protective effect.

As indicated above, the knowledge of the structures of DPRP enables synthetic
15 binding site homologues and analogues to be designed. Such molecules will facilitate greatly the use of the binding properties to target potential therapeutic agents, and they may also be used to screen potential therapeutic agents. Furthermore, they may be used as immunogens in the production of monoclonal antibodies, which antibodies may themselves be used in diagnosis and/or therapy as described hereinbefore.

Given the ubiquitous expression of several members of the prolyl oligopeptidase
20 S9B family, cell lines in which targeted gene disruption of DPPIV, DPRP-1, DPRP-2, DPRP-3, FAP and DPPVI to establish the null phenotype will be of great value to assist screening for selective and potent compounds. Accordingly, the invention provides such cell lines engineered with Lox-Neo IRES tk cassette and GFP-IRES-Neo Knock-in/out
25 cassette DNA element for constructing somatic gene targeting vectors.

Example 1

Cloning and Expression of DPRP genes Using the Mammalian Expression System

DNA fragments encoding the full-length polypeptide DPRP-1 were amplified
using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene,
30 i.e. SEQ ID NO:45 and NO:46. In addition, DNA fragments encoding the full length polypeptide DPRP-2 were amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of that gene, i.e. SEQ ID NO:50 and NO:51. Furthermore, DNA fragments encoding the full length polypeptide DPRP-3 were amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of that gene, i.e. SEQ
35 ID NO:55 and NO:56.

WO 02/31134

PCT/US01/31874

The three amplified sequences were respectively isolated from a 0.7% agarose gel using commercially available kit (GFX PCR DNA and Gel Band Purification Kit, Amersham Pharmacia Biotech Inc., Piscataway NJ, USA). The fragments were then ligated into cloning vector, pGEM-7Zf(-) (Promega Corporation, Madison WI, USA) and sequenced. The corresponding cloning constructs were respectively designated pGEM7-DPRP1, pGEM7-DPRP2 and pGEM7-DPRP3. The DNA sequences encoding the truncated DPRP-1 or DPRP-2 or DPRP-3 were amplified using pGEM7-DPRP1 or pGEM7-DPRP2 or pGEM7-DPRP3 as a template and PCR oligonucleotide primers. SEQ ID NO:45 and NO:47 were used for DPRP-1; SEQ ID NO:50 and NO:52 were used for DPRP-2; and SEQ ID NO:57 and NO:58 for DPRP-3. The amplified sequences were again isolated from a 0.7% agarose gel using the same purification kits and sub-cloned into pGEM-7Zf(-). The resulting constructs were designated pGEM7-DPRP1f, pGEM7-DPRP2f and pGEM7-DPRP3f.

To make the DPRP-1 mammalian expression construct, pGEM7-DPRP1 was digested with the restriction enzymes KpnI and NotI to release the full length DPRP-1 gene. The DNA fragment carrying the DPRP-1 gene was gel band purified using the above kit and then inserted into expression vector pcDNA3 (Invitrogen, Carlsbad CA, USA) to make the native DPRP-1 expression construct, which was designated pcDNA-DPRP1. pGEM7-DPRP1f was digested with the restriction enzymes XbaI and HindIII to release the truncated DPRP-1f gene. The DNA fragment carrying the DPRP-1f gene was gel band purified using the above kit and then inserted into expression vector pcDNA3.1(-)/myc-His A (Invitrogen, Carlsbad CA, USA) to make the tagged DPRP-1 expression construct pcDNA-MycHis-DPRP1.

To make the DPRP-2 mammalian expression construct, pGEM7-DPRP2 was digested with the restriction enzymes HindIII and BamHI to release the full length DPRP-2 gene. The DNA fragment carrying the DPRP-2 gene was gel band purified using the above kit and then inserted into expression vector pcDNA3 (Invitrogen, Carlsbad CA, USA) to make the native DPRP-2 expression construct, which was designated pcDNA-DPRP2. pGEM7-DPRP2f was digested with the restriction enzymes EcoRI and BamHI to release the truncated DPRP-2f gene. The DNA fragment carrying the DPRP-2f gene was gel band purified using the above kit and then inserted into expression vector pcDNA3.1(-)/myc-His B (Invitrogen, Carlsbad CA, USA) to make the tagged DPRP-2 expression construct designated pcDNA-MycHis-DPRP2.

To make the DPRP-3 mammalian expression construct, pGEM7-DPRP3 was digested with the restriction enzymes EcoRI and XhoI to release the full length DPRP-3

WO 02/31134

PCT/US01/31874

gene. The DNA fragment carrying the DPRP-3 gene was gel band purified using the above kit and then inserted into expression vector pcDNA3 (Invitrogen, Carlsbad CA, USA) to make the native DPRP-3 expression construct designated pcDNA-DPRP3. pGEM7-DPRP3f was digested with the restriction enzymes NheI and ApaI to release the truncated DPRP-3f gene. The DNA fragment carrying the DPRP-3f gene was gel band purified using the above kit and then inserted into expression vector pcDNA3.1(-)/myc-His B (Invitrogen, Carlsbad CA, USA) to make the tagged DPRP-3 expression construct pcDNA-MycHis-DPRP3.

Example 2

10 Expression Pattern of DPRP genes in human tissues

Quantitative PCR analysis was carried out to examine the levels of expression of the mRNAs for the polypeptides of the present invention in human tissues. RT PCR was also carried out on a number of human cell lines including but not limited to prostate cancer cells (LNCaP, PC3, DU145), the MLTC-1 line (mouse testis), and MDA-MB231 cells (breast cancer). Bands of the expected sizes for DPRP-1, DPRP-2 and DPPIV were all expressed in the various cancer cell lines, with FAP also being expressed at very low levels.

Northern Blot Analysis

Northern blot analysis was performed with 2 μ g poly(A)⁺ RNA isolated from eight different tissues using DPRP probes. Specifically, a human Multiple Tissue Northern (MTN) blot (Clontech, Palo Alto, Calif.) was probed with a 1 kb N-terminal fragment that had been radioactively labeled by random priming in the presence of a ³²PdCTP (A. P. Feinberg et al., *Anal. Biochem.*, 132, 6 (1983)). Hybridization was performed at 68°C overnight in ExpressHyb™ hybridization solution (Clontech, Palo Alto, Calif.). The blots were first washed at room temperature in 2 times SSC and 0.05% SDS, and then washed at 60°C (DPRP-1 & DPRP-2) and 50°C (DPRP-3) in 0.1 times SSC and 0.1% SDS.

Northern analysis showed expression of DPRP-1 in several tissues with the most abundant signal being in testis, prostate, muscle and brain. Testis showed 3 transcripts approximately 7.5, 4.5 and 2.5 kb in length. The shorter mRNA species was very abundant in testis but negligible in the other tissues tested. DPRP-2 was ubiquitously expressed in every tissue with highest levels in liver and muscle and a predominant transcript at 5kb. DPRP-3 expression was limited to brain and pancreas. Further analysis was conducted for the three proteases in specific brain regions (cerebellum, cortex, medulla, spinal cord, occipital lobe, frontal lobe temporal lobe and putamen).

WO 02/31134

PCT/US01/31874

DPRP-1 was expressed in all regions with low levels present in the spinal cord, while DPRP-2 was expressed in all brain regions tested.

Oligonucleotide primers SEQ ID NO:48 and NO:49 were used for DPRP-1 quantitative PCR, whereas oligonucleotide primers SEQ ID NO:53 and NO:54 were used for DPRP-2 quantitative PCR. Human Multiple Tissue cDNA (MTC™) Panel I and Panel II (Clontech, Palo Alto CA, USA) were used as normalized cDNA templates. 0.5 ng of each cDNA were used in a 25 µl PCR reaction, with each primer at a final concentration of 300 nM. The PCR reaction was performed using a SYBR Green PCR Core Reagents Kit (Applied Biosystems, Foster City CA, USA) and detected with an Applied Biosystems GeneAmp 5700 sequence detection system. Manufacturer's recommended thermal cycling parameter, e.g. 50°C for 2 min, 95°C for 10 min followed by 40 cycles of 95°C for 15 sec and 60°C for 1 min was used. Data obtained shows relatively high rates of expression for both DPRP-1 and DPRP-2 in the pancreas, ovary and testis, and a particularly high rate for DPRP-2 in the liver.

15 **Example 3 – Production of DPRP Polyclonal Antibodies and Western Blotting**

The amino acid sequence deduced from the cDNA encoding DPRP-1 was analyzed using DNASTAR software (DNASTAR, Inc.) to determine regions of high immunogenicity, and a corresponding oligopeptide was synthesized and used to raise anti-DPRP-1 antibodies. The procedure was repeated for DPRP-2 and DPRP-3. The selection of appropriate peptide sequences and the techniques for antibody production are methods well known to those of skill in the art. Selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions, is well known in this art.

Typically, oligopeptides that are about 15 to 20 residues in length, e.g. SEQ ID NO:59 for DPRP-1, SEQ ID NO:60 for DPRP-2 and SEQ ID NO:61 for DPRP-3, were synthesized using an Applied Biosystems Peptide Synthesizer Model 431 A. Fmoc-chemistry was used and the 19- or 15-residue peptides were respectively coupled to keyhole limpet hemocyanin (KLH, Sigma, St. Louis, Mo.) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS). Rabbits were immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. The resulting antisera were tested for antipeptide activity, e.g., by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radiiodinated, goat anti-rabbit IgG.

Western blotting was performed using normal human protein samples (Protein Medley) obtained from Clontech (about 36 µg of total proteins). Proteins were fractionated through 10% SDS-polyacrylamide gels, and transferred to 0.45 mm

WO 02/31134

PCT/US01/31874

nitrocellulose membranes. Membranes were blocked in Tris-buffered saline (TBS) with 0.05% Tween 20 and 1% BSA. Anti DPRP-1 or DPRP-2 specific antibodies were used as primary antibodies and were diluted 1:5,000 in Tris-buffered saline with 0.05% Tween 20 (TBST) and the Alkaline Phosphatase (AP) conjugated goat anti-Rabbit IgG (Promega) was diluted 1: 5,000 in the same buffer before use. The positive reaction was visualized by incubating the membrane in Western Blue Stabilized Substrate (Promega) for AP until the bands of interest have reached the desired intensity. DPRP-1 and DPRP-2 proteins were detected in brain, muscles, kidney, prostate, testis and ovary tissues. DPRP-1 and DPRP-2 were synthesized as approximately 101kDa and 100kDa forms, respectively, which are in good agreement with the molecular masses estimated from their primary structure as shown in Table 3.

Table 3. Predicted Molecular Weight, Number of potential N-linked glycosylation sites (Asn residues) and predicted pI values of DPRP-1, DPRP-2 and DPRP-3, based on sequence analysis using the method developed by Hopp and Woods, *Proc. Nat. Acad. Sci.* 78:3824-3828 (1981).

	M.W. (Da.)	No. of Asn	pI
DPRP1	101422	26	5.39
DPRP2	98263	27	6.01
DPRP3	90914	33	6.11

Several additional bands of similar molecular weight were observed. These are thought to be due to the presence of post-translational glycosylation of the proteins. Table 3 also shows the number of potential N-glycosylation sites for the DPRP proteins. The presence of glycosylated and unglycosylated forms of the proteins was evaluated using tunicamycin, an inhibitor of the oligosaccharide synthesis. It is evident that the smaller forms were unglycosylated forms. The correlation between mRNA (Northern analysis) and protein quantity (Western analysis) for DPRP-1 is shown in Table 4.

Table 4. Correlation of mRNA and protein expression of DPRP-1 in human tissues

	Heart	Brain	Placenta	Muscles	Kidney	Prostate	Testis	Ovary
Northern	++	+++	+	+++	++	+++	++++	+
Western	-	++++	-	+	++	+	+++	+++

WO 02/31134

PCT/US01/31874

Example 4**Immunohistochemical localization of DPRP proteins in human tissues**

Four-micron sections were prepared from a number of different formalin-fixed, paraffin-embedded human tissues. Tissue sections were deparaffined through 4
5 immersions in xylenes for 5 minutes, followed by a graded alcohol series to distilled water. Steam heat induced epitope recovery (SHIER) was used with several different SHIER solutions with and without enzyme digestion tissue in two different concentrations (Ladner et al, Cancer Res.; 60, p 3493-3503, 2000). The treatments and antibody dilutions employed are outlined below.

- 10 1. Blocking Reagent for 15 minutes (Normal Goat Serum)
2. Primary Antibody for 25, 60 min or overnight incubation
3. Secondary Antibody for 25 minutes (Biotinylated Goat-anti-rabbit IgG)
4. Endogenous Peroxidase Blocking for 3 x 1.5 minutes
5. ABC (avidin-biotin complex) / Horse Radish Peroxidase for 25 minutes
- 15 6. DAB Chromogen for 3 x 5 minutes (Brown reaction product)
7. Light Hematoxylin Counter Stain 1 minute

Positive controls were run to assure the detection chemistries and antigen pretreatments were working appropriately. Rabbit IgG was run as a negative control. An avidin-biotin based tissue staining system was used for the detection of the DPRP-1
20 antibody. Horseradish peroxidase was used as a reporter enzyme with DAB as chromogen. After staining, slides were dehydrated through an alcohol series to absolute ethanol followed by xylene rinses. Slides were permanently coverslipped with glass coverslips and permount. Digital images of representative staining, where positive staining was indicated by a dark brown chromogen (DAB-HRP reaction product), were captured
25 using a video camera from Olympus. Hematoxylin counterstain provides a blue nuclear stain to assess cell and tissue morphology.

DPRP-1 rabbit polyclonal antibody labels formalin-fixed, paraffin-embedded human tissues, including normal testis, prostate glands, endometrial glands, tonsils and pancreas. It was also present in endothelial cells of normal ovary, bladder and kidney.
30 Staining was localized in the cytoplasm in epithelial and some stromal cells such as fibroblasts, endothelial cells and lymphocytes. Interestingly in normal testis tested with DPRP-1 antibodies, there was distinctive expression in Leydig cells and multinucleated macrophages found in interstitial tissue, which is the space surrounding the seminiferous tubules. Tonsil B cells were stained with DPRP-1 antibody.

WO 02/31134

PCT/US01/31874

Example 5**Mammalian and Insect Cell Expression of DPRP Proteins and Purification**

Plasmid DNA of pcDNA-DPRP1, pcDNA-MycHis-DPRP1, pcDNA-DPRP-2 or pcDNA-MycHis-DPRP2 was transfected into PEAK (EdgeBioSystems, Gaithersburg MD, USA) or COS-1 (ATCC CRL-1650) using LipofectAmine (Life Technologies, Gaithersburg MD, USA) method recommended by the manufacturer. Transfected cells were maintained in DMEM with 5% FBS at 37°C with 5% CO₂ for 48 hours. Cells were then collected and used for recombinant protein extraction. Cells were harvested 48 hours after transfection, homogenized and then spun at 18,000 x g for 40 min. The supernata were collected as cytosolic fractions. This fraction was loaded on TALON spin column (Clontech), and His-tagged proteins were eluted with 50mM PBS, 150mM imidazole, pH 7. Recombinant proteins were then detected by western blotting with anti-myc antibody and visualized using a ProtoBlot II AP system (Promega). Recombinant affinity purified fusions of the DPRP-1 and DPRP-2 were detected by western blot, and DPRP-1 and DPRP-2 were synthesized as 112kDa and 109kDa forms as predicted.

Naturally occurring or recombinant DPRP proteins were substantially purified by immunoaffinity chromatography using antibodies specific for DPRP-1, DPRP-2 or DPRP-3. An immunoaffinity column was constructed by covalently coupling DPRP antibodies to an activated chromatographic resin, such as CNBr-activated Sepharose (Pharmacia & Upjohn). After the coupling, the resin was blocked and washed according to the manufacturer's instructions.

Media or cell extracts containing DPRP proteins were passed over the immunoaffinity column, and the column was washed under conditions that allow the preferential absorbance of DPRPs (e.g., high ionic strength buffers in the presence of detergent). The column was eluted under conditions that disrupt antibody/DPRP binding (e.g., a buffer of pH 2-3 or a high concentration of a chaotrope, such as urea or thiocyanate ion), and purified DPRP was collected.

Example 6**Enzymatic Activity of DPRP proteins and Methods of Screening for Inhibitors**

The kinetic properties of recombinant DPRP-1 and DPRP-2 were determined in a continuous fluorimetric assay. Buffer, pH and temperature dependence optimization led to the following assay conditions: Enzyme assays were performed in 50mM PBS, pH7.4 50 µl (50 µg/ml) of purified enzymes were mixed with 1 µl of different concentration of Ala-Pro-AMC (Enzyme Systems). Plates were then incubated at 37°C for 30 min, and

WO 02/31134

PCT/US01/31874

fluorescence was detected using a Wallac 1420 Fluorimeter with $\lambda_{ex}40355$ and $\lambda_{em}535$. The K_m values of DPRP-1 and DPRP-2 were similar (208 and 161 μ M respectively).

Further biochemical characterization reveals that DPRP-1 and DPRP-2 have similar profiles to DPPIV. The two purified proteases and DPPIV were preincubated with inhibitors at room temperature for 30 min. Substrate, Ala-Pro-AMC (100 μ M), was then added, and the fluorescence intensity was recorded as 60 readings during a 60 min period. The irreversible serine protease inhibitor AEBSF was the only inhibitor tested that showed strong inhibition of all three enzymes (Table 5). This confirms the structural and domain analysis prediction that these proteins belong to the serine protease superfamily.

Table 5. Inhibition of DPRP-1 and DPRP-2 by Protease Inhibitors

Inhibitor	Inhibitor Property	Concentration	Residual activity (% of control)		
			DPRP-1	DPRP-2	DPPIV
AEBSF	serine, irreversible	5mM	29.6	23.9	21.1
Aprotinin	serine, reversible	5 μ g/ml	77.5	63.2	80.2
Pepstatin	aspartic, reversible	2 μ g/ml	97.3	95.0	93.5
DTT	cysteine	2mM	100.1	94.8	98.3
B-Mercaptoethonal	cysteine	100mM	93.2	84.0	98.0
EDTA	metallo, reversible	2mM	91.5	86.0	93.5
Leupeptin	serine, reversible	50 μ g/ml	91.1	90.4	90.7

In addition to Ala-Pro-AMC, additional substrates tested also confirmed that DPRP-1 and DPRP-2 are dipeptidyl peptidases. The data were derived by determining the fluorescence change following a 30-minute incubation of the substrates (125 μ M) with enzymes as a percentage of the fluorescence measured at Ala-Pro-AMC and Gly-Pro-AMC were the only good substrates among those tested.

Table 6. DPRP-1 and DPRP-2 are dipeptidyl peptidases.

Substrate	% Change in Fluorescence at 30 minutes		
	DPRP-1	DPRP-2	DPPIV
Ala-Pro-AMC	239.0	127.5	379.0
Gly-Pro-AMC	341.5	205.0	444.0
Ala-Pro-pNA	45.5	44.0	29.5
Pro-pNA	-1	-2.5	0.0
Gly-Arg-pNA	-4.5	-0.5	0.0
Lys-Ala-pNA	2.5	0.5	0.5
Ala-Phe-Pro-pNA	-4	-0.5	2.0

WO 02/31134

PCT/US01/31874

Additional natural and non-natural amino acid di-, tri- and tetra-peptides were tested in order to find an optimal substrate for testing each of the DPRP proteins that will also show reduced activity when incubated DPPIV.

The enzyme assay method described here is one of a number of methods that can be utilized to screen for peptide and non-peptide inhibitors of the DPRP enzymes. Libraries of tetrapeptide inhibitors were tested to discover inhibitors of enzyme activity. Candidate inhibitors were prepared as 10-20 mM stock solutions in DMSO and stored at -20°C. Dilutions were made in assay buffer. Inhibition was determined by comparing the changes in fluorescence of the inhibited enzyme to the change in fluorescence of the control (vehicle) enzyme. $100 - (\text{fl units of sample} / \text{fl units of control} \times 100)$ gives percent inhibition value. The percent inhibition and the inhibitor concentration at which the enzyme was 50% inhibited (IC_{50}) was ascertained by plotting percent inhibition vs. inhibitor concentration on the log scale. As shown in Figure 3, several tetrapeptides amides inhibited enzyme activity, wherein data are expressed as the % of activity in the presence of vehicle (0.02% DMSO) alone. Compounds were added at 1 mM. Most interesting was the apparent differential activity of some tetrapeptides for DPRP-1 and DPRP-2, compared to DPPIV. While all three enzymes were inhibited by Peptide-1, only DPRP-1 and DPRP-2 were significantly inhibited by Peptide-4 and Peptide-5. This demonstrates that selective inhibition of the purified enzymes is achievable.

The assay described in this example can also be used to screen additional synthetic or naturally occurring compound libraries, including macromolecules, for agents that either inhibit or enhance DPRP activity. The DPRP-1 and DPRP-2 polypeptides to be used in the assay can be obtained by, for example, *in vitro* translation, recombinant expression (see Example 5) or biochemical procedures. Methods other than those described here can also be used to screen and identify compounds that inhibit DPRP-1, DPRP-2 or DPRP-3, which methods can include, for example, binding assays such as ELISAs and RIAs.

Example 7

Effect of DPRP Inhibitors on the Proliferation of Human Cancer Cells In Vitro

In an attempt to assess the effect that several inhibitors of DPRP-1 and DPRP-2 activity may have on the proliferation of human cancer cells, LNCap, PC3 and Du145, mouse testis line MLTC-1 and MDA-MB231 breast cancer cells were plated (10^4 per well) in 96-well tissue culture plates and allowed to grow and attach for 24 hours at 37°C in a CO₂ incubator. Compounds at various dilutions (final dilutions: 0.1 nM – 10 μM) were then added to the wells for various incubation periods from 24 hours to 96 hours,

WO 02/31134

PCT/US01/31874

- with fresh compound being replaced each day. Addition of the diluent DMSO alone served as the control. Following incubation with these compounds in triplicate, proliferation of the cells was determined using an XTT cell proliferation assay (Roche 1-465-015). The plates were read at 490 and 650nm 5 hours after the XTT mix was added.
- 5 An increase in cell proliferation was observed with three of the inhibitors at concentrations equal to 0.1, 1, 10 and 100 x IC₅₀, and the results are shown in FIGS. 4A, 4B and 4C for PC3 cells.

- Overall, the DPRPs are expressed in a wide variety of tissues as has been demonstrated by mRNA amplification, western blotting and immunohistochemistry.
- 10 DPRP-1 was most abundant in the testis by Northern blot and western blot. The large number of expressed sequence tags (ESTs) from testis cDNA sources that are homologous to DPRP-1 also confirms abundant expression of DPRP-1 in testis. Example 4 describes the immunohistochemical localization of DPRP-1 protein in human testis using a specific DPRP-1 antibody. DPRP-1 is strongly expressed in epitheloid
- 15 Leydig cells, and Leydig cells are the primary source of testicular androgens (male steroid hormones) in the mammalian male. In the interstitium of the testis, Leydig cells and macrophages are in close association with "digitation" of Leydig cell process extending onto macrophage surface. Multinucleated cells in close proximity to the Leydig cells were also stained with DPRP-1 antibody suggesting that the protease was
- 20 also expressed in macrophages, and macrophages in the testis play an important role in the paracrine regulation of Leydig cells. Cytokines secreted by the testicular macrophages are mitogenic to Leydig cells and play an important role in the differentiation of mesenchymal progenitor cell into mature Leydig cells. A clearer understanding of the proteins and pathways involved in the maturation of the testis is
- 25 important for the discovery of new treatments for precocious puberty. In addition, Leydig cells cause tumors such as sex cord-stromal tumors via sexual steroid production (predominantly testosterone). Testosterone is associated with several neoplasia and diseases such as breast carcinoma and uterine cancers, ovarian carcinoma and androgenic alopecia (hair loss). Further examination of the localization of DPRP
- 30 proteins in other glands in the body (e.g. adrenal glands) that produce testosterone and other androgenic hormones are currently under investigation. The possible association of DPRP-1 with steroid and polypeptide hormone biosynthetic pathways functions is being investigated, and Example 7 is relevant to understanding the role of DPRP proteins in prostate, testis and breast in vitro cell models.

WO 02/31134

PCT/US01/31874

Immunohistochemical analysis also localized DPRP-1 to endometrial glands in the uterus (see Example 4), pancreatic acini, glomeruli of the kidney, plasma cells in the bladder, a subset of B-cells in the tonsils, columnar epithelial cells of the prostate and poorly differentiated prostate squamous metaplasia, Gleason grade 4 prostatic carcinoma, and hyperplastic glands in benign prostatic hyperplasia. Positive staining in breast carcinoma, as well as in seminoma and prostate squamous metaplasia, suggests a general association of DPRP-1 with hormone-sensitive tissues, particularly in cells that become poorly differentiated. The presence of the DPRP-1 in specialized epithelial cells and in inflammatory plasma cells (lymphocytes) is also of interest. Inflammatory breast carcinoma has an abundance of infiltrating lymphocytes and an overall bad prognosis. DPRP-1 and other DPRP proteins appear in medullary carcinomas that typically have a constant infiltrating lymphoplasmacytic component at the periphery of the tumor, which is thought to represent a reaction of the host tissues to the neoplasm. Most of the lymphocytes are T Cells, and most of the plasma cells are of the IgG-producing type. Several antigens are abundant on B cells, a subgroup of breast-cancer cells, and other epithelial cancer cells, and these antigens are targets for a new class of therapeutic monoclonal antibodies with some notable success having been achieved with a humanized monoclonal antibody against the B-cell-specific antigen CD20. Accordingly, monoclonal antibodies to DPRP proteins are felt to be useful to diagnose and treat diseases in which they are involved, including cancer.

The expression of DPRP-1 in specialized epithelial cells of a number of tissues suggests that DPRP-1 and other DPRP proteins may be involved in growth and differentiation thereof. Testing using inhibitors described in Example 6 in *in vitro* models of prostate and testis cancer (Example 7) showed that DPRP-1/DPRP-2 inhibitors caused a 50-60% increase in proliferation of PC3 cells at nM concentrations as shown in FIGS. 4A-4C.

Although the invention has been described in accordance with its preferred embodiments, which constitute the best mode presently known to the inventors, it should be understood that changes and modifications as would be obvious to those skilled in this art may be made without departing from its scope which is set forth in the claims appended hereto. For example, although the disclosure focuses on DPRP-1 and DPRP-2 in certain instances, DPRP-3 and its fragments are considered to be similarly useful, as are nucleic acids encoding same. Particular features of the invention are emphasized in the claims that follow.

WO 02/31134

PCT/US01/31874

CLAIMS:

1. Isolated nucleic acid which encodes (a) a polypeptide, which includes the amino acid sequence of one of SEQ ID NOS:1, 3 and 5, or (b) a polypeptide having an amino acid sequence that is at least about 70% similar thereto and exhibits the same biological function; or which is an alternative splice variant of one of SEQ ID NOS:2, 4 and 6; or which is a probe comprising at least 14 contiguous nucleotides from said nucleic acid encoding (a) or (b); or which is complementary to any one of the foregoing.
2. The isolated nucleic acid of claim 1 which is DNA or RNA.
3. The isolated nucleic acid of claim 1 which is a DNA transcript that includes the entire length of any one of SEQ ID NOS:2, 4 and 6 or which is complementary to the entire coding region of one of SEQ ID NOS:2, 4 and 6.
4. An antisense oligonucleotide directed against the DNA of claim 3.
5. The isolated nucleic acid of claim 1 which is an RNA transcript which includes the entire length of any one of SEQ ID NOS:2, 4 and 6.
6. The isolated nucleic acid of claim 1 which is an alternative splice variant of one of SEQ ID NOS:2, 4 and 6.
7. A polypeptide encoded by the nucleic acid of claim 6.
8. The isolated nucleic acid of claim 1 which encodes a polypeptide having an amino acid that is at least about 90% similar to one of SEQ ID NOS:1, 3 and 5.
9. The isolated nucleic acid of claim 1 which encodes a polypeptide having an amino acid that is at least about 95% similar to one of SEQ ID NOS:1, 3 and 5.
10. The isolated nucleic acid of claim 1 which encodes a polypeptide that has at least about 90% identity with one of SEQ ID NOS:1, 3 and 5.
11. A nucleic acid probe according to claim 1, comprising at least 14 contiguous nucleotides from one of SEQ ID NOS:2, 4 and 6.
12. An isolated recombinant polynucleotide molecule comprising nucleic acid according to claim 1 plus expression-controlling elements linked operably with said nucleic acid to drive expression thereof.
13. An expression vector comprising the nucleic acid of claim 1 encoding a polypeptide having the entire amino acid sequence set forth in any one of SEQ ID NOS:1, 3 and 5 operably linked to a promoter, said expression vector being present in a compatible host cell.

WO 02/31134

PCT/US01/31874

14. A mammalian, insect or bacterial host cell that has been genetically engineered by the insertion of nucleic acid according to claim 1 which codes for at least the mature protein portion of the amino acid sequence of SEQ ID NO:1, 3 or 5.

15. A process for producing a polypeptide which includes the mature protein portion of one of SEQ ID NOS:1, 3 and 5, which process comprises culturing the host cell of claim 11 under conditions sufficient for the production of said polypeptide.

16. The process of claim 15 wherein said polypeptide is expressed at the surface of said cell and further includes the step of recovering the polypeptide or a fragment thereof from the culture.

17. A polypeptide which may be optionally glycosylated, and which (a) has the amino acid sequence of a mature protein set forth in any one of SEQ ID NOS:1, 3 and 5; (b) has the amino acid sequence of a mature protein having at least about 70% similarity to one of the mature proteins of (a) and which exhibits the same biological function; (c) has the amino acid sequence of a mature protein having at least about 90% identity with a mature protein of any of SEQ ID NOS:1, 3 and 5; or (d) is an immunologically reactive fragment of (a).

18. The polypeptide according to claim 14 which is a mature protein having at least about 95% similarity to a mature protein of (a).

19. The polypeptide according to claim 14 which is a mature protein having at least about 95% similarity to a mature protein of (a).

20. The polypeptide according to claim 14 having the amino acid sequence of the mature protein of one of SEQ ID NOS:1, 3 and 5, or is a fragment thereof which exhibits the same biological function as the respective mature protein.

21. A DPRP antagonist which inhibits the biological function of one of said mature proteins of claim 17, 18 and 19.

22. An antibody that recognizes a polypeptide or a fragment according to claim 17.

23. The antibody of claim 22 which recognizes a polypeptide having an amino acid sequence of SEQ ID NO:1 or 3 or 5.

24. A method for the screening for a compound capable of inhibiting the enzymatic activity of at least one mature protein of claim 17, which method comprises incubating said mature protein and a suitable substrate for said mature protein in the presence of one or more test compounds or salts thereof, measuring the enzymatic activity of said mature protein, comparing said activity with comparable activity determined in the

WO 02/31134

PCT/US01/31874

absence of a test compound, and selecting the test compound or compounds that reduce the enzymatic activity.

25. A method for the screening for a compound capable of inhibiting the enzymatic activity of DPPIV that does not inhibit the enzymatic activity of at least one of the mature proteins of claim 20, which method comprises incubating said mature protein and a suitable substrate for said mature protein in the presence of one or more inhibitors of DPPIV or salts thereof, measuring the enzymatic activity of said mature protein, comparing said activity with comparable activity determined in the absence of the DPPIV inhibitor, and selecting a compound that does not reduce the enzymatic activity of said mature protein.

WO 02/31134

PCT/US01/31874

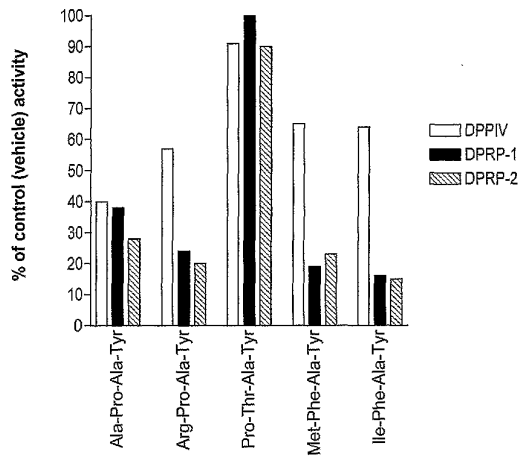


FIG. 3

WO 02/31134

PCT/US01/31874

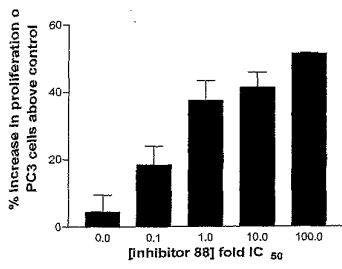


FIG. 4A

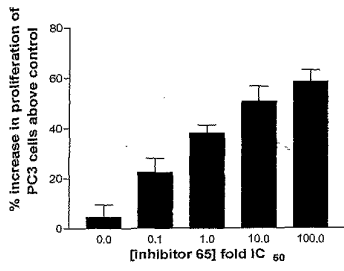


FIG. 4B

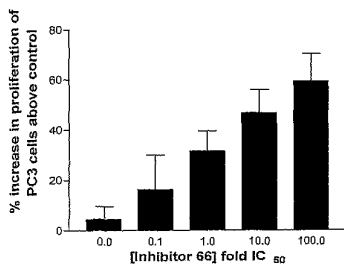


FIG. 4C

WO 02/31134

PCT/US01/31874

Sequence Listing Summary

SEQ ID.	
	1. DPRP1 a.a. sequence
	2. DPRP1 DNA sequence
5	3. DPRP2 a.a. sequence
	4. DPRP2 DNA sequence
	5. DPRP-3 a.a. sequence
	6. DPRP-3 DNA sequence
	7. DPRP-1 transcript 0 a.a. sequence
10	8. DPRP-1 transcript 0 DNA sequence
	9. DPRP-1 transcript 1 a.a. sequence
	10. DPRP-1 transcript 1 DNA sequence
	11. DPRP-1 transcript 2 a.a. sequence
	12. DPRP-1 transcript 2 DNA sequence
15	13. DPRP-1 transcript 3 a.a. sequence
	14. DPRP-1 transcript 3 DNA sequence
	15. DPRP-1 transcript 4 a.a. sequence
	16. DPRP-1 transcript 4 DNA sequence
	17. DPRP-1 transcript 5 a.a. sequence
20	18. DPRP-1 transcript 5 DNA sequence
	19. DPRP-1 transcript 6 a.a. sequence
	20. DPRP-1 transcript 6 DNA sequence
	21. DPRP-1 transcript 7 a.a. sequence
	22. DPRP-1 transcript 7 DNA sequence
25	23. DPRP-2 transcript 0 a.a. sequence
	24. DPRP-2 transcript 0 DNA sequence
	25. DPRP-2 transcript 1 a.a. sequence
	26. DPRP-2 transcript 1 DNA sequence
	27. DPRP-2 transcript 2 a.a. sequence
30	28. DPRP-2 transcript 2 DNA sequence
	29. DPRP-2 transcript 3 a.a. sequence
	30. DPRP-2 transcript 3 DNA sequence
	31. DPRP-2 transcript 4 a.a. sequence
	32. DPRP-2 transcript 4 DNA sequence
35	33. DPRP-2 transcript 5 a.a. sequence

WO 02/31134

PCT/US01/31874

- 34. DPRP-2 transcript 5 DNA sequence
- 35. DPRP-2 transcript 6 a.a. sequence
- 36. DPRP-2 transcript 6 DNA sequence
- 37. DPRP-2 transcript 7 a.a. sequence
- 5 38. DPRP-2 transcript 7 DNA sequence
- 39. DPRP-2 transcript 8 a.a. sequence
- 40. DPRP-2 transcript 8 DNA sequence
- 41. DPRP-3 transcript 0 a.a. sequence
- 42. DPRP-3 transcript 0 DNA. Sequence
- 10 43. DPRP-3 transcript 1 a.a. sequence
- 44. DPRP-3 transcript 1 DNA sequence
- 45. DPRP1 forward primer used for cloning
- 46. DPRP1 reverse primer used for cloning full length gene
- 47. DPRP1 reverse primer used for cloning fusion gene
- 15 48. DPRP1 forward primer used for expression profiling
- 49. DPRP1 reverse primer used for expression profiling
- 50. DPRP2 forward primer used for cloning
- 51. DPRP2 reverse primer used for cloning full length gene
- 52. DPRP2 reverse primer used for cloning fusion gene
- 20 53. DPRP2 forward primer used for expression profiling
- 54. DPRP2 reverse primer used for expression profiling
- 55. DPRP3 forward primer used for cloning
- 56. DPRP3 reverse primer used for cloning full length gene
- 57. DPRP3 forward primer used for cloning fusion gene
- 25 58. DPRP3 reverse primer used for cloning fusion gene
- 59. DPRP1 peptide antigen sequences
- 60. DPRP2 peptide antigen sequences
- 61. DPRP3 peptide antigen sequences

WO 02/31134

PCT/US01/31874

SEQUENCE LISTING

<110> Qi, Steve
Akinsanya, Karen
Riviere, Pierre
Junien, Jean-Louis

<120> NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

<130> 70669

<150> US 60/240,117

<151> 2000-10-12

<160> 61

<170> Patent In version 3.1

<210> 1
<211> 882
<212> PRT
<213> Homo sapiens
<400> 1

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

WO 02/31134

PCT/US01/31874

Leu Gln Glu Glu Phe Asp Arg Tyr Ser Gly Tyr Trp Trp Cys Pro Lys
 260 265 270
 Ala Glu Thr Thr Pro Ser Gly Gly Lys Ile Leu Arg Ile Leu Tyr Glu
 275 280 285
 Glu Asn Asp Glu Ser Glu Val Glu Ile Ile His Val Thr Ser Pro Met
 290 295 300
 Leu Glu Thr Arg Arg Ala Asp Ser Phe Arg Tyr Pro Lys Thr Gly Thr
 305 310 315 320
 Ala Asn Pro Lys Val Thr Phe Lys Met Ser Glu Ile Met Ile Asp Ala
 325 330 335
 Glu Gly Arg Ile Ile Asp Val Ile Asp Lys Glu Leu Ile Gln Pro Phe
 340 345 350
 Glu Ile Leu Phe Glu Gly Val Glu Tyr Ile Ala Arg Ala Gly Trp Thr
 355 360 365
 Pro Glu Gly Lys Tyr Ala Trp Ser Ile Leu Leu Asp Arg Ser Glu Thr
 370 375 380
 Arg Leu Gln Ile Val Leu Ile Ser Pro Glu Leu Phe Ile Pro Val Glu
 385 390 395 400
 Asp Asp Val Met Glu Arg Gln Arg Leu Ile Glu Ser Val Pro Asp Ser
 405 410 415
 Val Thr Pro Leu Ile Ile Tyr Glu Glu Thr Thr Asp Ile Trp Ile Asn
 420 425 430
 Ile His Asp Ile Phe His Val Phe Pro Gln Ser His Glu Glu Ile
 435 440 445
 Glu Phe Ile Phe Ala Ser Glu Cys Lys Thr Gly Phe Arg His Leu Tyr
 450 455 460
 Lys Ile Thr Ser Ile Leu Lys Glu Ser Lys Tyr Lys Arg Ser Ser Gly
 465 470 475 480
 Gly Leu Pro Ala Pro Ser Asp Phe Lys Cys Pro Ile Lys Glu Glu Ile
 485 490 495
 Ala Ile Thr Ser Gly Glu Trp Glu Val Leu Gly Arg His Gly Ser Asn
 500 505 510
 Ile Gln Val Asp Glu Val Arg Arg Leu Val Tyr Phe Glu Gly Thr Lys
 515 520 525
 Asp Ser Pro Leu Glu His His Leu Tyr Val Val Ser Tyr Val Asn Pro
 530 535 540
 Gly Glu Val Thr Arg Leu Thr Asp Arg Gly Tyr Ser His Ser Cys Cys
 545 550 555 560
 Ile Ser Gln His Cys Asp Phe Phe Ile Ser Lys Tyr Ser Asn Gln Lys
 565 570 575
 Asn Pro His Cys Val Ser Leu Tyr Lys Leu Ser Ser Pro Glu Asp Asp
 580 585 590
 Pro Thr Cys Lys Thr Lys Glu Phe Trp Ala Thr Ile Leu Asp Ser Ala
 595 600 605
 Gly Pro Leu Pro Asp Tyr Thr Pro Pro Glu Ile Phe Ser Phe Glu Ser
 610 615 620
 Thr Thr Gly Phe Thr Leu Tyr Gly Met Leu Tyr Lys Pro His Asp Leu
 625 630 635 640
 Gln Pro Gly Lys Lys Tyr Pro Thr Val Leu Phe Ile Tyr Gly Gly Pro
 645 650 655
 Gln Val Gln Leu Val Asn Asn Arg Phe Lys Gly Val Lys Tyr Phe Arg
 660 665 670
 Leu Asn Thr Leu Ala Ser Leu Gly Tyr Val Val Val Val Ile Asp Asn
 675 680 685
 Arg Gly Ser Cys His Arg Gly Leu Lys Phe Glu Gly Ala Phe Lys Tyr
 690 695 700
 Lys Met Gly Gln Ile Glu Ile Asp Asp Gln Val Glu Gly Leu Gln Tyr
 705 710 715 720

WO 02/31134

PCT/US01/31874

Leu Ala Ser Arg Tyr Asp Phe Ile Asp Leu Asp Arg Val Gly Ile His
725 730 735
Gly Trp Ser Tyr Gly Gly Tyr Leu Ser Leu Met Ala Leu Met Gln Arg
740 745 750
Ser Asp Ile Phe Arg Val Ala Ile Ala Gly Ala Pro Val Thr Leu Trp
755 760 765
Ile Phe Tyr Asp Thr Gly Tyr Thr Glu Arg Tyr Met Gly His Pro Asp
770 775 780
Gln Asn Glu Gln Gly Tyr Tyr Leu Gly Ser Val Ala Met Gln Ala Glu
785 790 795 800
Lys Phe Pro Ser Glu Pro Asn Arg Leu Leu Leu His Gly Phe Leu
805 810 815
Asp Glu Asn Val His Phe Ala His Thr Ser Ile Leu Leu Ser Phe Leu
820 825 830
Val Arg Ala Gly Lys Pro Tyr Asp Leu Gln Ile Tyr Pro Gln Glu Arg
835 840 845
His Ser Ile Arg Val Pro Glu Ser Gly Glu His Tyr Glu Leu His Leu
850 855 860
Leu His Tyr Leu Gln Glu Asn Leu Gly Ser Arg Ile Ala Ala Leu Lys
865 870 875 880
Val Ile

<210> 2
<211> 2671
<212> DNA
<213> Homo sapiens
<400> 2

cggtaccatg gcagcagcaa tggaaacaga acagctgggt gttgagatat ttgaaactgc 60
ggactgtgag gagaatattg aatcacagga tcggcctaaa ttggagcctt tttatgttga 120
gcggtattcc tggagtcagc ttaaaaaagc gcttgccgat accagaaaat atcatggcta 180
catgatggct aaggcaccac atgatttcat gtttgtgaag aggaatgato cagatggacc 240
tcattcagac agaactctatt accttggcat gtctggtagg aacagagaaa atacactgtt 300
ttattctgaa attcccacaaa ctatcaatag agcagcagtc ttaaltgctct cttggaagcc 360
tcctttggat ctttttcagg caacactgga ctatggaatg tattctcgag aagaagaact 420
attaagagaa agaaaaagca ttggaacagt cggaaatgct tcttaccgatt atcaccagg 480
aagtggaaaca tttctgtttc aagccggtag tggaaatgat cagtaaaag atggagggcc 540
acaagatttt ccgaacaac ctttaaggcc caatctagtg gaaatagtt gtcccacat 600
acgatattgg atacttaaca tcgtaaccag agaagaaagg agactcactt atgtgcacaa 660
tgaactagcc aacatggaag aagatgccc atcagctgga gtccctacct ttgtctcca 720
agaagaattt gatagatatt ctggctattg gtggtgtcca aaagctgaaa caactcccag 780
tggtggtaaa attcttagaa ttctatatga agaaatgat gaactcgagg tggaaattat 840
tcattgtaca tcccctatgt tggaaacaag gagggccgat tcattccgtt atcctaaaac 900
aggtacagca aatcctaagb tcaactttaa gatgtcagaa ataagattg atgtggaag 960
aaggatcata gatgtcatag ataaggaact aattcaacct tttgagattc tatttgaag 1020
agttgaatat attgccagag ctggatggac tcttgaggga aatattgctt ggtccatcct 1080
actagatcgc tcccagactc gccctgcgat agtgttgatc tcactgcaat tatttatccc 1140
agtagaagat gatgttatgg aaaggcagag actcattgag tcactgctct atctctgac 1200
gcactaatt atctatgaag aaacaacaga catctggata aatatccatg acatcttca 1260
tggttttccc caagtcacg aagaggaat tgagtttatt tttgctctg aatgcaaac 1320
aggtttccgt cattataca aaattacatc tattttaaag gaaagcaat ataacgatc 1380
cagtggtggg ctgctgctc caagtgtatt caagtgtcct atcaagagg agatagcaat 1440
taccagtggt gaatgggaag ttcttgccc goatggatct aatatccaag ttgatgaagt 1500
cagaagcctg gtatattttg aaggcaccac agactccctc tttagcatc accctgact 1560
agtcagttac gtaaatcctg gagaggtgac aaggctgact gaccgtggct actcacatc 1620
tgtctgcatc agtcagcact gtgacttctt tataagtaag tatagtaac agaagaatc 1680
acactgttg tccttttaca agctatcaag tctgtaagat gacccaact gcaaaaacaa 1740
ggaattttgg gccaccattt tggattcagc aggtcctctt cctgactata ctctccaga 1800
aattttctct tttgaaagta ctaactggatt tacattgtat gggatgctct acaagcctca 1860
1920

WO 02/31134

PCT/US01/31874

```

tgatctacaag cctggaaaga aatatcctac tgtgctgttc atatatggtg gtcctcaggt 1980
gcagttgggtg aataatcgat ttaaaggagt caagtatttc cgcttgaata ccctagcctc 2040
tctaggttat gtggtttag tagatagaca caggggatcc tgtcacgag ggcttaaat 2100
tgaaggcgc ttaaatata aatgggtca aatagaatt gcagtcagg tggaggact 2160
ccaatatcta gcttctcgat atgatttcat tgacttagat cgtgtggca tccacggctg 2220
gtcctatgga ggatacctct cctctgctgc attaatgcoag aggtcagata tcttcagggt 2280
tgctattgct ggggcccocag tcaactctgt gatcttttat gatcacagat acacggcaac 2340
ttatattgggt caccctgacc agaatagaaca gggctattac ttaggatctg tggccatgca 2400
agcagaaaag ttcccctctg acccaaatcg tttactgctc ttacatggtt tcctggatga 2460
gaatgtccat tttgcacata ccagtatatt actgagtttt ttagtgggg ctggaagcc 2520
atatgattta cagatctatc ctcaaggagag acacagcata agagtctctg aatcgggaga 2580
acattatgaa ctgcatcttt tgcactacct tcaagaaaac cttggatcac gtattgctgc 2640
tctaaaagtg atatgagcgg ccgagagctc c 2671

```

```

<210> 3
<211> 863
<212> PRT
<213> Homo sapiens
<400> 3

```

```

Met Ala Thr Thr Gly Thr Pro Thr Ala Asp Arg Gly Asp Ala Ala Ala
1 5 10 15
Thr Asp Asp Pro Ala Ala Arg Phe Gln Val Gln Lys His Ser Trp Asp
20 25 30
Gly Leu Arg Ser Ile Ile His Gly Ser Arg Lys Tyr Ser Gly Leu Ile
35 40 45
Val Asn Lys Ala Pro His Asp Phe Gln Phe Val Gln Lys Thr Asp Glu
50 55 60
Ser Gly Pro His Ser His Arg Leu Tyr Tyr Leu Gly Met Pro Tyr Gly
65 70 75
Ser Arg Glu Asn Ser Leu Leu Tyr Ser Glu Ile Pro Lys Lys Val Arg
85 90 95
Lys Glu Ala Leu Leu Leu Ser Trp Lys Gln Met Leu Asp His Phe
100 105 110
Gln Ala Thr Pro His His Gly Val Tyr Ser Arg Glu Glu Leu Leu
115 120 125
Arg Glu Arg Lys Arg Leu Gly Val Phe Gly Ile Thr Ser Tyr Asp Phe
130 135 140
His Ser Glu Ser Gly Leu Phe Leu Phe Gln Ala Ser Asn Ser Leu Phe
145 150 155
His Cys Arg Asp Gly Gly Lys Asn Gly Phe Met Val Ser Pro Met Lys
165 170 175
Pro Leu Glu Ile Lys Thr Gln Cys Ser Gly Pro Arg Met Asp Pro Lys
180 185 190
Ile Cys Pro Ala Asp Pro Ala Phe Phe Ser Phe Ile Asn Asn Ser Asp
195 200 205
Leu Trp Val Ala Asn Ile Glu Thr Gly Glu Glu Arg Arg Leu Thr Phe
210 215 220
Cys His Gln Gly Leu Ser Asn Val Leu Asp Asp Pro Lys Ser Ala Gly
225 230 235
Val Ala Thr Phe Val Ile Gln Glu Glu Phe Asp Arg Phe Thr Gly Tyr
245 250 255
Trp Trp Cys Pro Thr Ala Ser Trp Glu Gly Ser Glu Gly Leu Lys Thr
260 265 270
Leu Arg Ile Leu Tyr Glu Glu Val Asp Glu Ser Glu Val Glu Val Ile
275 280 285
His Val Pro Ser Pro Ala Leu Glu Glu Arg Lys Thr Asp Ser Tyr Arg
290 295 300
Tyr Pro Arg Thr Gly Ser Lys Asn Pro Lys Ile Ala Leu Lys Leu Ala
305 310 315

```

WO 02/31134

PCT/US01/31874

Glu Phe Gln Thr Asp Ser Gln Gly Lys Ile Val Ser Thr Gln Glu Lys
 325 330 335
 Glu Leu Val Gln Pro Phe Ser Ser Leu Phe Pro Lys Val Glu Tyr Ile
 340 345 350
 Ala Arg Ala Gly Trp Thr Arg Asp Gly Lys Tyr Ala Trp Ala Met Phe
 355 360 365
 Leu Asp Arg Pro Gln Gln Trp Leu Gln Leu Val Leu Leu Pro Pro Ala
 370 375 380
 Leu Phe Ile Pro Ser Thr Glu Asn Glu Glu Gln Arg Leu Ala Ser Ala
 385 390 395 400
 Arg Ala Val Pro Arg Asn Val Gln Pro Tyr Val Val Tyr Glu Glu Val
 405 410 415
 Thr Asn Val Trp Ile Asn Val His Asp Ile Phe Tyr Pro Phe Pro Gln
 420 425 430
 Ser Glu Gly Glu Asp Glu Leu Cys Phe Leu Arg Ala Asn Glu Cys Lys
 435 440 445
 Thr Gly Phe Cys His Leu Tyr Lys Val Thr Ala Val Leu Lys Ser Gln
 450 455 460
 Gly Tyr Asp Trp Ser Glu Pro Phe Ser Pro Gly Glu Asp Glu Phe Lys
 465 470 475 480
 Cys Pro Ile Lys Glu Glu Ile Ala Leu Thr Ser Gly Glu Trp Glu Val
 485 490 495
 Leu Ala Arg His Gly Ser Lys Ile Trp Val Asn Glu Glu Thr Lys Leu
 500 505 510
 Val Tyr Phe Gln Gly Thr Lys Asp Thr Pro Leu Glu His His Leu Tyr
 515 520 525
 Val Val Ser Tyr Glu Ala Ala Gly Glu Ile Val Arg Leu Thr Thr Pro
 530 535 540
 Gly Phe Ser His Ser Cys Ser Met Ser Gln Asn Phe Asp Met Phe Val
 545 550 555 560
 Ser His Tyr Ser Ser Val Ser Thr Pro Pro Cys Val His Val Tyr Lys
 565 570 575
 Leu Ser Gly Pro Asp Asp Asp Pro Leu His Lys Gln Pro Arg Phe Trp
 580 585 590
 Ala Ser Met Met Glu Ala Ala Ser Cys Pro Pro Asp Tyr Val Pro Pro
 595 600 605
 Glu Ile Phe His Phe His Thr Arg Ser Asp Val Arg Leu Tyr Gly Met
 610 615 620
 Ile Tyr Lys Pro His Ala Leu Gln Pro Gly Lys Lys His Pro Thr Val
 625 630 635 640
 Leu Phe Val Tyr Gly Gly Pro Gln Val Gln Leu Val Asn Asn Ser Phe
 645 650 655
 Lys Gly Ile Lys Tyr Leu Arg Leu Asn Thr Leu Ala Ser Leu Gly Tyr
 660 665 670
 Ala Val Val Val Ile Asp Gly Arg Gly Ser Cys Gln Arg Gly Leu Arg
 675 680 685
 Phe Glu Gly Ala Leu Lys Asn Gln Met Gly Gln Val Glu Ile Glu Asp
 690 695 700
 Gln Val Gln Gly Leu Gln Phe Val Ala Glu Lys Tyr Gly Phe Ile Asp
 705 710 715 720
 Leu Ser Arg Val Ala Ile His Gly Trp Ser Tyr Gly Gly Phe Leu Ser
 725 730 735
 Leu Met Gly Leu Ile His Lys Pro Gln Val Phe Lys Val Ala Ile Ala
 740 745 750
 Gly Ala Pro Val Thr Val Trp Met Ala Tyr Asp Thr Gly Tyr Thr Glu
 755 760 765

WO 02/31134

PCT/US01/31874

Arg Tyr Met Asp Val Pro Glu Asn Asn Gln His Gly Tyr Glu Ala Gly
 770 775 780
 Ser Val Ala Leu His Val Glu Lys Leu Pro Asn Glu Pro Asn Arg Leu
 785 790 795 800
 Leu Ile Leu His Gly Phe Leu Asp Glu Asn Val His Phe Phe His Thr
 805 810 815
 Asn Phe Leu Val Ser Gln Leu Ile Arg Ala Gly Lys Pro Tyr Gln Leu
 820 825 830
 Gln Ile Tyr Pro Asn Glu Arg His Ser Ile Arg Cys Pro Glu Ser Gly
 835 840 845
 Glu His Tyr Glu Val Thr Leu Leu His Phe Leu Gln Glu Tyr Leu
 850 855 860

<210> 4
 <211> 2617
 <212> DNA
 <213> Homo sapiens
 <400> 4

```

caagcttacc atggccacca ccgggacccc aacggccgac cgaggcgacg cagccgccac 60
agatgacccg gcccccgcct tccaggtgca gaagcaactg tgggacgggc tccgggacat 120
catccaacgg agccgcaagt actcgggcct cattgtcaac aaggcgcccc acgacttcca 180
gtttgtgcag aagacggatg agtctgggccc ccactcccac cgcctctact acctgggaat 240
gccatattgg agccogagaga actccctcct ctactctgag attcccaga aggtccggaa 300
agaggctctg ctgctcctgt cctggaagca gatgctggat catttccagg ccacgcccga 360
ccatggggct tactctcggg aggaggagct gctgaggagc cggaaacgcc tgggggtctt 420
cggcatcaac tectacgact tccacagcga gactggcctc ttctcttccc aggcagcaaa 480
cagcctcttc cactgtcggc acggcggcaa gaacggcttc atggtgtccc ctatgaaacc 540
gctggaaatc aagaccagat gctcagggcc ccggatggac cccaaatctc gcctgcccga 600
cctgcctctc ttctccttca tcaataacag cgaactgtgg gtggccaaca tcgagacagg 660
cgaggagcgg cggctgacct tctgccaaca aggtttatcc aatgtcctgg atgaccocaa 720
gtctgcgggt gtggccacct tctgcataca ggaagagttc gaccgcttca ctgggtactg 780
gtggtgcccc acagcctcct gggaaagttc agagggcctc aagacgctgc gaatcctgta 840
tgaggaagtc gatgagtcgg aggtggaggt cattcaactc cctctcctg cgtatagaag 900
aaggaagcgg gactcgtatc ggtaccccag cacaggcagc aagaatccca agattgcctt 960
gaaactggct gacttccaga ctgacagcca gggcaagatc gtctcgacc aggagaagga 1020
gctgggtgcag ccttccagct cgtcgttccc gaaggtggag tacatggcca gggccgggtg 1080
gaccocggat ggcaatacag cctgggcaat gttcctggac cggcccacag agtggctcca 1140
gctcgtcttc ctcccocgg cctgttccat ccggagcaca gagaatgagg agcagcggct 1200
agcctctgoc agactgtcc ccaggaatgt ccagccttat gtaggtgacg aggaggtcac 1260
caacgtctgg atcaatgttc atgacatctt ctatcccttc ccccaatag agggagagga 1320
cgagctctgc ttctcogcg ccaatgaatg caagacgggc ttctgcoatt tgtacaaagt 1380
caccgcctgt ttaaaatccc agggctacga ttggagtgag cccttcagcc cgggggaaga 1440
tgaatttaag tgcccattta aggaagagat tgcctcgacc agcgggtaat gggaggtttt 1500
ggcagggcac ggctccaaga tctgggtcaa tgaggagacc aagctggtg acttccaggg 1560
caccaggac acgcccctgg agcaccacct ctactggctc agctatgagg cggccggcca 1620
gatcgtacgc ctaccacgcg ccggctcttc ccatagctgc tccatgagcc agaacttcca 1680
catgttctgc agccaactaca gcaagctgag caacgcoccc tgggtgcaag tctacaagct 1740
gagcggcccc gaccagcacc ccttgcacaa gcaaccccgc ttctgggta gcatgatgga 1800
ggcagccagc tgcccocggg attatgttcc tccagagatc ttccatttc acacgcgctc 1860
ggatgtgccc ctctacggca tgatctcaaa gcccccagcc ttgacgcaa ggaagaagca 1920
ccccaccctc ctctttgat atggaggccc ccagggtgag ctggtgaata actccttcaa 1980
agccatcaag tacttccggc tcaacacact ggcctccctg ggtcacgocg tggttgtgat 2040
tgacggcagg ggctcctgtc agcaggggct tcggttcgaa ggggcctgga aaaaccaaat 2100
ggccaggggt gagatcgagg accaggtgga gggcctgcag ttctggccc agaagtatgg 2160
cttcaatcac ctgagccgag ttgccatcca tggctggctc tacgggggct tcctctcgtc 2220
catggggcta atccacaagc ccaggtggtt caaggtggcc atcgggggtg ccccggtcac 2280
cgtctgggatg gcttacgaca cagggtacac tgagcgtcac atggaactc ctgagaacaa 2340
ccagcagcgc tatgagggcg gttccgtggc cctgcacgtg gagaagctgc ccaatgagcc 2400
caaccgctg cttatcctcc acggcttctt ggacgaaaac gtgcactttt tccacacaaa 2460

```

WO 02/31134

PCT/US01/31874

```

cttctcgtc toccaactga tccgagcagg gaaaccttac cagctccaga totaccocaa 2520
cgagagacac agtattcgtc gccccgagtc gggcgagcac tatgaagtca cgttgctgca 2580
cttctacag gaataactct gagcggcgc ggcacccg 2617

```

```

<210> 5
<211> 796
<212> PRT
<213> Homo sapiens
<400> 5

```

```

Met Asn Gln Thr Ala Ser Val Ser His His Ile Lys Cys Gln Pro Ser
1 5 10 15
Lys Thr Ile Lys Glu Leu Gly Ser Asn Ser Pro Pro Gln Arg Asn Trp
20 25 30
Lys Gly Ile Ala Ile Ala Leu Leu Val Ile Leu Val Val Cys Ser Leu
35 40 45
Ile Thr Met Ser Val Ile Leu Leu Ser Pro Asp Glu Leu Thr Asn Ser
50 55 60
Ser Glu Thr Arg Leu Ser Leu Glu Asp Leu Phe Arg Lys Asp Phe Val
65 70 75 80
Leu His Asp Pro Glu Ala Arg Trp Ile Asn Asp Thr Asp Val Val Tyr
85 90 95
Lys Ser Glu Asn Gly His Val Ile Lys Leu Asn Ile Glu Thr Asn Ala
100 105 110
Thr Thr Leu Leu Leu Glu Asn Thr Thr Phe Val Thr Phe Lys Ala Ser
115 120 125
Arg His Ser Val Ser Pro Asp Leu Lys Tyr Val Leu Leu Ala Tyr Asp
130 135 140
Val Lys Gln Ile Phe His Tyr Ser Tyr Thr Ala Ser Tyr Val Ile Tyr
145 150 155 160
Asn Ile His Thr Arg Glu Val Trp Glu Leu Asn Pro Pro Glu Val Glu
165 170 175
Asp Ser Val Leu Gln Tyr Ala Ala Trp Gly Val Gln Gly Gln Gln Leu
180 185 190
Ile Tyr Ile Phe Glu Asn Asn Ile Tyr Tyr Gln Pro Asp Ile Lys Ser
195 200 205
Ser Ser Leu Arg Leu Thr Ser Ser Gly Lys Glu Glu Ile Ile Phe Asn
210 215 220
Gly Ile Ala Asp Trp Leu Tyr Glu Glu Glu Leu Leu His Ser His Ile
225 230 235 240
Ala His Trp Trp Ser Pro Asp Gly Glu Arg Leu Ala Phe Leu Met Ile
245 250 255
Asn Asp Ser Leu Val Pro Thr Met Val Ile Pro Arg Phe Thr Gly Ala
260 265 270
Leu Tyr Pro Lys Gly Lys Gln Tyr Pro Tyr Pro Lys Ala Gly Gln Val
275 280 285
Asn Pro Thr Ile Lys Leu Tyr Val Val Asn Leu Tyr Gly Pro Thr His
290 295 300
Thr Leu Glu Leu Met Pro Pro Asp Ser Phe Lys Ser Arg Glu Tyr Tyr
305 310 315 320
Ile Thr Met Val Lys Trp Val Ser Asn Thr Lys Thr Val Val Arg Trp
325 330 335
Leu Asn Arg Pro Gln Asn Ile Ser Ile Leu Thr Val Cys Glu Thr Thr
340 345 350
Thr Gly Ala Cys Ser Lys Lys Tyr Glu Met Thr Ser Asp Thr Trp Leu
355 360 365
Ser Gln Gln Asn Glu Glu Pro Val Phe Ser Arg Asp Gly Ser Lys Phe
370 375 380
Phe Met Thr Val Pro Val Lys Gln Gly Arg Gly Glu Phe His His
385 390 395 400

```

WO 02/31134

PCT/US01/31874

```

Ile Ala Met Phe Leu Ile Gln Ser Lys Ser Glu Gln Ile Thr Val Arg
405 410 415
His Leu Thr Ser Gly Asn Trp Glu Val Ile Lys Ile Leu Ala Tyr Asp
420 425 430
Glu Thr Thr Gln Lys Ile Tyr Phe Leu Ser Thr Glu Ser Ser Pro Arg
435 440 445
Gly Arg Gln Leu Tyr Ser Ala Ser Thr Glu Gly Leu Leu Asn Arg Gln
450 455 460
Cys Ile Ser Cys Asn Phe Met Lys Glu Gln Cys Thr Tyr Phe Asp Ala
465 470 475 480
Ser Phe Ser Pro Met Asn Gln His Phe Leu Leu Phe Cys Glu Gly Pro
485 490 495
Arg Val Pro Val Val Ser Leu His Ser Thr Asp Asn Pro Ala Lys Tyr
500 505 510
Phe Ile Leu Glu Ser Asn Ser Met Leu Lys Glu Ala Ile Leu Lys Lys
515 520 525
Lys Ile Gly Lys Pro Glu Ile Lys Ile Leu His Ile Asp Asp Tyr Glu
530 535 540
Leu Pro Leu Gln Leu Ser Leu Pro Lys Asp Phe Met Asp Arg Asn Gln
545 550 555 560
Tyr Ala Leu Leu Leu Ile Met Asp Glu Glu Pro Gly Gly Gln Leu Val
565 570 575
Thr Asp Lys Phe His Ile Asp Trp Asp Ser Val Leu Ile Asp Met Asp
580 585 590
Asn Val Ile Val Ala Arg Phe Asp Gly Arg Gly Ser Gly Phe Gln Gly
595 600 605
Leu Lys Ile Leu Gln Glu Ile His Arg Arg Leu Gly Ser Val Glu Val
610 615 620
Lys Asp Gln Ile Thr Ala Val Lys Phe Leu Leu Lys Leu Pro Tyr Ile
625 630 635 640
Asp Ser Lys Arg Leu Ser Ile Phe Gly Lys Gly Tyr Gly Tyr Ile
645 650 655
Ala Ser Met Ile Leu Lys Ser Asp Glu Lys Leu Phe Lys Cys Gly Ser
660 665 670
Val Val Ala Pro Ile Thr Asp Leu Lys Leu Tyr Ala Ser Ala Phe Ser
675 680 685
Glu Arg Tyr Leu Gly Met Pro Ser Lys Glu Glu Ser Thr Tyr Gln Ala
690 695 700
Ala Ser Val Leu His Asn Val His Gly Leu Lys Glu Glu Asn Ile Leu
705 710 715 720
Ile Ile His Gly Thr Ala Asp Thr Lys Val His Phe Gln His Ser Ala
725 730 735
Glu Leu Ile Lys His Leu Ile Lys Ala Gly Val Asn Tyr Thr Met Gln
740 745 750
Val Tyr Pro Asp Glu Gly His Asn Val Ser Glu Lys Ser Lys Tyr His
755 760 765
Leu Tyr Ser Thr Ile Leu Lys Phe Phe Ser Asp Cys Leu Lys Glu Glu
770 775 780
Ile Ser Val Leu Pro Gln Glu Pro Glu Glu Asp Glu
785 790 795

```

```

<210> 6
<211> 2583
<212> DNA
<213> Homo sapiens
<400> 6

```

```

gctggggatt gtgcactgtc cagggtcctg aaacatgaac caaactgccca gcgtgtccca 60
tcaatcaag tgtcaacct caaaacaat caaggaactg ggaagtaaca gcctccaca 120
gagaactgg aagggaattg ctattgtctt gctgggtgatt ttagttgtat gctcactcat 180

```

WO 02/31134

PCT/US01/31874

```

cactatgtca gtcctcctct taagcccaga tgaactcaca aattcgtcag aaaccagatt 240
gtcctttggaa gacctcttta ggaagagact tgtgcttcaac gatccagagg ctcggtggat 300
caatgataca gatgtgggtg ataaaagcga gaatggaacat gtcattaaac tgaatataga 360
aacaatagct accacattat tattggaaaa cacaaactttt gtaaccttca aagcatcaag 420
acattcagtt tcaccagatt taaaatagt cctcttgcca tatgatgtca aacagatttt 480
tcattatcgg tatactgctt catatgtgat ttacaacata cacactaggg aagtttggga 540
gttaaatccc ccagaagtag aggaactccgt cttgcagtac gcgccctggg gtgtccaagg 600
gcacgagctg atttatattt ttgaaaaata tatctactat caacctgata taaagagcag 660
ttcattgcca ctgacatctt ctggaaaaga agaataaatt ttaaatgga ttgtcagctg 720
gttatatgaa gaggaactcc tgcattctca catcgcccac tgggtggtcac cagatggaga 780
aagacttgcc ttctgatgga taaatgactc tttggtaccc accatgggta tcctcggtt 840
tactggagcg ttgtatccca aaggaagaca gtatcogtat cctaaggcag gtcaagttaa 900
ccccacaata aaattatag ttgtaaacct gtatggacca actcacactt tggagctcat 960
gccacctgac agctttaaat caagagaata ctatcacact atggttaaat gggtaagcaa 1020
taccaagact ggttaagat ggttaaacct acctcagaac atctccatcc tcacagctcg 1080
tgagaccact acaggtgctt gtatgaaaaa atatgagatg acatcagata cgtggctctc 1140
tcacgagaaat gaggagcccg tgtttcttag agacggcagc aaattcttta tgacagtgcc 1200
tgtaagcaaa gggggagctg gagaatttca ccacatagct atgttctcca tccagagtaa 1260
aagtggacaa attaccgtgc ggcactgac atcaggaaac tgggaagtga taaagatctt 1320
ggcatacagat gaactactc aaaaaattta cttctgagc actgactctt ctcccagagg 1380
aagcagctg tacagtgctt ctactgaagg attattgaat cgccaatgca tttcatgtaa 1440
tttcataaaa gaacaatgta catattttga tgcagtttt agtccatgca atcaacattt 1500
cttatattc tgtgaaggtc caaggttccc agtggctcag ctacatagta cggacaaccc 1560
agcaaaat tttatattgg aaagcaatct batgctgaag gaagctatcc tgaagaagaa 1620
gataggaaag ccagaataa aaatccttca tattgacgac tatgaaactc ctttacagtt 1680
gtcccttccc aaagatttta tggaccgaaa ccagatgct cttctgttaa taatggatga 1740
agaaccagga ggcagctgg ttacagataa gttccatatt gactgggatt ccgactcat 1800
tgacatggat aatgtcattg tagcaagatt tgatggcaga ggaagtggat tccagggctc 1860
gaaaattttg caggagattc atcgaagatt aggttcagta gaagtaagg accaaataac 1920
agctgtgaaa tttttgctga aactgcctta cattgactcc aaaaatttaa gcattttgg 1980
aaaggttat ggtggtcaba ttgcatcaat gatcttaaaa tcagatgaaa agctttttaa 2040
atgtggatcc ggtggtgac ctatcacaga cttgaaattg tatgctcag ctttctctga 2100
aagatacctt gggatgccat ctaaggaaga aagcacttac caggcagcca gttgtctaca 2160
taattgtcat ggttgaaag aagaaaatatt atataaatt catgggaactg ctgacacaaa 2220
agttcatttc caacactcag cagaattaat caagcaccta ataaaagctg gagtgaatta 2280
tactatgoag gtctaccag atgaaggtca laactatct gagaagagca agtatcatct 2340
ctacagcaca atcctcaaat tcttcagtga ttgttgaag gaagaaalat ctgtctacc 2400
acaggaacca gaagaagatg aataatggac cgtatttata cagaactgaa gggaaatttg 2460
aggctcaatg aaacctgaca aagagactgt aatattgtag ttgctccaga atgtcaagg 2520
cagcttacgg agatgtcaat gtagcagcac gctcagagac agtgaactag catttgaata 2580
cac 2583

```

```

<210> 7
<211> 690
<212> PRT
<213> Homo sapiens
<400> 7

```

```

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

```

WO 02/31134

PCT/US01/31874

Leu Phe Tyr Ser Glu Ile Pro Lys Thr Ile Asn Arg Ala Ala Val Leu
 100 105 110
 Met Leu Ser Trp Lys Pro Leu Leu Asp Leu Phe Gln Ala Thr Leu Asp
 115 120 125
 Tyr Gly Met Tyr Ser Arg Glu Glu Glu Leu Leu Arg Glu Arg Lys Arg
 130 135 140
 Ile Gly Thr Val Gly Ile Ala Ser Tyr Asp Tyr His Gln Gly Ser Gly
 145 150 155 160
 Thr Phe Leu Phe Gln Ala Gly Ser Gly Ile Tyr His Val Lys Asp Gly
 165 170 175
 Gly Pro Gln Gly Phe Thr Gln Gln Pro Leu Arg Pro Asn Leu Val Glu
 180 185 190
 Thr Ser Cys Pro Asn Ile Arg Met Asp Pro Lys Leu Cys Pro Ala Asp
 195 200 205
 Pro Asp Trp Ile Ala Phe Ile His Ser Asn Asp Ile Trp Ile Ser Asn
 210 215 220
 Ile Val Thr Arg Glu Glu Arg Arg Leu Thr Tyr Val His Asn Glu Leu
 225 230 235 240
 Ala Asn Met Glu Glu Asp Ala Arg Ser Ala Gly Val Ala Thr Phe Val
 245 250 255
 Leu Gln Glu Glu Phe Asp Arg Tyr Ser Gly Tyr Trp Trp Cys Pro Lys
 260 265 270
 Ala Glu Thr Thr Pro Ser Gly Gly Lys Ile Leu Arg Ile Leu Tyr Glu
 275 280 285
 Glu Asn Asp Glu Ser Glu Val Glu Ile Ile His Val Thr Ser Pro Met
 290 295 300
 Leu Glu Thr Arg Arg Ala Asp Ser Phe Arg Tyr Pro Lys Thr Gly Thr
 305 310 315 320
 Ala Asn Pro Lys Val Thr Phe Lys Met Ser Glu Ile Met Ile Asp Ala
 325 330 335
 Glu Gly Arg Ile Ile Asp Val Ile Asp Lys Glu Leu Ile Gln Pro Phe
 340 345 350
 Glu Ile Leu Phe Glu Gly Val Glu Tyr Ile Ala Arg Ala Gly Trp Thr
 355 360 365
 Pro Glu Gly Lys Tyr Ala Trp Ser Ile Leu Leu Asp Arg Ser Gln Thr
 370 375 380
 Arg Leu Gln Ile Val Leu Ile Ser Pro Glu Leu Phe Ile Pro Val Glu
 385 390 395 400
 Asp Asp Val Met Glu Arg Gln Arg Leu Ile Glu Ser Val Pro Asp Ser
 405 410 415
 Val Thr Pro Leu Ile Ile Tyr Glu Glu Thr Thr Asp Ile Trp Ile Asn
 420 425 430
 Ile His Asp Ile Phe His Val Phe Pro Gln Ser His Glu Glu Glu Ile
 435 440 445
 Glu Phe Ile Phe Ala Ser Glu Cys Lys Thr Gly Phe Arg His Leu Tyr
 450 455 460
 Lys Ile Thr Ser Ile Leu Lys Glu Ser Lys Tyr Lys Arg Ser Ser Gly
 465 470 475 480
 Gly Leu Pro Ala Pro Ser Asp Phe Lys Cys Pro Ile Lys Glu Glu Ile
 485 490 495
 Ala Ile Thr Ser Gly Glu Trp Glu Val Leu Gly Arg His Gly Ser Asn
 500 505 510
 Ile Gln Val Asp Glu Val Arg Arg Leu Val Tyr Phe Glu Gly Thr Lys
 515 520 525
 Asp Ser Pro Leu Glu His His Leu Tyr Val Val Ser Tyr Val Asn Pro
 530 535 540
 Gly Glu Val Thr Arg Leu Thr Asp Arg Gly Tyr Ser His Ser Cys Cys
 545 550 555 560
 Ile Ser Gln His Cys Asp Phe Phe Ile Ser Lys Tyr Ser Asn Gln Lys
 565 570 575

WO 02/31134

PCT/US01/31874

Asn Pro His Cys Val Ser Leu Tyr Lys Leu Ser Ser Pro Glu Asp Asp
 580 585 590
 Pro Thr Cys Lys Thr Lys Glu Phe Trp Ala Thr Ile Leu Asp Ser Ala
 595 600 605
 Gly Pro Leu Pro Asp Tyr Thr Pro Pro Glu Ile Phe Ser Phe Glu Ser
 610 615 620
 Thr Thr Gly Phe Thr Leu Tyr Gly Met Leu Tyr Lys Pro His Asp Leu
 625 630 635 640
 Gln Pro Gly Lys Lys Tyr Pro Thr Val Leu Phe Ile Tyr Gly Gly Arg
 645 650 655
 Leu Leu Leu Leu Gly Pro Gln Ser Leu Cys Gly Ser Ser Met Ile Gln
 660 665 670
 Asp Thr Arg Asn Val Ile Trp Val Thr Leu Thr Arg Met Asn Arg Ala
 675 680 685
 Ile Thr
 690

<210> 8
 <211> 4523
 <212> DNA
 <213> Homo sapiens
 <400> 8

aagtgcataaa gctccgagg ccaagggcgc tgctactgcc ggcgctgctt cttagtgcgc 60
 cgttgcgcgc ctgggttgtc accggcgccg ccgcccaggga agccactgca accgggaccg 120
 gagtggaggc ggccgagcat gaagcggcgc aggcccgctc catagcgcac gtcgggaccg 180
 tccggcgccg gccgggggga aggaaaatgc aacatggcag cagcaatgga aacagaacag 240
 ctgggtgttg agatatttga aactgcccgc tgtgaggaga atattgaatc acaggatcgg 300
 cctaaattgg agccttttta tgttgacggg tattcctgga gtcagcttaa aaagctcctt 360
 gccgatacca gaaaatatca tggctacatg atggctaagg caccacatga tttcatgttt 420
 gtgaagagga atgatccaga tggacctcat tcagacagaa tctattacct tggcatgtct 480
 ggtgagaaca gagaaaatc actgttttat tctgaaatcc ccaaaactat caatagagca 540
 gcagctctaa tgcctctctg gaagcctctt ttggatcttt ttcaggcaac actggactat 600
 ggaatgtatt ctcgagaaga agaactatta agagaaaaga aacgcatggy aacagtcgga 660
 attgctctct acgattatca ccaaggaagt ggaacatttc tgtttcaagc cggtagtggg 720
 atttatcaog taaaagatgy agggccacaa ggaattacgc aacaaccttt aagpcccaat 780
 ctagtggaaa ctagtgttcc caacatacgg atggatccaa aattatgccc tgcctgatca 840
 gactggattg cttttataca tagcaacgat atttggatat ctaacatgt aaccagagaa 900
 gaaaggagac tcaactatgt gcaacatgag ctgcccacaa tggagaaga tgcagatca 960
 gctggagtcg ctacctttgt tctccaagaa gaatttgata gatattctgg ctattgttgy 1020
 tgtccaaaag ctgaacacac tcccagtggt ggtaaaatc ttagaattct atatgagaa 1080
 aatgatgaat ctgaggtgga aattattcat gttacatccc ctatgttggg aacaaggagg 1140
 gcagattcat tccgttatcc taaaacaggt acagcaaatc ctaaatgca ttttaagatg 1200
 tcagaaataa tgattgatgc tgaaggagg atcatagatg tcatagataa ggaactaat 1260
 caacctttg agattctat tgaaggagtt gaabatattg ccagagctgg atggactcct 1320
 gagggaaaat atgcttggtc catcctaata gatcgctccc agactcgctt acagatagt 1380
 ttgatctcac ctgaattatt tatcccagta gaagatgatg ttatggaaag gcagagactc 1440
 attgatcag tgcctgatcc tgtgacgcc aataattatc atgagaagac aacagacatc 1500
 tggataaata tccatgacat ctttcatgtt ttcccocaaa gtcacgaaga ggaattgag 1560
 tttatttttg cctctgatg caaaacaggt ttccgctcatt tatacaaaat tcaactat 1620
 ttaaaggaaa gcaaatataa acgatccagt ggtggcctgc ctgctccaag tgatttcaag 1680
 tgtcctatca aagaggagat agcaattacc agtggatgat gggagttct tggcggcat 1740
 ggaatataa tccaagtga tgaagtcaag agcctggtat atttgaag caccagaagc 1800
 tccccttag agcatcaact gtcagtatgc agttactgaa atctcggaga ggtgacaagg 1860
 ctgactgaoc gtggctactc acattcttgc tgcactcagc agcactgtga cttcttata 1920
 agtaagtata gtaaccagaa gaatccacac tgtgtgtccc tttacaagct atcaagctc 1980
 gaagatgacc caacttgcaa acaaaaggaa ttttgggcca ccattttgga ttcagcaggt 2040
 cctctctcgc actatactcc tccagaat tctctctttg aaagtaacta tggatttaca 2100
 ttgatggga tgcctacaaa gctcctgat ctacagcctg gaaagaata tctactgtg 2160

WO 02/31134

PCT/US01/31874

```

ctgttcatat atgggtggtcg gttgctattg ctggggcccc agtcaactctg tggatcttct 2220
atgatacagg atacacggaa cgttatatgg gtcacacctga ccagaatgaa cagggtctatt 2280
acttaggacc tgbggccatg caagcagaaa agttccocctc tgaacocaaat cgtttactgc 2340
tcttacatgg tttcctggat gagaatgtcc attttgcaca taccagmata ttaactgagtt 2400
tttttagtgg gctgggaaag ccatatgatt tacagatcta tctctcaggag agacacagca 2460
taagagtccc tgaatcgga gaaacatgat aactgcactct tttgcaactac cttcaagaaa 2520
accttggatc acgtatttgc gctctaaaag tgatataaatt ttgacctgtg tagaactctc 2580
tgggtatacac tggctattta accaaatgag gaggttfaat caacagaaaa cacagaattg 2640
atcaatcacat tttgataacct gccatgtaac atctactcct gaaaataaat gtggtgccat 2700
cgagggtctc acggtttgtg gtgatatact aataccttaa cccocacatgc tcaaaatcaa 2760
atgatacata ttccctgagag acccagcaat accataagaa ttaactaaaa aaaaaaaaaa 2820
aaaaagacat tagcacocat tattcatact accctatttt cacttttaat agtattataa 2880
acctcatgaa cttaattagt gtatttttac agtatacttt tgagtttgtt aaaatatgat 2940
gatattagtg attggtttgg ttcagttcca gaattottga ctagtaccag atttgatagc 3000
acttaaatgt aattgaatag cttatgcttc attgcttggg catatccagc atgttatgaa 3060
ctaataacta ttaaacctga cttaacccagt cattcattaa taatttttca aggataaact 3120
agtggcctcc taaagacact tgttttggca ctgaccagtt ttagccaat ttaactctga 3180
ctagataaaa ataattctca tttttctttg atgatattaa cagagtgggc ttttctcttt 3240
gcataaaggc tagtaactgt atatgtagca tggatttaat tagtcatgat attgataatt 3300
acaggcagaa aatttttaat caaatgatta gagcttaaat atttgcagcc aagttttttt 3360
ttttctctta agaaaaggaa aaagtacaca ttaactagaa tctctcagaa aatttagtgg 3420
tgccagtttc catttgggat ttccttatta aaatattcta gaattttaag gagattgaag 3480
ggatcacacg tgggtggggg agacctgggt ttggggaaat acagagagaa gaggtggtaa 3540
gggcctgatt aaaaactaag cagaagtagt ttaacaaaaa atactcatga aaatgtttgg 3600
aaactgaaat ttaaaccaact gtaattataa ggaaaccaga atcaataaat cactgtcttg 3660
ccagcacagc tacagagtaa catgattcag gggaggaaaa gttccttaga gttactttta 3720
taattctttt ttttttctct cttaggttta gaaatcttac aaatttaaac tttatccttt 3780
taaaattatt tgaacataat ttagatattg taagcttaaa atacaatagt tttatagata 3840
ctcttttaac taaactaat cctggcagc coagggctct cttttttttt ttgggtgtaa 3900
aagcctgtaa acagtttttc tgaatgatca tgaacttttc ttggtttagc actaggattt 3960
agctatgaag agagctcaata ggctttcagg tgctaattga gatctgacct gttagagctc 4020
tgggttgcta gattggtcac attgacaoca gtggcaggga aggcactatb gagtttgatg 4080
ctttttatca cacacttcag tgtttagaaa gttattacca atacttttaa acaacactcc 4140
aagaaaattt gctataattc tttctcaoca ctacagagag agtagatttc cccatagaga 4200
gcacagccct cattagtaag gttgggtgact attggtaaga ggtggaactc attgacacca 4260
agtggaggtt agggaaagcc cagaaatggc aggatgatat ggtggttctg tctgtgggaa 4320
aggtattggg ttttctgttt tctatttata ctgtataata gataccacgc tttttcttat 4380
tatctgtata tctatgtctt ttcatgtttg atattttccc atgccaagat ttgtttatat 4440
atattttcaa tgtlaaatta aattgatttg ggttaactttc ttcccaaga aagttatttc 4500
cccttaagt ataaactga ctg 4523

```

```

<210> 9
<211> 241
<212> PRT
<213> Homo sapiens
<400> 9

```

```

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

```

WO 02/31134

PCT/US01/31874

100 105 110
Met Leu Ser Trp Lys Pro Leu Leu Asp Leu Phe Gln Ala Thr Leu Asp
115 120 125
Tyr Gly Met Tyr Ser Arg Glu Glu Glu Leu Leu Arg Glu Arg Lys Arg
130 135 140
Ile Gly Thr Val Gly Ile Ala Ser Tyr Asp Tyr His Gln Gly Ser Gly
145 150 155 160
Thr Phe Leu Phe Gln Ala Gly Ser Gly Ile Tyr His Val Lys Asp Gly
165 170 175
Gly Pro Gln Gly Phe Thr Gln Gln Pro Leu Arg Pro Asn Leu Val Glu
180 185 190
Thr Ser Cys Pro Asn Ile Arg Met Asp Pro Lys Leu Cys Pro Ala Asp
195 200 205
Pro Asp Trp Ile Ala Phe Ile His Ser Asn Asp Ile Trp Ile Ser Asn
210 215 220
Ile Val Thr Arg Glu Glu Arg Arg Leu Thr Tyr Val His Asn Gly Lys
225 230 235 240
Ala

<210> 10
<211> 1356
<212> DNA
<213> Homo sapiens
<400> 10

```

aagtgctaaa goctccgagg ccaaggccgc tgctaetgcc gccgctgctt cttagtcccg 60
cgttcgccgc ctgggttgtc accggcgcgc ccgcccagga agccactgca accaggaccg 120
gagtggaggc ggcgcagcat gaagcggcgc agcccccctc catagcgcac gtcgggaccg 180
tccggcgggg gccgggggga aggaaaaatgc aacatggcag cagcaatgga aacggaaacg 240
ctgggtgttg agatatttga aactgcccgc tgtgaggaga atattgaatc acaggatcgg 300
cctaaattgg agccttttta tgttgaggcg tattcctgga gtcagcttaa aaagctgctt 360
gccgatacca gaaaatatca tggctacatg atggctaagg caccacatga tttcatgttt 420
gtgaagagga atgatccaga tggacctcat tcagacagaa tctattacct tgccatgtct 480
ggtgagaaca gagaaaatca actgttttat tctgaaatcc ccaaaactat caatagagca 540
gcagtcctaa tgcctccttg gaagcctctt ttggatcttt ttcaggcaac actggactat 600
ggaatgtatt ctgagaaga agaactatta agagaagaaa aacgcatgag aacgatcgga 660
attgtctctt acgattatca ccaaggaagt ggaacatttc tgtttcaagc cggtagtgga 720
attatcacg taaaagatgg agggccacaa ggaattacgc aacaactttt aaggccaat 780
ctagtgaaa ctagtgtcc caacatacgg atggatccaa aattatgcc tgcgtgtcca 840
gactggatgg cttttataca tagcaacgat atttggatct ctaacatcgt aaccagagaa 900
gaaagagagc tcaattatgt gcaaatggt agggcgtagt ctttcagatt tactttctg 960
aacagtattt tttgaagtat aatttctgct tgcattttg aaattagatt accaagttgg 1020
gtgatcttta tatttgaat tcaagctttt aaaattttta aaaatggag aaaagtcag 1080
aggataactt gtagtacc caatgtataat attcatttta atgttttaat gttcatttc 1140
aaacagtgaa acaaaagaac ctctgacatg attgtctttt tagcttgeta agactgccag 1200
aattttccca aaactgttct tattaaata aaattttagg ctaggcatgg tggtccatgc 1260
ctgtaacctc agcactctgg gaggctgagg caggcagatt gtttgagccc agaagtcaa 1320
gatcaggatg gccaacatgg tgacacctcg tttgac 1356

```

<210> 11
<211> 661
<212> PRT
<213> Homo sapiens
<400> 11

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

WO 02/31134

PCT/US01/31874

```

35          40          45
Leu Ala Asp Thr Arg Lys Tyr His Gly Tyr Met Met Ala Lys Ala Pro
50
His Asp Phe Met Phe Val Lys Arg Asn Asp Pro Asp Gly Pro His Ser
65
Asp Arg Ile Tyr Tyr Leu Ala Met Ser Gly Glu Asn Arg Glu Asn Thr
85
Leu Phe Tyr Ser Glu Ile Pro Lys Thr Ile Asn Arg Ala Ala Val Leu
100
Met Leu Ser Trp Lys Pro Leu Leu Asp Leu Phe Gln Ala Thr Leu Asp
115
Tyr Gly Met Tyr Ser Arg Glu Glu Glu Leu Leu Arg Glu Arg Lys Arg
130
Ile Gly Thr Val Gly Ile Ala Ser Tyr Asp Tyr His Gln Gly Ser Gly
145
Thr Phe Leu Phe Gln Ala Gly Ser Gly Ile Tyr His Val Lys Asp Gly
165
Gly Pro Gln Gly Phe Thr Gln Gln Pro Leu Arg Pro Asn Leu Val Glu
180
Thr Ser Cys Pro Asn Ile Arg Met Asp Pro Lys Leu Cys Pro Ala Asp
195
Pro Asp Trp Ile Ala Phe Ile His Ser Asn Asp Ile Trp Ile Ser Asn
210
Ile Val Thr Arg Glu Glu Arg Arg Leu Thr Tyr Val His Asn Glu Leu
225
Ala Asn Met Glu Glu Asp Ala Arg Ser Ala Gly Val Ala Thr Phe Val
245
Leu Gln Glu Glu Phe Asp Arg Tyr Ser Gly Tyr Trp Trp Cys Pro Lys
260
Ala Glu Thr Thr Pro Ser Gly Gly Lys Ile Leu Arg Ile Leu Tyr Glu
275
Glu Asn Asp Glu Ser Glu Val Glu Ile Ile His Val Thr Ser Pro Met
290
Leu Glu Thr Arg Arg Ala Asp Ser Phe Arg Tyr Pro Lys Thr Gly Thr
305
Ala Asn Pro Lys Val Thr Phe Lys Met Ser Glu Ile Met Ile Asp Ala
325
Glu Gly Arg Ile Ile Asp Val Ile Asp Lys Glu Leu Ile Gln Pro Phe
340
Glu Ile Leu Phe Glu Gly Val Glu Tyr Ile Ala Arg Ala Gly Trp Thr
355
Pro Glu Gly Lys Tyr Ala Trp Ser Ile Leu Leu Asp Arg Ser Gln Thr
370
Arg Leu Gln Ile Val Leu Ile Ser Pro Glu Leu Phe Ile Pro Val Glu
385
Asp Asp Val Met Glu Arg Gln Arg Leu Ile Glu Ser Val Pro Asp Ser
405
Val Thr Pro Leu Ile Ile Tyr Glu Glu Thr Thr Asp Ile Trp Ile Asn
420
Ile His Asp Ile Phe His Val Phe Pro Gln Ser His Glu Glu Glu Ile
435
Glu Phe Ile Phe Ala Ser Glu Cys Lys Thr Gly Phe Arg His Leu Tyr
450
Lys Ile Thr Ser Ile Leu Lys Glu Ser Lys Tyr Lys Arg Ser Ser Gly
465
Gly Leu Pro Ala Pro Ser Asp Phe Lys Cys Pro Ile Lys Glu Glu Ile
485
Ala Ile Thr Ser Gly Glu Trp Glu Val Leu Gly Arg His Gly Ser Asn
500
Ile Gln Val Asp Glu Val Arg Arg Leu Val Tyr Phe Glu Gly Thr Lys

```

WO 02/31134

PCT/US01/31874

515 520 525
 Asp Ser Pro Leu Glu His His Leu Tyr Val Val Ser Tyr Val Asn Pro
 530 535 540
 Gly Glu Val Thr Arg Leu Thr Asp Arg Gly Tyr Ser His Ser Cys Cys
 545 550 555 560
 Ile Ser Gln His Cys Asp Phe Phe Ile Ser Lys Tyr Ser Asn Gln Lys
 565 570 575
 Asn Pro His Cys Val Ser Leu Tyr Lys Leu Ser Ser Pro Glu Asp Asp
 580 585 590
 Pro Thr Cys Lys Thr Lys Glu Phe Trp Ala Thr Ile Leu Asp Ser Ala
 595 600 605
 Gly Pro Leu Pro Asp Tyr Thr Pro Pro Glu Ile Phe Ser Phe Glu Ser
 610 615 620
 Thr Thr Gly Phe Thr Leu Tyr Gly Met Leu Tyr Lys Pro His Asp Leu
 625 630 635 640
 Gln Pro Gly Lys Lys Tyr Pro Thr Val Leu Phe Ile Tyr Gly Gly Leu
 645 650 655
 Leu Arg Cys Ser Trp
 660

<210> 12
 <211> 4829
 <212> DNA
 <213> Homo sapiens
 <400> 12

aagtgcataa gctccgagg ccaaggccgc tgctactgcc gccgctgctt cttagtgcgc 60
 cgttcgocgc ctgggttgtc accggcgccg ccgccgagga agccactgoa accaggaccg 120
 gactggagcc ggcgcagcat gaagcggcgc aggcccgctc catagcgcac gtcgggaccg 180
 tccggcgagg gccgggggga agaaaaatgc aacatggcag cagcaatgga aacagaaacag 240
 ctgggtgttg agatatttga aactgcccgc tgtgaggaga atattgaaic acaggatcgg 300
 cctaaattgg agccttttta tgttgaggcg tattcctgga gtcagcttaa aaagctcctt 360
 gccgatacca gaaaatatca tggctacatg atggctaagg caccacatga tttcatgttt 420
 gtgaagagga atgatccaga tggacctcat tcagacagaa tctattacct tgccatgtct 480
 ggtgagaaca gagaaaatca actgttttat cctgaaatcc ccaaaactat caatagagca 540
 gcaatcctaa tgcctctctg gaagcctcct ttggatcttt ttcaggcaac actggactat 600
 ggaatgtatt ctcgagaaga agaactatta agaaaaagaa aacgcattgg aacagtcgga 660
 attgctctct acgattatca ccaaggaagt ggaacatttc tgtttcaagc cggtagtggg 720
 atttatcacg taaaagatgg agggccacaa ggaattacgc aaccaacttt aaggccaat 780
 ctagtggaaa ctagtgtccc caacatacgg atggatccaa aattatgccc tgcgtatcca 840
 gaactggattg cttttataca tagcaacgat atttggatct ctaaatcgtt aaccagagaa 900
 gaaaggagac tcaactatgt gcaaatgag ctgcccacaa tggagaaga tgccagatca 960
 gctggagtcg ctaactttgt tctccaagaa gaatttgata gatattctgg ctattgtggg 1020
 tgtccaaaag ctgaacaac tcccagtggt ggtaaaatcc ttagaattct atatgaagaa 1080
 aatgatgaat ctgaggtgga aattattcat gttacatccc ctatgttggg aacaaggagg 1140
 gcagatcat tccgttatcc taaaacaggt acagcaaatc ctaaatgtaac ttttaagatg 1200
 tcagaataaa tgattgatgc tgaaggaagg atcatagatg tcatagataa ggaactaatt 1260
 caaccttttg agattctatt tgaaggagtt gaatatattg ccagagctgg atggactcct 1320
 gagggaaaat atgcttggtc catcctacta gatcgtctcc agactcgcct acagatagtg 1380
 ttgatctcac ctgaattatt tatcccagta gaagatgatg ttatggaaag gcagagactc 1440
 attgatcag tcctgatctc tgtgacgcca ctaattatct atgagaagaa aacagacatc 1500
 tggataaata tccatgacat ctttcatgtt tttcccacaa gtcacgaaga ggaattgag 1560
 tttatttttg cctctgatg caaaacaggt ttccgctcatt tatacaaaat tacatctatt 1620
 ttaaaggaaa gcaaatataa acgatccagt ggtgggctgc ctgctccaag tgatttcaag 1680
 tgcctatata aagggagat agcaattacc agtggatgat gggagattct tggccggcat 1740
 ggaatataa tccaagtga tgaagtcaga agcctggtat attttgaag caaccaagac 1800
 tccccttagc agcatcaact gtaactagtc agttactgaa atcctggaga ggtgacaagg 1860
 ctgactgacc gtggctactc acattcttgc tgcactcgtc agcaactgga cttctttata 1920
 agtaagtata gtaaccagaa gaatccacac tgtgtgtccc tttacaagct atcaagctct 1980
 gaagatgacc caacttgcaa aacaaggaa ttttgggcca ccattttgga ttcagcaggt 2040

WO 02/31134

PCT/US01/31874

```

cctcttctcg actabactcc tccagaiaatt ttctcttttg aaagtactac tggatttaca 2100
ttgtatggga tgctctacaa gcctcatgat ctacagcctg gaaagaaata tctactgtg 2160
ctgttcatat agtgggtgct cctcagggtc agttgggtgaa taatcggttt aaaggagtca 2220
agtatctcgg ctggaatacc ctgacctctc taggttatgt ggttgtagtg atagacaaca 2280
gggatctctg tcaccgaggg cttaaatttg aagcgccctt taatatataa atgggtcaaa 2340
tggaaattga cgtacaggtg gaaggactcc aatctatagc ttctcgatat gatttcattg 2400
acctagatcg tgtgggcctc cccggctggt cctatggagg atacctctcc ctgatggcat 2460
taatgocagag gtcagatctc ttccagggtg ctattgtctg ggccccagtc actctgtgga 2520
ctctctatga tacaggatag acggaacggt atatgggtca ccctgaccag aatgaacagg 2580
gctattactt aggatctgtg gccatgcaag cagaaaagt cccctctgaa ccaaatcgtt 2640
tactgctctt acatgggttc ctggatgaga atgtccattt tgcaatacc agtatattac 2700
tgagtttttt agtgagggtc ggaagccat atgatttaca gatctatctc caggagagac 2760
acagcataag agttcctgaa tccgggagac attatgaaat gcatcttttg cactaccttc 2820
aagaaaacct tggatcacgt attgtctctc taaaagtgat ataattttga cctgtgtaga 2880
actctctggt atacactggc tattaacca aatgaggagg ttaatacaac agaaaaacaca 2940
gaattgatca tcacattttg atacctgcca tgaacatctc actcctgaaa ataatgtgg 3000
tgccatgcag gggctcacgg ttgtggttag taatctaata ccttaacccc acatgctcaa 3060
aatcaaatga tacatattcc tgagagaccg agcaatacca taagaattac taaaaaaaaa 3120
aaaaaaaaaa agacattagc accatgtatt catactacc tattttcaact ttaaatagta 3180
ttataaactt catgaactta attagtgtat ttttacagta tacttttgag tttgttaaaa 3240
tatgtatgata tcagtgtatt gtttggttca gttccagaat cttgactagg ttacagatgt 3300
gatagcactt aaatgtaatt gaatagctta tgcttcattg cttgggcata tccagcatgt 3360
tatgaactaa taactattaa acttgactta accagtcatt cattaataat ttttcaagga 3420
taacttagtg gctcctctaa gacacttggt ttggcactga ccagttttaa gccaatltaa 3480
tctgtatcta gtataaataa ttctcatttt tctttgtaga tattaacaga gtgggctttt 3540
ccttttgcat aaaggctagt aactgtatat gtatgcatgga ttaatttagt catgatattg 3600
ataattacag gcagaaaaatt ttaaatcaaa tgatttagagc ttaaatattt gcaggcaagt 3660
tttttttttt ccttcaagaa aaggaaaaag tacacattca ctagaattct tcagaaaatt 3720
tagtgggtcc agtttccatt tgggtatttc ttatataaat attctagaaat ttaaggaga 3780
ttgaagggaa tcacagtggt gttgggagac ctgggttttg ggaatgacag agagaagagg 3840
tggtgagggc ctgattaaaa actaagcaga agtagtttta acaaaaatca tcatgaaat 3900
gtttggaaac tgaattttaa acaactgtaa tattaaggaa accagaatca ataatcaact 3960
gtcttgccag cacagctaca gagtaacatg attcagggga ggaaaagttc cttagagtta 4020
cttttataat cctttttttt ttctctotta ggtttagaaa tcttaacaat ttaaacctta 4080
tccttttaaa atbatttgaa cataatttag atattgtaag cttaaaatca aaatgttat 4140
agataacctc tttaccataa actaatccct ggcaagccat ggtctctctt ttttttttg 4200
tgtttaagc ctgtaaacag tttttctgaa tgatcatgaa cttttcttgg ttttagcacta 4260
ggatttagct atgaagagag ctcataggct ttcaggtgct aattgagatc tggcctgta 4320
gagctctggg gtctagatt ggtcacattg acaccagtg cagggaaggc atctatgagt 4380
ttgatgctt ttatcaaca cttcaggtt tagaaagtta ttaccatac ttttaacaa 4440
cactcaaga aaatttgcta tatttcttc tcatcactac agagagagta gatttcccc 4500
tagagagcac agctccatt agtaaggtg gtgactattg gtaagaggtg gacttcatg 4560
acaccagtg ggaggtaggg aaagcccaga aatggcagga tgataggtg gttctgtctg 4620
tgggaaaggt attgggtttt gctgtttgta ttatactgt ataatagata ccacgtttt 4680
tcttattatc tgbatagta ttgcttttca tgtttgatat tttccatgc caagattgt 4740
ttatatatat tttcaatggt aaattaattt gatttgggta actttotctc ccaagaaagt 4800
attttcccc ttaagtataa atctgactg 4829

```

- <210> 13
- <211> 358
- <212> PRT
- <213> Homo sapiens
- <400> 13

```

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

```

WO 02/31134

PCT/US01/31874

Leu Ala Asp Thr Arg Lys Tyr His Gly Tyr Met Met Ala Lys Ala Pro
50 55 60
His Asp Phe Met Phe Val Lys Arg Asn Asp Pro Asp Gly Pro His Ser
65 70 75 80
Asp Arg Ile Tyr Tyr Leu Ala Met Ser Gly Glu Asn Arg Glu Asn Thr
85 90 95
Leu Phe Tyr Ser Glu Ile Pro Lys Thr Ile Asn Arg Ala Ala Val Leu
100 105 110
Met Leu Ser Trp Lys Pro Leu Leu Asp Leu Phe Gln Ala Thr Leu Asp
115 120 125
Tyr Gly Met Tyr Ser Arg Glu Glu Glu Leu Leu Arg Glu Arg Lys Arg
130 135 140
Ile Gly Thr Val Gly Ile Ala Ser Tyr Asp Tyr His Gln Gly Ser Gly
145 150 155 160
Thr Phe Leu Phe Gln Ala Gly Ser Gly Ile Tyr His Val Lys Asp Gly
165 170 175
Gly Pro Gln Gly Phe Thr Gln Gln Pro Leu Arg Pro Asn Leu Val Glu
180 185 190
Thr Ser Cys Pro Asn Ile Arg Met Asp Pro Lys Leu Cys Pro Ala Asp
195 200 205
Pro Asp Trp Ile Ala Phe Ile His Ser Asn Asp Ile Trp Ile Ser Asn
210 215 220
Ile Val Thr Arg Glu Glu Arg Leu Thr Tyr Val His Asn Glu Leu
225 230 235 240
Ala Asn Met Glu Glu Asp Ala Arg Ser Ala Gly Val Ala Thr Phe Val
245 250 255
Leu Gln Glu Glu Phe Asp Arg Tyr Ser Gly Tyr Trp Trp Cys Pro Lys
260 265 270
Ala Glu Thr Thr Pro Ser Gly Gly Lys Ile Leu Arg Ile Leu Tyr Glu
275 280 285
Glu Asn Asp Glu Ser Glu Val Glu Ile Ile His Val Thr Ser Pro Met
290 295 300
Leu Glu Thr Arg Arg Ala Asp Ser Phe Arg Tyr Pro Lys Thr Gly Thr
305 310 315 320
Ala Asn Pro Lys Val Thr Phe Lys Met Ser Glu Ile Met Ile Asp Ala
325 330 335
Glu Gly Arg Ser Lys Leu Met Lys Ser Glu Gly Trp Tyr Ile Leu Lys
340 345 350
Ala Pro Lys Thr Pro Leu
355

<210> 14
<211> 4309
<212> DNA
<213> Homo sapiens
<400> 14

aagtgtctaaa gcttccgagg ccaaggccgc tgetactgcc gccgctgctt cttagtgccg 60
cgttcgccgc ctgggttgct accggccgcg ccgcccaggga agccactgca accaggaccg 120
gagtgagggc gccgcagcat gaagcggcgc aggcocgctc catagcgcac gtcgggacgg 180
tccggggcgg gccgggggga aggaaaatgc aacatggcag cagcaatgga aacagaacag 240
ctgggtgttg agatatttga aactggcgac tgtgaggaga atattgaatc acaggatcgg 300
cctaaattgg agccttttta tgttgacggg tattcctgga gtcagcttaa aaagctgctt 360
gocgatacca gaaaatatca tggctacatg atggctaagg cacoacatga tttcatggtt 420
gtgaagagga atgatccaga tggacctcat tccagacagaa tctattacct tgcctatctt 480
ggtgagaaaca gagaaaatac actgttttat tctgaaattc ccaaaaactat caatagagca 540
gcagctctaa tgctctcttg gaagctcttt ttggatcttt ttcaggcaac actggactat 600
ggaatgtatt ctcgagaaga agaactatta agagaaagaa aacgcattgg aacagtcgga 660
attgctcttt acgattatca ccaaggaagt ggaacatttc tgtttcaagc cggtagtgga 720
atttatcagc taaaagatgy agggccacaa ggatttaagc aacaaccttt aaggccaat 780

WO 02/31134

PCT/US01/31874

```

ctagtggaaa ctagttgtcc caacatacgg atggatcoaa aattatgccc tgetgatcca 840
gactggattg cttttatata tagcaacgat atttggatag atcaacatcgt aaccagagaa 900
gaaaggagac tcaactatgt gcaaatagag ctagccaaca tggagaaga tgcagatca 960
gctggagtcg ctaactttgt tctccaagaa gaatttgata gatattctgg ctattggtgg 1020
tgtccaaaag ctgaaacaac tccocagtggt ggtaaaatc ttagaattct atatgaaga 1080
aatgatgaat ctgaggtgga aattattcat gttacatccc ctatgttga aacaaggagg 1140
cagatlcgat tccgttatcc taaaacaggt acagcaaatc ctaagtcac ttttaagatg 1200
tcagaataaa tgattgatgc tgaaggaga tccaagtga tgaagtcaga aggcctgat 1260
atlttgaagc caccaaaagc tcccccttag agcatcacct gtacgtatgc agttacgtaa 1320
atcctggaga ggtgacaagg ctgactgacc gtggctactc acattcttgc tgcactagtc 1380
agcactgtga cttcttata agtaagtata gtaacoagaa gaatccacac tgtgtgtccc 1440
tttaaacgct atcaagtcot gaagatgacc caacttgcaa aacaaggaa ttttgggcca 1500
ctcttttggg ttcagocaggt cctcttctgt actatactcc tccagaatct tctcttttgg 1560
aaagtactac tggatttaca ttgtatggga tgcctocaaa gcoctatgat ctacagcctg 1620
gaaagaataa tctactctgt ctgttcatat atgggtgtct cctcaggtgc agttggtgaa 1680
taactcggtt aaaggagtcg agtatttccg cttgaatacc cttagctctc taggttatgt 1740
ggtgtgtagt atagacaaca ggggatcctg tcaccgaggg cttaaatttg aaggcgcctt 1800
taataataaa atgggtcaaa tagaattga cgtacaggtg gaaggactcc aatatctagc 1860
ttctcgatat gatctcatg acttagatcg tgtgggcatc cacggctggt cctatggagg 1920
atcctctccc ctgattggcat taatgoagag gtcagatata ttcagggctg ctattgctgg 1980
ggccccagtc actctgtgga tcttctatga tacaggatcc acggaacgct atatgggtca 2040
ccttgaccag aatgaacagg gctattactt aggatctgtg gccatgcaag cagaaaagtt 2100
ccccctgaaa ccaaatcgtt taactgctctt acatggttcc ctggatgaga atgtccattt 2160
tgcaacatcc agtatattac tgagtttttt agtgagggtc ggaagcccat atgatctaca 2220
gatctatcct caggagagac acagcataag agttcctgaa tccgggagaa atbtatgaac 2280
goactctttg cactaccttc aagaaaacct tggtaocgt atgtgtctc taaaagtgat 2340
ataatttga cctgtgtaga actctctggt atacactggc tatttaacca aatgaggagg 2400
tttaactaac agaaaacaca gaattgatca tcaacttttg atacctgca tgtaacatct 2460
actcctgaaa ataaatgtgg tgccatgocg gggctcaccg tttgtgtag taatcaata 2520
ccttaacccc acatgctcaa aatcaaatga tacatattcc tgagagacc agcaatacca 2580
taagaattac taaaaaaaaa aaaaaaaaaa agacattagc accatgtatt catactacc 2640
tatttcaact ttaaatagta ttataaactt catgaaacta attagtattt ttttacagta 2700
tacttttgag tttgttaaaa tatgatgata ttatgtattg gtttggttca gttccagaat 2760
ctttgactag ttacagattt gatagcactt aaatgtaatt gaatagctta tgettoactg 2820
cttgggcata tccagcatgt tatgaactaa taactattaa acttgactta accagtcatt 2880
cattaaatatt ttttcaagga taacttagtg gcctcotaaa gacacttgtt ttggcactga 2940
ccagtttita gccaatltaa tctgtatcta gtataaataa ttctcatttt cctttgatga 3000
tatlaacaga gtgggctttt ccttttgcct aaaggctagt aactgtatat gtagcatgga 3060
tttaattagt catgatattg ataatcacag gcagaaaatt tttaactcaa tgattagagc 3120
ttaaatattt gcaggcaagt tttttttttt ccttaagaa aaggaaaag taacatcca 3180
ctagatctct tcaaaaattt tagtgggtcc agtttccatt tggatattcc ttatataat 3240
aktctagaat ttaaggaga ttgaaaggaa tcaactgagg gtggggagac ctgggtttgg 3300
ggaatgacag agagaagagg tggtaggggc ctgattaaaa actaagcaga agtagttta 3360
acaaaaatcc tcatgaaat gtttggaaac tgaatttaa acaactgtaa tattaggaa 3420
accagaatca ataaatcact gttctggccag cacagataca gagtaactg attcaggga 3480
gaaaagttic cttagagtta cttttataat tctttttttt tttctctta gttttagaa 3540
ctttacaat ttaaacttta tcttttaaaa atbtattgaa cataatttag atattgtaag 3600
cttaaaatcc aaatgtttat agataacctc tttacoataa actaatccct ggoagocat 3660
ggctctcttt ttttttttgg tgtttaaago ctgtaaacag tttttctgaa tgcactgaa 3720
cttttcttgg tttagactta ggatttagct atgaagagag ctcataggct ttcaggtgct 3780
aattgagatc tgcctgtta gactcttggg gtgctagatt ggtcacattg acaccagttg 3840
caggggaagg abctatgagt ttgatgcttt ttatcacaca cttcaggtt tagaaagtta 3900
ttaccaatcc ttttaaaoca caactcaaga aaatttgcta tatttcttct tcatcactac 3960
agagagagta gatttcccca tagagagcac agcctccatt agtaagttg gtgactattg 4020
tgaagagtg gacttcaatg acaccaagtg ggaggtaggg aaagccaga aatggcagga 4080
tgatctgtg gctctctgtg tgggaaaggt atgggtttt gctgtttgta tttactctg 4140
ataatagata ccacgctttt tcttatatcc tgtatagta ttgctttca tgtttgatat 4200
ttccactgc caagatttgt ttatataat tttcaatggt aaattaaatt gatttgggta 4260
actttcttcc caagaaagt attttcccc ttaagtataa atctgactg 4309

```

WO 02/31134

PCT/US01/31874

<210> 15
<211> 108
<212> PRT
<213> Homo sapiens
<400> 15

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

<210> 16
<211> 620
<212> DNA
<213> Homo sapiens
<400> 16

aagtgctaaa gctccgagc ccaaggccgc tgctaactgcc gccgctgctt cttagtgcgc 60
cgttcgccgc ctgggttgtc accggcgccg ccgcccagga agccactgca accaggaccg 120
gagtgaggcc ggcgcagcat gaagcggcgc agcccccctc catagcgcac gtcgggacgg 180
tcggggcggg gccgggggga aggaaaaatgc aacatggcag cagcaatgga aacagaacag 240
ctgggtgttg agatatttga aactgcccac tctgaggaga atattgaatc acaggatcgg 300
cctaaattgg agccttttta tggtagcggg tattcctgga gtcagcttaa aaagctgctt 360
gccgatacca gaaaatatca tggctacatg atggctaagg caccacatga ttctatgttt 420
gtgaagagga atgatccaga tggacctcat tcagacagaa tctattacct tggtaacca 480
tcattaattg atcatgatcg tttttcaaaa tcgaagatgc cagaaatgct ttcttctcaa 540
agctagcttg aaatgccttt cttttgatgg tctgattagg aaaacaaca ataaaccat 600
tagtttgctc ccaactcaaca 620

<210> 17
<211> 194
<212> PRT
<213> Homo sapiens
<400> 17

WO 02/31134

PCT/US01/31874

Met Ala Ala Ala Met Glu Thr Glu Gln Leu Gly Val Glu Ile Phe Glu
1 5 10 15
Thr Ala Asp Cys Glu Glu Asn Ile Glu Ser Gln Asp Arg Pro Lys Leu
20 25 30
Glu Pro Phe Tyr Val Glu Arg Tyr Ser Trp Ser Gln Leu Lys Lys Leu
35 40 45
Leu Ala Asp Thr Arg Lys Tyr His Gly Tyr Met Met Ala Lys Ala Pro
50 55 60
His Asp Phe Met Phe Val Lys Arg Asn Asp Pro Asp Gly Pro His Ser
65 70 75 80
Asp Arg Ile Tyr Tyr Leu Ala Met Ser Gly Glu Asn Arg Glu Asn Thr
85 90 95
Leu Phe Tyr Ser Glu Ile Pro Lys Thr Ile Asn Arg Ala Ala Val Leu
100 105 110
Met Leu Ser Trp Lys Pro Leu Leu Asp Leu Phe Gln Ala Thr Leu Asp
115 120 125
Tyr Gly Met Tyr Ser Arg Glu Glu Glu Leu Leu Arg Glu Arg Lys Arg
130 135 140
Ile Gly Thr Val Gly Ile Ala Ser Tyr Asp Tyr His Gln Gly Ser Gly
145 150 155 160
Thr Phe Leu Phe Gln Ala Gly Ser Gly Ile Tyr His Val Lys Asp Gly
165 170 175
Gly Pro Gln Gly Phe Thr Gln Pro Leu Arg Pro Asn Leu Val Glu Thr
180 185 190
Cys Ala

<210> 18
<211> 832
<212> DNA
<213> Homo sapiens
<400> 18

```

aagtgcataa gctccgagg ccaaggccgc tgctactgcc gccgtgctt cttagtccg 60
cgttcgcgcg ctgggttgtc accggcgcgc ccgcccggga agccactgca accaggaccg 120
gagtggaagg ggcgcagcat gaagcggcgc agcctccgctc catagcgcac gtcgggaccg 180
tccgggcggg gccgggggga agaaaaatgc aacatggcag cagcaatgga aacagaaacg 240
ctgggtgttg agatattga aactgcggac tgtgaggaga atattgaac acaggatcgg 300
cctaaattgg agccttttta tgttgaggcg tattcctgga gtcagctaac aaagotgctt 360
gocgatacca gaaaatata tggctacatg atggctaagg caccacatga ttcatgttt 420
gtgagagaga atgatccaga tggacctcat tcagacagaa tctattaact tgcocatgct 480
ggtgagaaca gaaaaatcac actgttttat totgaaattc ccaaaactat caatagagca 540
gcagtcttaa tctctctctg gaagcctctt ttggatcttt tcaggcaac actggactat 600
ggaatgtatt ctgcagaaga agaactatta agagaagaa aacgcatgg aacagtcgga 660
attgctcttt acgattatca ccaaggaagt gaaacatttc tgtttoaagc cggtagtgga 720
attatcacgc taaaagatgg agggccacaa gbatattacg wacaaccttt aaggcccaat 780
ctagtggaaa ctasttgtsc caracytga tgaccaatc agatcctgta ga 832

```

<210> 19
<211> 658
<212> PRT
<213> Homo sapiens
<400> 19

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

WO 02/31134

PCT/US01/31874

```

50          55          60
His Asp Phe Met Phe Val Lys Arg Asn Asp Pro Asp Gly Pro His Ser
65          70          75          80
Asp Arg Ile Tyr Tyr Leu Ala Met Ser Gly Glu Asn Arg Glu Asn Thr
85          90          95
Leu Phe Tyr Ser Glu Ile Pro Lys Thr Ile Asn Arg Ala Ala Val Leu
100         105         110
Met Leu Ser Trp Lys Pro Leu Leu Asp Leu Phe Gln Ala Thr Leu Asp
115         120         125
Tyr Gly Met Tyr Ser Arg Glu Glu Glu Leu Leu Arg Glu Arg Lys Arg
130         135         140
Ile Gly Thr Val Gly Ile Ala Ser Tyr Asp Tyr His Gln Gly Ser Gly
145         150         155         160
Thr Phe Leu Phe Gln Ala Gly Ser Gly Ile Tyr His Val Lys Asp Gly
165         170         175
Gly Pro Gln Gly Phe Thr Gln Gln Pro Leu Arg Pro Asn Leu Val Glu
180         185         190
Thr Ser Cys Pro Asn Ile Arg Met Asp Pro Lys Leu Cys Pro Ala Asp
195         200         205
Pro Asp Trp Ile Ala Phe Ile His Ser Asn Asp Ile Trp Ile Ser Asn
210         215         220
Ile Val Thr Arg Glu Glu Arg Arg Leu Thr Tyr Val His Asn Glu Leu
225         230         235         240
Ala Asn Met Glu Glu Asp Ala Arg Ser Ala Gly Val Ala Thr Phe Val
245         250         255
Leu Gln Glu Glu Phe Asp Arg Tyr Ser Gly Tyr Trp Trp Cys Pro Lys
260         265         270
Ala Glu Thr Thr Pro Ser Gly Gly Lys Ile Leu Arg Ile Leu Tyr Glu
275         280         285
Glu Asn Asp Glu Ser Glu Val Glu Ile Ile His Val Thr Ser Pro Met
290         295         300
Leu Glu Thr Arg Arg Ala Asp Ser Phe Arg Tyr Pro Lys Thr Gly Thr
305         310         315         320
Ala Asn Pro Lys Val Thr Phe Lys Met Ser Glu Ile Met Ile Asp Ala
325         330         335
Glu Gly Arg Ile Ile Asp Val Ile Asp Lys Glu Leu Ile Gln Pro Phe
340         345         350
Glu Ile Leu Phe Glu Gly Val Glu Tyr Ile Ala Arg Ala Gly Trp Thr
355         360         365
Pro Glu Gly Lys Tyr Ala Trp Ser Ile Leu Leu Asp Arg Ser Gln Thr
370         375         380
Arg Leu Gln Ile Val Leu Ile Ser Pro Glu Leu Phe Ile Pro Val Glu
385         390         395         400
Asp Asp Val Met Glu Arg Gln Arg Leu Ile Glu Ser Val Pro Asp Ser
405         410         415
Val Thr Pro Leu Ile Ile Tyr Glu Glu Thr Thr Asp Ile Trp Ile Asn
420         425         430
Ile His Asp Ile Phe His Val Phe Pro Gln Ser His Glu Glu Glu Ile
435         440         445
Glu Phe Ile Phe Ala Ser Glu Cys Lys Thr Gly Phe Arg His Leu Tyr
450         455         460
Lys Ile Thr Ser Ile Leu Lys Glu Ser Lys Tyr Lys Arg Ser Ser Gly
465         470         475         480
Gly Leu Pro Ala Pro Ser Asp Phe Lys Cys Pro Ile Lys Glu Glu Ile
485         490         495
Ala Ile Thr Ser Gly Glu Trp Glu Val Leu Gly Arg His Gly Ser Asn
500         505         510
Ile Gln Val Asp Glu Val Arg Arg Leu Val Tyr Phe Glu Gly Thr Lys
515         520         525
Asp Ser Pro Leu Glu His His Leu Tyr Val Val Ser Tyr Val Asn Pro

```

WO 02/31134

PCT/US01/31874

530 535 540
 Gly Glu Val Thr Arg Leu Thr Asp Arg Gly Tyr Ser His Ser Cys Cys
 545 550 555 560
 Ile Ser Gln His Cys Asp Phe Phe Ile Ser Lys Tyr Ser Asn Gln Lys
 565 570 575
 Asn Pro His Cys Val Ser Leu Tyr Lys Leu Ser Ser Pro Glu Asp Asp
 580 585 590
 Pro Thr Cys Lys Thr Lys Glu Phe Trp Ala Thr Ile Leu Asp Ser Ala
 595 600 605
 Gly Pro Leu Pro Asp Tyr Thr Pro Pro Glu Ile Phe Ser Phe Glu Ser
 610 615 620
 Thr Thr Gly Phe Thr Leu Tyr Gly Met Leu Tyr Lys Pro His Asp Leu
 625 630 635 640
 Gln Pro Gly Lys Lys Tyr Pro Thr Val Leu Phe Ile Tyr Gly Gly Arg
 645 650 655
 Val Lys

<210> 20
 <211> 4676
 <212> DNA
 <213> Homo sapiens
 <400> 20

aagtgtctaaa gctctccgagg ccaaggccgc tgctactgcc gccgctgctt cttagtccc 60
 cgttcgcccg ctgggtttgc accggcgcgc ccgcccagga agccactgca accaggaccg 120
 gagtggaggg ggcgcagcat gaagcggcgc agcccgctc catagcgcao gtcgggacgg 180
 tccggcgggg gccgggggga aggaaaaatg aacatggcag cagcaatgga aacagaacag 240
 ctgggtgttg agatatttga aactcgggac tgtgaggaga atattgaaac acaggatcgg 300
 cctaaattgg agccttttta tgttgagcgg tattcctgga gtcagcttaa aaagctgctt 360
 gccgatccca gaaaatatca tggctacatg atggctaagg caccacatga ttctatgttt 420
 gtgaagagga atgatccaga tggacctcat tcagacagaa tctattacct tgccatgtct 480
 ggtgagaaca gagaaaatca actgttttat tctgaaatcc ccaaaactat caatagagca 540
 gcagtcttaa tgcctccttg gaagcctctt ttggatcttt ttcaggcaac actggactat 600
 ggaatgtatt ctgcagaaga agaactatta agagaaga aaecgattgg aacagtccga 660
 attgcttctt acgattatca ccaagggaag ggaacatttc tgtttcaagc cggtagtggga 720
 atttatcaag taaaagatgg agggcccaaa ggatttacgc aacaaccttt aaggccaat 780
 ctagtggaaa ctagtgtgcc caacatacgg atggatccaa aattatgccc tctgatcca 840
 gactggattg cttttatata tagcaacgat atttggatat ctaacatcgt aocagagaa 900
 gaaaggagac tcaacttatg gcacaatgag ctagccaaca tggagaaga tgcagatca 960
 gctggagctg ctacctttgt tctccaagaa gaatttgata gatattctgg ctattggtgg 1020
 tgtccaaaag ctgaaacaaac tcccagtggt ggtaaaatc ttagaattct atatgagaa 1080
 aatgatgaat ctgaggtgga aattattcat gttacatccc ctatgttggg aacaaggagg 1140
 gcagattcat tccgttatcc taaaacaggt acagcaaatc ctaagtcac ttttaagatg 1200
 tcaagaataa tgattgatgc tgaaggaggg atcatagatg tcatagataa ggaactaat 1260
 caaccttttg agattctatt tgaaggaggt gaatatattg ccagagctgg atggactcct 1320
 gagggaaaat atgcttggtc catcctacta gatcgtccc agactcgcct acagatagt 1380
 ttgatctcac ctgaattatt tatcccagta gaagatgatg ttatggaaag gcagagactc 1440
 attgagtcag tgcotgattc tgtgacgcca ctaattatct atgagaaac aacagacatc 1500
 tggataataa tccatgacat ctttcatggt tttcccacaa gtcacaaga ggaattgag 1560
 tttatttttg octctgaatg caaaaacaggt ttcctcatt tatacaaat tacatctatt 1620
 ttaaaggaaa gcaaatataa acgatccagt ggtgggtggo ctgtccaag tgatttcaag 1680
 tgtcctatca aagaggagat agcaattacc agtggatgat gggagttct tggccggcat 1740
 ggatctataa tccaagtga tgaagtccga aggcctggtat atttggaag caocaaagac 1800
 tcccttttag agcatcactc gtacgtagtc agttacgtaa atcctggaga ggtgacaagg 1860
 ctgactgacc gggctactc acattcttgc tgcactagtc agcactgtga ctcttttata 1920
 agtaagtata gtaaccagaa gaatccaac tgtgtgtccc tttacaagct atcaagctct 1980
 gaagtgacc caactgcaa acaaaagaa ttttggcca ccaatttggg ttcagcaggt 2040
 cctcttctcg ctatactcc tccagaatt ttctcttttg aaagtaactac tggatttaca 2100
 ttgtatggga tgcctcaaa gctcatgat ctacagcctg gaaagaata tctactgtg 2160
 ctgttcatat atggtggtc ggtcaaatg aaattgacga tcaggtggaa ggactccaat 2220

WO 02/31134

PCT/US01/31874

```

atctagcttc tcgatatgat ttcattgact tagatcgtgt gggcatccac ggtgggtcct 2280
atggaggata cctctccctg atggcaatga tgcagaggtc agatatttc agggttgcta 2340
ttgtgggggc cccagtcact ctgtggatct tctatgatac aggatacaag gaacgttata 2400
tgggtcaccc tgaccagaat gaacagggct attacttagg atctgtggcg atgcaagcag 2460
aaaagttccc ctctgaaaca aatcgtttac tgcctttaca tggttctcgt gatgagaatg 2520
tccattttgc acataccagc atattactga gttttttagt gagggtctga aagccatattg 2580
atttacagat ctatccctcag gagagacaca gcataagagt tccotgaatcg ggagaacatt 2640
atgaactgca tcttttgcac taccctcaag aaaaccttgg atcagctatt gctgctctaa 2700
aagtgatata attttgacct gtgtagaact ctctggtata caactggctat ttaacaaat 2760
gagggggttt aatcaacaga aaacacagaa ttgactatca cattttgata cctgccaatg 2820
aacatctact cotgaaata aatgtgggtc catgcaagggg totacgggtt gtggtagtaa 2880
tctaatacct taaccccaca tgcctcaaat caaatgatac atattcctga gagaccocagc 2940
aataccataa gaattactaa aaaaaaaaaa aaaaaaaga cattagcaco atgtatcoat 3000
actaccctat tttcactttt aatagtatta taaactcoat gaacttaatt agtgtatttt 3060
tacagtatac ttttgagttt gttaaaatgt gatgatatta gtgattggtt tggttcagtt 3120
cagaatcctt tgactagtta cagatttgat agcacttaaa tgaattgaa tagcttatgc 3180
ttcaatgctt gggcatatcc agcatgttat gaactaataa ctattaact tgacttaacc 3240
agtcattcat taataatttt tcaaggataa cttagtggcc tctaaagac actgtttttg 3300
gcactgacca gtcttttagcc aatttaactc gtatctagta taaataatc tcatttttct 3360
ttgatgatat taacagagtg ggtctttcct tttgataaaa ggctagtaac tgbatattgta 3420
gcaatgattt aattagtcat gatattgata attacaggca gaaaattttt aatcaaatga 3480
ttagagctta aatatttcca ggaagttttt ttttttccct ttaagaaaag gaaaaagtac 3540
acatcactca gaattcttca gaaaatttag tgggtccagt ttccattgg tatttcccta 3600
ttaaataatt ctagaatttt aaggagatbg aagggaatca cagtggggtg gggagacctg 3660
ggtttgggga atgacagaga gaagaggtgg tgagggcctg attaaaaact aagcagaagt 3720
agttttaaca aaaactca tgaaaatggt tggaaactga aatttaaca actgtaatat 3780
taaggaacc agaatacaata aatcactgtc ttgccagcac agctacagag taacatgatt 3840
caggggagga aaagtctctt agagttactt ttataactct tttttttttt cctcttaggt 3900
ttagaatctt tacaatttta aactttatcc ttttaaat atttgaact aatttagata 3960
ttgtaagctt aaaaacaaa tgtttataga taacctctt accataact aatcctgctg 4020
aagccatggc tctctttttt tttttgggtt taaagcctg taacagttt ttctgaaatga 4080
tcatgaactt tctttgggtt agcactagga tttagctatg aagagagctc ataggcttct 4140
aggtgctaat tgagatctgc cctgttagag tcttgggggt ctgattggtt cacattgaca 4200
ccagttggcag ggaaggaatc tatgagtttg atgcttttta tcaacactt cagtgtttag 4260
aaagttatta ccaatacttt taaacaacac tccaagaaaa ttgctatat ttctttctca 4320
tcaactacaga gagagtagat ttccccatag agagcacagc ctccattagt aaggttgggt 4380
actattggta agaggtggac ttcattgaca ccaagtggga ggtagggaag gccagaaaat 4440
ggcagatga tatgggtggt ctgtcgttgg gaaaggtatt gggttttgct gtttctattt 4500
atactgtata atagatacca cgtcttttct ttttatctgt atatgtatg ottttcoatg 4560
ttgatatttt cccatgcaa gattgttata tatataattt caatgttaaa ttaatttgat 4620
ttgggtaact ttcttcccca agaaagtatt ttcccccta agtataaatc tgactg 4676

```

```

<210> 21
<211> 613
<212> PRT
<213> Homo sapiens
<400> 21

```

```

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

```

WO 02/31134

PCT/US01/31874

Leu Phe Tyr Ser Glu Ile Pro Lys Thr Ile Asn Arg Ala Ala Val Leu
100 105 110
Met Leu Ser Trp Lys Pro Leu Leu Asp Leu Phe Gln Ala Thr Leu Asp
115 120 125
Tyr Gly Met Tyr Ser Arg Glu Glu Glu Leu Leu Arg Glu Arg Lys Arg
130 135 140
Ile Gly Thr Val Gly Ile Ala Ser Tyr Asp Tyr His Gln Gly Ser Gly
145 150 155 160
Thr Phe Leu Phe Gln Ala Gly Ser Gly Ile Tyr His Val Lys Asp Gly
165 170 175
Gly Pro Gln Gly Phe Thr Gln Gln Pro Leu Arg Pro Asn Leu Val Glu
180 185 190
Thr Ser Cys Pro Asn Ile Arg Met Asp Pro Lys Leu Cys Pro Ala Asp
195 200 205
Pro Asp Trp Ile Ala Phe Ile His Ser Asn Asp Ile Trp Ile Ser Asn
210 215 220
Ile Val Thr Arg Glu Glu Arg Arg Leu Thr Tyr Val His Asn Glu Leu
225 230 235 240
Ala Asn Met Glu Glu Asp Ala Arg Ser Ala Gly Val Ala Thr Phe Val
245 250 255
Leu Gln Glu Glu Phe Asp Arg Tyr Ser Gly Tyr Trp Trp Cys Pro Lys
260 265 270
Ala Glu Thr Thr Pro Ser Gly Gly Lys Ile Leu Arg Ile Leu Tyr Glu
275 280 285
Glu Asn Asp Glu Ser Glu Val Glu Ile Ile His Val Thr Ser Pro Met
290 295 300
Leu Glu Thr Arg Arg Ala Asp Ser Phe Arg Tyr Pro Lys Thr Gly Thr
305 310 315 320
Ala Asn Pro Lys Val Thr Phe Lys Met Ser Glu Ile Met Ile Asp Ala
325 330 335
Glu Gly Arg Ile Ile Asp Val Ile Asp Lys Glu Leu Ile Gln Pro Phe
340 345 350
Glu Ile Leu Phe Glu Gly Val Glu Tyr Ile Ala Arg Ala Gly Trp Thr
355 360 365
Pro Glu Gly Lys Tyr Ala Trp Ser Ile Leu Leu Asp Arg Ser Gln Thr
370 375 380
Arg Leu Gln Ile Val Leu Ile Ser Pro Glu Leu Phe Ile Pro Val Glu
385 390 395 400
Asp Asp Val Met Glu Arg Gln Arg Leu Ile Glu Ser Val Pro Asp Ser
405 410 415
Val Thr Pro Leu Ile Ile Tyr Glu Glu Thr Thr Asp Ile Trp Ile Asn
420 425 430
Ile His Asp Ile Phe His Val Phe Pro Gln Ser His Glu Glu Ile
435 440 445
Glu Phe Ile Phe Ala Ser Glu Cys Lys Thr Gly Phe Arg His Leu Tyr
450 455 460
Lys Ile Thr Ser Ile Leu Lys Glu Ser Lys Tyr Lys Arg Ser Ser Gly
465 470 475 480
Gly Leu Pro Ala Pro Ser Asp Phe Lys Cys Pro Ile Lys Glu Glu Ile
485 490 495
Ala Ile Thr Ser Gly Glu Trp Glu Val Leu Gly Arg His Gly Ser Asn
500 505 510
Ile Gln Val Asp Glu Val Arg Arg Leu Val Tyr Phe Glu Gly Thr Lys
515 520 525
Asp Ser Pro Leu Glu His His Leu Tyr Val Val Ser Tyr Val Asn Pro
530 535 540
Gly Glu Val Thr Arg Leu Thr Asp Arg Gly Tyr Ser His Ser Cys Cys
545 550 555 560
Ile Ser Gln His Cys Asp Phe Phe Ile Ser Lys Tyr Ser Asn Gln Lys
565 570 575

WO 02/31134

PCT/US01/31874

Asn Pro His Cys Val Ser Leu Tyr Lys Leu Ser Ser Pro Glu Asp Asp
 580 585 590
 Pro Thr Cys Lys Thr Lys Glu Phe Trp Ala Thr Ile Leu Asp Ser Val
 595 600 605
 Leu Arg Cys Ser Trp
 610

<210> 22
 <211> 4685
 <212> DNA
 <213> Homo sapiens
 <400> 22

```

aagtgtctaa gctccgagg ccaaggccgc tgotactgcc gccogtctt cttagtcccg 60
cgttcgccc cttgggtgtc accggcggcg ccgcccagga agccactgca accaggaccg 120
gagtgaggcg ggcgcagcat gaagcggcgc agcccgcctc catagcgcac gtcgggaccg 180
tccggcggcg gccgggggga agaaaaatgc aacatggcag cagcaatgga aacagaaacg 240
ctgggtgttg agatatttga aactgcggac tgtgaggaga atattgaatc acaggatcgg 300
cctaaattgg agccttttta tgttgagcgg tattcctgga gtcagcttaa aaagctgctt 360
ccgatataca gaaaatatac tggctacatg atggctaagg caccacatga tttcatgttt 420
gtgaaagaga atgatccaga tggacatcat tcagacagaa tctattacct tgcctatgct 480
ggtgagaaca gagaaaatc actgttttat tctgaaatct ccaaaactat caatagagca 540
gcaagcttaa tgcctctctg gaagcctctt ttggatcttt ttcaggcaac actggactat 600
ggaatgtatt ctgcagaaga agaactatta agaaaaagaa aacgcattgg aacagctcga 660
attgcttctt acgatataca ccaaggaaat ggaacatttc tgtttcaagc cggtagtgga 720
atattatcac taaaagatgg agggccacaa ggaattacgc aacaaccttt aaggcccaat 780
ctagtgaaaa ctagtgttcc caacatacgg atggatccaa aattatgccc tgctgatcca 840
gactggattg cttttataca tagcaacgat atttggatca ctaacatcgt aaccagagaa 900
gaaaggagag taacttatgt gcaacaatgag ctagccaaca tggagaaga tgccagatca 960
gctggagtgc ctactttgt tctccaaaga gaatttgata gatattctgg ctattgtgtg 1020
tgtccaaaag ctgaaacaac tcccagtggt ggtaaaatct ttagaattct atatgaagaa 1080
aatgatgaat ctgaggtgga aattattcat gttacatccc ctatgttggg aacaaggagg 1140
gocagatcat tccgttatcc taaaacaggt acagcaaatc ctaaatgccc ttttaagatg 1200
tcagaataaa tgattgtatg tgaaggaagg atcatagatg tcatagataa ggaactaatt 1260
caaccttttg agattctatt tgaaggagtt gaatataatg ccagagctgg atggactcct 1320
gagggaaaaat atgctgtgtc catcctacta gatcgtctcc agactcgcct acagatagtg 1380
ttgatctcac ctgaaattatt latcccagta gaagatgatg ttatggaaag gcagagactc 1440
atbgatccag tgcctgattc tgtgagccca ctaattatct atgaaagaac aacagacatc 1500
tggataaata tccatgacat ctttcatggt ttccccaaa gtcacgaga ggaatttgag 1560
tttatttttg cctctgaatg caaaaacaggt ttccgtcatt tacaacaaat tacatctatt 1620
tlaaaggaaa gcaaatataa acgatccagt ggtgggtggo ctgctccaag tgatttaag 1680
tgtcctatac aagaggagat agcaattacc agtggatgat ggaagttct tggcgggat 1740
ggatcctaata tccaagtga tgaagtcaga agtctggtat atttgagg caccaaagac 1800
tcccctttag agcatcaact gtaactgagc agttacgtaa atcctggaga ggtgacaagg 1860
ctgactgacc gggctactc acattcttgc tgcactcgtc agcactgtga cttctttata 1920
agtaagtata gtaaccagaa gaatccacac tgtgtgtccc ttacaagct atcaagctct 1980
gaaagtacc caacttgcaa acaaaaggaa ttttgggcca ccattttgga ttactcctc 2040
aggtgcagtt ggtgaataat cggtttaaag gactcaagta tttccgcttg aataccctag 2100
cctctctagg ttatgtggtt gtagtगतग acaacagggg atcctgtcac cgagggtta 2160
aattgaaag cgcctttaa tataaaatgg gtaaataga aattgacgat caggtggaag 2220
gactccatac tctagctctc cgatgatgat tcatgtactt agatcgtgtg ggcactcacc 2280
gctggctcta tggaggatc actctccctga tggcattaat ccagaggtca gatattctca 2340
gggttgctat tctggggccc ccagctcact tgtggatctt ctatgataca ggatcacagg 2400
aacgttatat gggctcaccct gaccagaatg aacagggcta ttaacttaga tctgtggcca 2460
tgaacagaga aaagtcccc tctgaaccaa tactgttact gctcttacct ggttctctgg 2520
atgagaatgt caattttgca cataccagta tattaactgag ttttttagtg agggctggaa 2580
agccaatatg tttacagatc tatcctcagg agagacacag cataagatt cctgaatcgg 2640
gagaaacatg tgaactgoat cttttgcact acctcaaga aaaccttggg tccagtatg 2700
ctgctctaaa agtगतataa ttttgacctg tगतagaatc tctgttatc actggctatt 2760
taaccaaaag aggaggttta atcaacagaa aacacagaat tगतatcac atttgatac 2820

```

WO 02/31134

PCT/US01/31874

```

ctgccatgta acatctactc ctgaaaataa atgtggtgcc atgcaggggt ctacgggttg 2880
tggttagtaat ctaatacctt aaccocccat gctcaaaatc aaatgataca tattcctgag 2940
agaccocaga ataccataag aatctactaa aaaaaaaaaa aaaaaaagac attagcaccs 3000
tgtattcaba ctaccctatf ttcactttta atagatttat aaactctatg aacttaatta 3060
gtgtattttt acagataact tttgagtttg ttaaaatag atgattatag tgattgggtt 3120
ggctcagttc csagaacttt gactagttac agatttgata gcactaaat gtaattgaat 3180
agcttagctt cacttgcttg ggcatacca gcaatgtatg aactaaatc tattaactt 3240
gacttaacca gtcattcatt aataatttt caagataaac ttatggcctt cctaaagaca 3300
ctgtttttgg cactgaccag tttttagcca atttaotog tatctagtat aaataattot 3360
catttttctt tgatgatatt aacagatggg gcttttccct ttgcataaag gctagtaact 3420
atcataatag catggattta attagtcag atattgataa ttacagccag aaaaatttta 3480
atcaaatgat tagagcttaa atattgocag gcaagttttt ttttttccct taagaaaagg 3540
aaaaagtaoa cactcactag aattcctcag aaaaatttag ggtgccagtt tccatttggt 3600
atctccttat taaaatattc tagaatttta aggagattga agggatcac agtgggggtg 3660
ggagacctgg gtttgggaa tgacagagag aagaggtggg gagggcctga ttaaaaacta 3720
agcagaagta gttttaacaa aaatactcat gaaaatgttt ggaactgaa atttaacaa 3780
ctgtaatatt aaggaaacaa gaatcaataa atcactgtot tgccagcaca gctcacagat 3840
aacatgattc agggaggaa aagtccctta gagttaactt tataattctt ttttttttct 3900
ctcttaggtt tagaaatctt acaaatttaa actttatcct tttaaaatta ttgaaacata 3960
atctagatag tgtaagctta aaatacaaat gtttatagat aacctcttta ccataaacta 4020
ctccctggca agccatggct ctcttttttt ttttggtgtt taaagcctgt aaacagtttt 4080
tctgaatgat catgaaactt tcttggttta gcactaggat ttatgctatg agagagctca 4140
taggctttca ggtgctaatt gagatctgcc ctggttagagt cttgggtgca tagattggtc 4200
acatgacac cagtgccagg gaagcctctc atgagtttga tgctttttat cacacacttc 4260
agtgatttag aagttattac caatactttt aaacaacact ccaagaaaat ttgctatat 4320
tctttctcat cactcacagag agagttagatt tccccataga gagcacagcc tccattagta 4380
aggttggtga ctattggtaa gagtgagact tcattgacac caagtgagag gttaggaaag 4440
cccagaatg gcagatgat abggtggttc tgcgtgggg aaaggtattg ggttttgctg 4500
tttgatttta tactgtataa tagataccac gctttttctt attatctgta tatgtattgc 4560
ttttcatggt tgatattttc coatgcccaag atttgtttat atatatctt aatgttaaat 4620
taaattgatt tgggtaactt tcttcccocaa gaaagtattt tcccccttaa gtataaatct 4680
gactg 4685

```

```

<210> 23
<211> 892
<212> PRT
<213> Homo sapiens
<400> 23

```

```

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

```


WO 02/31134

PCT/US01/31874

	645		650		655										
Pro	His	Ala	Leu	Gln	Pro	Gly	Lys	Lys	His	Pro	Thr	Val	Leu	Phe	Val
	660		665		670										
Tyr	Gly	Gly	Pro	Gln	Val	Gln	Leu	Val	Asn	Asn	Ser	Phe	Lys	Gly	Ile
	675		680		685										
Lys	Tyr	Leu	Arg	Leu	Asn	Thr	Leu	Ala	Ser	Leu	Gly	Tyr	Ala	Val	Val
	690		695		700										
Val	Ile	Asp	Gly	Arg	Gly	Ser	Cys	Gln	Arg	Gly	Leu	Arg	Phe	Glu	Gly
	705		710		715										
Ala	Leu	Lys	Asn	Gln	Met	Gly	Gln	Val	Glu	Ile	Glu	Asp	Gln	Val	Glu
	725		730		735										
Gly	Leu	Gln	Phe	Val	Ala	Glu	Lys	Tyr	Gly	Phe	Ile	Asp	Leu	Ser	Arg
	740		745		750										
Val	Ala	Ile	His	Gly	Trp	Ser	Tyr	Gly	Gly	Phe	Leu	Ser	Leu	Met	Gly
	755		760		765										
Leu	Ile	His	Lys	Pro	Gln	Val	Phe	Lys	Val	Ala	Ile	Ala	Gly	Ala	Pro
	770		775		780										
Val	Thr	Val	Trp	Met	Ala	Tyr	Asp	Thr	Gly	Tyr	Thr	Glu	Arg	Tyr	Met
	785		790		795										
Asp	Val	Pro	Glu	Asn	Asn	Gln	His	Gly	Tyr	Glu	Ala	Gly	Ser	Val	Ala
	805		810		815										
Leu	His	Val	Glu	Lys	Leu	Pro	Asn	Glu	Pro	Asn	Arg	Leu	Leu	Ile	Leu
	820		825		830										
His	Gly	Phe	Leu	Asp	Glu	Asn	Val	His	Phe	Phe	His	Thr	Asn	Phe	Leu
	835		840		845										
Val	Ser	Gln	Leu	Ile	Arg	Ala	Gly	Lys	Pro	Tyr	Gln	Leu	Gln	Ile	Tyr
	850		855		860										
Pro	Asn	Glu	Arg	His	Ser	Ile	Arg	Cys	Pro	Glu	Ser	Gly	Glu	His	Tyr
	865		870		875										
Glu	Val	Thr	Leu	Leu	His	Phe	Leu	Gln	Glu	Tyr	Leu				
	885		890												

<210> 24
 <211> 4302
 <212> DNA
 <213> Homo sapiens
 <400> 24

```

caggccgccc cctgggtcgc tcaacttcgg ggtcaaaggt gcctgagccg gcggggtccc 60
tbtgtccgcc ggggtgtcgc tccccgcctc ccgccaactc cggggtcgca gtcccgggca 120
tggagccgcg accgtgaggc gccctgggac ccgggagcac ctgcccagtc cggccgcccg 180
cccagtcgcc ggtctgtgtc ccacgcctgc agctggaatg gagctctctt ggacccttta 240
gaaggcaccg ctgccctcct gaggctcagc gacggttaa tgggaaaggt taagaaactg 300
cgcttggaac agagaacac cggagttgg agaagcttct cgtgaalto cgagggggct 360
gagagatgg ccaccaccgg gaccccaacg gccgaccgag gcgacgcagc gccacacagc 420
gaccggcccg ccgccttcca ggtgcagaag cactcgtggg acgggctccg gagcatcctc 480
caccggagcc gcaagtactc gggcctcatt gtcaacaagg ccgcccaacg ctctcagttt 540
gtgcagaaga cggatgagtc tgggccccac tcccaccgcc tctactacct gggaaatgca 600
tatggcagcc gagagaactc cctcctctac tctgagatcc ccaagaaagt ccggaagag 660
gctctgtcgc tctgtcctg gaagcagatg ctggatcatt tccaggccac gccccaacct 720
ggggtctact ctgggagga gaggctcctg agggagcgga aacgctcggg ggtcttcggc 780
atcaccctct acgacttcca cagcgagagt ggcctctctc tcttccaggg cagcaacagc 840
ctcttccact gccgcgacgg cggcaagaac ggcctctatg tgtcccctat gaaaccgctg 900
gaaatcaaga cccagtgtc agggccccgg atggacccca aaactcgccc tgcggaccct 960
gcctctctct ccttcatcaa taacagcgac ctgtgggtgg ccaacatcga gaccagcgag 1020
gagcggcgcg tgacctctg ccaccaaggt ttatccaatg tcttgatga ccccaagtct 1080
cggggtgtgg ccacctctg cacaagaa gagttcagc gcttcaactg gtactgtgtg 1140
tgccccacag cctcctggga aggttcagag ggcctcaaga cgctgcgaat cctgtatgag 1200
gaagtcgatg agtccgaggt ggaggtcatt cagctcccct ctctgcgct agaagaaggy 1260
aagacggact cgtatcgta ccccaggaca ggcagcaaga atcccaagat tgccttgaaa 1320

```

WO 02/31134

PCT/US01/31874

```

ctggctgagt tccagactga cagccagggc aagatcgtct cgaccagga gaaggagctg 1380
gtgaoaccct tcaactccct gtcccgaag gtggagtaca tccaccaggc cgggtggacc 1440
cggatggcca aatagcctcg ggcacatgtc ctggaccggc cccagcagtg gctccagctc 1500
gtcctcctcc ccccgccctt gtccatcccg agcacagaga atgagagaga cgggtagacc 1560
tctgccagag ctgtcccocag gaatgtccag ccgtatgtgg tgtacagaga ggtcacooac 1620
gtctggatca atgttcatga catctcttat cccttcccc aatcacaggg agaggacagag 1680
ctctgttctc tcccgcccaa tgaatgcaag accggctctc gccatttcta caaagtcaac 1740
cggcttttaa aatcccaggg ctacgattgg agtgagctcg tccagcccgg ggaagtatgaa 1800
tttaagtccc caatlaagga agagattgct ctgaccagcg gtgaatggga ggttttggcg 1860
aggaacggct ccaagatctg ggtcaatgag gagaccaagc tgggtactat ccaggggacc 1920
aaggaacagc cgtggagaga ccaactctac gtggtcagct atgaggcggo cggcgagatc 1980
gtacgctca ccaagccggg ctctcccat agctgctcca tgagccagaa ctctgacatg 2040
tctgtcagcc actacagcag cgtgagcag ccgcccctgg tgcagctca caagctgagc 2100
ggcccagcag acgaccocct gcaacaagcag ccccctctc gggctagcat gatggaggaa 2160
gcccagctgc ccccgattta tgttctcca gagatcttcc atttccacac gcgctcggat 2220
gtggcgctct accgcatgat ctacaagccc caccgcttc agccagggaa gaagcaccoc 2280
accgtcctct ttgtatatgg aggccccagc gtgcagctgg tgaataactc cttcaaggc 2340
atcaagtaact tgcggctcaa cacactggcc tccctgggct acgcccgtgt tgtgattgac 2400
ggcaggggct cctgtcagcg agggcttcgg ttgcaagggg cccctgaaaa ccaaatggcg 2460
caggtggaga tccagagcca ggtggagggg ctgagctcgg tggccagaaa gatggcttc 2520
atccaactga gccaggttgc catccatggc ttgtctaacg ggggcttctc ctgctcatg 2580
gggtaatcc acaagcccca ggtgttcaag gtggccatcg cgggtgcccc ggtcacctc 2640
tggatccctc acgacacagg gtacactgag cgtacatgag acgtccctga gaacaaccag 2700
cacggctatg aggggggttc cgtggcccct cactggaga agctgcccaa tgagcccaac 2760
cgctgtctta tctccacagg ctctctggac gaaaaogtc actttttcca caaaaactc 2820
ctcgtctccc aactgatccg agcagggaaa ccttaccagc tccagatca ccccaagcag 2880
agacaacagta tctcgtgcc ctgactgggc gagcaatag aagtcagtt gctgcactt 2940
ctacagaaat acctctgagc ctgcccacgg ggagccgcca catcacaga caagtggtg 3000
cagctccgc ggggaaccag ccgggagggga ctgagtgccc cggggcccc agtgaggcac 3060
tttatccgc ccagcgtcg ccagcccaga ggagccgctg ccttcaacgc cccagcctc 3120
tttatccctt tttaaacgct cttgggtttt atgtccgctg ctttttggtt gccgagacag 3180
agagatgggt gtctcgggoc agcccctcct ccccccgcct tctgggaga gtaggtcaca 3240
cgctgatggg cactggagag gccagaagag actcagagga gggggctgcc ttccgctg 3300
ggctccctgt gacctctcag tcccctggcc cggcagcca ccgtcccag cacccaagca 3360
tgciaattgct tctccccccc ggcacgcctc cccaaactga tgtttgtgt ttgtttggg 3420
ggatattttt cataattatt taaaagacag gccggcgcg gtggctcag tctgtaatcc 3480
cagpactttg gggggctgag ccggggcgat cactgaggt tgggagttca agaccagct 3540
ggccaacatg gggaaacccc gtctctacta aaaaacaaa aaatlagcc ggtgtggtg 3600
cgcgtgccta taatcccagc taactcgggag gctgagccag gagaatcgt tgaacccggg 3660
aggtggaggt tgcggtgagc caagatgca ccatlgcact ccagctggg caacaagagc 3720
gaaactctgt ctcaaaataa taaaaaata aagacagaa agcaaggggt gctcaatct 3780
agacttgggg tccacaccgg ccagcggggt tgcacaaccag cactggtag gctccattc 3840
ttcccagcc ccagcagagc gtcatgcccg ccccaacagg gaagcggcca gggcccggg 3900
ggggcaccac ctgtggacag cctcctctgc cccaagctt caggcagca ctgaaacgca 3960
ccgaacttcc acgctctgct ggtcagttgc ggctgtcccc tcccagccc agcccaccg 4020
ccaatgtgt ctgctgacc cgtacacacc aggggttcgg gggttggag ctgaaccatc 4080
cccacotcag gtttatatt cctctcccc ttccctccc gccaaagct ctgcccggg 4140
cgggcaaaaa aaaaagttaa aagaaaagaa aaaaaaaaaa aagaacaaa ccacctctc 4200
atattatgga aaaaaaat ttttctgat tottatctt ttataattat cgtggaaga 4260
agtacacaca ttaaaccgatt ccagttgaa acatgtcacc tg 4302

```

```

<210> 25
<211> 518
<212> PRT
<213> Homo sapiens
<400> 25

```

```

Met Arg Lys Val Lys Lys Leu Arg Leu Asp Lys Glu Asn Thr Gly Ser
1 5 10 15
Trp Arg Ser Phe Ser Leu Asn Ser Glu Gly Ala Glu Arg Met Ala Thr

```

WO 02/31134

PCT/US01/31874

20 25 30
 Thr Gly Thr Pro Thr Ala Asp Arg Gly Asp Ala Ala Thr Asp Asp
 35 40 45
 Pro Ala Ala Arg Phe Gln Val Gln Lys His Ser Trp Asp Gly Leu Arg
 50 55 60
 Ser Ile Ile His Gly Ser Arg Lys Tyr Ser Gly Leu Ile Val Asn Lys
 65 70 75 80
 Ala Pro His Asp Phe Gln Phe Val Gln Lys Thr Asp Glu Ser Gly Pro
 85 90 95
 His Ser His Arg Leu Tyr Tyr Leu Gly Met Pro Tyr Gly Ser Arg Glu
 100 105 110
 Asn Ser Leu Leu Tyr Ser Glu Ile Pro Lys Lys Val Arg Lys Glu Ala
 115 120 125
 Leu Leu Leu Leu Ser Trp Lys Gln Met Leu Asp His Phe Gln Ala Thr
 130 135 140
 Pro His His Gly Val Tyr Ser Arg Glu Glu Glu Leu Leu Arg Glu Arg
 145 150 155 160
 Lys Arg Leu Gly Val Phe Gly Ile Thr Ser Tyr Asp Phe His Ser Glu
 165 170 175
 Ser Gly Leu Phe Leu Phe Gln Ala Ser Asn Ser Leu Phe His Cys Arg
 180 185 190
 Asp Gly Gly Lys Asn Gly Phe Met Val Ser Pro Met Lys Pro Leu Glu
 195 200 205
 Ile Lys Thr Gln Cys Ser Gly Pro Arg Met Asp Pro Lys Ile Cys Pro
 210 215 220
 Ala Asp Pro Ala Phe Phe Ser Phe Ile Asn Asn Ser Asp Leu Trp Val
 225 230 235 240
 Ala Asn Ile Glu Thr Gly Glu Glu Arg Arg Leu Thr Phe Cys His Gln
 245 250 255
 Gly Leu Ser Asn Val Leu Asp Asp Pro Lys Ser Ala Gly Val Ala Thr
 260 265 270
 Phe Val Ile Gln Glu Glu Phe Asp Arg Phe Thr Gly Tyr Trp Trp Cys
 275 280 285
 Pro Thr Ala Ser Trp Glu Gly Ser Glu Gly Leu Lys Thr Leu Arg Ile
 290 295 300
 Leu Tyr Glu Glu Val Asp Glu Ser Glu Val Glu Val Ile His Val Pro
 305 310 315 320
 Ser Pro Ala Leu Glu Glu Arg Lys Thr Asp Ser Tyr Arg Tyr Pro Arg
 325 330 335
 Thr Gly Ser Lys Asn Pro Lys Ile Ala Leu Lys Leu Ala Glu Phe Gln
 340 345 350
 Thr Asp Ser Gln Gly Lys Ile Val Ser Thr Gln Glu Lys Glu Leu Val
 355 360 365
 Gln Pro Phe Ser Ser Leu Phe Pro Lys Val Glu Tyr Ile Ala Arg Ala
 370 375 380
 Gly Trp Thr Arg Asp Gly Lys Tyr Ala Trp Ala Met Phe Leu Asp Arg
 385 390 395 400
 Pro Gln Gln Trp Leu Gln Leu Val Leu Leu Pro Pro Ala Leu Phe Ile
 405 410 415
 Pro Ser Thr Glu Asn Glu Glu Gln Arg Leu Ala Ser Ala Arg Ala Val
 420 425 430
 Pro Arg Asn Val Gln Pro Tyr Val Val Tyr Glu Glu Val Thr Asn Val
 435 440 445
 Trp Ile Asn Val His Asp Ile Phe Tyr Pro Phe Pro Gln Ser Glu Gly
 450 455 460
 Glu Asp Glu Leu Cys Phe Leu Arg Ala Asn Glu Cys Lys Thr Gly Phe
 465 470 475 480
 Cys His Leu Tyr Lys Val Thr Ala Val Leu Lys Ser Gln Gly Tyr Asp
 485 490 495
 Trp Ser Glu Pro Phe Ser Pro Gly Glu Gly Glu Gln Ser Leu Thr Asn

WO 02/31134

PCT/US01/31874

500 505 510
Ala Val Asp Ser Ser Arg
515

<210> 26
<211> 2411
<212> DNA
<213> Homo sapiens
<400> 26

caggcgcgcg cctgggtcgc tcaacttcg ggtcaaaagt gectgagccg ggggtccccc 60
tgtctccgcc gcggtgtcgc tccccgcctc ccgccacttc cggggtcgca gtcccgggca 120
tggagccgcg accgtgaggg gccgctggac ccgggacgac ctgcccagtc cggccgcgcg 180
cccacgtccc ggctctgtct ccacgcctgc agctggaatg gaggctctct ggacccttta 240
gaagggacc ccgccctcct gaggctcagc gagggttaa tgcggaagt taagaaactg 300
gcctggaca aggagaacac cgaagtgg agaaagtct cgtgaaatc cgagggggct 360
gagagatgg ccaccaccg gaccacaag gccgaccgag gcgacgcagc cgccacagat 420
gaccggccgc ccgcttcca ggtgcagaag cactcgtgg acgggtccgc gaggcatctc 480
cacggcagcc gcaagtactc gggcctcatt gtccacaagg cgccccaga cttccagttt 540
gtgcagaaga cggatgagtc tggccccacc tcccaccgcc tctactacct gggaaatgcc 600
tatggagcgc gagagaactc cctcctctac tctgagattc ccaagaagt ccgaaagag 660
gctctctctc tctctcctg gaagcagatg ctggatcatt tccaggccac gccccaacct 720
gggggtctact ctccggaggga ggagctgctg agggagcggg aacgctggg ggtctctggc 780
atcacctcct acgactcca cagcagagat ggcctcttcc tcttccaggc cagcaacagc 840
ctcttccact gccgcgagcg cggcaagaac ggcttcagtg tgtcccctat gaaaccgctg 900
gaaatcaaga ccagtgctc agggccccgg atggacccca aaatctgcc tgcgacacct 960
gccttctctc ccttcacaaa taacagcgac ctgtgggtgg ccaacatoga gacagcgag 1020
gagcggcgcc tgacctctg ccaccaaggt ttatccaatg tcttggatg ccccaagtct 1080
gggggtgtgg caacctctg catacaggaa gaggtcgaac gcttcaactg gtactgggtg 1140
tgccccacag cctcctggga aggttcagag ggcctcaaga cgtctgcaat cctgtatgag 1200
gaagtcatg agtccgaggt ggaggtcatt cacgtcccct ctctctgctc agaagaaagg 1260
aagaccgact cgtatcgta ccccaggaca ggcgaaga atccccagat tgccttgaag 1320
ctggctgagt tccagactga cagccagggc aagatcgtct cagccaggga gaaggagctg 1380
gtgcagccct cagctcctg gtcccgaag gtggagtaca tgcgaccggc cgggtggacc 1440
cgggatggca aatacgcctg ggccatgttc ctggaccgca cccagcagtg gctccagctc 1500
gtctcctccc cccggccctt gttcatccc agcacagaga atgaggagca gggctagcc 1560
tctcccagag ctgtcccag gaatgtccag ccgtatgtgg tgtacgagga ggtcaccaac 1620
gtctggatca atgttcatga catctcttat cccttcccc aatcagaggg agaggcagag 1680
ctctgtcttc tccgcccaca tgaatgcaag accgctctct gccatttga caaagtcaac 1740
gocgttttaa aatcccaggg ctacgattgg agtgagcct tcagcccgg ggaaggtgag 1800
cagagcctga cgaatgctgt cgaactcagc cgttagtca cgtgtggtca atatgctgt 1860
tgttcattgg tggccccccc cactcagcca gcacaccctg cgggagaggg aacagggatc 1920
ggcaggaagc cagccttccc cagtgaetgc atgatctgg agggcttaga gcaccaact 1980
gttggcttat tcaggcagca gatttactga gcactcccc tgtgcccagg ccttagaaca 2040
accaggggtt gccacactac ggcaccaggg tcaaatccgg ccccaacct ggttccataa 2100
ataagtttt atggcactg agcccacagc acttgtttac agagactgtc tgggtcgtc 2160
tttgtctgc agcagcagaa ctgggtgctc ccagcagaaa ctgttgcga agggcaagat 2220
tactgtctca gccctttgta gaaacatttg ccagctcctg ctgtaggtag ctgtgatgga 2280
attgtcact gtaataaag aaaaaggaaa atccctgctc ttggacctt ctatggagg 2340
aggcagttt ccagaaacag ttagaggtgc tgctctggt gtgctgtggg tggcagatgc 2400
agatcctagt c

<210> 27
<211> 892
<212> PRT
<213> Homo sapiens
<400> 27

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

WO 02/31134

PCT/US01/31874

Trp Arg Ser Phe Ser Leu Asn Ser Glu Gly Ala Glu Arg Met Ala Thr
 20 25 30
 Thr Gly Thr Pro Thr Ala Asp Arg Gly Asp Ala Ala Thr Asp Asp
 35 40 45
 Pro Ala Ala Arg Phe Gln Val Gln Lys His Ser Trp Asp Gly Leu Arg
 50 55 60
 Ser Ile Ile His Gly Ser Arg Lys Tyr Ser Gly Leu Ile Val Asn Lys
 65 70 75 80
 Ala Pro His Asp Phe Gln Phe Val Gln Lys Thr Asp Glu Ser Gly Pro
 85 90 95
 His Ser His Arg Leu Tyr Tyr Leu Gly Met Pro Tyr Gly Ser Arg Glu
 100 105 110
 Asn Ser Leu Leu Tyr Ser Glu Ile Pro Lys Lys Val Arg Lys Glu Ala
 115 120 125
 Leu Leu Leu Leu Ser Trp Lys Gln Met Leu Asp His Phe Gln Ala Thr
 130 135 140
 Pro His His Gly Val Tyr Ser Arg Glu Glu Leu Leu Arg Glu Arg
 145 150 155 160
 Lys Arg Leu Gly Val Phe Gly Ile Thr Ser Tyr Asp Phe His Ser Glu
 165 170 175
 Ser Gly Leu Phe Leu Phe Gln Ala Ser Asn Ser Leu Phe His Cys Arg
 180 185 190
 Asp Gly Gly Lys Asn Gly Phe Met Val Ser Pro Met Lys Pro Leu Glu
 195 200 205
 Ile Lys Thr Gln Cys Ser Gly Pro Arg Met Asp Pro Lys Ile Cys Pro
 210 215 220
 Ala Asp Pro Ala Phe Phe Ser Phe Ile Asn Asn Ser Asp Leu Trp Val
 225 230 235 240
 Ala Asn Ile Glu Thr Gly Glu Glu Arg Arg Leu Thr Phe Cys His Gln
 245 250 255
 Gly Leu Ser Asn Val Leu Asp Asp Pro Lys Ser Ala Gly Val Ala Thr
 260 265 270
 Phe Val Ile Gln Glu Glu Phe Asp Arg Phe Thr Gly Tyr Trp Trp Cys
 275 280 285
 Pro Thr Ala Ser Trp Glu Gly Ser Glu Gly Leu Lys Thr Leu Arg Ile
 290 295 300
 Leu Tyr Glu Glu Val Asp Glu Ser Glu Val Glu Val Ile His Val Pro
 305 310 315 320
 Ser Pro Ala Leu Glu Glu Arg Lys Thr Asp Ser Tyr Arg Tyr Pro Arg
 325 330 335
 Thr Gly Ser Lys Asn Pro Lys Ile Ala Leu Lys Leu Ala Glu Phe Gln
 340 345 350
 Thr Asp Ser Gln Gly Lys Ile Val Ser Thr Gln Glu Lys Glu Leu Val
 355 360 365
 Gln Pro Phe Ser Ser Leu Phe Pro Lys Val Glu Tyr Ile Ala Arg Ala
 370 375 380
 Gly Trp Thr Arg Asp Gly Lys Tyr Ala Trp Ala Met Phe Leu Asp Arg
 385 390 395 400
 Pro Gln Gln Trp Leu Gln Leu Val Leu Leu Pro Pro Ala Leu Phe Ile
 405 410 415
 Pro Ser Thr Glu Asn Glu Glu Gln Arg Leu Ala Ser Ala Arg Ala Val
 420 425 430
 Pro Arg Asn Val Gln Pro Tyr Val Val Tyr Glu Glu Val Thr Asn Val
 435 440 445
 Trp Ile Asn Val His Asp Ile Phe Tyr Pro Phe Pro Gln Ser Glu Gly
 450 455 460
 Glu Asp Glu Leu Cys Phe Leu Arg Ala Asn Glu Cys Lys Thr Gly Phe
 465 470 475 480
 Cys His Leu Tyr Lys Val Thr Ala Val Leu Lys Ser Gln Gly Tyr Asp
 485 490 495

WO 02/31134

PCT/US01/31874

Trp Ser Glu Pro Phe Ser Pro Gly Glu Asp Glu Phe Lys Cys Pro Ile
 500 505 510
 Lys Glu Glu Ile Ala Leu Thr Ser Gly Glu Trp Glu Val Leu Ala Arg
 515 520 525
 His Gly Ser Lys Ile Trp Val Asn Glu Glu Thr Lys Leu Val Tyr Phe
 530 535 540
 Gln Gly Thr Lys Asp Thr Pro Leu Glu His His Leu Tyr Val Val Ser
 545 550 555 560
 Tyr Glu Ala Ala Gly Glu Ile Val Arg Leu Thr Thr Pro Gly Phe Ser
 565 570 575
 His Ser Cys Ser Met Ser Gln Asn Phe Asp Met Phe Val Ser His Tyr
 580 585 590
 Ser Ser Val Ser Thr Pro Pro Cys Val His Val Tyr Lys Leu Ser Gly
 595 600 605
 Pro Asp Asp Asp Pro Leu His Lys Gln Pro Arg Phe Trp Ala Ser Met
 610 615 620
 Met Glu Ala Ala Ser Cys Pro Pro Asp Tyr Val Pro Pro Glu Ile Phe
 625 630 635 640
 His Phe His Thr Arg Ser Asp Val Arg Leu Tyr Gly Met Ile Tyr Lys
 645 650 655
 Pro His Ala Leu Gln Pro Gly Lys Lys His Pro Thr Val Leu Phe Val
 660 665 670
 Tyr Gly Gly Pro Gln Val Gln Leu Val Asn Asn Ser Phe Lys Gly Ile
 675 680 685
 Lys Tyr Leu Arg Leu Asn Thr Leu Ala Ser Leu Gly Tyr Ala Val Val
 690 695 700
 Val Ile Asp Gly Arg Gly Ser Cys Gln Arg Gly Leu Arg Phe Glu Gly
 705 710 715 720
 Ala Leu Lys Asn Gln Met Gly Gln Val Glu Ile Glu Asp Gln Val Glu
 725 730 735
 Gly Leu Gln Phe Val Ala Glu Lys Tyr Gly Phe Ile Asp Leu Ser Arg
 740 745 750
 Val Ala Ile His Gly Trp Ser Tyr Gly Gly Phe Leu Ser Leu Met Gly
 755 760 765
 Leu Ile His Lys Pro Gln Val Phe Lys Val Ala Ile Ala Gly Ala Pro
 770 775 780
 Val Thr Val Trp Met Ala Tyr Asp Thr Gly Tyr Thr Glu Arg Tyr Met
 785 790 795 800
 Asp Val Pro Glu Asn Asn Gln His Gly Tyr Glu Ala Gly Ser Val Ala
 805 810 815
 Leu His Val Glu Lys Leu Pro Asn Glu Pro Asn Arg Leu Leu Ile Leu
 820 825 830
 His Gly Phe Leu Asp Glu Asn Val His Phe Phe His Thr Asn Phe Leu
 835 840 845
 Val Ser Gln Leu Ile Arg Ala Gly Lys Pro Tyr Gln Leu Gln Ile Tyr
 850 855 860

WO 02/31134

PCT/US01/31874

Pro	Asn	Glu	Arg	His	Ser	Ile	Arg	Cys	Pro	Glu	Ser	Gly	Glu	His	Tyr
865					870					875					880
Glu	Val	Thr	Leu	Leu	His	Phe	Leu	Gln	Glu	Tyr	Leu				
			885						890						

<210> 28
 <211> 4219
 <212> DNA
 <213> Homo sapiens
 <400> 28

```

caggccgcgc cctgggtcgc tcaacttcgc ggtcaaaggc gccctgagcc ggggtcccc 60
tgtgtccgcc gggctgtcgc tccccgcctc ccgccaattc cggggtcgca gtcccgggca 120
tgagccgcgc accgtgagcc gccctgggac ccgggaacac ctgcccagtc cggcccgccg 180
cccaagtcgc ggtctgtgtc ccaagcctgc agctggaatg gaggtctctc ggacccttta 240
gaagccaccg ctgcccctct gaggtcagct gagcgggtaa tgcggaaggt taagaaactg 300
ggctgggaca aggagaacac cggaaagttg agaagctctc cgtggaattc cgagggggct 360
gagaggtatg ccaaccaccg gaccccaacg gccgaccgag gcgacgcagc cgcaccagat 420
gaccggggcg ccgcttcca ggtgcagaag cactcgtggg acgggctccg gagcatcctc 480
cacggagacc gcaagtaact gggcctcatt gtccaacagg ccgcccaaga cttccagttt 540
gtgcagaaga cggatgagtc tgggcccacc tcccaccgcc tctactacct gggaaatcca 600
tatggagacc gagagaactc cctcctctac tctgagattc ccaagaaggt ccggaagag 660
gctctgctgc tctgctcctg gaagcagatg ctggatcatt tccaggccac gcccaccact 720
ggggtctact ctgggggaga ggagctgctg agggagcggg aacgcctggg ggtcttggcc 780
atcaactcct ccgacttcca cagcagagat ggctcttccc tcttccagcc cagcaacagc 840
ctcttccact gccgcgagcg cggcaagaac ggctctatgg tgtcccctat gaaaccgctg 900
gaaatcaaga cccagtgctc agggcccggg atggacccca aaatctgccc tgcgaccctc 960
gccttctctc cctctatcaa taacagcgac ctgtgggtgg ccaacatcga gaccggcgag 1020
gagcggcgcc aggactctct caaccaaggt ttatccaatg tctggatgta cccaaggtct 1080
gggggtgtgg ccacctctgt catacagaaa gacttgcacc gctcactcgg gctactgttg 1140
tgcccacaag cctcctggga aggttcagag ggctccaaga cgtctgcaat cctgtatgag 1200
gaagtcgatg agtccaggtt ggaggtcatt cacgtcccct ctctgcgctc agaagaaaag 1260
aagacggact cgtatcggta ccccaggaca ggcaagcaaga atcccagat tgccctgaaa 1320
ctggctgagt tccagactga cagccaggcc aagatcgtct cgcaccagga gaaggagctg 1380
gtgcagccct tcagctcgtc gttcccgaag gtggagtaca tgcgaccggc cgggtgggac 1440
cgggatggca aatacgcctg ggtccatgtc ctggaccggc cccagcagtg gctccagctc 1500
gtcctctccc ccccggccct gttccatccc agcaacagaga atgaggagca gctccagctc 1560
tctcccaagc ctgtcccagc gaatgtccag ccgtatgtgg tgtacgagga ggtccacca 1620
gtctggatca atgttcatga catctctcat cctctcccct aatccagagg agggagcgag 1680
ctctgctttc tccggcccaa tgaatgcaag acggcctctc gccatttcta caaagtcaac 1740
gcccgtttac aatcccaggc ctacgattgg agtgagccct tccagcccgg ggaagatgaa 1800
tlttaagtgc ccattaaagg agagattgct ctgaccagcg gtgaatggga ggttttggcg 1860
agggaccgct ccaagatctg ggtcaatgag gagaccagc tgggtgactt ccagggcacc 1920
aaggacacgc cgtcggagca ccaactctac gtagtcagct atgagcggc cggcgagatc 1980
gtacgcctca caccgcccgg ctctcccact agctgctcca tgaccagaa ctctgacatg 2040
ttcgtcagcc actacagcag cgtgagcacg ccgcccctgc tgacagctca caagctgagc 2100
ggcccagcag agcaccocct gcacaagcag ccccgcttct gggctagcat gatggagca 2160
gccagctgcc cccgggatta gtttctcca gagatcttcc atttccacac gcgctcggat 2220
gtgcccgtct agggcatgat ctacaagccc caeccttgc agccagggaa gaagcacccc 2280
accgtcctct tgtatatgag agggcccagc gtgcagctgg tgaataactc cttcaaggcc 2340
atcaagtaact tccgctcaa cacactggcc tccctgggct accgctggtt gttgatgac 2400
ggcaggggct cctgtcagcg agggcttogg ttcgaagggg cctgaaaaa ccaaatgggc 2460
caggtggaga tccaggaaca ggtggagggc ctgcagttcg tggccagaaa gtaggtctt 2520
atcgacctga gccaggttgc catccatggc tggctctacg ggggttctct ctgctcatg 2580
ggctaatcc ccaagcccga ggtgttcaag gtagccatcg cgggtgcccc ggtcaccctc 2640
tggatggcct acgacacagg gtaactcagc cgtctacatg acgtccctga gaacaaccag 2700
cacggctatg aggggggttc cgtggcctcg cacgtggaga agctgcccac tgaccacca 2760
cgtctgctta tectccaagg ctctctggac gaaaaactgc actttttcca caaaaacttc 2820
ctcgtctccc aactgatccg agcagggaaa ccttaccagc tccagatcta cccaaccag 2880
agacacagta ttcgctgccc cagctggggc gagcactatg aagtaactgt gctgcaactt 2940

```

WO 02/31134

PCT/US01/31874

```

ctcacaggaat acctctgagc ctgcccaccg ggagccgcca catcacagca caagtggctg 3000
cagcctccgc ggggaaccag gcgggaggga ctgagtgccg cgcgggcccc agtgaggcac 3060
tttgtccgc ccagcctgg ccagcccga ggagccgctg ccttcaccgc cccagcct 3120
tttatacctt tttaaacgct cttgggttt atgtccgctg cttcttggtt ccgagacag 3180
agagatggtg gtctcgggcc agcccctct ctcccgcct tctgggagga ggaggtcaca 3240
cgctgatggg cactggagag gccagaagag actcagagga gcgggtgccc ttcgctcctg 3300
ggctccctgt gacctctcag tcccctggcc cggccagcca ccgtcccag caccacaaga 3360
tgcaattgcc tgtccccccc ggcagcctc cccaactga tgtttgtgtt ttgtttgggg 3420
ggatattttt cataattatt taaaagacag gccgggccc gtggctcacg tctgtaatcc 3480
cagcactttg ggaggtgag gcgggaggat caectgaggt tgggagttca agaccagcct 3540
ggccaacatg ggaaacccc gtctctacta aaaatacaaa aaattagccg ggtgtggtgg 3600
cgctgccta taatcccagc tactcgggag gctgagcag gagaatcgct tgaacccggg 3660
aggtggaggt tgcggtgagc caagatgca ccattgcact ccagcctggg caacaagagc 3720
gaaactctgt ctcaaaaata aaaaaaata aaagacagaa agcaagggtt gcctaaatct 3780
agacttggg tccacaccgg gcagcggggt tgcaaccag cacctggtag gctccatttc 3840
ttcccagcc cgactttcag gcagcactg aaacgcaccg aacttccag cctctgctgt 3900
cagtggcggc tgtcccctcc ccagcccagc cggccagcca catgtgtctg cctgaccctg 3960
acacaccag gggtccgggg ttgggagctg aacctcccc acctcagggt tatatttccc 4020
ctccccttc cctcccgcg aagagctctg ccagggcgg gcaaaaaaaa aagtaaaaag 4080
aaaaaaaaa aaaaaaaaag aaacaaacca cctctacata ttatggaaa aaaaatattt 4140
tgtcgattct tattctttta taattatgct tggaagaagt agacacatta aacgattcca 4200
gttggaaaca tgtcacctg 4219

```

```

<210> 29
<211> 832
<212> PRT
<213> Homo sapiens
<400> 29

```

```

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

```


WO 02/31134

PCT/US01/31874

	725		730		735										
Gly	Leu	Gln	Phe	Val	Ala	Glu	Lys	Tyr	Gly	Phe	Ile	Asp	Leu	Ser	Arg
	740							745							750
Val	Ala	Ile	His	Gly	Trp	Ser	Tyr	Gly	Gly	Phe	Leu	Ser	Leu	Met	Gly
	755							760							765
Leu	Ile	His	Lys	Pro	Gln	Val	Phe	Lys	Ala	Gln	Pro	Leu	Ala	Tyr	Pro
	770							775							780
Pro	Arg	Leu	Pro	Gly	Arg	Lys	Arg	Ala	Leu	Phe	Pro	His	Lys	Leu	Pro
	785							790							800
Arg	Leu	Pro	Thr	Asp	Pro	Ser	Arg	Glu	Thr	Leu	Pro	Ala	Pro	Asp	Leu
			805						810						815
Pro	Gln	Arg	Glu	Thr	Gln	Tyr	Ser	Leu	Pro	Arg	Val	Gly	Arg	Ala	Leu
			820					825							830

<210> 30
 <211> 4159
 <212> DNA
 <213> Homo sapiens
 <400> 30

```

caggccgcg cctgggtcgc tcaactccg ggtcaaaagt gcctgagcc ggggtcccc 60
tgtgtccgc gggctgtcg tccccgcct ccgcoacttc cggggtcgca gtcccggga 120
tggagccgc accgtgagg gccgtggac ccgggacgac ctgcccagtc cggccgcgc 180
cccacgtccc ggtctgtgtc ccacgcctgc agctggaatg gaggctctct ggaccttta 240
gaaggcaacc ctgccctcct gaggtcagct gagcggtaaa tgcggaaggt taagaaactg 300
gccttgaca aggagacac cgaagtctg agaagctct cgtgaatc cgagggggct 360
gagaggatgg ccaccaccg gaccocaaag gccgacgag gcgacgcagc cgccacagat 420
gaccggcccg ccgcttcca ggtgcagaag cactcgtggg acgggtcccg gaggatcatc 480
cacggcaacc ccaagtactc gggcctcatt gtcaacaagg cgcgccacga cttccagttt 540
gtgcagaaga cgaatgagtc tgggcccacc tcccacgcc tctactactc gggaatgcca 600
tatggcaacc gagagaactc cctcctctac tctgagatc ccaagaaggt ccggaagag 660
gctcctctgc tctgtcctg gaagcagatg ctggaboaat tccagccca ccccaaccat 720
gggtctact ctccggagga ggaagctgct agggagcggg aacgctggg ggtctctggc 780
atcaactcct acgacttcca cagcagaggt ggcctcttcc tcttccaggc cagcaacagc 840
ctcttccact gccgcgacgg cggcaagaac ggcctcaatg tgtcccctat gaaaocgctg 900
gaaatcaaga ccagtgctc agggccccc atggacccca aaactctccc tgccgacct 960
gecttctct ccttoatcaa taacagcgac ctgtgggtgg ccaacatcga gacaggcgag 1020
gagcgggtgg tgacctctg ccaccaagg tlatccaatg tcttgatga ccccaagtct 1080
gcgggtgtgg caacctctg cacaaggaa gacttcgacc gcttcaactg gtactggtgg 1140
tgccccacag cctcctggga aggttcagag ggcctcaaga cgttcgaat cctgtatgag 1200
gaagtcatg agtccgaggt ggaagtcatt cagtcacct ctctgagct agaagaagg 1260
aagaccgact cgtatcggta ccccaggaca ggcagcaaga atccaagat tgccttga 1320
ctggctgagt tccagactga cagccagggc aagatcgtct cgaccacaga gaaggagctg 1380
gtgcagccct tcaactcgtc gtcocgaaag gtggagtaca tgcgccaggg cgggtggacc 1440
cggatggaca aatacactg ggcctatgtc ctggacgggc cccagcagtg gctccagctc 1500
gtcctcctcc ccccgccctc gttcatccc agcacagaga atgaggaga cggctagcc 1560
tctgcaagag ctgcccag gaatgtccag ccgtatgtg tgtacagaga ggtcaccac 1620
gtctggatca atgttcatga catctctat cccttcccc aatcagagg agaggagag 1680
ctctgctttc tccgcccaca tgaatgcaag accggtctct gccatttga caaagtacc 1740
gccgttttaa aatcccagg ctacgattgg agtgagccct tcagccccg ggaagtga 1800
tttaagtgcc ccaatgaaga agagattgct ctgaccagc gtgatggga ggttttggc 1860
aggcaaggct ccaagatctg ggtcaatgag gagaccaagc tgggtactt ccagggcacc 1920
aaggacacgc cgtggagca ccaactctac gttgctagct atgagggcc cggcgagatc 1980
gtagcctca ccaagcccg ottctccat agctgctcca tgaccagaa cttcgacatg 2040
ttcgtcagcc actacagcag cgtgagcagc ccgcccctgc tgcaactca caagctgagc 2100
ggcccagcag acagccccct gcaacaagcag ccccgttctc gggctagcat gatggaggca 2160
gccagctgac ccccgatta tgttctcca gagatcttcc atttccacac gcgctcgat 2220
gtcggctct accgcatgat ctacaagccc cagccttgc agccaggaa gaagcacc 2280
accgtctct tgtataggg agggcccag gtcagctgg tgaataact cttcaagcc 2340
atcaagatct tgccgctcaa cacactggcc tccctgggct acgccgtggt tgtgatgac 2400

```

WO 02/31134

PCT/US01/31874

```

ggcaggggct cctgtcagcg agggcttcgg ttcgaagggg ccotgaaaaa ccaaatgggc 2460
caggtggaga tccgaggaoca ggtggagggc ctgcagttcg tggccgagaa gtatggcttc 2520
atcagactga gccgagttgc catccatggc tggctctacg ggggcttctc ctgctcatg 2580
gggttaatcc acaagcccca ggtgttcaag gcccaacccg ttgcttatcc tccagggctt 2640
cctggacgaa aacgtgcaact tttccacac aaacttctct gtctcccaac tgatccgagc 2700
agggaaacct taccagctcc agatctaccc caacagagaa cacagatctc gctgccccga 2760
gtcagggcagc cactatgaag tcaactgtct gcactttcta caggaatacc tctgagctcg 2820
cccaccggga gccgccacat cacagacaaa gtggctgcag cctcccgggc gaaccagggc 2880
ggaggaactg agtggcccgcc gggcccccag gaggcacttt gtcccgccca gcgctggcca 2940
gccccgagga gccctgctct tcaaccgccc gaagcctttt atcctttttt aaacgctctt 3000
gggttttatg tccgctgctt cttggttgcc gagacagaga gatggtggtc tcgggcccagc 3060
ccctctctct cccgctctct gggaggagga ggtcaacaga tgatgggacg tggagaggcc 3120
agaagagact cagaggagcg ggtgctcttc cgcctggggc tccctgtgac ctctcagctc 3180
cctggcccg ggcagccacc tcccagccac ccaagcatgc aattgctctg cccccccggc 3240
cagctcccc aacttgatgt ttgtgttttg ttgggggga tatttttcat aattatttaa 3300
aagacaggcc gggcgggtg gctcactctc gtaatcccag cactttggga ggctgaggcg 3360
ggcggatcac ctgaggttgg gattcaaga ccagcctggc caacatgggg aaaccocgctc 3420
tctactaaaa atacaaaaaa ttagccgggt ggtgtggcgc gtgctctata tcccagctac 3480
tcgggagcct gaggcaggag aatgctctga acccgggagg tggaggttgc ggtgagccaa 3540
gatccacca ttgactcca gccctgggcaa caagagcgaa actctgtctc aaaataaata 3600
aaaaataaaa gacagaaagc aaggggtgcc taaatctaga cttggggtcc acaccgggca 3660
goggggttgc aaccagcac ctggtaggct ccaattcttc ccaagcccca gcagagggtc 3720
gtcggggccc cacaggagaa gccggccagg ccccgggggc gcaccacctg tggacagccc 3780
tctctcccc aagcttccag gcagccactg aaacgoccc aacttccact ctctgctggt 3840
cagtgggcgc tgtccctcc ccagcccagc cgcaccagca catgtgtctg cctgaccctg 3900
acacaccagg ggttccgggg ttgggagctg aaccatccc acctcagggt tabatttccc 3960
tctccccttc cctcccggcc aagagctctg ccagggggcg gcaaaaaaaa aagtaaaaag 4020
aaaagaaaaa aaaaaaaag aaacaaacca cctctacata ttatggaag aaaatatttt 4080
tgtcgattct tattctttta taattatgcg tggagaagt agacacatta aacgattcca 4140
gttggaaaca tgtcacctg 4159

```

```

<210> 31
<211> 832
<212> PRT
<213> Homo sapiens
<400> 31

```

```

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

```


WO 02/31134

PCT/US01/31874

	660		665		670										
Tyr	Gly	Gly	Pro	Gln	Val	Gln	Leu	Val	Asn	Asn	Ser	Phe	Lys	Gly	Ile
	675		680		685		690		695		700		705		710
Lys	Tyr	Leu	Arg	Leu	Asn	Thr	Leu	Ala	Ser	Leu	Gly	Tyr	Ala	Val	Val
	715		720		725		730		735		740		745		750
Val	Ile	Asp	Gly	Arg	Gly	Ser	Cys	Gln	Arg	Gly	Leu	Arg	Phe	Glu	Gly
	755		760		765		770		775		780		785		790
Ala	Leu	Lys	Asn	Gln	Met	Gly	Gln	Val	Glu	Ile	Glu	Asp	Gln	Val	Glu
	795		800		805		810		815		820		825		830
Gly	Leu	Gln	Phe	Val	Ala	Glu	Lys	Tyr	Gly	Phe	Ile	Asp	Leu	Ser	Arg
	835		840		845		850		855		860		865		870
Val	Ala	Ile	His	Gly	Trp	Ser	Tyr	Gly	Gly	Phe	Leu	Ser	Leu	Met	Gly
	875		880		885		890		895		900		905		910
Leu	Ile	His	Lys	Pro	Gln	Val	Phe	Lys	Ala	Gln	Pro	Leu	Ala	Tyr	Pro
	915		920		925		930		935		940		945		950
Pro	Arg	Leu	Pro	Gly	Arg	Lys	Arg	Ala	Leu	Phe	Pro	His	Lys	Leu	Pro
	955		960		965		970		975		980		985		990
Arg	Leu	Pro	Thr	Asp	Pro	Ser	Arg	Glu	Thr	Leu	Pro	Ala	Pro	Asp	Leu
	995		1000		1005		1010		1015		1020		1025		1030
Pro	Gln	Arg	Glu	Thr	Gln	Tyr	Ser	Leu	Pro	Arg	Val	Gly	Arg	Ala	Leu

<210> 32
 <211> 4076
 <212> DNA
 <213> Homo sapiens
 <400> 32

```

caggccgccc cctgggtgct tcaacttcgc ggtcaaaagt gcctgagccg ggggtcccc 60
tgtgtccgcc ggggtgtctg tccccgcctc ccgccacttc cggggtgcca gtcccgggca 120
tggagccgcg accgtgaggg gccgctggac ccgggacgac ctgcccagtc cggccgccc 180
ccccacgccc ggtctgtgtc ccacgcctgc agctggaatg gaggctctct ggacccttta 240
gaaggcaccc ctgcccctct gaggctcagc gagcggttaa tgcggaaggt taagaaactg 300
cgcttgagca aggagaacac cggaaattgg agaagctctc cgctgaattc cgagggggct 360
gagaggatgg ccaccacggg gaccccaacg gccgaccgag gcgagcgagc cgccacagat 420
gacccgagcc ccgcttccca ggtgcagaag cactcgtggg acgggctccc gaggatcctc 480
cacggcagcc gcaagtactc gggcctcatt gtcaacaagg cgcaccaaga ctccagttt 540
gtgcagaaga cggatgagtc tgggcccacc tcccaccgcc tctactacct gggaatgcca 600
tatggcagcc gagagaactc cctcctctac tctgagatc ccagaaggt ccggaagag 660
gctctgtgct tctgtcctg gaagcagatg ctggatcatt tccaggccac gcccaacct 720
gggtctact ctgggagga gaggctgctg agggagcaga aacgctggg ggtctctggc 780
atcacctact acgacttcca cagcagaggt ggctctctcc tctccaggc cagcaacagc 840
ctctccact ccgcgcaggg ccgcaagaac ggtctaatgg tgtcccctat gaaacgctg 900
gaaatcaaga ccagtgctc agggcccggg atggaoccca aaatctgcc tgcgaccct 960
gecttctct cctcactcaa taacagcgac ctgtgggtgg ccaacatcga gaccagcgag 1020
gagcggcggc tgacctctg ccaccaaggt ttatccaatg tcttggatga cccaagtct 1080
cggggtgtag ccactctgt catacagaaa gagttagacc gttcactggt gtagctgtgg 1140
tgcccacag cctcctggga aggttcagag ggctoaaga cgtcgaat cctgtatgag 1200
gaagtcgatg agtccaggt gagggtcatt cactcccct ctctcgctg agaagaagg 1260
aagcaggact cgtatcggtt cccacggaca gccagcaaga atccaagat tgcttgaaa 1320
ctggctgagt tccagactga cagccagggo aagatcgtct cgaccaggga gaaggagctg 1380
gtggagccct tcagctcct gtcccogaag gtggagtaca tgcaccgggc cgggtggacc 1440
cgggatggca aatacgcctg ggcctatgtc ctggaccggc ccagcagtg gctccagctc 1500
gtctcctcc ccccgccct gtctatccc agcacagaga atgaggagca gcgctagcc 1560
tctccagag ctgtcccag gaatgtccag ccgtatgtg tgtcagagga ggtcaccac 1620
gtctggatca atgttcatga catctctat cccttcccc aatcagagg agaggacgag 1680
ctctgcttcc tccgcgcaa tgaatgcaag accgctctc gccatttga caaagtcaac 1740
gccgttttaa aatcccagg ctacgattgg agtgagccct tcagcccgg ggaagatgaa 1800
tttaagtgcc caattaagga agagattgct ctgaccagcg gtgaatggga ggttttggcg 1860
aggcacgctc ccaagatctg ggtcaatgag gagaccaagc tgggtgactt ccagggcacc 1920

```

WO 02/31134

PCT/US01/31874

```

aaggacaacgc cgtctggagca ccaacctctac gtggtcaagct atgaggcggc cggcgagatc 1980
gtacgcctca ccaagcccggt cttctcccat agctgctcca tgagccagaa cttcgacatg 2040
ttcgtcaagcc actacagcag cgtgagcagc ccgcccctgag tgcaagctca caaagtgagc 2100
ggcccggagc agcaccocct gcaacaagcag ccccctctct gggctagcat gatggaggca 2160
gcccagctgc ccccgatga ttgtccotcca gagatcttcc atttccacac gcgctcggat 2220
gtggcgctct acggcatgat ctacaagccc caagccttgc agccaggaaa gaagcaccoc 2280
accgtctctt ttgtatatgg aggcccocag gtgcaagctgg tgaataacte cttcaaggc 2340
atcaagtact tgcgctcaa cacactggcc tccctgggct acgcccgtgt gtgatgtgac 2400
ggcaggggct cctgtcagcg agggctctcg ttcaagggg ccctgaaaaa ccaaatgggc 2460
cagtgaggaga tccaggacca ggtggagggc ctgcaagctc tggccagaaa gtabggctc 2520
atcgactga gccaggttgc catccatggc tggctctacg gggcttctct ctgcctcatg 2580
gggttaatcc acaagcccca ggtgttcaag gcccaaccgc ttgttatcc tccacggctt 2640
cctggacgaa aacgtgcact ttttccacac aaacttctct gtctcccaac tgatccgagc 2700
agggaaacct taccagctcc agatctaccc caacgagaga cacagtattc ggtccccca 2760
gtcggcgagc cactatgaag tcaagttgct gcacttteta caggaatacc tctgagctg 2820
cccaccggga gcccccacat cacagcacia gtggctgcag cctccggggg gaaccaggcg 2880
ggagggactg agtgcccgcc gggcccacgt gaggcacttt gtcccggcca gcgctggcca 2940
gcccggagga gccctgcct tcaagccccc gacgctttt atcctttttt aaacgctctt 3000
gggttttatg tccgctgctt ctgtgttgcg gagacagaga gatggtggtc tccggccagc 3060
cctctctctc cccgctctct gggaggagga ggtcacacgc tgatgggca c tggaggggcc 3120
agaagagact cagaggagcg ggctgccttc cgcctggggc tccctgtgac ctctcagctc 3180
cctggccggc ccagccaccg tcccagcacc ccaagcatgc aatgctgtc cccccggc 3240
cagctccccc aacttgatgt ttgtgttttg tttgggggga tatttttoat aatatttaa 3300
aagacaggcc gggcggggtg gctcagctct gtaatcccag cactttggga ggtgaggcg 3360
ggcggataca ctgaggttgg gagtcaaga ccagcctggc caacatgggg aaacccgctc 3420
tctactaaaa atacaaaaaa ttgcccgggt gtgtggcgc gtgcctataa tcccagctac 3480
tggggagctc gaggcaggag aatcgcttga acccggggag tggaggttgc ggtgagccaa 3540
gatcgaccca ttgcactcca gctgggcaa caagagcgaa actctgtctc aaaaataata 3600
aaaaataaaa gcagaaagc aaggggtgct taaatctaga ctggggctc acaccggcca 3660
gggggttgc aaccagcac ctgtaggct ccaattcttc ccaagccoga ctttcaggca 3720
ggcactgaaa cgcaccgaac ttccaagctc tgcctgtcag tggcgctgt cccctcccca 3780
gcccagccgc ccagccacat gtgtctgctt gaccgtaca caccaggggt tccggggtg 3840
ggagctgaac catccccacc tcaagggttat attccctct ccccttccot ccccgccaag 3900
agctctgcca gggcggggca aaaaaaaaaa taaaaagaaa agaaaaaaaa aaaaaagaaa 3960
caaacaccct ctacatatta tggaaagaaa atatttttgt cgattcttat tctttataa 4020
ttatgctggg aagaagtaga cacattaaac gattccagtt ggaacaatgt cacctg 4076

```

```

<210> 33
<211> 879
<212> FRT
<213> Homo sapiens
<400> 33

```

```

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

```


WO 02/31134

PCT/US01/31874

Gly Phe Ser His Ser Cys Ser Met Ser Gln Asn Phe Asp Met Phe Val
 565 570 575
 Ser His Tyr Ser Ser Val Ser Thr Pro Pro Cys Val His Val Tyr Lys
 580 585 590
 Leu Ser Gly Pro Asp Asp Asp Pro Leu His Lys Gln Pro Arg Phe Trp
 595 600 605
 Ala Ser Met Met Glu Ala Ala Ser Cys Pro Pro Asp Tyr Val Pro Pro
 610 615 620
 Glu Ile Phe His Phe His Thr Arg Ser Asp Val Arg Leu Tyr Gly Met
 625 630 635
 Ile Tyr Lys Pro His Ala Leu Gln Pro Gly Lys Lys His Pro Thr Val
 645 650 655
 Leu Phe Val Tyr Gly Gly Pro Gln Val Gln Leu Val Asn Asn Ser Phe
 660 665 670
 Lys Gly Ile Lys Tyr Leu Arg Leu Asn Thr Leu Ala Ser Leu Gly Tyr
 675 680 685
 Ala Val Val Val Ile Asp Gly Arg Gly Ser Cys Gln Arg Gly Leu Arg
 690 695 700
 Phe Glu Gly Ala Leu Lys Asn Gln Met Gly Gln Val Glu Ile Glu Asp
 705 710 715 720
 Gln Val Glu Gly Leu Gln Phe Val Ala Glu Lys Tyr Gly Phe Ile Asp
 725 730 735
 Leu Ser Arg Val Ala Ile His Gly Trp Ser Tyr Gly Gly Phe Leu Ser
 740 745 750
 Leu Met Gly Leu Ile His Lys Pro Gln Val Phe Lys Val Ala Ile Ala
 755 760 765
 Gly Ala Pro Val Thr Val Trp Met Ala Tyr Asp Thr Gly Tyr Thr Glu
 770 775 780
 Arg Tyr Met Asp Val Pro Glu Asn Asn Gln His Gly Tyr Glu Ala Gly
 785 790 795 800
 Ser Val Ala Leu His Val Glu Lys Leu Pro Asn Glu Pro Asn Arg Leu
 805 810 815
 Leu Ile Leu His Gly Phe Leu Asp Glu Asn Val His Phe Phe His Thr
 820 825 830
 Asn Phe Leu Val Ser Gln Leu Ile Arg Ala Gly Lys Pro Tyr Gln Leu
 835 840 845
 Gln Ile Tyr Pro Asn Glu Arg His Ser Ile Arg Cys Pro Glu Ser Gly
 850 855 860
 Glu His Tyr Glu Val Thr Leu Leu His Phe Leu Gln Glu Tyr Leu
 865 870 875

<210> 34
 <211> 4263
 <212> DNA
 <213> Homo sapiens
 <400> 34

caggcgcgcg cctgggtcgc tcaacttccg ggtcaaaagg gcctgagccg ggggtcccc 60
 tgtgtccgcc ggggtgtcgc tcccgcctc ccgccaactc cggggtcgca gtcccgggca 120
 tggagccgcg accgtgaggc gcgctgggac ccgggacgac ctgcccagtc cgcccgccgc 180
 cccactgtccc ggtctgtgtc ccacgcctgc agctggaatg gaggctctct ggacccttta 240
 gaaggcaacc ctgcctctct gaggtcagct gagcggttaa tggggaaggt taagaaactg 300
 ggcctggaca aggagaacac cgaagtgtg agaagcttct cgtggaatc cgaaggggct 360
 gagagatgg caaccaccgg gaccccaacg gccgaccgag gcgacgcagc cgccacagat 420
 gaccggccgc ccgcttcca ggtgcagaag cactcgtggg acgggctcgc gaggcatatc 480
 cccggcagcc gcaagtactc gggcctcatt gtcaacaagg cgcaccaaga ctccagttt 540
 gtgcagaaga cggatgagtc tggccccac tcccaccgcc tctactactc ggggatgcca 600
 tatggcagcc gagagaactc cctcctctac tctgagattc ccaagaaggt ccggaagag 660
 gctctgctgc tctgtcctg gaagcagatg ctggatcatt tccaggccac gcccccacct 720
 ggggtctact ctccggagga ggagctgctg agggagcggg aacgcctggg ggtcttcggc 780

WO 02/31134

PCT/US01/31874

```

atcacctcct acgacttcca cagcagaggt ggcctcttcc tcttccagge cagcaacagc 840
ctcttccact gcgcgcaggg cggcaagaao ggcctcatgg tgtccocctat gaaaccgctg 900
gaaatacaaga cccagtgctc agggcccggg atggacccca aaatctgccc tgccgaccct 960
gccttcttct cctcctcaaa taacagcgac ctgtgggtgg ccaacatcga gacagggcag 1020
gagcggcggc tgaactcttg ccaccaaggt ttatccaatg tcctggatga cccaagctct 1080
gggggtgtgg ccaactctgt catacagaaa gaggctcgacc gcttcaactg gtactcgtgg 1140
tgcccacaag cctcctggga aggttcagag ggctcaaga cgtcgcgaat cctgtatgag 1200
gaagtcgatg agtcocaggt ggaggtcatt cactccctct cctcctcgct agaagaaagg 1260
aagccggact cgtatcggta ccccaggaca ggcaagaaga atcccaagat tgcttgaaa 1320
ctggctgagt tccagactga cagccagggc aagatcgtct cgcaccagga gaaggagctg 1380
gtgcagccct tcagctcgtt gttccogaag gtggagtaca tcgccagggc cgggtggacc 1440
gggatggca aatagcctgt ggcctatgtc ctggaccggc cccagcagtg gctccagctc 1500
gtcctctcct ccccgccctt gttcatcccg agcacagaga atgaggagca gcggtagcc 1560
tctccagag ctgtcccag gaatgtccag cccttcccc aatcagaggg agagcagag 1620
gtctggatca atgttcagta catctctat ccttcccc aatcagaggg agagcagag 1680
ctctgctttc tcgcgcgcaa tgaatgcaag accggctctt gccattgta caaagtccc 1740
gcccgtttaa aatcccagg ctacgattgg agtgagccct tcagcccgg ggaagtgaa 1800
tttaagtccc ccaatgaaga agagattgct ctgaccagga gtgaatggga ggttttggcg 1860
aggccagctc ccaagggcac caaggacacg ccgctggagc accacctcta cgtggtcagc 1920
tatgagcggc cggcgagat cgtacgcctc accacggccg gcttctcca tagctgctcc 1980
atgagccaaga acttcgacat gttcgtcagc cactacaaga gcgtgagcac gcccccctg 2040
gtgcagctct acaagctgag cggcccggac gacgaccccc tgcaagaaga gcccccctc 2100
tgggttagca tgatggagg agccagctgc ccccggatt atgttctcc agagatctc 2160
catttccaca cgcctcggga tgtcggctc tacggcatga tctacaagcc ccacgcttg 2220
cagccaggga agaagcacc caccgtctc ttgtatatg gaggcccaca ggtgcagctg 2280
tgaataaact cctcaagg catcaagtac ttgggctca acacactgg cctccctggc 2340
tacccgtgg atgttcagta cggcaggggc tctgtcagc gagggtctc gttcgaagg 2400
gccctgaaaa accaaatgg ccaggtggag atcaggacc aggtggagg cctgagctc 2460
gtggccgaga agtatggctt catcgacctg agcagagttg ccatccatgg ctggtcctac 2520
ggggcttcc tctcgtcat ggggctaate cacaagcccc aggtgttcaa ggtggcctc 2580
gcccgtgccc cgttcacgt ctggatggcc taagacaag ggtacactga gcgtacatg 2640
gagctccctg agaacaacca ccaagctat gaggcgggt ccgttccctg gcagctggag 2700
aagctgccc atgagcccac cctgtctcc caactgatcc gagcagggaa accottaacc 2760
cacttttcc acacaaact cctcgtctcc caactgatcc gagcagggaa accottaacc 2820
ctccagatct accccaacga gagacacagt attcgtgccc ccgagtcggg ccgagcactat 2880
gaagtcagct tgcctgactt tctacaggaa tacctctgag cctgcccacc gggagcccct 2940
acalcaacag caaagtggct gcagcctccg cgggaaacca gggggaggg actgagtgcc 3000
ccggggccc cagtgaggca ctttgcctcg cccagcctg gccagcccg aggaaccgct 3060
gccttcaacc ccccgagccc ttttatcctt ttttaacgc tcttgggtt tatgtccgct 3120
gtctctggg tgcgagaca gtagatggt ggtctgggc cagccctcc tctcccggc 3180
ttctggagg agggatcac agcgtgatgg gcactgaga ggcagaga gactcaagg 3240
agcggctgc cttcgcctg gggctccctg tgactctca gtccttggc ccggccagcc 3300
accctccca gacccaagg atgcaattgc ctgtcccc ccggcagct cccaacttg 3360
atgttgtgt tttgtttgg gggatattt tcaataatt ttaaaagca gccggggcc 3420
ggtggctcac gtcgtaatc ccagcaactt gggagctga gggcgggga tcaactgag 3480
tggggagttc aagaccagcc tggccaact ggggaaaccc cgtctact aaaaataca 3540
aaaattagcc ggtgtggtg gcgctgctt ataactccag ctactcggga ggtgagga 3600
ggagaatcgc tfgaacccgg gagggtgagg ttgcggtag ccaagatcc accattgac 3660
tccagctgg gcaacaagg cgaactctg tctcaaaata aataaaaaat aaaaagaga 3720
aagcaaggg tgcctaaatc tagacttgg gtccacacgg ggcagcggg ttgcaacca 3780
gcacttgta ggtccattt cttcccagc ccgagcagag ggtcatcgg gccccaagg 3840
agaagcggc agggcccgg gggggcaaca cctgtggaca gcctcctgt cccaagctt 3900
tcaggaggg actgaaacgc accgaaactt cactctctgc tggctagtg ccgctgctc 3960
ctcccagcc cagcccccga gccacatgt tctgctgac cctgacac cagggttcc 4020
ggggttggga gctgaacct ccccaactca ggttatatt tctctccc ctctcctcc 4080
cggcaagac cctgcccagg gggggcaaaa aaaaaagtaa aaagaaaga aaaaaaaaa 4140
aaagaaaca accacctca catattatg aaaaaataa ttttctcga tcttattct 4200
ttataatta tgcgtggaag aagtagacac attaaacgat tccagtgga aacatgtcac 4260
ctg

```

WO 02/31134

PCT/US01/31874

<210> 35
<211> 879
<212> PRT
<213> Homo sapiens
<400> 35

```

Met Arg Lys Val Lys Lys Leu Arg Leu Asp Lys Glu Asn Thr Gly Ser
1      5      10      15
Trp Arg Ser Phe Ser Leu Asn Ser Glu Gly Ala Glu Arg Met Ala Thr
      20      25      30
Thr Gly Thr Pro Thr Ala Asp Arg Gly Asp Ala Ala Thr Asp Asp
35      40      45
Pro Ala Ala Arg Phe Gln Val Gln Lys His Ser Trp Asp Gly Leu Arg
50      55      60
Ser Ile Ile His Gly Ser Arg Lys Tyr Ser Gly Leu Ile Val Asn Lys
65      70      75      80
Ala Pro His Asp Phe Gln Phe Val Gln Lys Thr Asp Glu Ser Gly Pro
      85      90      95
His Ser His Arg Leu Tyr Tyr Leu Gly Met Pro Tyr Gly Ser Arg Glu
100      105      110
Asn Ser Leu Leu Tyr Ser Glu Ile Pro Lys Lys Val Arg Lys Glu Ala
115      120      125
Leu Leu Leu Leu Ser Trp Lys Gln Met Leu Asp His Phe Gln Ala Thr
130      135      140
Pro His His Gly Val Tyr Ser Arg Glu Glu Leu Leu Arg Glu Arg
145      150      155      160
Lys Arg Leu Gly Val Phe Gly Ile Thr Ser Tyr Asp Phe His Ser Glu
165      170      175
Ser Gly Leu Phe Leu Phe Gln Ala Ser Asn Ser Leu Phe His Cys Arg
180      185      190
Asp Gly Gly Lys Asn Gly Phe Met Val Ser Pro Met Lys Pro Leu Glu
195      200      205
Ile Lys Thr Gln Cys Ser Gly Pro Arg Met Asp Pro Lys Ile Cys Pro
210      215      220
Ala Asp Pro Ala Phe Phe Ser Phe Ile Asn Asn Ser Asp Leu Trp Val
225      230      235      240
Ala Asn Ile Glu Thr Gly Glu Glu Arg Arg Leu Thr Phe Cys His Gln
245      250      255
Gly Leu Ser Asn Val Leu Asp Asp Pro Lys Ser Ala Gly Val Ala Thr
260      265      270
Phe Val Ile Gln Glu Glu Phe Asp Arg Phe Thr Gly Tyr Trp Trp Cys
275      280      285
Pro Thr Ala Ser Trp Glu Gly Ser Glu Gly Leu Lys Thr Leu Arg Ile
290      295      300
Leu Tyr Glu Glu Val Asp Glu Ser Glu Val Glu Val Ile His Val Pro
305      310      315      320
Ser Pro Ala Leu Glu Glu Arg Lys Thr Asp Ser Tyr Arg Tyr Pro Arg
325      330      335
Thr Gly Ser Lys Asn Pro Lys Ile Ala Leu Lys Leu Ala Glu Phe Gln
340      345      350
Thr Asp Ser Gln Gly Lys Ile Val Ser Thr Gln Glu Lys Glu Leu Val
355      360      365
Gln Pro Phe Ser Ser Leu Phe Pro Lys Val Glu Tyr Ile Ala Arg Ala
370      375      380
Gly Trp Thr Arg Asp Gly Lys Tyr Ala Trp Ala Met Phe Leu Asp Arg
385      390      395      400
Pro Gln Gln Trp Leu Gln Leu Val Leu Leu Pro Pro Ala Leu Phe Ile
405      410      415

```

WO 02/31134

PCT/US01/31874

Pro Ser Thr Glu Asn Glu Glu Gln Arg Leu Ala Ser Ala Arg Ala Val
 420 425 430
 Pro Arg Asn Val Gln Pro Tyr Val Val Tyr Glu Glu Val Thr Asn Val
 435 440 445
 Trp Ile Asn Val His Asp Ile Phe Tyr Pro Phe Pro Gln Ser Glu Gly
 450 455 460
 Glu Asp Glu Leu Cys Phe Leu Arg Ala Asn Glu Cys Lys Thr Gly Phe
 465 470 475 480
 Cys His Leu Tyr Lys Val Thr Ala Val Leu Lys Ser Gln Gly Tyr Asp
 485 490 495
 Trp Ser Glu Pro Phe Ser Pro Gly Glu Asp Glu Phe Lys Cys Pro Ile
 500 505 510
 Lys Glu Glu Ile Ala Leu Thr Ser Gly Glu Trp Glu Val Leu Ala Arg
 515 520 525
 His Gly Ser Lys Gly Thr Lys Asp Thr Pro Leu Glu His His Leu Tyr
 530 535 540
 Val Val Ser Tyr Glu Ala Ala Gly Glu Ile Val Arg Leu Thr Thr Pro
 545 550 555 560
 Gly Phe Ser His Ser Cys Ser Met Ser Gln Asn Phe Asp Met Phe Val
 565 570 575
 Ser His Tyr Ser Ser Val Ser Thr Pro Pro Cys Val His Val Tyr Lys
 580 585 590
 Leu Ser Gly Pro Asp Asp Asp Pro Leu His Lys Gln Pro Arg Phe Trp
 595 600 605
 Ala Ser Met Met Glu Ala Ala Ser Cys Pro Pro Asp Tyr Val Pro Pro
 610 615 620
 Glu Ile Phe His Phe His Thr Arg Ser Asp Val Arg Leu Tyr Gly Met
 625 630 635 640
 Ile Tyr Lys Pro His Ala Leu Gln Pro Gly Lys Lys His Pro Thr Val
 645 650 655
 Leu Phe Val Tyr Gly Gly Pro Gln Val Gln Leu Val Asn Asn Ser Phe
 660 665 670
 Lys Gly Ile Lys Tyr Leu Arg Leu Asn Thr Leu Ala Ser Leu Gly Tyr
 675 680 685
 Ala Val Val Val Ile Asp Gly Arg Gly Ser Cys Gln Arg Gly Leu Arg
 690 695 700
 Phe Glu Gly Ala Leu Lys Asn Gln Met Gly Gln Val Glu Ile Glu Asp
 705 710 715 720
 Gln Val Glu Gly Leu Gln Phe Val Ala Glu Lys Tyr Gly Phe Ile Asp
 725 730 735
 Leu Ser Arg Val Ala Ile His Gly Trp Ser Tyr Gly Gly Phe Leu Ser
 740 745 750
 Leu Met Gly Leu Ile His Lys Pro Gln Val Phe Lys Val Ala Ile Ala
 755 760 765
 Gly Ala Pro Val Thr Val Trp Met Ala Tyr Asp Thr Gly Tyr Thr Glu
 770 775 780
 Arg Tyr Met Asp Val Pro Glu Asn Asn Gln His Gly Tyr Glu Ala Gly
 785 790 795 800
 Ser Val Ala Leu His Val Glu Lys Leu Pro Asn Glu Pro Asn Arg Leu
 805 810 815
 Leu Ile Leu His Gly Phe Leu Asp Glu Asn Val His Phe His Thr
 820 825 830
 Asn Phe Leu Val Ser Gln Leu Ile Arg Ala Gly Lys Pro Tyr Gln Leu
 835 840 845
 Gln Ile Tyr Pro Asn Glu Arg His Ser Ile Arg Cys Pro Glu Ser Gly
 850 855 860
 Glu His Tyr Glu Val Thr Leu Leu His Phe Leu Gln Glu Tyr Leu
 865 870 875

WO 02/31134

PCT/US01/31874

<210> 36
<211> 4180
<212> DNA
<213> Homo sapiens
<400> 36

```

caggccgcgc cctgggtcgc tcaacttccg ggtcaaaagt gcctgagccg gcgggtcccc 60
tgtgtccgcc gcggtgtcgc tcccocgctc ccgcaacttc cggggtcgca gtcccgggca 120
tggagccgcg accgtgaggg gccgctggac ccgggacgac ctgccagtc cggccgcgcg 180
cccacgtccc ggtctgtctc cccgcctgc agctggaatg gaggctctct ggacccttta 240
gaaggcaacc ctgccctcct gaggtcagct gagcggtaa tgccgaaggt taagaaactg 300
cgctcgaca aggagaacac cggaaagtgg agaagcttct cgctgaatc cgagggggct 360
gagagatgg ccaccaccgg gaccccaacg gccgaccgag ccgacgcgag cgcacacgat 420
gaccocggcg ccgcttccca ggtgcagaag cactcgtggg aggggtccg gacgatcatc 480
cacggcagcc gcaagtactc gggcctcatt gtccaacaag ccgcccacga cttccagttt 540
gtgcagaaga cggatgagtc tggccccacc tcccaccgcc tctactact gggaatgcca 600
tatggcagcc gagagaactc cctcctctac tctgagattc ccaagaaggt ccggaagag 660
gctctgtctc tctgtcctcg gaagcagatg ctggatcatt tccagccacc gccccaccat 720
ggggtctact ctccgggagga ggagctgctg agggagcgga aacgctggg ggtctctggc 780
atcacctcct acgacttcca cagcagagat ggcctcttcc tcttccaggc cagcaaacgc 840
ctcttccact gccgcgacgg cggcaagaac ggcctctatg tgtcccctat gaaaccgctg 900
gaaatcaaga cccatgtctc agggccccgg atggacccca aaatctgccc tgcgcacctt 960
gccttctctc cctctcatca taacagcgac ctgtgggtgg ccaacatcga gacggcgag 1020
gagcggcgcc tgacctctcg ccaaccaagt ttatccaatg tcctggatga cccaagtct 1080
gggggtctgg cacacctcgt cacaaggaag gacttgacc gottcactgg gtactggtgg 1140
tgccccacag cctcctggga aggttcagag ggcctcaaga cgtcggaaat cctgtatgag 1200
aaagcagatg agtccgaggt ggaggtcatt cacgtcccct ctctcgctc agaagaaagg 1260
aagaccgact cgtatcgta ccccaggaca ggagcaaga atcccagat tgcctgaaa 1320
gtgcagccct tcagctcgct gttcccgaag gtggagtaca tgcgcaaggc cgggtggacc 1380
cgggatggca aatagcctcg ggcctatgtc ctggaccggc cccagcagtg gctccagctc 1440
gtcctctccc ccccgccctt gttcctcccg agcacagaga atgaggagca gcgctagcc 1500
tetcccagag ctgtcccacg gaatgtccag ccctatgtgg tgtacgagga ggtcaccaac 1560
gtctggatca atgtctatga catctcttat cccttcccc aatcagaggg agaggacgag 1620
ctctgtcttc tccgcccaca tgaatgcaag accgctctct gccatttcta caaagtcacc 1680
gccgttttaa aatcccaggg ctacgattgg agtgagccct tcagcccggg ggaagtggaa 1740
tttaagtgcc ccattaagga agagattgct ctgaccagcg gtgaatggga ggttttggcg 1800
aggcagccct ccaagggcac caaggacacg ccgctggagc accacctcta cgtggtcagc 1860
tatgagccgg ccggcgagat cgtacgcctc accacgcccg gcttctcca tagctgctcc 1920
atgagccaga acttcgacat gttcgtcagc cactacagca gcctgagcac gccgctctcc 1980
gtgcaagtcct acaagctgag ccgcccggac gacgacccc tgcacaagca gcccgctctc 2040
tggctagaca tgatggaggc agcccagctgc ccccggatt atgttctcc agagatcttc 2100
catttccaca cgcgctcggc tbtgagctct taaggcaatg tctacaagc ccagccttg 2160
cagccaggga agaagcaacc cacctctctc tttglatatg gagggcccca ggtgcagctg 2220
gtgaaatact ccttcaagg catcaagtac ttgctggctc acacactgac ctccctgggc 2280
tacgcgctgg ttgtgattga cggcaggggc tctgtcagc gagggcttcg gttcgaagg 2340
gccctgaaaa accaaatggg ccaggtggag atcgagacc aggtggagg cctgcagttc 2400
gtggccgaga agtatggctt catcgacctg agccagttg ccatccatgg ctggtctcac 2460
ggggcttccc tctcgtcat ggggtcaatc cacagcccc aggtgtcaa ggtggcctac 2520
gcgggtgccc cgttccagct ctggatggcc tacgaacacg ggtacaotga gcctacatg 2580
gacgtccctg agaacaacca gccagctat gaggcgggtt ccgtgcccct gcagctggag 2640
aagttgccca atgagcccaa ccgcttgctt atcctccacg gcttctgga cgaaaaactg 2700
cacttttccc accaaaactt cctcgtctcc caactgatcc gagcagggaa acctaccag 2760
ctccagatct accccaacga gagacacagt attcgtgccc ccgagtcggg cgagcactat 2820
gaagtcaact tgcctgactt tctacagaaa taactctgag cctgccacc gccgaccgcc 2880
acatcacagc acaagtggct gcagcctccg cggggaacca gggggaggg actgagtgcc 2940
ccgccccccc cagtgggcca ctttctccc cccagcctg gccaccccc aggagccgct 3000
gcttcaacc ccccgaccgc ttttatcctt ttttaaccgc tcttgggttt tatgtccgct 3060
gcttcttggg tgccgagaca gagagatggt ggtctcgggc cagcccctcc tctcccggcc 3120

```

WO 02/31134

PCT/US01/31874

```

ttctggagg aggggtcac acgctgatgg gcactggaga ggccagaaga gactcagagg 3240
agcgggtcgc ctccgcctg gggctccctg tgacctctca gtcccctggc cggccagacc 3300
accgtcccga gcccccgaag atgcaattgc ctgtcccctc cggccagcct ccccacttg 3360
atgtttgtgt ttgtttggg gggatatttt tcataattat tlaaaagaca gggcggggcg 3420
gggtgctcac gtctgtaatc ccagcaacttt gggaggctga ggccggcgga tcacctgagg 3480
ttggggagttc aagaccagcc tggccaacct ggggaaacct cgtctctact aaaatacaa 3540
aaaattagcc ggggtgtggtg gcgcgtgacct ggggaaacct ctactcggga ggtgaggca 3600
ggagaatcgc tgaaccocgg gagggtgagc ttgcgggtgag ccaagatcgc accattgac 3660
tccagctgg gcaacaagag cgaactctg tctcaaaaata aataaaaaat aaagacaga 3720
aagcaaggg tgctcaaatc tagacttggg gtccacacgc ggcagcgggg ttgcaaccga 3780
gcactgtga ggtctcattt ctcccgaagc ccgactttca ggcaggcact gaaacgcacc 3840
acatgtgtct gctgaccocg tacacaccag gggttccggg gtggggagct gaaccatccc 3900
cacctcaggg ttatatttcc ctctcccctt cctcccctc caagagctct gacagggggc 3960
ggcaaaaaaa aaagtaaaaa gaaaagaaaa aaaaaaaaaa gaaacaaacc acctctacat 4020
attatggaaa gaaaatattt ttgtcgattc ttattctttt ataattatgc gtggaagagg 4080
tagacacatt aaacgattcc agttggaacc atgtcaactg 4140

```

```

<210> 37
<211> 819
<212> PRT
<213> Homo sapiens
<400> 37

```

```

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

```

WO 02/31134

PCT/US01/31874

Phe Val Ile Gln Glu Glu Phe Asp Arg Phe Thr Gly Tyr Trp Trp Cys
 275 280 285
 Pro Thr Ala Ser Trp Glu Gly Ser Glu Gly Leu Lys Thr Leu Arg Ile
 290 295 300
 Leu Tyr Glu Glu Val Asp Glu Ser Glu Val Glu Val Ile His Val Pro
 305 310 315 320
 Ser Pro Ala Leu Glu Glu Arg Lys Thr Asp Ser Tyr Arg Tyr Pro Arg
 325 330 335
 Thr Gly Ser Lys Asn Pro Lys Ile Ala Leu Lys Leu Ala Glu Phe Gln
 340 345 350
 Thr Asp Ser Gln Gly Lys Ile Val Ser Thr Gln Glu Lys Glu Leu Val
 355 360 365
 Gln Pro Phe Ser Ser Leu Phe Pro Lys Val Glu Tyr Ile Ala Arg Ala
 370 375 380
 Gly Trp Thr Arg Asp Gly Lys Tyr Ala Trp Ala Met Phe Leu Asp Arg
 385 390 395 400
 Pro Gln Gln Trp Leu Gln Leu Val Leu Leu Pro Pro Ala Leu Phe Ile
 405 410 415
 Pro Ser Thr Glu Asn Glu Glu Gln Arg Leu Ala Ser Ala Arg Ala Val
 420 425 430
 Pro Arg Asn Val Gln Pro Tyr Val Val Tyr Glu Glu Val Thr Asn Val
 435 440 445
 Trp Ile Asn Val His Asp Ile Phe Tyr Pro Phe Pro Gln Ser Glu Gly
 450 455 460
 Glu Asp Glu Leu Cys Phe Leu Arg Ala Asn Glu Cys Lys Thr Gly Phe
 465 470 475 480
 Cys His Leu Tyr Lys Val Thr Ala Val Leu Lys Ser Gln Gly Tyr Asp
 485 490 495
 Trp Ser Glu Pro Phe Ser Pro Gly Glu Asp Glu Phe Lys Cys Pro Ile
 500 505 510
 Lys Glu Glu Ile Ala Leu Thr Ser Gly Glu Trp Glu Val Leu Ala Arg
 515 520 525
 His Gly Ser Lys Gly Thr Lys Asp Thr Pro Leu Glu His His Leu Tyr
 530 535 540
 Val Val Ser Tyr Glu Ala Ala Gly Glu Ile Val Arg Leu Thr Thr Pro
 545 550 555 560
 Gly Phe Ser His Ser Cys Ser Met Ser Gln Asn Phe Asp Met Phe Val
 565 570 575
 Ser His Tyr Ser Ser Val Ser Thr Pro Pro Cys Val His Val Tyr Lys
 580 585 590
 Leu Ser Gly Pro Asp Asp Asp Pro Leu His Lys Gln Pro Arg Phe Trp
 595 600 605
 Ala Ser Met Met Glu Ala Ala Ser Cys Pro Pro Asp Tyr Val Pro Pro
 610 615 620
 Glu Ile Phe His Phe His Thr Arg Ser Asp Val Arg Leu Tyr Gly Met
 625 630 635 640
 Ile Tyr Lys Pro His Ala Leu Gln Pro Gly Lys Lys His Pro Thr Val
 645 650 655
 Leu Phe Val Tyr Gly Gly Pro Gln Val Gln Leu Val Asn Asn Ser Phe
 660 665 670
 Lys Gly Ile Lys Tyr Leu Arg Leu Asn Thr Leu Ala Ser Leu Gly Tyr
 675 680 685
 Ala Val Val Val Ile Asp Gly Arg Gly Ser Cys Gln Arg Gly Leu Arg
 690 695 700
 Phe Glu Gly Ala Leu Lys Asn Gln Met Gly Gln Val Glu Ile Glu Asp
 705 710 715 720
 Gln Val Glu Gly Leu Gln Phe Val Ala Glu Lys Tyr Gly Phe Ile Asp
 725 730 735

WO 02/31134

PCT/US01/31874

Leu Ser Arg Val Ala Ile His Gly Trp Ser Tyr Gly Gly Phe Leu Ser
 740 745 750
 Leu Met Gly Leu Ile His Lys Pro Gln Val Phe Lys Ala Gln Pro Leu
 755 760 765
 Ala Tyr Pro Pro Arg Leu Pro Gly Arg Lys Arg Ala Leu Phe Pro His
 770 775 780
 Lys Leu Pro Arg Leu Pro Thr Asp Pro Ser Arg Glu Thr Leu Pro Ala
 785 790 795 800
 Pro Asp Leu Pro Gln Arg Glu Thr Gln Tyr Ser Leu Pro Arg Val Gly
 805 810 815
 Arg Ala Leu

<210> 38
 <211> 4120
 <212> DNA
 <213> Homo sapiens
 <400> 38

caggccgcgc cctgggtcgc tcaactccg ggtcaaaagt gcctgagcgc ggggtcccc 60
 tgtgtccgcc gggctgtcgc tccccgcctc ccgccacttc cggggtcgca gtcccgggca 120
 tggagccgcg accgtgaggc gccgctggac ccgggacgac ctgcccagtc cggccgcgcc 180
 cccacgtccc ggtctgtgtc ccacgcctgc agctggaatg gaggctctct ggaccttcta 240
 gaaggaaccc ctgcccctct gaggtcagct gaggcgttaa tgcggaaggt taagaaactg 300
 gcctgggaca aggagaacac cgaagtctgg agaagctctc cgtgaatto cgagggggct 360
 gagaggatgg ccaccacggg gaccocaaag gccgaccgag gcgacgcgag ccgacacagat 420
 gaccocggcc cccgcttcca ggtgcagaag cactcgtggg acgggctccg gaggcatctc 480
 cacgcagcgc gcaagtactc gggcctcatt gtcaacaagg cgcocccaga cttccagttt 540
 gtgcagaaga cggatgagtc tgggcccacc tcccaccgoc tctaactact gggaatgcca 600
 tatggcagcc gagagaactc cctcctctac tctgagattc ccaagaaggt ccggaagagag 660
 gctctgtctc tctgtcctg gaagcagatg ctggatcatt tccagccacc gcccccacat 720
 ggggtctact ctccggagga ggagctgctg agggagcggg aacgcctggg ggtctctcgc 780
 atcaactcct acgacttcca cagcgagagt ggcctcttcc tcttccaggc cagcaacagc 840
 ctcttccact gccgcagcgg cggcaagaac ggcttccatg tgtcccctat gaaaccgctg 900
 gaaatcaaga ccagtgctc agggcccggg atggaccoca aaactctccc tgcccaccct 960
 gocttcttct ccttcabcaa taacagcgac ctgtgggtgg ccaacatcga gacaggcgag 1020
 gagcggcggc tgacctctct ccaccaaggt ttatccaatg tctctgagga ccccaagctc 1080
 ggggtgtggt ccacctctgt catacaggaa gacttcaacc gcttcaactg gactgtgtgg 1140
 tgccccacag cctcctggga aggttcagag ggcctcaaga cgttcgaat cctgtatgag 1200
 gaagtcgat agtccgaggt ggaagtcatt caggtccctc ctctcgctc agaagaagg 1260
 ctggctgagt tccagactga cagccagggc aagatogtct cgcaccagga gaaggagctg 1320
 gtgcagccct tcaagctcgt gttcccgaag gtggagtaca tgcaccagga cgggtggacc 1380
 cgggatggca aatacgcctg ggcctatgttc ctggaccggc ccagcagtg gctccagctc 1440
 gctcctctcc ccccgccctc gttcatcccg agcacagaga atgaggaga gcggtagacc 1500
 tctgcoagag ctgtcccag gaatgtccag ccgatgtgg tgtacagga ggtcaccac 1560
 gtctgatatc atgttcatga catctctat cccttcccc aatcagagg agaggacgag 1620
 ctctgtttc tccggcccaa tgaatgcaag accggtctct gccatttga caaagtacc 1680
 gccgttttaa aatcccagg ctacgattgg agtgagccct tcagcccgg ggaagatgaa 1740
 ttttaagtcc caatgaaga agagattgct ctgacccgg gtgaatggg ggttttgccg 1800
 aggcaggctc ccaaggacac caaggacacg ccgctggagc accactota cgtgtcagc 1860
 tatgaggcgg ccgagagat cgtacgctc accaccggc gcttctcca tagctgtcc 1920
 atgagccaga acttcgacat gttgtcagc cactacagca gcgtgagcgc gccccctgc 2040
 gtgcaactct acaagctgag cggcccggac gaccaccoco tgcacaagca gccccgctc 2100
 tggcgtagca tgatggaggc agccagctgc ccccggatt atgttctcc agagatctc 2160
 catttccaca cgcctcggg tgtgcgctc tacggctatg tctcaagcc ccacgcttg 2220
 cagccaggga agaagcacc caccgtctc tttgtatatg gaggcccca ggtcagctg 2280
 gtgaataact ccttcaagg catcaagtac ttgcgctca acacactgc cctccctggc 2340
 taocccgtgg ttgtgattga cggcaggggc tctgtcagc gagggtctcg gttcgaagg 2400
 gcctgaaaa accaatggg ccaggtggag atcgaggacc aggtggagg cctgcagttc 2460

WO 02/31134

PCT/US01/31874

```

gtggccgaga agtatggctt catcgacctg agccgagttg ccattccatgg ctggctcctac 2520
ggggggttcc totcgctcat ggggctaato cacaagcccc aggtgttcaa ggcocaaaccg 2580
ctgtcttacc ctccacggct tcttggacga aaactgtcac tttttccaca caaacttctct 2640
cgtctcccaa ctgatccgag cagggaacc ttaccagctc cagatctacc ccaacgagag 2700
acaacgatatt cgctgccccg agtccggcga gcactatgaa gtcaactgtgc tgcactttct 2760
acaggaatac ctctgacctt gccacccggg agccgcccaca tcacagcaca agtggctgca 2820
gaactccggg ggaaccaggc gggaggact gactggcccg cgggcccag tgaggcaact 2880
tgtccgcccc agcctgccc agccccgag agccctgccc ttaccgcccc cgaagccttt 2940
tatccttttt taaacgctct tgggttttat gtccgctgct tottggttgc cgaagacagag 3000
agatgttgtt ctggggccag cccctcctct ccccgcttct tggggaggag aggtcaacag 3060
ctgatgggca ctggagaggc cagaagagac tcagaggagc gggctgcctt ccgctggggg 3120
ctccctgtga cctctcagtc cctggcccgc gccaccacc gtcccagca ccaagcagtg 3180
caattgcctg tccccccggg ccagcctccc caactgatg tttgtgttt gtttgggggg 3240
atatttttca taattattta aaagacaggg cggggcggtt ggtcagctc tgaatccca 3300
gcactttggg aggtgagggc gggcggatca cctgaggttg ggagttcaag accagcctgg 3360
ccaacatggg gaaacccctt ctctactaaa aatacaaaaa attagccggg tgggtggcg 3420
cgtgcctata atcccagcta ctggggaggc tgaggcagga gaatcgcttg aaccggggag 3480
gtggaggttg cggtgagcca agatcgacc attgcactcc agcctgggca acaagagcga 3540
aactctgtct caaaataaat aaaaaataa agacagaaag caaggggtgc ctaaatctag 3600
acttgggttc cacaccgggc agcgggttg caaccagca cctggtaggc tccatttctt 3660
cccaagcccc agcagagggg catcggggcc ccacaggaga agcggccagg gcccgcgggg 3720
ggcaccacct gtggacagcc ctctgtccc caagcttca gccagccact gaaacgcacc 3780
gaacttccac gctctgctgg tcagtggcgg ctgtccccto cccagcccag ccgcccagcc 3840
acatgtgtct gctcagccc tacacaccg gggttcccgg gttgggagct gaacctccc 3900
cacctcaggg ttabattttc ctctcccctt ccctcccgc caagagctct gccagggggc 3960
ggcaaaaaaa aaagtataaa gaaaagaaa aaaaaaaaaa gaaacaaacc acctctacat 4020
attatggaaa gaaaatattt ttgtcgattc ttattctttt ataattatgc gtggaagaag 4080
tagacacatt aaacgattcc agttggaaac atgtcacctg 4120

```

```

<210> 39
<211> 819
<212> PRT
<213> Homo sapiens
<400> 39

```

```

Met Arg Lys Val Lys Leu Arg Leu Asp Lys Glu Asn Thr Gly Ser
1 5 10 15
Trp Arg Ser Phe Ser Leu Asn Ser Glu Gly Ala Glu Arg Met Ala Thr
20 25 30
Thr Gly Thr Pro Thr Ala Asp Arg Gly Asp Ala Ala Thr Asp Asp
35 40 45
Pro Ala Ala Arg Phe Gln Val Gln Lys His Ser Trp Asp Gly Leu Arg
50 55 60
Ser Ile Ile His Gly Ser Arg Lys Tyr Ser Gly Leu Ile Val Asn Lys
65 70 75 80
Ala Pro His Asp Phe Gln Phe Val Gln Lys Thr Asp Glu Ser Gly Pro
85 90 95
His Ser His Arg Leu Tyr Tyr Leu Gly Met Pro Tyr Gly Ser Arg Glu
100 105 110
Asn Ser Leu Leu Tyr Ser Glu Ile Pro Lys Lys Val Arg Lys Glu Ala
115 120 125
Leu Leu Leu Leu Ser Trp Lys Gln Met Leu Asp His Phe Gln Ala Thr
130 135 140
Pro His His Gly Val Tyr Ser Arg Glu Glu Glu Leu Leu Arg Glu Arg
145 150 155 160
Lys Arg Leu Gly Val Phe Gly Ile Thr Ser Tyr Asp Phe His Ser Glu
165 170 175
Ser Gly Leu Phe Leu Phe Gln Ala Ser Asn Ser Leu Phe His Cys Arg
180 185 190

```

WO 02/31134

PCT/US01/31874

Asp Gly Gly Lys Asn Gly Phe Met Val Ser Pro Met Lys Pro Leu Glu
195 200 205
Ile Lys Thr Gln Cys Ser Gly Pro Arg Met Asp Pro Lys Ile Cys Pro
210 215 220
Ala Asp Pro Ala Phe Phe Ser Phe Ile Asn Asn Ser Asp Leu Trp Val
225 230 235 240
Ala Asn Ile Glu Thr Gly Glu Glu Arg Arg Leu Thr Phe Cys His Gln
245 250 255
Gly Leu Ser Asn Val Leu Asp Asp Pro Lys Ser Ala Gly Val Ala Thr
260 265 270
Phe Val Ile Gln Glu Glu Phe Asp Arg Phe Thr Gly Tyr Trp Trp Cys
275 280 285
Pro Thr Ala Ser Trp Glu Gly Ser Glu Gly Leu Lys Thr Leu Arg Ile
290 295 300
Leu Tyr Glu Glu Val Asp Glu Ser Glu Val Glu Val Ile His Val Pro
305 310 315 320
Ser Pro Ala Leu Glu Glu Arg Lys Thr Asp Ser Tyr Arg Tyr Pro Arg
325 330 335
Thr Gly Ser Lys Asn Pro Lys Ile Ala Leu Lys Leu Ala Glu Phe Gln
340 345 350
Thr Asp Ser Gln Gly Lys Ile Val Ser Thr Gln Glu Lys Glu Leu Val
355 360 365
Gln Pro Phe Ser Ser Leu Phe Pro Lys Val Glu Tyr Ile Ala Arg Ala
370 375 380
Gly Trp Thr Arg Asp Gly Lys Tyr Ala Trp Ala Met Phe Leu Asp Arg
385 390 395 400
Pro Gln Gln Trp Leu Gln Leu Val Leu Leu Pro Pro Ala Leu Phe Ile
405 410 415
Pro Ser Thr Glu Asn Glu Glu Gln Arg Leu Ala Ser Ala Arg Ala Val
420 425 430
Pro Arg Asn Val Gln Pro Tyr Val Val Tyr Glu Glu Val Thr Asn Val
435 440 445
Trp Ile Asn Val His Asp Ile Phe Tyr Pro Phe Pro Gln Ser Glu Gly
450 455 460
Glu Asp Glu Leu Cys Phe Leu Arg Ala Asn Glu Cys Lys Thr Gly Phe
465 470 475 480
Cys His Leu Tyr Lys Val Thr Ala Val Leu Lys Ser Gln Gly Tyr Asp
485 490 495
Trp Ser Glu Pro Phe Ser Pro Gly Glu Asp Glu Phe Lys Cys Pro Ile
500 505 510
Lys Glu Glu Ile Ala Leu Thr Ser Gly Glu Trp Glu Val Leu Ala Arg
515 520 525
His Gly Ser Lys Gly Thr Lys Asp Thr Pro Leu Glu His His Leu Tyr
530 535 540
Val Val Ser Tyr Glu Ala Ala Gly Glu Ile Val Arg Leu Thr Thr Pro
545 550 555 560
Gly Phe Ser His Ser Cys Ser Met Ser Gln Asn Phe Asp Met Phe Val
565 570 575
Ser His Tyr Ser Ser Val Ser Thr Pro Pro Cys Val His Val Tyr Lys
580 585 590
Leu Ser Gly Pro Asp Asp Asp Pro Leu His Lys Gln Pro Arg Phe Trp
595 600 605
Ala Ser Met Met Glu Ala Ala Ser Cys Pro Pro Asp Tyr Val Pro Pro
610 615 620
Glu Ile Phe His Phe His Thr Arg Ser Asp Val Arg Leu Tyr Gly Met
625 630 635 640
Ile Tyr Lys Pro His Ala Leu Gln Pro Gly Lys Lys His Pro Thr Val
645 650 655

WO 02/31134

PCT/US01/31874

Leu Phe Val Tyr Gly Gly Pro Gln Val Gln Leu Val Asn Asn Ser Phe
 660 665 670
 Lys Gly Ile Lys Tyr Leu Arg Leu Asn Thr Leu Ala Ser Leu Gly Tyr
 675 680 685
 Ala Val Val Val Ile Asp Gly Arg Gly Ser Cys Gln Arg Gly Leu Arg
 690 695 700
 Phe Glu Gly Ala Leu Lys Asn Gln Met Gly Gln Val Glu Ile Glu Asp
 705 710 715 720
 Gln Val Glu Gly Leu Gln Phe Val Ala Glu Lys Tyr Gly Phe Ile Asp
 725 730 735
 Leu Ser Arg Val Ala Ile His Gly Trp Ser Tyr Gly Gly Phe Leu Ser
 740 745 750
 Leu Met Gly Leu Ile His Lys Pro Gln Val Phe Lys Ala Gln Pro Leu
 755 760 765
 Ala Tyr Pro Pro Arg Leu Pro Gly Arg Lys Arg Ala Leu Phe Pro His
 770 775 780
 Lys Leu Pro Arg Leu Pro Thr Asp Pro Ser Arg Glu Thr Leu Pro Ala
 785 790 795 800
 Pro Asp Leu Pro Gln Arg Glu Thr Gln Tyr Ser Leu Pro Arg Val Gly
 805 810 815
 Arg Ala Leu

<210> 40
 <211> 4037
 <212> DNA
 <213> Homo sapiens
 <400> 40

caggccgcgc cctgggtcgc tcaacttcgc ggtcaaaagg gcctgagccg gcgggtcccc 60
 tgtgtccgcc gggctgtcgc tccccgcctc ccgccaactc cggggtcgca gtccccggca 120
 tggagccgcg accgtgaggc gccgctggac ccgggaacac ctgcccagtc cggccgcccc 180
 cccacgtccc ggtctgtgtc ccacgcctgc agctggaatg gaggctctct ggacccttta 240
 gaaggcaccc ctgcccctct gaggtcagct gacgggttaa tggggaaggt taagaaactg 300
 ccctctggaca aggagaacac cggaaattgg agaagcttct cgctgaattc ctagggggct 360
 gagaggatgg ccaccaccgg gaccocaaag ccgacccgag gcgacgcagc cgcacacgat 420
 gaccggcgcc ccgcttccca ggtgcagaag cactcgtggg acgggctccg gacacatc 480
 cacggcagcc gcaagtactc yggcctcatt gtcaacaagg cgcaccaaga cttccagttt 540
 gtgcagaaga cggatgagtc tgggccccac tcccccgcc tctactacct gggaaatgca 600
 talggcagcc gagagaactc cctcctctac tctgagattc ccaagaaggt ccggaaagg 660
 gctctctcgc tctgtcctg gaagcagatg ctggatcatt tccaggccac gcccaacct 720
 gggctctact ctccggagga gagctgctg agggagcggg aacgctggg ggtctctggc 780
 atcaactcct agcacttcca cagcgagagt ggcctcttcc tcttccagc cagcaaacac 840
 ctcttccact gccgtagcgg cggcaagaac ggccttctgg tgtcccctat gaaaccctg 900
 gaaatcaaga ccagtgctc agggccccgg atggaccoca aaatctgccc tgcgcacct 960
 gctctctctc cttcatcaa taaccggcag ctgtgggtgg ccaacatcga gaccagcgag 1020
 gagcgggggc tgaacttctg ccaccaaggt tlatccaatg tctctgatga cccaagtct 1080
 gcgggtgtgg ccactctgt caticagaaa gactctgacc gcttactggt gactgtgtgg 1140
 tgcaccaacg cctcctggga aggttcagag ggcctcaaga cgtcgaat cctgtatgag 1200
 gaagtcgatg agtccgaggt ggaggtcatt cacgtccct ctctcgctg agaagaaagg 1260
 aagacggact cgtatcggta ccccaggaca ggacagcaaga atccaagat tgcctgaaa 1320
 ctggctgagt tccagactga cagccagggc aagatcgtct ccagccagga gaaggagctg 1380
 gtgcagccct tcaactcgtc gttcccgaag gtggagtaca tgcgcaaggc cgggtggacc 1440
 cgggatggca aatacgcctg ggcctatgtc ctggaccggc ccagcagtg gctccagctc 1500
 gtctctctcc ccccgccctc gttcatccc agcaacagaga atgaggagca gcgctagcc 1560
 tctgccaagc ctgtcccagc gaatgtccag ccgtatgtgg tgtcagagga ggtcaccac 1620
 gtctggatca atgttcatga catottctat cccttcccc aatcagagg agaggcag 1680
 ctctgtcttc tcccgccaa tgaattcaag accggtctct gccatttga caaagtcacc 1740
 gccgttttaa aatccaggc ctaagattgg agtgagccct ttagccccgg ggaagatgaa 1800
 ttttaagtc ccattaagga agagattgct ctgaccagcg gtgaatggga ggttttggcg 1860

WO 02/31134

PCT/US01/31874

```

agggcaggct ccaagggcac caaggacag ccgctggagc accacctcta cgtggtcagc 1920
tatgagggcg cggcgagat cgtacgcctc acccagcccg gcttctccca tagctgctcc 1980
atgagccaga acttcgacat gttcgtcagc cactacagca gcgtgagcac gccgcctgc 2040
gtgcaagctc acaagctgag cggcccgagc gaogcaccoc tgcacaagca gccocgcttc 2100
tgggttagca tgatggaggc agccaagctc ccccgatt atgttctcc agagatcttc 2160
catttccaca cgcgctcggg tgtgcgctc taecggatga tctacaagc ccaocgcttg 2220
cagccagggg agaagcacc caccgtcctc tttgtatag gaggcccca ggtgcagctg 2280
gtgataaact ccttcaagg catcaagtac ttccggctca acacactggc ctccctgggc 2340
tacccgctgg tggtagtga cggcaggggc tctgtcagc gagggttcg gttcgaagg 2400
gccctgaaaa accaaatggg ccaggtggag atcagaggcc aggtggagg cctgagcttc 2460
gtggccgaga agtatggctt catcgacctg agccagttg ccatocatgg ctggtcctac 2520
gggggttcc tctgctcat ggggttaac cacaagccc aggtgttcaa ggccaaccg 2580
cttcttatc ctccaaggct tctggagca aaactgcaac ttttccaca caaactctct 2640
cgtctccaa ctgatccgag caggaaacc ttacagctc cagatctacc ccaacgagag 2700
acacagtatt cgtgccccg agtcgggca gcaactgaa gtcacgttc tgcacttct 2760
acaggaatc ctctgagcct gccaccggg agccgacca tcacagcaca agtggctga 2820
gctccgctgg gaaaccaggc gggaggact gagtggccc cggcccccag tgaggcact 2880
tgtccgccc agcgtggcc agcccagag agccgctgc ttaccgcc cgacgcttt 2940
tatcctttt taaacgctc tgggtttat gtcgctgct tcttggttg cgagacagag 3000
agatggtgg ctggggccg cccctcctc ccccgcttc tgggagagg aggtcacaag 3060
ctgatgggca ctggagagg cagaagagac tcagaggagc gggctgctt ccgctgggg 3120
ctccctgtga cctctcagtc cctggcccg gccagccacc tctccagca ccaagcatg 3180
caattgcctg tccccccgg ccagcctcc caactgatg tttgtgttt gttgggggg 3240
atattttca taattattta aaagacagc cggcgcggt ggtcacctg ttaatacca 3300
gactttggg aggtcaggc gggcggatca cctgagttg gtagttcaag accgctgctg 3360
ccaactggg gaaacccct ctctactaaa aatcaaaaa attagccggg tgggtggcg 3420
cgtcctata atcccagcta ctcgaggc tgaggcagga gaatcgctg aaccggggg 3480
gtggagttg cgtgagcca agatcgacc attgactcc agcctgggca acaagagca 3540
aactctgtc caaaataaat aaaaaataa agacagaaa caaggggtg ctaaatctag 3600
acttgggtc cacaccggc agcgggttg caaccagca cctgtaggc tccattctt 3660
cccagcccg actttcaggc agcactgaa acgcaccgaa ctccaagct ctgctgctc 3720
gtggggctg tccctcccc agcccagcc cccagccaca tgtgtctgc tgaccgtac 3780
acaccaggg tccgggggtt gggagctgaa ccatccccac ctcagggtta tatctctc 3840
tcccttccc tcccggccaa gagctctgcc agggcgggc aaaaaaaaa gtaaaaagaa 3900
aagaaaaaaa aaaaaagaa aaaaaccacc tctacatatt atggaaagaa aatattttg 3960
tcgattotta tcttttata attatgctg gaagaagtag acacattaaa cgattccag 4020
tggaaacatg tcaactg 4037

```

```

<210> 41
<211> 706
<212> FRT
<213> Homo sapiens
<400> 41

```

```

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

```

WO 02/31134

PCT/US01/31874

Gln Pro Asp Ile Lys Ser Ser Ser Leu Arg Leu Thr Ser Ser Gly Lys
 115 120 125
 Glu Glu Ile Ile Phe Asn Gly Ile Ala Asp Trp Leu Tyr Glu Glu Glu
 130 135 140
 Leu Leu His Ser His Ile Ala His Trp Trp Ser Pro Asp Gly Glu Arg
 145 150 155 160
 Leu Ala Phe Leu Met Ile Asn Asp Ser Leu Val Pro Thr Met Val Ile
 165 170 175
 Pro Arg Phe Thr Gly Ala Leu Tyr Pro Lys Gly Lys Gln Tyr Pro Tyr
 180 185 190
 Pro Lys Ala Gly Gln Val Asn Pro Thr Ile Lys Leu Tyr Val Val Asn
 195 200 205
 Leu Tyr Gly Pro Thr His Thr Leu Glu Leu Met Pro Pro Asp Ser Phe
 210 215 220
 Lys Ser Arg Glu Tyr Tyr Ile Thr Met Val Lys Trp Val Ser Asn Thr
 225 230 235 240
 Lys Thr Val Val Arg Trp Leu Asn Arg Pro Gln Asn Ile Ser Ile Leu
 245 250 255
 Thr Val Cys Glu Thr Thr Thr Gly Ala Cys Ser Lys Lys Tyr Glu Met
 260 265 270
 Thr Ser Asp Thr Trp Leu Ser Gln Gln Asn Glu Glu Pro Val Phe Ser
 275 280 285
 Arg Asp Gly Ser Lys Phe Phe Met Thr Val Pro Val Lys Gln Gly Gly
 290 295 300
 Arg Gly Glu Phe His His Ile Ala Met Phe Leu Ile Gln Ser Lys Ser
 305 310 315 320
 Glu Gln Ile Thr Val Arg His Leu Thr Ser Gly Asn Trp Glu Val Ile
 325 330 335
 Lys Ile Leu Ala Tyr Asp Glu Thr Thr Gln Lys Ile Tyr Phe Leu Ser
 340 345 350
 Thr Glu Ser Ser Pro Arg Gly Arg Gln Leu Tyr Ser Ala Ser Thr Glu
 355 360 365
 Gly Leu Leu Asn Arg Gln Cys Ile Ser Cys Asn Phe Met Lys Glu Gln
 370 375 380
 Cys Thr Tyr Phe Asp Ala Ser Phe Ser Pro Met Asn Gln His Phe Leu
 385 390 395 400
 Leu Phe Cys Glu Gly Pro Arg Val Pro Val Val Ser Leu His Ser Thr
 405 410 415
 Asp Asn Pro Ala Lys Tyr Phe Ile Leu Glu Ser Asn Ser Met Leu Lys
 420 425 430
 Glu Ala Ile Leu Lys Lys Lys Ile Gly Lys Pro Glu Ile Lys Ile Leu
 435 440 445
 His Ile Asp Asp Tyr Glu Leu Pro Leu Gln Leu Ser Leu Pro Lys Asp
 450 455 460
 Phe Met Asp Arg Asn Gln Tyr Ala Leu Leu Ile Met Asp Glu Glu
 465 470 475 480
 Pro Gly Gly Gln Leu Val Thr Asp Lys Phe His Ile Asp Trp Asp Ser
 485 490 495
 Val Leu Ile Asp Met Asp Asn Val Ile Val Ala Arg Phe Asp Gly Arg
 500 505 510
 Gly Ser Gly Phe Gln Gly Leu Lys Ile Leu Gln Glu Ile His Arg Arg
 515 520 525
 Leu Gly Ser Val Glu Val Lys Asp Gln Ile Thr Ala Val Lys Phe Leu
 530 535 540
 Leu Lys Leu Pro Tyr Ile Asp Ser Lys Arg Leu Ser Ile Phe Gly Lys
 545 550 555 560
 Gly Tyr Gly Gly Tyr Ile Ala Ser Met Ile Leu Lys Ser Asp Glu Lys
 565 570 575

WO 02/31134

PCT/US01/31874

Leu Phe Lys Cys Gly Ser Val Val Ala Pro Ile Thr Asp Leu Lys Leu
580 585 590
Tyr Ala Ser Ala Phe Ser Glu Arg Tyr Leu Gly Met Pro Ser Lys Glu
595 600 605
Glu Ser Thr Tyr Gln Ala Ala Ser Val Leu His Asn Val His Gly Leu,
610 615 620
Lys Glu Glu Asn Ile Leu Ile Ile His Gly Thr Ala Asp Thr Lys Val
625 630 635 640
His Phe Gln His Ser Ala Glu Leu Ile Lys His Leu Ile Lys Ala Gly
645 650 655
Val Asn Tyr Thr Met Gln Val Tyr Pro Asp Glu Gly His Asn Val Ser
660 665 670
Glu Lys Ser Ser Lys Tyr His Leu Tyr Ser Thr Ile Leu Lys Phe Phe Ser
675 680 685
Asp Cys Leu Lys Glu Glu Ile Ser Val Leu Pro Gln Glu Pro Glu Glu
690 695 700
Asp Glu
705

<210> 42
<211> 4541
<212> DNA
<213> Homo sapiens
<400> 42

gkctykgkkg wtsmagatac agatgtggbg tataaaaagcg agaatggaca tgcattataa 60
ctgaatatag aaacaaatgc taccacatta ttattggaaa acacaacttt tgtaaccttc 120
aaagcatcaa gacattcagc ttcacacgat ttaaaatagc tcctctgggc atatgatgtc 180
aaacagattt ttcattattc gtatactgct tcataatgta tttacaacat acacactagg 240
gaagtttggg agttaaattc tccagaagta gaggactccg tcttgcagta cggggcctgg 300
gggtgtccaag ggcagcagct gatttatatt ttgaaaaata atatctacta tcaacctgat 360
ataaagagca gttcattgcg actgacatct tctggaaaag aagaataaat ttttaattgg 420
attgtgact ggttatatga agaggaactc ctgcattctc acatcgcccc ctgggtgtca 480
ccagatggag aaagacttgc cttcctgatg ataaatgact ctttggtaac caccatggtt 540
atcctcctgt ttaactggagc gttgtatccc aaaggaagc agtatccgta tctaaggcca 600
ggtoagctga acccaacaat aaaattatat gctgtaaac tglatggacc aactcacact 660
ttggagctca tgcacaactga cagcttataa tcaagagact actatatac tatggtaaa 720
tgggtaagca ataccaagac tgggttaaga tgggttaaac gaactcagaa catctccatc 780
ctcacagctc ctgagacacac tacagtgctc tgtatgtaaa aatgatgagat gacatcagat 840
acgtggctct ctgagcagaa tgaggagccc gtgttttcta gagacggcag caaattcttt 900
atgacagtcg ctgttaagca aggggagcgt ggagaatttc accacatagc tatgttcttc 960
atccagagta aaagttagca aattaccgtg cggcatctga catcagaaa ctgggaagtg 1020
ataaagatct tggcatakga tgaactact caaaaaattt actttctgag cactgaatct 1080
tctcccagag gaagcagct gtacagtgct totactgaag gattattgaa tcgccaatgc 1140
athtcatgta athtcatgaa agaacaatgt acatattttg atgccagtt tagtcccattg 1200
aatcaacatt tottattatt ctgtgaaggt ccaaggtccc cagtggtcag cctacatagt 1260
acggacaacc cagcaaaata ttttatattg gaaagcaatt ctatgctgaa ggaagctatc 1320
ctgaagaaga agataggaaa gccagaaatt aaaatccttc atattgacga ctatgaactt 1380
cctttacagt tgtoccttcc caaagathtt atggaccgaa accagatagc tcttctgtta 1440
ataatggatg aagaaccagg aggcacagtg gtlacagata agttccatag tgaactgggat 1500
tccgtactca tgcacatgga taatgtcatt gttagcaagat ttagtgccag aggaagtggg 1560
ttccagggtc tgaaaatttt gcaggagatt catcgagat taggttcagt agaagtaag 1620
gaccaataaa cagctgtgaa attttctgctg aaactgcctt acattgactc caaagatta 1680
agcatttttg gaaggggta tgggtgctat attgcatcaa tgatcttaaa atcagatgaa 1740
agctttcttg aaatgacact cgtgggtgca cctatcacag acttgaaatt gtagcctca 1800
gcttctctcg aaagatacct tgggatgcca totaaggaag aaagcactta ccaggcagcc 1860
agtgctctac ataattgtca tggcttgaaa gaagaaaata tatataaat tcatggaact 1920
gctgacacaa aagttcattt ccaactcoa gcagaattaa tcaagcactt aataaagct 1980
ggagtgaatt atactatgca ggtctaccca gatgaagctc ataacgtatc tgagaagagc 2040

WO 02/31134

PCT/US01/31874

```

aagtatcatc tctacagcac aatcctcaaa ttcttcagtg attgtttgaa ggaagaaata 2100
tctgggttac cacaggaacc agaagaagat gaataatgga ccgattttat acagaactga 2160
agggatattt gaggctcaat gaaacctgac aaagagactg taatattgta gttgctccag 2220
aatgtcaagg gcagcttacg gagatgtcac tggagcagca cgctcagaga cagtgaacta 2280
gcaattggaat acacaagtcg aagtctactg tgttcttagg ggtgcagaac cgtttctttt 2340
gtatggagaga ggcaaaaggg ttggtttcct ggggaaattt agttttgcat taagttagga 2400
gtagtgcatg ttttctctct ttatccccct gtttggctct taactagtgt ctctcatttt 2460
aatttcaotg gccaccatca tctttgcata taatgcacaa tctatcatct gtctcactgt 2520
ccttgatctt tcatggctga gctgcaatct aacactttac tgtaccttta taataagtgc 2580
aattctttca ttgtctatta tlatgcttaa gaaaatattc agttaataaa aaacagagta 2640
ttttatgtaa tttctgtttt taaaagaca ttattaaatg ggtcaaaagga catatagaaa 2700
tgtggatttc agcaccttcc aaagtccagc cagttatcag tagatacaat atctttaaata 2760
gaacaacaga gtgtatgtct cacaatatat atacacaagt gtgcataac agttaatgaa 2820
actatcttta aatgttattc atgctataaa gagtaaacgt ttgatgaatt agaagagatg 2880
ctcttttcca agctataabg gatgctttgt ttaatgacc aaatatgatg aaacattttt 2940
tccaattcaa attctagcta ttgctttcct ataatgtttt gggttgtgtt tggattgttt 3000
tttagtggtt aatagttttc cagttgcatt taattttttg aatagtagac ctgtcacat 3060
gtaaataga tacttaata ttaaatata gttctogata aagaattttt gtaacaatg 3120
caatgcoact gatgctatt ttgctctttt ggtggagaag gcttttttca aaactctgg 3180
tcttttact tttctctctc agtcgcaaat caattctcat ttctactgta aaagcaata 3240
agtggtattt ttcatttgcc agtttctatt tagtattcca tgcctgcca attcatctgt 3300
tactgtttaa ttcaattct tctggtgaga attagaaatg aaatattttt tattcatctg 3360
cctgtagtaa gtcaatttaa gtgattgaa catttaactt aattgaaagg acaaaatgca 3420
tcaattcctg tctgtgtgtg actgtagta gtaagtaact gagagcataa aataaactg 3480
actgtagtaa gtcaatttaa gtgattgaa catttaactt tggtagctaa agtcagaata 3540
tcttctcact tcaatttagg gatcttcag aagatactca aaagtctgta ataagcttag 3600
aagtcaaat aatcttaggc aggatactgc atttttgtgg ttttaaaaaa gtccttagga 3660
cagatgaat tctataact tatggcatca ggaggaaact ttaaaatato aaggaatcac 3720
cagtcaccc tctctgtttg ttgaaggatc aaccccaaat tctgggtatt tgagtagatg 3780
tgaatcagtg atttgttatt caacttttcc cctggatgct ttggaatcgt gtcttccatg 3840
ctccactggg tccaatttaa aataggagag gctttctctt ctgaaagatc cattttaggt 3900
ctttttcaag aatagtgaac acatttttta acaaaataag ttgtaatttt aaaaggaag 3960
ttttgctat tttattaaga tggaaatttc tttttaggtt aatttgaat ccaactgaag 4020
cttttaacc aatattttaa attgaaacca ctatgttttt tbatgtgca aatgattatg 4080
ttgtctgaaa ggtgtggttt tattgaaatg ctatttgagt atcattttaa aagattttgc 4140
cttttactgt catcatttct ctgtttttat tattattatc aatgtttatc tatttttcaa 4200
ttaattaat acagtttcta atgtgaaaga catttttctg gaaccctgtt tccctttaa 4260
cactaaagag acctcaagtg aaagcatatt gcttagtagg aaggtagaaa atgttaatcc 4320
ctgcatctct ttgagtttta atgacagggc cattttcagt aaaggaatg ctcaaccaaca 4380
catagtcacc aactatttaa ggaatcatgt gatggatttt tccccgtat acatgacc 4440
ttggtcataa tccactatt tcatcatat ttatgcatg ctgattttc ctaggactcc 4500
aatagcatgc tttccaagtg ttattatttc cttaatgta a 4541

```

```

<210> 43
<211> 691
<212> PRT
<213> Homo sapiens
<400> 43

```

```

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

```

WO 02/31134

PCT/US01/31874

Asn Pro Pro Glu Val Glu Asp Ser Val Leu Gln Tyr Ala Ala Trp Gly
 85 90 95
 Val Gln Gly Gln Gln Leu Ile Tyr Ile Phe Glu Asn Asn Ile Tyr Tyr
 100 105 110
 Gln Pro Asp Ile Lys Ser Ser Ser Leu Arg Leu Thr Ser Ser Gly Lys
 115 120 125
 Glu Glu Ile Ile Phe Asn Gly Ile Ala Asp Trp Leu Tyr Glu Glu Glu
 130 135 140
 Leu Leu His Ser His Ile Ala His Trp Trp Ser Pro Asp Gly Glu Arg
 145 150 155 160
 Leu Ala Phe Leu Met Ile Asn Asp Ser Leu Val Pro Thr Met Val Ile
 165 170 175
 Pro Arg Phe Thr Gly Ala Leu Tyr Pro Lys Gly Lys Gln Tyr Pro Tyr
 180 185 190
 Pro Lys Ala Gly Gln Val Asn Pro Thr Ile Lys Leu Tyr Val Val Asn
 195 200 205
 Leu Tyr Gly Pro Thr His Thr Leu Glu Leu Met Pro Pro Asp Ser Phe
 210 215 220
 Lys Ser Arg Glu Tyr Tyr Ile Thr Met Val Lys Trp Val Ser Asn Thr
 225 230 235 240
 Lys Thr Val Val Arg Trp Leu Asn Arg Pro Gln Asn Ile Ser Ile Leu
 245 250 255
 Thr Val Cys Glu Thr Thr Thr Gly Ala Cys Ser Lys Lys Tyr Glu Met
 260 265 270
 Thr Ser Asp Thr Trp Leu Ser Gln Gln Asn Glu Glu Pro Val Phe Ser
 275 280 285
 Arg Asp Gly Ser Lys Phe Phe Met Thr Val Pro Val Lys Gln Gly Gly
 290 295 300
 Arg Gly Glu Phe His His Ile Ala Met Phe Leu Ile Gln Ser Lys Ser
 305 310 315 320
 Glu Gln Ile Thr Val Arg His Leu Thr Ser Gly Asn Trp Glu Val Ile
 325 330 335
 Lys Ile Leu Ala Tyr Asp Glu Thr Thr Gln Lys Ile Ser Ala Ser Thr
 340 345 350
 Glu Gly Leu Leu Asn Arg Gln Cys Ile Ser Cys Asn Phe Met Lys Glu
 355 360 365
 Gln Cys Thr Tyr Phe Asp Ala Ser Phe Ser Pro Met Asn Gln His Phe
 370 375 380
 Leu Leu Phe Cys Glu Gly Pro Arg Val Pro Val Val Ser Leu His Ser
 385 390 395 400
 Thr Asp Asn Pro Ala Lys Tyr Phe Ile Leu Glu Ser Asn Ser Met Leu
 405 410 415
 Lys Glu Ala Ile Leu Lys Lys Lys Ile Gly Lys Pro Glu Ile Lys Ile
 420 425 430
 Leu His Ile Asp Asp Tyr Glu Leu Pro Leu Gln Leu Ser Leu Pro Lys
 435 440 445
 Asp Phe Met Asp Arg Asn Gln Tyr Ala Leu Leu Ile Met Asp Glu
 450 455 460
 Glu Pro Gly Gly Gln Leu Val Thr Asp Lys Phe His Ile Asp Trp Asp
 465 470 475 480
 Ser Val Leu Ile Asp Met Asp Asn Val Ile Val Ala Arg Phe Asp Gly
 485 490 495
 Arg Gly Ser Gly Phe Gln Gly Leu Lys Ile Leu Gln Glu Ile His Arg
 500 505 510
 Arg Leu Gly Ser Val Glu Val Lys Asp Gln Ile Thr Ala Val Lys Phe
 515 520 525
 Leu Leu Lys Leu Pro Tyr Ile Asp Ser Lys Arg Leu Ser Ile Phe Gly
 530 535 540

WO 02/31134

PCT/US01/31874

Lys Gly Tyr Gly Gly Tyr Ile Ala Ser Met Ile Leu Lys Ser Asp Glu
545 550 555 560
Lys Leu Phe Lys Cys Gly Ser Val Val Ala Pro Ile Thr Asp Leu Lys
565 570 575
Leu Tyr Ala Ser Ala Phe Ser Glu Arg Tyr Leu Gly Met Pro Ser Lys
580 585 590
Glu Glu Ser Thr Tyr Gln Ala Ala Ser Val Leu His Asn Val His Gly
595 600 605
Leu Lys Glu Glu Asn Ile Leu Ile Ile His Gly Thr Ala Asp Thr Lys
610 615 620
Val His Phe Gln His Ser Ala Glu Leu Ile Lys His Leu Ile Lys Ala
625 630 635 640
Gly Val Asn Tyr Thr Met Gln Val Tyr Pro Asp Glu Gly His Asn Val
645 650 655
Ser Glu Lys Ser Lys Tyr His Leu Tyr Ser Thr Ile Leu Lys Phe Phe
660 665 670
Ser Asp Cys Leu Lys Glu Glu Ile Ser Val Leu Pro Gln Glu Pro Glu
675 680 685
Glu Asp Glu
690

<210> 44
<211> 4496
<212> DNA
<213> Homo sapiens
<400> 44

gkctykgtkg wtsmagatag agatgtggtg tataaaagcg agaatggaca tgtcattaaa 60
ctgaaatag aaccaaagc taccacatta ttattggaaa acacaacttt tgtaaccttc 120
aaagcatcaa gacattcagt ttcaccagat ttaaaatag tcctctggc atatgatgtc 180
aaacagattt ttcattatc gtatactgct tcatatgtga tttacaacat acacactagg 240
gaagtttggg agttaaattc tccagaagta gaggactccg tcttgcagta cggggcctgg 300
gggtccaag ggcagcagct gatttatatt ttgaaaaa atactacta tcaacctgat 360
ataaagaca gtccattgcy actgacatct tctgaaaag aagaataat ttttaatggg 420
atgtcagact ggttatatga agaggaactc ctgcattctc acatggcca ctgggtgtca 480
ccagatggag aaagacttgc ctctcctgat aaaaatgact ctttggacc caccatggtt 540
atccctcggg ttactggagc gttgtatccc aaaggaagc agtatccga tccaaaggca 600
ggtcaagtga acccaacaat aaaaatataat gttgtaaac tgtatggacc aactcaactc 660
ttggagctca tggccactga cagcttataa tcaagagact actatctac batggttaa 720
tggtaagca ataccagac tgtggttaaga tgggttaaac gactcagaa catctcaatc 780
ctcacagtct gtgagaccac tacaggtgct tgtagttaaa aatagagat gacatcagat 840
acgtgctctc ctacagcagaa tgaaggagcc gtttttcta gagacggcag caaattcttt 900
atgacagtgc ctgttaagca agggggagct ggagaatttc accacatagc tatgttcttc 960
atccagagta aaagttagca aatfaccgtg cggcatctga catcagaaa ctgggaagtg 1020
ataaagatct tggcatcaga tgaactactc caaaaatca gtgctctac tgaaggatta 1080
ttgaatcgcc aatgcatttc atgtaatttc atgaaagAAC aatgtacata ttttgatgcc 1140
agtttttagtc ccatgaatca acatttctta tlatctgtg aaggtccaag ggtccagtg 1200
gtcagcttac atagtacgga caaccagca aatatttta tattgaaag caattctatg 1260
ctgaaggaaag ctactctgaa gaagaagata gaaagccag aatataaat ccttcatatt 1320
gacgactatg aacttcttt acagttgtcc ctcccaagc attttatgga cggaaaccag 1380
catattgact gggatccgt actcattgac atggataatg tcatgtagc aagatttgat 1440
ggcagagaaa gttgattcca gggctcgaat atttgcagc agatcactc aagattaggt 1500
tcagtgaag taaaggacca aataacagct gtgaaatttt tgcctgaaact gcctacatt 1560
gactccaaaa gattaagcat ttttgaagc ggttatgtg gctatattgc atcaatgatc 1620
ttaaatcag atgaaaagct ttttaaatgt ggtccctggg ttgcacctat cacagacttg 1680
aaattgtatg cctcagcttt ctctgaaaga tacctgggga tgcactataa ggaagaaagc 1740
acttaccagg cagccagtg gctacataat gttcattggc tgaagaaga aatatatta 1800
ataattcatg gaactgctga cacaaaagt ctttccaac actcagcaga ataatcaag 1860

WO 02/31134

PCT/US01/31874

```

cacctaataa aagctggagt gaattatact atgcaggtct acccagatga aggtcataac 1980
gtatctgaga agagcaagta tcaatctctac agcacaatcc tcaaatctct cagtgattgt 2040
ttgaaggaag aaatatctgt gotaccaoaq gaaccagaag aagatgaata atggaccgta 2100
tttatacaga actgaagggga atattgaggc tcaatgaaac ctgacaaaga gactgtaata 2160
ttgtagttgc tccagaatgt caagggcagc ttacggagat gtcactggag cagcaagctc 2220
agagacagtg aactagcatt tgaatacacaa agtccaagtc tactgtgttg ctagggtg 2280
agaacccctt tctttgtatg agagaggtca aagggttggg ttctggggag aaattagttc 2340
tgcattaaaq taggagtagt gcatgttttc ttctgttalc cccctgtttg ttctgtaact 2400
agttgctctc attttaattt cactggccac catcatcttt gaataaatg cacaactcat 2460
catctgtcct acagccctgt atcttctatg gctgagctgc aatctaacac ttactgtac 2520
ctttataaba agtgcaattc ttcaattgtc tattattatg cttaagaaaa tattcagtta 2580
ataaaaaaca gagtatttta tghtaattct gtttttaaaa agacattatt aaatgggtca 2640
aaggaacatq agaaatgtgg atttcagcac ctccaagat tcagccagtt atcagtagat 2700
acaatctctt taaatgaaca caccagtgta ttctcaca taatataca caagtggtca 2760
tatcacgtha atgaaactat cttaaatgt tattcatgct ataaagagta aacgtttgat 2820
gaattagaag agatgctctt ttccaagcta taatggatct ttgtttaat gagccaata 2880
tqatgaaaca ttttttccaa ttcaaatct agctattgct ttctataaaa tgtttgggtt 2940
gtgtttggta ttgtttttag tggtaaatag ttttccagtt gcatttaatt ttttgaatat 3000
gataccctgt cacatgtaaa ttgatactt aaatatata ttatagtttc tgataaagaa 3060
attttgttaa caatgcaatg cactgagtg ctattttgt cttttggtg agaggcttt 3120
tttcaaacct ctgtgtcctt ttacttcttt ctctcagtc agaactcaat ctcattttca 3180
tqtaaaaagc aaatagctgg attattctat ttgccagtt ctatttagta ttccatgct 3240
gcccaattca ttgttactg ttaatttca atctctctg tgagaattag aaatgaata 3300
ttttttattc atggccaaa aagtctcacg acagcagtg ttgtctatta ctttgaatg 3360
aagccaaaa atgcatcaat tctgtgctg tgttgactg cagtgttag taactgagag 3420
cataaaaaa acctgactgt atgaagtcaa ttaagtatg gagaacatt aactttggtg 3480
actaaagtc a gaatctctt tcacttcaat taaggatct tccagaagat atctaaaag 3540
ctgaataag cttagaagtt cagataaatc taggcagat actgctatt tgtggttta 3600
aaaaagctct taggacagac tgaattatca taactatgg catcaggagg aaactttaa 3660
atataagga atcactcagt caccctctg ttttttgaa ggaacaccc caaattctgg 3720
gtatttgagt acatgtgaat catggattg gtattcaact tttccctgg atgotttgg 3780
atcgtctctt ccatgctcca ctggttcaa tttaaaatg gagaggcttt ctctctgaa 3840
agatccattt taggtctttt tcaagaatag tgaacacatt ttttaaaaa ataagttgta 3900
attttaaaag taaagctttt cctattttat taagatgaa atttctttt aggtcaattt 3960
gaaatccaac tgaagctttt taaccaatat tttaaattg aaccactaga gttttttatg 4020
atgcaaatga ttatgtttgc tgaaggtgt gttttttatg aatgtctatt tgagatcat 4080
ttaaaaagca tttgctttt actgtcatca ttctctgtg ttattatta ttatcaatgt 4140
ttatctattt ttcaattaat ttaatacagt ttctaatgt aaagacatt ttctggaacc 4200
cgttttcccc ttaaacacta aagagacctc aagtgaagc atattgetta gtaggagggt 4260
agaaaatggt aatccctgag atcttttgag ttttaatgac agggctatt tcagtaaaag 4320
aaatgctcac caacacatag tcaaccaata ttaaggaat catgtgattg gattttcccc 4380
tgtatacagb taccttgggt cataatccca ctatttcaata catatttatg cattgctaga 4440
ttttctcagg actccaatag catgctttcc aagtgttatt attccccaa tgttaa 4496

```

```

<210> 45
<211> 29
<212> DNA
<213> Homo sapiens
<400> 45

```

```

cggtagcatg gcagcagcaa tggaaacag 29

```

```

<210> 46
<211> 39
<212> DNA
<213> Homo sapiens
<400> 46

```

```

ggagctcgg cccgtccta tcaattttag agcagcaat 39

```

WO 02/31134

PCT/US01/31874

```

<210> 47
<211> 27
<212> DNA
<213> Homo sapiens
<400> 47
caagctttat cacttttaga gcagcaa                27

<210> 48
<211> 22
<212> DNA
<213> Homo sapiens
<400> 48
cacattcttg ctgcatcagt ca                      22

<210> 49
<211> 22
<212> DNA
<213> Homo sapiens
<400> 49
ttgggtcaatc ttcaggactt ga                   22

<210> 50
<211> 27
<212> DNA
<213> Homo sapiens
<400> 50
caagcttacc atggccacca cggggac              27

<210> 51
<211> 37
<212> DNA
<213> Homo sapiens
<400> 51
cggatccgcg gccgctcaga ggtattcctg tagaaag   37

<210> 52
<211> 27
<212> DNA
<213> Homo sapiens
<400> 52
cggatccagg tattcctgta gaaagtg              27

<210> 53
<211> 20
<212> DNA
<213> Homo sapiens
<400> 53
tacgccgtgg ttgtgattga                       20

<210> 54
<211> 20
<212> DNA

```

WO 02/31134

PCT/US01/31874

```

<213> Homo sapiens
<400> 54
ccatacttct cggccacgaa 20

<210> 55
<211> 19
<212> DNA
<213> Homo sapiens
<400> 55
gcctgggatt gtgcactgt 19

<210> 56
<211> 29
<212> DNA
<213> Homo sapiens
<400> 56
gtgtattcaa atgctagttc actgtctct 29

<210> 57
<211> 22
<212> DNA
<213> Homo sapiens
<400> 57
agctagcact gtccagggtc ct 22

<210> 58
<211> 25
<212> DNA
<213> Homo sapiens
<400> 58
agggcccttc atcttcttct gggtc 25

<210> 59
<211> 19
<212> PRT
<213> Homo sapiens
<400> 59
Val Glu Asp Asp Val Met Glu Arg Gln Arg Leu Ile Glu Ser Val Pro
1 5 10 15
Asp Ser Val

<210> 60
<211> 19
<212> PRT
<213> Homo sapiens
<400> 60
Ser Thr Glu Asn Glu Glu Gln Arg Leu Ala Ser Ala Arg Ala Val Pro
1 5 10 15
Arg Asn Val

<210> 61
<211> 15

```

WO 02/31134

PCT/US01/31874

<212> PRT
<213> Homo sapiens
<400> 61

Lys Glu Ala Ile Leu Lys Lys Lys Ile Gly Lys Pro Glu Ile Lys
1 5 10 15

【国際公開パンフレット(コレクション)】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
18 April 2002 (18.04.2002)

PCT

(10) International Publication Number
WO 02/031134 A3

- (51) International Patent Classification: C12N 9/48
- (21) International Application Number: PCT/US01/31874
- (22) International Filing Date: 12 October 2001 (12.10.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
50/240,117 12 October 2000 (12.10.2000) US
- (71) Applicant (for all designated States except US): **FERRING BV** [US/US]; Polaris Avenue 144, NL-2132 JX Hoofddorp (NL).
- (72) Inventors; and
(75) Inventors/Applicants (for US only): **QI, Steve** [GB/US]; 11034 West Ocean Air Drive, Apt. 24B, San Diego, CA 92130 (US). **AKINSANYA, Karen, O.** [GB/US]; 7675 Palmilla Drive, Apt. 642B, San Diego, CA 92122-5094 (US). **RIVIERE, Pierre, J.-M.** [FR/US]; 3993 Via Cangejo, San Diego, CA 92130 (US). **JUNEN, Jean-Louis** [FR/FR]; Ferring SAS, 1-75507 Paris (FR).
- (74) Agents: **SAMPLES, Kenneth, H.** et al.; Fitch, Even, Tabin & Flannery, 120 South LaSalle Street, Suite 1600, Chicago, IL 60603 (US).
- (81) Designated States (national): AF, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GI, GM, HR, HU, ID, IL, IN, IS, JP, KL, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPPO patent (GI, GM, KE, LS, MW, MZ, 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, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— with international search report
- (88) Date of publication of the international search report:
17 July 2002
- 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 02/031134 A3

(54) Title: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

(57) Abstract: Novel proteins or polypeptides having significant sequence homology to DPPIV, nucleic acids coding therefor, cells which have been modified with such nucleic acid so as to express these proteins, antibodies to these proteins, screening methods for the discovery of new therapeutic agents which are inhibitors of the activity of these proteins or of related proteins, and therapeutic agents discovered by such screening methods, as well as new therapeutic treatments, are all provided.

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/31874
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N9/48		
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		
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) EPO-Internal, EMBL, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ABBOTT CATHERINE A ET AL: "Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8" EUROPEAN JOURNAL OF BIOCHEMISTRY, BERLIN, DE, vol. 267, no. 20, October 2000 (2000-10), pages 6140-6150, XP002204858 ISSN: 0014-2956 figure 1	1-25
X	DATABASE EMBL 'Online! 22 February 2000 (2000-02-22) retrieved from EBI Database accession no. AK000290 XP002223531 abstract	1-3,9-11
--- -/--		
<input checked="" type="checkbox"/>	Further documents are listed in the continuation of box C.	<input checked="" type="checkbox"/> 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 doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)		
O document referring to an oral disclosure, use, exhibition or other means		
P document published prior to the international filing date but later than the priority date claimed		
T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principles 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, each 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	
5 March 2003	26.03.03	
Name and mailing address of the ISA European Patent Office, P.O. 5818 Palantian 2 NL - 2280 HV The Hague Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Grosskopf, R	

INTERNATIONAL SEARCH REPORT		International Application No PCT/US 01/31874
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 'Online! 12 March 1999 (1999-03-12) retrieved from EBI Database accession no. AL040398 XP002223532 abstract	1-3,9-11
P, X	WO 01 19866 A (ABBOTT CATHERINE ANNE ;GORELL MARK DOUGLAS (AU); UNIV SYDNEY (AU)) 22 March 2001 (2001-03-22) see claims and Figure 2	1-25
E	WO 01 98468 A (ELLIOTT WICKI S;INCYTE GENOMICS INC ; YUE HENRY (US)) 27 December 2001 (2001-12-27) see claims and SEQ ID NO: 9	1-25
X	WO 00 58473 A (CURAGEN CORP ;LEACH MARTIN (US); SHINKETS RICHARD A (US)) 5 October 2000 (2000-10-05) SEQ ID NO: 2780 has 97.9% identity with SEQ ID NO: 3 in the range from aa 1 to aa 643	1,2,4, 7-12, 17-20
X	WO 00 42201 A (INCYTE PHARMA INC ;AZIMZAI YALDA (US); YUE HENRY (US); BANDMAN OLG) 20 July 2000 (2000-07-20) SEQ ID NO: 16 has 99.3% identity with SEQ ID NO: 3 in the range from aa 1 to aa 480	1,2,4, 8-12, 17-20
X	NAGASE T ET AL: "PREDICTION OF THE CODING SEQUENCES OF UNIDENTIFIED HUMAN GENES. XVII. THE COMPLETE SEQUENCES OF 100 NEW CDNA CLONS FROM BRAIN WHICH CODE FOR LARGE PROTEINS IN VITRO" DNA RESEARCH, UNIVERSAL ACADEMY PRESS, JP, vol. 7, no. 2, 28 April 2000 (2000-04-28), pages 143-150. XP008010899 ISSN: 1340-2838 see the whole document	1,2,4, 7-12, 17-20
X	-& DATABASE SMALLPROT 'Online! 1 October 2000 (2000-10-01) retrieved from EBI Database accession no. Q9P236 XP002233481 The sequence has 100% identity with SEQ ID NO:5 in the range from aa 87 to 796 abstract	1,2,4, 7-12, 17-20
	--- -/--	

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/31874
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	YOKOTANI N ET AL: "NON-CONSERVATION OF A CATALYTIC RESIDUE IN A DIPEPTIDYL AMINOPEPTIDASE IV-RELATED PROTEIN ENCODED BY A GENE ON HUMAN CHROMOSOME 7" HUMAN MOLECULAR GENETICS, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 2, no. 7, July 1993 (1993-07), pages 1037-1039, XP001026522 ISSN: 0964-6906	
A	-& DATABASE SWALLPROT 'Online! 1 November 1995 (1995-11-01) retrieved from EBI Database accession no. P42658 XP002233482 The sequence has 52.8% identity with SEQ ID NO: 5 in the region from aa 22 to aa 796 abstract	
P;X	WO 01 55399 A (MAO YUMIN ;XIE YI (CN); BIODOR GENE TECHNOLOGY LTD SH (CN)) 2 August 2001 (2001-08-02) SEQ ID NO: 2 has 100 % identity with SEQ ID NO: 5 in the region from aa 570 to aa 796	1,2,4, 8-12

INTERNATIONAL SEARCH REPORT	International Application No. PCT/US 01/31874
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:	
<p>1. <input type="checkbox"/> Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:</p> <p>2. <input checked="" type="checkbox"/> Claims Nos.: 21 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</p> <p>3. <input type="checkbox"/> Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).</p>	
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)	
This International Searching Authority found multiple inventions in this International application, as follows:	
see additional sheet	
<p>1. <input checked="" type="checkbox"/> As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.</p> <p>2. <input type="checkbox"/> As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.</p> <p>3. <input type="checkbox"/> 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.:</p> <p>4. <input type="checkbox"/> 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.:</p>	
<p>Remark on Protest</p> <p><input type="checkbox"/> The additional search fees were accompanied by the applicant's protest.</p> <p><input checked="" type="checkbox"/> No protest accompanied the payment of additional search fees.</p>	

International Application No. PCT/US 01 81874

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 21

Claim 21 refers to an agonist/antagonist of the polypeptide without giving a true technical characterization. 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. 83 and 84 EPC). No search can be carried out for 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 01 81874

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-20, 22-25 (all partially)

Claims relating to the polypeptide having the SEQ ID NO: 1 and amino acid sequences having at least 70% similarity thereto and the DNA encoding the protein having SEQ ID NO: 2, or encoding proteins having at least 90% similarity with SEQ ID NO: 1, splice variants of said protein, probes for the DNA, antisense oligonucleotides directed against the DNA, host cells comprising the DNA antibodies directed against the protein and methods for screening for a compound capable of inhibiting the enzymatic activity of the protein.

2. Claims: 1-20, 22-25 (all partially)

Claims relating to the polypeptide having the SEQ ID NO: 3 and amino acid sequences having at least 70% similarity thereto and the DNA encoding the protein having SEQ ID NO: 4, or encoding proteins having at least 90% similarity with SEQ ID NO: 1, splice variants of said protein, probes for the DNA, antisense oligonucleotides directed against the DNA, host cells comprising the DNA antibodies directed against the protein and methods for screening for a compound capable of inhibiting the enzymatic activity of the protein.

3. Claims: 1-20, 22-25 (all partially)

Claims relating to the polypeptide having the SEQ ID NO: 5 and amino acid sequences having at least 70% similarity thereto and the DNA encoding the protein having SEQ ID NO: 6, or encoding proteins having at least 90% similarity with SEQ ID NO: 1, splice variants of said protein, probes for the DNA, antisense oligonucleotides directed against the DNA, host cells comprising the DNA antibodies directed against the protein and methods for screening for a compound capable of inhibiting the enzymatic activity of the protein.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/31874

* Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0119866	A	22-03-2001	WO 0119866 A1 22-03-2001
			AU 7394600 A 17-04-2001
			EP 1214344 A1 19-06-2002
WO 0198468	A	27-12-2001	AU 6983601 A 02-01-2002
			WO 0198468 A2 27-12-2001
WO 0058473	A	05-10-2000	AU 3774500 A 16-10-2000
			EP 1165784 A2 02-01-2002
			WO 0058473 A2 05-10-2000
WO 0042201	A	20-07-2000	AU 2501200 A 01-08-2000
			CA 2360464 A1 20-07-2000
			EP 1151113 A2 07-11-2001
			WO 0042201 A2 20-07-2000
WO 0155399	A	02-08-2001	CN 1307128 A 08-08-2001
			AU 2997301 A 07-08-2001
			WO 0155399 A1 02-08-2001

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 9/12	A 6 1 P 9/12	4 C 0 8 4
A 6 1 P 17/06	A 6 1 P 17/06	4 H 0 4 5
A 6 1 P 21/04	A 6 1 P 21/04	
A 6 1 P 25/08	A 6 1 P 25/08	
A 6 1 P 25/14	A 6 1 P 25/14	
A 6 1 P 31/12	A 6 1 P 31/12	
A 6 1 P 35/02	A 6 1 P 35/02	
A 6 1 P 43/00	A 6 1 P 43/00	1 0 1
C 0 7 K 16/40	A 6 1 P 43/00	1 1 1
C 1 2 N 1/21	C 0 7 K 16/40	
C 1 2 N 5/10	C 1 2 N 1/21	
C 1 2 N 9/64	C 1 2 N 9/64	
C 1 2 Q 1/37	C 1 2 Q 1/37	
G 0 1 N 33/15	G 0 1 N 33/15	Z
G 0 1 N 33/50	G 0 1 N 33/50	Z
G 0 1 N 33/53	G 0 1 N 33/53	D
G 0 1 N 33/566	G 0 1 N 33/53	M
	G 0 1 N 33/566	
	C 1 2 N 5/00	B

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

(72) 発明者 カレン オー・アキンサンヤ

アメリカ合衆国 9 2 1 2 2 - 5 0 9 4 カリフォルニア州 サン ディエゴ パルミラ ドライブ 7 6 7 5 アpartment 6 4 2 ビー

(72) 発明者 ピエール ジェイ・エム・リビエール

アメリカ合衆国 9 2 1 3 0 カリフォルニア州 サン ディエゴ ヴィア キャングレジョー 3 9 9 3

(72) 発明者 ジャン・ルイ ユニアン

フランス エフ - 7 5 5 0 7 パリ フェリング エスエイエス (番地なし)

F ターム(参考) 2G045 AA34 AA35 AA40 BA11 BB50 DA12 DA13 DA14 DA36 FB02
 4B024 AA01 AA11 BA14 BA41 CA01 CA09 CA11 DA02 DA05 EA04
 GA11 HA12
 4B050 CC01 CC04 DD11 EE10 LL01 LL03
 4B063 QA01 QA18 QQ36 QQ95 QR41 QR57 QR67
 4B065 AA26X AA50X AA90X AA91X AA93Y AB02 BA02 BA08 CA25 CA33
 CA44 CA46
 4C084 AA17 NA14 ZA221 ZA291 ZA421 ZA451 ZA541 ZA891 ZA941 ZB211
 ZB212 ZB271 ZB331 ZC201 ZC211
 4H045 AA10 AA20 AA30 BA10 CA40 DA75 DA89 EA20 EA23 EA29
 EA50 FA72 FA73 FA74

专利名称(译)	与DPPIV相关的新型丝氨酸蛋白酶基因		
公开(公告)号	JP2004528812A	公开(公告)日	2004-09-24
申请号	JP2002534503	申请日	2001-10-12
[标]申请(专利权)人(译)	凡林有限公司		
申请(专利权)人(译)	辉凌奋Besuroten笔记本闭嘴		
[标]发明人	ステーブキー カレンオーアキンサンヤ ピエールジェイエムリビエール ジャンルイユニアン		
发明人	ステーブ キー カレン オー.アキンサンヤ ピエール ジェイ.-エム.リビエール ジャン-ルイ ユニアン		
IPC分类号	G01N33/50 A61K45/00 A61P3/14 A61P7/02 A61P9/10 A61P9/12 A61P17/06 A61P21/04 A61P25/08 A61P25/14 A61P31/12 A61P35/02 A61P43/00 C07K16/40 C12N1/21 C12N5/10 C12N9/48 C12N9/64 C12N15/09 C12Q1/37 G01N33/15 G01N33/53 G01N33/566		
CPC分类号	A61P3/14 A61P17/06 A61P21/04 A61P25/08 A61P25/14 A61P31/12 A61P35/02 C12N9/48 C12N2799 /026		
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P3/14 A61P7/02 A61P9/10 A61P9/12 A61P17/06 A61P21/04 A61P25/08 A61P25/14 A61P31/12 A61P35/02 A61P43/00.101 A61P43/00.111 C07K16/40 C12N1/21 C12N9/64 C12Q1/37 G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00. B		
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/AA40 2G045/BA11 2G045/BB50 2G045/DA12 2G045/DA13 2G045 /DA14 2G045/DA36 2G045/FB02 4B024/AA01 4B024/AA11 4B024/BA14 4B024/BA41 4B024/CA01 4B024/CA09 4B024/CA11 4B024/DA02 4B024/DA05 4B024/EA04 4B024/GA11 4B024/HA12 4B050 /CC01 4B050/CC04 4B050/DD11 4B050/EE10 4B050/LL01 4B050/LL03 4B063/QA01 4B063/QA18 4B063/QQ36 4B063/QQ95 4B063/QR41 4B063/QR57 4B063/QR67 4B065/AA26X 4B065/AA50X 4B065/AA90X 4B065/AA91X 4B065/AA93Y 4B065/AB02 4B065/BA02 4B065/BA08 4B065/CA25 4B065 /CA33 4B065/CA44 4B065/CA46 4C084/AA17 4C084/NA14 4C084/ZA221 4C084/ZA291 4C084 /ZA421 4C084/ZA451 4C084/ZA541 4C084/ZA891 4C084/ZA941 4C084/ZB211 4C084/ZB212 4C084 /ZB271 4C084/ZB331 4C084/ZC201 4C084/ZC211 4H045/AA10 4H045/AA20 4H045/AA30 4H045 /BA10 4H045/CA40 4H045/DA75 4H045/DA89 4H045/EA20 4H045/EA23 4H045/EA29 4H045/EA50 4H045/FA72 4H045/FA73 4H045/FA74		
代理人(译)	谷义 安倍晋三和夫		
优先权	60/240117 2000-10-12 US		
其他公开文献	JP2004528812A5		
外部链接	Espacenet		

摘要(译)

与DPPIV具有显著序列同源性的新型蛋白质或多肽，为其编码的核酸，已经用这种核酸修饰以表达这些蛋白质的细胞，这些蛋白质的抗体，用于发现新的治疗剂的筛选方法 提供了这些蛋白质或相关蛋白质的活性，以及通过这种筛选方法发现的治疗剂，以及新的治疗方法。

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/00	C 1 2 N 15/00 Z N A A	2 G 0 4 5
A 6 1 K 45/00	A 6 1 K 45/00	4 B 0 2 4
A 6 1 P 3/14	A 6 1 P 3/14	4 B 0 5 0
A 6 1 P 7/02	A 6 1 P 7/02	4 B 0 6 3
A 6 1 P 9/10	A 6 1 P 9/10	4 B 0 6 5
審査請求 未請求 予備審査請求 有 (全 157 頁)		最終頁に続く

(21) 出願番号	特願2002-534503 (P2002-534503)	(71) 出願人	500297535
(86) (22) 出願日	平成13年10月12日 (2001.10.12)		フレリッング ベスローテン フェンノート
(85) 翻訳文提出日	平成15年4月14日 (2003.4.14)		シャップ
(86) 国際出願番号	PCT/US2001/031874		オランダ エヌエル-2 1 3 2 ジェイエ
(87) 国際公開番号	W02002/031134		ックス ホーフドルフ ポラリス アベニ
(87) 国際公開日	平成14年4月18日 (2002.4.18)		ュー 1 4 4
(31) 優先権主張番号	60/240,117	(74) 代理人	100077481
(32) 優先日	平成12年10月12日 (2000.10.12)		弁理士 谷 義一
(33) 優先権主張国	米国 (US)	(74) 代理人	100088915
			弁理士 阿部 和夫
		(72) 発明者	スティーブ キー
			アメリカ合衆国 9 2 1 3 0 カリフォル
			ニア州 サン ディエゴ ウェスト オー
			シャン エア ドライブ 1 1 0 3 4 ア
			パートメント 2 4 ビー
			最終頁に続く