

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2005-500012

(P2005-500012A)

(43) 公表日 平成17年1月6日(2005.1.6)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
<b>C 1 2 N 15/09</b>	C 1 2 N 15/00 Z N A A	2 G O 4 5
<b>A O 1 K 67/027</b>	A O 1 K 67/027	4 B O 2 4
<b>A 6 1 K 31/7088</b>	A 6 1 K 31/7088	4 B O 6 3
<b>A 6 1 K 35/76</b>	A 6 1 K 35/76	4 B O 6 5
<b>A 6 1 K 38/00</b>	A 6 1 K 39/00 H	4 C O 8 4
審査請求 未請求 予備審査請求 未請求 (全 193 頁) 最終頁に続く		

(21) 出願番号	特願2002-569878 (P2002-569878)	(71) 出願人	502334043
(86) (22) 出願日	平成14年3月5日 (2002.3.5)		インファーマティカ リミテッド
(85) 翻訳文提出日	平成15年9月5日 (2003.9.5)		イギリス ロンドン ダブリュー1ティー
(86) 国際出願番号	PCT/GB2002/000948		2 エヌユー チャーロット ストリート
(87) 国際公開番号	W02002/070559		6 0
(87) 国際公開日	平成14年9月12日 (2002.9.12)	(74) 代理人	100059959
(31) 優先権主張番号	0105402.2		弁理士 中村 稔
(32) 優先日	平成13年3月5日 (2001.3.5)	(74) 代理人	100067013
(33) 優先権主張国	英国 (GB)		弁理士 大塚 文昭
		(74) 代理人	100082005
			弁理士 熊倉 禎男
		(74) 代理人	100065189
			弁理士 宍戸 嘉一
		(74) 代理人	100074228
			弁理士 今城 俊夫
最終頁に続く			

(54) 【発明の名称】 核内ホルモンレセプターリガンド結合ドメイン

## (57) 【要約】

本発明は、新規なタンパク質 (LBDG1およびLBDG4と称され、本明細書では核内ホルモンレセプターリガンド結合ドメインと同定された)、並びに疾患の診断、予防および治療における前記タンパク質およびコード遺伝子由来の核酸配列の使用に関する。

**【特許請求の範囲】****【請求項1】**

以下のポリペプチド：

(i) 配列番号：2に記載のアミノ酸配列を含むかまたは前記アミノ酸配列からなるポリペプチド；

(ii) 核内ホルモンレセプターリガンド結合ドメインの活性を有するか、または(i)のポリペプチドと共通の抗原決定基を有する、前記のフラグメントであるポリペプチド；または

(iii) (i)または(ii)の機能的等価物であるポリペプチド。

**【請求項2】**

LBDG1ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域を含む、請求項1の(ii)に記載のフラグメントであるポリペプチドであって、前記核内ホルモンレセプターリガンド結合ドメイン領域が配列番号：2に記載のアミノ酸配列の残基822から1020を含むと定義され、前記フラグメントが“LBDモチーフ”残基LEU878、ASP885、GLN886、Leu889およびLEU890、または等価な残基を保有し、さらに核内ホルモンレセプターリガンド結合ドメインの活性を保有する前記ポリペプチド。

10

**【請求項3】**

請求項1の(iii)に記載の機能的等価物であり、配列番号：2に記載のアミノ酸配列と相同であり、触媒性残基LEU878、ASP885、GLN886、Leu889およびLEU890、または等価な残基を保有し、さらに核内ホルモンレセプターリガンド結合ドメインの活性を保有するポリペプチド。

20

**【請求項4】**

以下のポリペプチド：

(i) 配列番号：4に記載されたアミノ酸配列を含むか、または前記アミノ酸配列から成るポリペプチド；

(ii) 核内ホルモンレセプターリガンド結合ドメインの活性を有するか、または(i)のポリペプチドと共通の抗原決定基を有する、そのフラグメントであるポリペプチド；または

(iii) (i)または(ii)の機能的等価物であるポリペプチド。

**【請求項5】**

LBDG4ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域を含む、請求項4の(ii)に記載のフラグメントであるポリペプチドであって、前記核内ホルモンレセプターリガンド結合ドメイン領域が配列番号：4に記載のアミノ酸配列の残基805から1005を含むと定義され、前記フラグメントが“LBDモチーフ”残基ILE863、ASP870、GLN871およびLEU875、または等価な残基を保有し、さらに核内ホルモンレセプターリガンド結合ドメインの活性を保有する前記ポリペプチド。

30

**【請求項6】**

請求項4の(iii)に記載の機能的等価物であり、配列番号：4に記載のアミノ酸配列と相同であり、触媒性残基ILE863、ASP870、GLN871およびLEU875、または等価な残基を保有し、さらに核内ホルモンレセプターリガンド結合ドメインの活性を保有するポリペプチド。

40

**【請求項7】**

前記機能的等価物が、LBDG1ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域と相同である請求項3または6に記載のポリペプチド。

**【請求項8】**

BLASTバージョン2.1.3でNCBI(the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>)によって同定されたデフォルトパラメーター{Bl o s u m 6 2 マトリックス; ギャップ開放ペナルティ = 1 1 およびギャップ伸長ペナルティ = 1 }を用いて決定したとき、配列番号：2もしくは配列番号：4に記載のアミノ酸配列、または核内ホルモンレセプターリガンド結合ドメイン活性を保有するそのフラグメン

50

トと80%を超える配列同一性、好ましくは85%、90%、95%、98%または99%を超える配列同一性を有する、請求項1-7のいずれか1項に記載のフラグメントまたは機能的等価物。

【請求項9】

配列番号：2もしくは配列番号：4のいずれかに記載のアミノ酸配列を有するポリペプチド、または核内ホルモンレセプターリガンド結合ドメイン活性を保有するそのフラグメントと顕著な構造相同性を示す請求項1-8のいずれか1項に記載の機能的等価物。

【請求項10】

配列番号：2または配列番号：4の配列に由来する7つもしくはそれより多い(例えば8、10、12、14、16、18、20またはそれより多い)アミノ酸残基からなる、請求項1の(i)のポリペプチドと共通の抗原決定基を有する請求項1、2、4または5に記載のフラグメント。

10

【請求項11】

請求項1-10のいずれか1項に記載のポリペプチドをコードする精製核酸分子。

【請求項12】

配列番号：1もしくは配列番号：3に記載の核酸配列を有するか、または重複等価物もしくはそのフラグメントである請求項11に記載の精製核酸分子。

【請求項13】

配列番号：1のヌクレオチド2732から3328または配列番号：3のヌクレオチド2913から3515を含むか、またはその重複等価物である請求項11または請求項12に記載の精製核酸分子のフラグメント。

20

【請求項14】

高いストリンジェンシー条件下で請求項11-13のいずれか1項に記載の核酸分子とハイブリダイズする精製核酸分子。

【請求項15】

請求項11-14のいずれか1項に記載の核酸分子を含むベクター。

【請求項16】

請求項15に記載のベクターで形質転換された宿主細胞。

【請求項17】

請求項1-10のいずれか1項に記載のポリペプチドと特異的に結合し、さらに好ましくは前記ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン活性を抑制するリガンド。

30

【請求項18】

抗体である請求項17に記載のリガンド。

【請求項19】

請求項1-10のいずれか1項に記載のポリペプチドの発現レベルまたは活性レベルを増加または低下させる化合物。

【請求項20】

請求項1-10のいずれか1項に記載のポリペプチドと結合し、前記ポリペプチドの生物学的作用のいずれをも誘発することがない請求項19に記載の化合物。

40

【請求項21】

天然または改変された基質、リガンド、酵素、レセプターまたは構造的もしくは機能的模倣体である請求項19または20に記載の化合物。

【請求項22】

疾患の治療または診断で使用することを目的とする請求項1-10のいずれか1項に記載のポリペプチド、請求項11-14のいずれか1項に記載の核酸分子、請求項15に記載のベクター、請求項17または18に記載のリガンド、または請求項19-21のいずれか1項に記載の化合物。

【請求項23】

患者の疾患を診断する方法であって、前記患者の組織で請求項1-10のいずれか1項に

50

記載のポリペプチドをコードする天然の遺伝子の発現レベルを評価するか、または請求項 1 - 10 のいずれか 1 項に記載のポリペプチドの活性を評価し、さらに前記発現または活性レベルをコントロールレベルと比較することを含み、前記コントロールレベルと異なるレベルが疾患を示唆する前記疾患の診断方法。

【請求項 24】

*in vitro*で実施される請求項 23 に記載の方法。

【請求項 25】

(a) 請求項 17 または請求項 18 に記載のリガンドをリガンド - ポリペプチド複合体の形成に適した条件下で生物学的サンプルと接触させ；さらに (b) 前記複合体を検出する工程を含む、請求項 23 または請求項 24 に記載の方法。

10

【請求項 26】

a) 患者由来の組織サンプルを、請求項 11 - 14 のいずれか 1 項に記載の核酸分子と前記プローブとの間でハイブリッド複合体を形成し得るストリンジェントな条件下で核酸プローブと接触させる工程；

b) コントロールサンプルと前記プローブを工程 a) で用いた条件と同じ条件下で接触させる工程；さらに

c) 前記サンプル中のハイブリッド複合体の存在を検出する工程；

を含み、コントロールサンプル中のハイブリッド複合体レベルと異なるハイブリッド複合体レベルが患者サンプルで検出されることが疾患を示唆する、請求項 23 または請求項 24 に記載の方法。

20

【請求項 27】

a) 患者の組織由来の核酸サンプルを、請求項 11 - 14 のいずれか 1 項に記載の核酸分子と前記プライマーとの間でハイブリッド複合体を形成し得るストリンジェントな条件下で核酸プライマーと接触させる工程；

b) コントロールサンプルと前記プライマーを工程 a) で用いた条件と同じ条件下で接触させる工程；

c) 前記サンプル化された核酸を増幅する工程；および、

d) 患者サンプルおよびコントロールサンプルの両方の増幅された核酸レベルを検出する工程；

を含み、コントロールサンプル中の増幅核酸レベルと顕著に異なる増幅核酸レベルが患者サンプルで検出されることは疾患を示唆する、請求項 23 または請求項 24 に記載の方法。

30

【請求項 28】

以下の工程を含む請求項 23 または請求項 24 に記載の方法：

a) 疾患について検査しようとする患者から組織サンプルを得る工程；

b) 請求項 11 - 14 のいずれかの項に記載の核酸分子を前記組織サンプルから単離する工程；および、

c) 核酸分子中で患に関連する変異の存在を疾患の徴候として検出することによって患者を疾患について診断する工程。

【請求項 29】

核酸分子を増幅して増幅生成物を形成し、前記増幅生成物中で変異の有無を検出することをさらに含む請求項 28 の方法。

40

【請求項 30】

核酸分子をストリンジェントな条件下で前記核酸分子とハイブリダイズする核酸プローブと接触させてハイブリッド二本鎖分子を形成させ、前記ハイブリッド二本鎖分子は疾患に関連する変異に対応するいずれかの部分でハイブリダイズしない核酸プローブ鎖部分を有するハイブリッドであり、および、疾患関連変異の有無を表示するものとして前記プローブ鎖の非ハイブリダイズ部分の有無を検出することによって患者で変異の有無を検出する請求項 28 または 29 のいずれかの方法。

【請求項 31】

50

疾患が以下から選択される請求項 23 - 30 のいずれか1項に記載の方法：

細胞増殖性疾患（新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍並びに他の固形腫瘍を含む）、骨髄増殖性疾患（例えば白血病、非ホジキンリンパ腫を含む）、白血球減少症、血小板減少症、血管形成疾患、カポジ肉腫、自己免疫/炎症性疾患（アレルギー、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む）、心脈管系疾患（高血圧、浮腫、アングINA、アテローム性動脈硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む）、神経学的疾患（中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む）、発達障害、代謝性障害（真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシウム血症、高コレステロール血症、高脂血症および肥満を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（アクネ、湿疹および創傷治癒を含む）、加齢の負の作用、エイズ、感染（ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む）、並びに他の病的状態（特に核内ホルモンレセプターが関与するもの）。

10

【請求項 32】

核内ホルモンレセプターリガンド結合ドメインとしての請求項 1 - 10 のいずれか1項に記載のポリペプチドの使用。

【請求項 33】

核内ホルモンレセプターリガンド結合ドメイン活性を保有するタンパク質の発現を目的とする請求項 11 - 14 のいずれか1項に記載の核酸の使用。

20

【請求項 34】

請求項 1 - 10 のいずれか1項に記載のポリペプチドを利用する細胞-細胞接着を行わせる方法。

【請求項 35】

請求項 1 - 10 のいずれか1項に記載のポリペプチド、請求項 11 - 14 のいずれか1項に記載の核酸分子、請求項 15 に記載のベクター、請求項 17 または 18 に記載のリガンド、または請求項 19 - 21 のいずれか1項に記載の化合物を含む医薬組成物。

【請求項 36】

請求項 1 - 10 のいずれか1項に記載のポリペプチドまたは請求項 11 - 14 のいずれか1項に記載の核酸分子を含むワクチン組成物。

30

【請求項 37】

以下の疾患の治療用医薬の製造で使用される請求項 1 - 10 のいずれか1項に記載のポリペプチド、請求項 11 - 14 のいずれか1項に記載の核酸分子、請求項 15 に記載のベクター、請求項 17 または 18 に記載のリガンド、請求項 19 - 21 のいずれか1項に記載の化合物、または請求項 35 に記載の医薬組成物：

細胞増殖性疾患（新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍並びに他の固形腫瘍を含む）、骨髄増殖性疾患（例えば白血病、非ホジキンリンパ腫を含む）、白血球減少症、血小板減少症、血管形成疾患、カポジ肉腫、自己免疫/炎症性疾患（アレルギー、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む）、心脈管系疾患（高血圧、浮腫、アングINA、アテローム性動脈硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む）、神経学的疾患（中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む）、発達障害、代謝性障害（真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシウム血症、高コレステロール血症、高脂血症および肥満を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（アクネ、湿疹および創傷治癒を含む）、加齢の負の作用、エイズ、感染（ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む）、並びに他の病的状態（特に核内ホルモンレセプターが関与するもの）。

40

【請求項 38】

請求項 1 - 10 のいずれか1項に記載のポリペプチド、請求項 11 - 14 のいずれか1項

50

に記載の核酸分子、請求項 15 に記載のベクター、請求項 17 または 18 に記載のリガンド、請求項 19 - 21 のいずれか 1 項に記載の化合物、または請求項 35 に記載の医薬組成物を患者に投与することを含む、患者において疾患を治療する方法。

【請求項 39】

天然の遺伝子の発現またはポリペプチドの活性が、健常な対象者における発現または活性レベルと比較したとき罹患対象者で低下する疾患に対して、前記対象者に投与される前記ポリペプチド、核酸分子、ベクター、リガンド、化合物、または組成物がアゴニストである請求項 38 に記載の方法。

【請求項 40】

天然の遺伝子の発現またはポリペプチドの活性が、健常な対象者における発現または活性レベルと比較したとき罹患対象者で上昇する疾患に対して、前記対象者に投与される前記ポリペプチド、核酸分子、ベクター、リガンド、化合物または組成物がアンタゴニストである請求項 38 に記載の方法。

10

【請求項 41】

患者の疾患の治療的処置をモニターする方法であって、一定期間にわたって請求項 1 - 10 のいずれか 1 項に記載のポリペプチドの発現もしくは活性レベル、請求項 11 - 14 のいずれか 1 項に記載の核酸分子の発現レベルを前記患者由来の組織でモニターすることを含み、前記期間にわたってコントロールレベルに向かって前記発現または活性レベルが変化することが前記疾患の軽減の徴候である、前記患者の疾患の治療的処置をモニターする方法。

20

【請求項 42】

疾患の治療および/または診断で有効な化合物の同定方法であって、前記方法が、請求項 1 - 10 のいずれか 1 項に記載のポリペプチド、請求項 11 - 14 のいずれか 1 項に記載の核酸分子、または請求項 16 に記載の宿主細胞を、前記ポリペプチドまたは核酸に対する結合親和性を保有すると思われる 1 つまたは 2 つ以上の化合物と接触させ、さらに前記核酸分子またはポリペプチドと特異的に結合する化合物を選択することを含む前記有効な化合物の同定方法。

【請求項 43】

ストリンジェントな条件下で請求項 11 - 14 のいずれか 1 項に記載の核酸分子とハイブリダイズする核酸プローブを含む第一の容器；前記核酸分子の増幅に有用なプライマーを含む第二の容器；および疾患の診断を容易にすることを目的とする前記プローブおよびプライマーの使用のための指示書を含む疾患の診断に有用なキット。

30

【請求項 44】

ハイブリダイズしていない RNA を消化するための薬剤を保持する第三の容器をさらに含む請求項 43 のキット。

【請求項 45】

核酸分子のアレイを含むキットであって、前記核酸分子の少なくとも一つが請求項 11 - 14 のいずれか 1 項に記載の核酸分子である前記キット。

【請求項 46】

請求項 1 - 10 のいずれかの項に記載のポリペプチドと結合する 1 つまたは 2 つ以上の抗体；および前記抗体と前記ポリペプチドとの間の結合反応の検出に有用な試薬を含むキット。

40

【請求項 47】

請求項 1 - 10 のいずれか 1 項に記載のポリペプチドをより高レベルまたはより低レベルで発現するように、または前記ポリペプチドを発現しないように形質転換したトランスジェニックまたはノックアウト非ヒト動物。

【請求項 48】

請求項 47 に記載の非ヒトトランスジェニック動物を候補化合物と接触させ、さらに前記動物の疾患に対する前記化合物の影響を決定することによって、疾患の治療に有効な化合物をスクリーニングする方法。

50

## 【発明の詳細な説明】

## 【0001】

本発明は、本明細書において核内ホルモンレセプターのリガンド結合ドメインとして同定された、075385およびBAA31598.1と称される、新規なタンパク質並びに疾患の診断、予防および治療における前記タンパク質およびそのコード遺伝子に由来する核酸配列の使用に関する。

本明細書に引用した全ての刊行物、特許および特許出願は、引用により完全に本明細書に含まれるものとする。

## 【0002】

## (背景技術)

薬剤の発見プロセスにおいて、機能ゲノム学 (functional genomics) の時代の到来につれて根幹的な革命が現在進行している。“機能ゲノム学”という用語は、対象のタンパク質配列に機能を帰属させるためにバイオインフォマティクスツールを利用するアプローチに適用される。そのようなツールは、配列データの生成速度が、これらタンパク質配列に機能を割り当てる研究室の能力をはるかに越えるために、ますます必要性を増している。バイオインフォマティクスツールの潜在能力および精度が高まっているために、前記ツールは通常の生化学的特徴付け技術と急速に置き換えられつつある。実際、本発明の同定に用いた高度なバイオインフォマティクスツールは、今や、高い信頼性をもつ結果を産出する能力を有する。

配列データが利用可能になるにつれ、種々の研究機関および企業の組織がそれらを調査し、重要な発見が絶え間なく達成され続けている。しかしながら、研究および薬剤の発見のための標的として更に新たな遺伝子およびそれらがコードするポリペプチドを同定し特徴付ける必要性は引き続き存在している。

最近、未知の機能をもつ配列を評価するための注目すべきツールが本発明の出願人によって開発された。このツールは、同時係属国際特許出願第PCT/GB01/01105号の主題であるデータベースシステム (バイオベンジウム (Biopendium) 検索データベースと称される) である。このデータベースシステムは、独占的技術を用いて作製され、利用可能な全てのタンパク質または核酸配列の完全な比較から作製された情報を含む集積データリソースから成る。

## 【0003】

別個のデータリソースからこれら配列データを一体化させたその背後の目的は、可能な限り多くのデータを、配列それ自体と各配列に関連する情報の両方に関して1つの完全なリソースにまとめることである。各配列と関係を有する全ての利用可能なデータ (入手可能な場合はコードするタンパク質の三次元構造に関するデータを含む) を一緒に統合し、各配列について知られている情報の最大限の利用を可能にし、したがって最も多くの知見を含む予測がこれら配列の比較から入手することができる。前記データベースで作製され、各配列のエントリーに付随する注釈は、生物学的に関連がある事柄を前記配列情報に付与する。

このデータリソースは、配列のみをベースにしてタンパク質の正確な機能を予測することを可能にした。通常の場合、このような予測は、同じ機能ファミリーに属する他のタンパク質に対して高度な配列同一性 (約20% - 30%の同一性) を示すタンパク質についてのみ可能である。既知の機能を有する他の近縁なタンパク質と低度な配列相同性を示すタンパク質に対しては、正確な予測が不可能である。

本件では、その配列が公的に利用可能なデータベースに075385として記録されているタンパク質 (NCBI GenBankヌクレオチドアクセッション番号AF045458およびGenBankタンパク質アクセッション番号075385) が、核内ホルモンレセプターリガンド結合ドメインファミリーの新規なメンバーとして意味付けられる。

## 【0004】

## I. 核内ホルモンレセプターリガンド結合ドメインに関する導入部

核内ホルモンレセプター遺伝子スーパーファミリー (表1参照) は、標的遺伝子の転写を

10

20

30

40

50

調節する構造的に関連するタンパク質をコードする。前記タンパク質には、ステロイドホルモンおよび甲状腺ホルモン、ビタミン並びにリガンドが未だ発見されていない他のタンパク質に対するレセプターが含まれる。核内レセプターは2つの主要なドメイン、DNA結合ドメイン(DBD)およびリガンド結合ドメイン(LBD)から構成されている。DBDは、前記レセプターがモノマー、ホモダイマーまたはヘテロダイマーとして特異的なDNA配列と結合するように指令する。DBDは、核内レセプターでのみ見出される特殊なタイプのジンクフィンガーである。DBDを有する核内レセプターは、PROSITEコンセンサス配列(PS00031)とのマッチングについて検索することによって容易に配列レベルで同定することができる。

リガンド結合ドメイン(LBD)は同系のホルモンと結合しこれに应答する。LBDと結合したリガンドは、すでに結合していた“核内レセプターコリプレッサー”を排斥する構造的変化をひき起こす。続いて前記コリプレッサーによって以前に占有されていた部位は空席になり、“核内レセプターコアクチベーター”で補充される。このリガンドによってひき起こされるコアクチベーターによるコリプレッサーの交換が、リガンド結合が標的遺伝子の転写活性化をもたらすメカニズムである。全てのリガンド結合ドメインはコンセンサス配列、“LBDモチーフ”(表2参照)を含み、前記モチーフは、コリプレッサー結合およびコアクチベーター結合を仲介する。LBDは、今日まで全ての核内ホルモンレセプター標的薬剤のための結合部位であり、したがって新規なリガンド結合ドメインは魅力的な薬剤標的であるので、それらの同定が所望される。リガンド結合ドメインは低い配列同一性(約15%)しか共有しないが、非常に類似する構造を有し、したがって、ゲノムスレッダー(Genome Threader)のような構造に基づく関係付けツール(structure-based relationship tool)のための理想的なターゲットである。

PROSITEのような基本的な検索ツールを用いて、DBDが存在することによって、さらにそれをもとに推測されたLBDによって、多くのタンパク質配列が核内ホルモンレセプターとしてパブリックドメインですでに注釈付けされている。このために、ゲノムスレッダーによって同定される、核内レセプターとして注釈されてないどの新規LBDも完全にDBDを欠くであろうということが予期される。LBDを有するがDBDを欠くタンパク質の先例は、DAX1によって提供される。したがって、我々は、これらDBDのないヒットを“核内ホルモンレセプター”ではなく“核内ホルモンレセプターリガンド結合ドメイン”を含むものと注釈付けする。

【0005】

【表1】

10

20

30

表 1 : 核内ホルモンレセプタースーパーファミリー	
ファミリー :	ステロイドホルモンレセプター
サブファミリー :	グルココルチコイドレセプター プロゲステロンレセプター アンドロゲンレセプター エストロゲンレセプター
ファミリー :	甲状腺ホルモンレセプター様因子
サブファミリー :	レチノイン酸レセプター (RAR) レチノイドXレセプター (RXR) 甲状腺ホルモンレセプター ビタミンDレセプター NGFI-B FTZ-F1 ペルオキシソーム増殖因子活性化レセプター (PPAR) クジソンレセプター レチノイドオーファンレセプター (ROR) タイレス (Tailless) / COUP HNF-4 CF1 Knirps
ファミリー :	DAX1
サブファミリー :	DAX1

10

20

## 【0006】

下表 2 : “LBDモチーフ”。最上段の数字はモチーフ内の残基の位置を示す。文字は 1 文字コードによるアミノ酸を示す。1 つの縦の欄内の文字は全てモチーフ内のその位置について許容される。例えば、L、I、A、V、M、F、Y または W は “LBDモチーフ” の最初の位置を占めることができる。位置 4 と 8 の間、および位置 9 と 12 の間で見出される残基の数には変動が観察されることを述べる。“LBDモチーフ” は、核内ホルモンレセプターリガンド結合ドメインの 681 の配列を整理化し、残基の保存パターンを同定することによって構築された。

30

## 【0007】

## 【表 2】

1	2	3	4	5	6	7	8	9	10	11	12	13
L	任意の2 残基		L	任意の3残基 (又は、2残基、 または4残基)			D	Q	任意の2 残基 (又は、1 残基、 または3 残基)		L	L
I			I				E	N			I	I
A			A					R			A	A
V			V					H			V	V
M			M					K			M	M
F			F					S			F	F
Y			Y					T			Y	Y
W			W								W	W

10

20

## 【0008】

II. 核内ホルモンレセプターおよび疾患

核内ホルモンレセプターは、多様な生理学的機能において役割を果たすことが示されている。それらの多くは疾患プロセスにおいて役割を果たし得る（表3参照）

## 【0009】

表3：核内ホルモンレセプターと疾患

## 【表3】

核内ホルモンレセプター	疾患	
アンドロゲンレセプター	<p>アンドロゲン非感受性症候群 (Lubahn et al. 1989, Proc. Natl. Acad. Sci. USA 86: 9534-9538)</p> <p>ライフエンスタイン症候群 (Wooster et al. 1992, Nat. Genet. 2: 132-134)</p> <p>X連鎖劣性脊髄および延髄筋萎縮 (MacLean et al. 1995 Mol. Cell. Endocrinol. 112: 133-141)</p> <p>男性乳癌 (Wooster et al. 1992, Nat. Genet. 2: 132-134)</p>	10
グルココルチコイドレセプター	<p>ネルソン症候群 (Karl et al. 1996, J. Clin. Endocrinol. Metab. 81: 124-129)</p> <p>グルココルチコイド耐性急性T細胞白血病 (Hala et al. 1996, Int. J. Cancer 68: 663-668)</p>	
鉱質コルチコイドレセプター	<p>仮性低アルドステロン症 (Chung et al. 1995, J. Clin. Endocrinol. Metab. 80: 3341-3345)</p>	20
エストロゲンレセプター $\alpha$	<p>ER<math>\alpha</math>発現はヒト乳癌サブセットで上昇する。タモキシフェンの適用は乳癌進行を予防する主要な治療法である。残念ながら、ER<math>\alpha</math>陽性乳癌の35%はタモキシフェン耐性である (Petrangeli et al. 1994, J. Steroid Biochem. Mol. Biol. 49: 327-331)</p>	
ビタミンD <sub>3</sub> レセプター	<p>ビタミンD<sub>3</sub>レセプターの変異は、ビタミンD<sub>3</sub>欠乏に対する表現型 (Rickets) に類似する遺伝性疾患をもたらす (Hughes et al. 1988, Science 242: 1702-1725)</p>	30
レチノイン酸レセプター $\alpha$	<p>急性骨髄性白血病 (Lavau &amp; Dejean 1994, Leukemia 8: 9-15)</p>	
甲状腺ホルモンレセプター $\beta$	<p>“甲状腺ホルモンに対する普遍的耐性” (GRTH) (Refetoff 1994, Thyroid 4: 345-349)</p>	
DAX1	<p>X連鎖先天性副腎形成不全 (AHC) および性機能不全 (Ito et al. 1997, Mol. Cell. Biol. 17: 1476-1483)</p>	40

## 【0010】

したがって、核内ホルモンレセプターのLBDと結合するリガンドによる核内ホルモンレセプターの変化は、疾患の表現型を変化させる手段を提供する。したがって、新規核内ホルモンレセプターのリガンド結合ドメインの同定は、それらタンパク質が上記で同定した疾患や他の症状において役割を果たし得るために強く希求される。したがって新規核内ホルモンレセプターのリガンド結合ドメインの同定は、疾患（特に表3で同定したような疾患

)の治療および診断と密接な関連を有する。

【0011】

(発明の詳細な説明)

本発明は、075385タンパク質およびBAA31598.1タンパク質が核内ホルモンレセプターリガンド結合ドメインとして機能するという発見を基にしている。

075385タンパク質の場合、前記タンパク質配列の残基822 - 1020を含む領域は、ヒトエストロゲンレセプター (PDBコード1ERR:A)の残基1 (ALA307)から残基201 (MET517)と等価なフォールドを採用することが見出された。ヒトエストロゲンレセプターは、核内ホルモンレセプターのリガンド結合ドメインとして機能することが知られている。さらにヒトエストロゲンレセプターの“LBDモチーフ”残基、PHE367、ASP374、GLN375、LEU378およびLEU379は、075385ではそれぞれLEU878、ASP885、GLN886、LEU889およびLEU890として保存されている。前記の関係は単にヒトエストロゲンレセプターとの関係ではなく、むしろ全体として核内ホルモンレセプターリガンド結合ドメインファミリー全体との関係である。したがって、ヒトエストロゲンレセプター (1ERR:A)と075385とのゲノムスレッダー (Genome Threader) (商標)アラインメントの参照によって、075385のLEU878、ASP885、GLN886、LEU889およびLEU890は“LBDモチーフ”残基を形成すると予測される。

等価なフォールドおよび“LBDモチーフ”残基保存が組み合わされて、075385の前記領域の機能的注釈付けを可能とし、したがってこの領域を含むタンパク質が核内ホルモンレセプターリガンド結合ドメイン活性を有するという機能的注釈付けを可能とする。

【0012】

本発明の第一の特徴の態様ではポリペプチドが提供され、前記ポリペプチドは、

(i) 配列番号：2に記載のアミノ酸配列を含む；

(ii) 核内ホルモンレセプターリガンド結合ドメインの活性を有するか、または(i)のポリペプチドと共通の抗原決定基を有する、配列番号：2のフラグメントである；または

(iii) (i)または(ii)の機能的等価物である。

好ましくは、前記ポリペプチドは、

(i) 配列番号：2に記載のアミノ酸配列から成る；

(ii) 核内ホルモンレセプターリガンド結合ドメインの活性を有するか、または(i)のポリペプチドと共通の抗原決定基を有するそのフラグメントである；または

(iii) (i)または(ii)の機能的等価物である。

配列番号：2に記載の配列を有するポリペプチドは、以下では“LBDG1ポリペプチド”と称する。

本発明の前記特徴にしたがえば、上記(ii)に記載の好ましいポリペプチドフラグメントは、核内ホルモンレセプターリガンド結合ドメインの活性を担うと予測される、LBDG1ポリペプチドの領域(以下では“LBDG1核内ホルモンレセプターリガンド結合ドメイン領域”と称する)を含むか、またはその変種であって“LBDモチーフ”(LEU878、ASP885、GLN886、LEU889およびLEU890、または等価な残基)を保有する。本明細書で明確にされるように、LBDG1核内ホルモンレセプターリガンド結合ドメイン領域は、LBDG1ポリペプチド配列の残基822と残基1020の間に広がっていると考えられる。

本発明の前記特徴はまた、ポリペプチドのフラグメントおよび上記で定義したこれらポリペプチドフラグメントの変種を取り込んだ融合タンパク質も含むが、ただし前記融合タンパク質は核内ホルモンレセプターリガンド結合ドメインとしての活性を保有することを条件とする。

【0013】

075385 (LBDG1) のホモサピエンスパラログもまた同定され、本明細書ではLBDG4と称される。このポリペプチドはアクセッションコードBAA31598.1を有する。BAA31598.1は、075385 (LBDG1) と51%の配列同一性を示し、さらに075385のリガンド結合ドメインフォールドにおいて主要な役割を果たすと予測される残基がBAA31598.1 (LBDG4、図22参照)

で保存されている。075385 (LBDG1) に対する高い相溶性および予測される主要な残基の保存に基づいて、BAA31598.1 (LBDG4) もまた核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けされる。

等価なフォールドおよび“LBDモチーフ”残基保存の組み合わせにより、BAA31598.1の前記領域の注釈付けを可能にし、したがって前記領域を含むタンパク質が核内ホルモンレセプターのリガンド結合ドメイン活性を有するという機能的注釈付けを可能にする。

【0014】

したがって、本発明の第一の特徴による第二の態様では、ポリペプチドが提供され、そのポリペプチドは、

- (i) 配列番号：4に記載のアミノ酸配列を含む；
- (ii) 核内ホルモンレセプターリガンド結合ドメインの活性を有するか、または(i)のポリペプチドと共通の抗原決定基を有するそのフラグメントである；または
- (iii) (i)または(ii)の機能的等価物である。

好ましくは前記ポリペプチドは、

- (i) 配列番号：4に記載のアミノ酸配列から成る；
- (ii) 核内ホルモンレセプターリガンド結合ドメインの活性を有するか、または(i)のポリペプチドと共通の抗原決定基を有するそのフラグメントである；または
- (iii) (i)または(ii)の機能的等価物である。

配列番号：4に記載の配列を有するポリペプチドは以下では“LBDG4ポリペプチド”と称する。

本発明の前記特徴にしたがえば、上記(ii)に記載の好ましいポリペプチドフラグメントは、核内ホルモンレセプターリガンド結合ドメインの活性に必要なものと予測される、LBDG4ポリペプチドの領域(以下では“LBDG4核内ホルモンレセプターリガンド結合ドメイン領域”)を含むか、またはその変種であって“LBDモチーフ”(ILE863、ASP870、GLN871およびLEU875、または等価な残基)を保有する。本明細書で明確にされるように、LBDG4核内ホルモンレセプターリガンド結合ドメイン領域は、LBDG4ポリペプチド配列の残基805と残基1005の間に広がっていると考えられる。

本発明の前記特徴はまた、ポリペプチドのフラグメントおよび上記で定義したこれらポリペプチドフラグメントの変種を取り込んだ融合タンパク質を含むが、ただし前記融合タンパク質は核内ホルモンレセプターリガンド結合ドメインとしての活性を保有することを条件とする。

【0015】

第二の特徴では、本発明は、本発明の第一の特徴に記載のポリペプチドをコードする精製核酸分子を提供する。好ましくは前記精製核酸分子は、配列番号：1(LBDG1ポリペプチドをコードする)または配列番号：3(LBDG4ポリペプチドをコードする)に記載の核酸配列を有するか、または前記配列の重複等価物またはフラグメントである。好ましい核酸フラグメントは、上記(ii)に記載のポリペプチドフラグメント、好ましくはLBDG1またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域を含むポリペプチドフラグメントをコードするものであるか、上記に定義したこれらフラグメントの変種をコードするものである。

第三の特徴では、本発明は、本発明の第二の特徴の核酸分子と高ストリンジェンシー条件下でハイブリダイズする精製核酸分子を提供する。

第四の特徴では、本発明は、発現ベクターのように本発明の第二または第三の特徴の核酸分子を含むベクターを提供する。

第五の特徴では、本発明は、本発明の第四の特徴のベクターで形質転換された宿主細胞を提供する。

第六の特徴では、本発明は、本発明の第一の特徴のポリペプチドと特異的に結合し、核内ホルモンレセプターリガンド結合ドメインの活性を阻害するリガンドを提供する。

第七の特徴では、本発明は、本発明の第一の特徴のポリペプチドをコードする天然の遺伝子の発現を変化させるか、または本発明の第一の特徴のポリペプチドの活性を調節するた

10

20

30

40

50

めに有効な化合物を提供する。

本発明の第七の特徴の化合物は、前記ポリペプチドの遺伝子の発現レベルまたは活性レベルを増加（作働）させるか、または低下（拮抗）させる。重要なことには、それぞれLBDG1またはLBDG4ポリペプチドのLBDG1またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域として本明細書で明確にされる領域の機能を同定することによって、核内ホルモンレセプターリガンド結合ドメインが関与する疾患の治療および/または診断に有効な化合物を同定することが可能なスクリーニング方法をデザインすることができる。

【0016】

第八の特徴では、本発明は、診断または治療で使用するための本発明の第一の特徴のポリペプチド、または本発明の第二もしくは第三の特徴の核酸分子、または本発明の第四の特徴のベクター、または本発明の第五の特徴のリガンド、または本発明の第六の特徴の化合物を提供する。前記分子はまた以下を含む疾患の治療を目的とする医薬品の製造においても用いることができる：細胞増殖性疾患（新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍並びに他の固形腫瘍を含む）、骨髄増殖性疾患（例えば白血病、非ホジキンリンパ腫、白血球減少症、血小板減少症、血管形成疾患、カポジ肉腫）、自己免疫/炎症性疾患（アレルギー、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む）、心脈管系疾患（高血圧、浮腫、アングINA、アテローム性動脈硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む）、神経学的疾患（中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む）、発達障害、代謝性障害（真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシウム血症、高コレステロール血症、高脂血症および肥満を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（アクネ、湿疹および創傷治癒を含む）、加齢の負の作用、エイズ、感染（ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む）、並びに他の病的状態（特に核内ホルモンレセプターが関与するもの）。

【0017】

第九番目の特徴では、本発明は患者で疾患を診断する以下の工程を含む方法を提供する：本発明の第一の特徴のポリペプチドをコードする天然の遺伝子の発現レベル、または本発明の第一の特徴のポリペプチドの活性レベルを前記患者由来の組織で評価し、さらに前記発現または活性レベルをコントロールレベルと比較する工程であって、この場合前記コントロールレベルと異なるレベルは疾患を示唆する。前記の方法は好ましくは *in vitro* で実施されるであろう。同様な方法は患者での疾患治療のモニタリングに使用することができる。この場合、時間の経過にしたがってポリペプチドまたは核酸分子の発現もしくは活性レベルがコントロールレベルに向かって変化するのは疾患の緩解の指標となる。

本発明の第一の特徴のポリペプチドを検出する好ましい方法は以下の工程を含む：（a）本発明の第六の特徴のリガンド（例えば抗体）を生物学的サンプルとリガンド-ポリペプチド複合体の形成に適した条件下で接触させる工程；および、（b）前記複合体を検出する工程。

本発明の第九番目の特徴に記載の方法には、例えば短いプローブによる核酸ハイブリダイゼーション法、点変異分析、ポリメラーゼ連鎖反応（PCR）増幅、および異常なタンパク質レベルを検出する抗体を用いる方法といった種々の異なる方法が存在することは、当業者には明らかであろう。同様な方法を短期または長期ベースで用いて、モニターされる疾患の治療を可能にすることができる。本発明はまた前記疾患診断方法で有用なキットも提供する。

【0018】

第十番目の特徴では、本発明は、核内ホルモンレセプターリガンド結合ドメインとしての本発明の第一の特徴のポリペプチドの使用を提供する。本発明はまた、核内ホルモンレセプターリガンド結合ドメインの活性をもつタンパク質の発現のために本発明の第二または第三の特徴に記載の核酸分子の使用を提供する。本発明はまた核内ホルモンレセプターリガンド結合ドメイン活性を作用させる方法を提供し、前記方法は本発明の第一の特徴のポ

10

20

30

40

50

リペプチドを利用する。

第十一番目の特徴では本発明は医薬組成物を提供し、前記医薬組成物は、本発明の第一の特徴のポリペプチド、または本発明の第二もしくは第三の特徴の核酸分子、または本発明の第四の特徴のベクター、または本発明の第六の特徴のリガンド、または本発明の第七の特徴の化合物を医薬的に許容できる担体と組合わせて含有する。

第十二番目の特徴では、本発明は、疾患の診断または治療を目的とする医薬品の製造で使用するために、本発明の第一の特徴のポリペプチド、または本発明の第二のもしくは第三の特徴の核酸分子、または本発明の第四の特徴のベクター、本発明の第六の特徴のリガンド、または本発明の第七の特徴の化合物を提供する。前記診断または治療される疾患は、例えば、細胞増殖性疾患（新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍並びに他の固形腫瘍、骨髄増殖性疾患（例えば白血病、非ホジキンリンパ腫、白血球減少症、血小板減少症、血管形成疾患、カボジ肉腫）、自己免疫/炎症性疾患（アレルギー、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む）、心脈管系疾患（高血圧、浮腫、アングナ、アテローム性動脈硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む）、神経学的疾患（中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む）、発達障害、代謝性障害（真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシウム血症、高コレステロール血症、高脂血症および肥満を含む）、腎疾患（糸球体腎炎、腎血管性高血圧を含む）、皮膚疾患（アクネ、湿疹および創傷治療を含む）、加齢の負の作用、エイズ、感染（ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む）、並びに他の病的状態（特に核内ホルモンレセプターが関与するもの）である。

#### 【0019】

第十三番目の特徴では、本発明は患者の疾患を治療する方法を提供し、前記方法は、本発明の第一の特徴のポリペプチド、または本発明の第二もしくは第三の特徴の核酸分子、または本発明の第四の特徴のベクター、または本発明の第六の特徴のリガンド、または本発明の第七の特徴の化合物を患者に投与することを含む。

本発明の第一の特徴のポリペプチドをコードする天然の遺伝子の発現、または本発明の第一の特徴のポリペプチドの活性が、健常な対象者の発現または活性レベルと比較したとき罹患対象者で低下する疾患の場合、前記患者に投与される前記ポリペプチド、核酸分子、リガンドまたは化合物はアゴニストであろう。逆に、前記天然の遺伝子の発現、または前記ポリペプチドの活性が、健常な対象者の発現または活性レベルと比較したとき罹患対象者で上昇する疾患の場合、前記患者に投与される前記ポリペプチド、核酸分子、リガンドまたは化合物はアンタゴニストであろう。前記アンタゴニストの例にはアンチセンス核酸分子、リボザイムおよびリガンド（例えば抗体）が含まれる。

第十四番目の特徴では、本発明は、本発明の第一の特徴のポリペプチドを高レベルで、または低レベルで発現させるために、または全く発現させないために形質転換したトランスジェニックまたは遺伝子ノックアウト非ヒト動物を提供する。前記トランスジェニック動物は、疾患の研究用モデルとして非常に有用であり、さらに前記疾患の治療または診断に有効な化合物の同定を目的とするスクリーニング方法で用いることができる。

#### 【0020】

本発明を利用するために用いることができる標準的な技術および方法の要旨は下記で提供される。本発明は、記載した同定の方法論、プロトコル、細胞株、ベクターおよび試薬に限定されないことは理解されよう。本明細書で用いられる専門用語は単に個々の態様を説明するためのものであり、前記用語によって本発明の範囲を限定しようとするものではないこともまた理解されよう。本発明の範囲は添付の請求の範囲の用語によってのみ限定される。

本明細書では、ヌクレオチドおよびアミノ酸についての標準的な略語が用いられる。

本発明の実施では別に指示がなければ、分子生物学、微生物学、リコンビナントDNA技術および免疫学の通常の技術が用いられるであろう。前記技術は当業者の技術範囲内である

。前記のような技術は文献で完全に説明されている。特に適切な解説書の例には以下が含まれる：Sambrook Molecular Cloning； A Laboratory Manual, Second Edition (1989)； DNA Cloning, Vol. I and II ( D.N. Glover ed. 1985)； Oligonucleotide Synthesis (M .J. Gait ed. 1984)； Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984)； Transcription and Translation (B.D. Hames & S.J. Higgins eds. 1984)； Animal Cell Culture (R.I. Freshney ed. 1986)； Immobilized Cells and Enzymes (IRL Press, 1986)； B. Perbal, A Practical Guide to Molecular Cloning (1984)； the Methods in Enzymology series (Academic Press, Inc.)特にVol. 154 & 155； Gene Transfer Vectors for Mammalian Cells (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory)； Immunochemical Methods in Cell and Molecular Biology (Mayer and Walker, eds. 1987, Academic Press, London)； Scopes, (1987) Protein Purification : Principles and Practice, Second Edition (Springer Verlag, NY)； および Handbook of Experimental Immunology, Vols. I - IV (D.M. Weir and C.C. Blackwell eds. 1986)。

10

20

30

40

50

#### 【 0 0 2 1 】

本明細書において用いる“ポリペプチド”という用語は、ペプチド結合または改変ペプチド結合によって互いに結合した2つまたは3つ以上のアミノ酸を含む任意のペプチドまたはタンパク質（すなわちペプチドイソスター）が含まれる。この用語は、短鎖（ペプチドおよびオリゴペプチド）および長鎖（タンパク質）の両方を指す。

本発明のポリペプチドは成熟タンパク質の形態を有するものでもよく、またプレ - 、プロ - またはプレプロ - タンパク質であってプレ - 、プロ - またはプレプロ - 部分の切断によって活性化され、活性成熟ポリペプチドを生じるタンパク質でもよい。そのようなポリペプチドでは、プレ - 、プロ - またはプレプロ - 配列はリーダー配列もしくは分泌配列であっても、または成熟ポリペプチド配列の精製のために用いられる配列であってもよい。

本発明の第一の特徴のポリペプチドは融合タンパク質の一部を形成することができる。例えば、1つまたは2つ以上の付加アミノ酸配列を含むことがしばしば有利である。前記付加アミノ酸配列は、例えばリコンビナント形成時に、分泌もしくはリーダー配列、プロ - 配列、精製を促進する配列、またはより高いタンパク質安定性を付与する配列を含んでもよい。あるいは、または前記に加えて、前記成熟ポリペプチドを別の化合物、例えば前記ポリペプチドの半減期を増加させるような化合物（例えばポリエチレングリコール）を融合させることができる。

#### 【 0 0 2 2 】

ポリペプチドは、天然のプロセス（例えば翻訳後プロセッシング）によって、または当業者に周知の化学的改変技術によって改変された、20の遺伝子コードアミノ酸以外のアミノ酸を含んでいてもよい。本発明のポリペプチドに一般的に存在する公知の改変にはグリコシル化、脂質付加、硫化、 $\alpha$ -カルボキシル化（例えばグルタミン酸残基の）、ヒドロキシル化およびADP - リボシル化がある。他の可能な改変には、アセチル化、アシル化、アミド化、フラビンの共有結合付加、ヘム部分の共有結合付加、ヌクレオチドまたはヌクレオチド誘導体の共有結合付加、脂質誘導体の共有結合付加、ホスファチジルイノシトールの共有結合付加、架橋、環状化、ジスルフィド結合形成、脱メチル化、共有結合架橋の形成、システインの形成、ピログルタメートの形成、ホルミル化、GPIアンカー形成、ヨード化、メチル化、ミリストイル化、酸化、タンパク分解性プロセッシング、リン酸化、プレニル化、ラセミ化、セレノイル化、タンパク質へのトランスファーRNA媒介性アミノ酸付加（例えばアルギニル化）およびユビキチン結合が含まれる。

改変は、ペプチド骨格、アミノ酸側鎖およびアミノまたはカルボキシ末端を含むポリペプチド内のいずれの場所に生じてもよい。実際、共有結合改変によるポリペプチドのアミノまたはカルボキシ末端またはその両端の妨害(blockage)は、天然に存在するポリペプチドおよび合成ポリペプチドで一般的であり、そのような改変は本発明のポリペプチドにも存在し得る。

## 【0023】

ポリペプチド内に生じる改変は多くの場合ポリペプチドが生成される方法の機能であろう。組換えによって生成されるポリペプチドの場合、大部分の改変の性質および程度は、個々の宿主細胞の改変能力および問題のポリペプチドのアミノ酸配列に存在する改変シグナルによって決定されるであろう。例えば、グリコシル化パターンは異なる種類の宿主細胞間で変動するのである。

本発明のポリペプチドは任意の適切な様式で調製することができる。そのようなポリペプチドには、単離された天然に存在するポリペプチド（例えば細胞培養から精製）、組換えにより生成されたポリペプチド（融合タンパク質を含む）、合成により生成されたポリペプチド、または前記方法を併用して生成されたポリペプチドが含まれる。

10

本発明の第一の特徴の機能的に等価なポリペプチドは、LBDG1またはLBDG4ポリペプチドと相同なポリペプチドであり得る。2つのポリペプチドは、前記ポリペプチドの一方の配列が他方のポリペプチドの配列に対して十分な同一性または類似性を有する場合、本明細書で用いられる用語のように“相同である”と称される。“同一性”とは、整列化した配列のどの特定の場所においても、アミノ酸残基が前記配列間で同一であることを示す。“類似性”は、整列化した配列のいずれの特定の場所においても、アミノ酸残基が前記配列間で類似の種類であることを示す。同一性および類似性の度合いは容易に計算できる（Computational Molecular Biology, A.M. Lesk ed., Oxford University Press, New York, 1988; Biocomputing. Informatics and Genome Projects, D.W. Smith ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part 1, A.M. Griffin and H.G. Griffin eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, G. von Heinje, Academic Press, 1987; および Sequence Analysis Primer, M. Gribskov and J. Devereux eds., M. Stockton Press, New York, 1991)。

20

## 【0024】

したがって、相同なポリペプチドには、LBDG1ポリペプチドまたはLBDG4ポリペプチドの天然の生物学的変種（例えば前記ポリペプチドが由来した種における対立形質変種または地理的変種）および変異体（例えばアミノ酸置換、挿入または欠失を含む変異体）が含まれる。前記変異体は、1つまたは2つ以上のアミノ酸残基が保存的または非保存的アミノ酸残基（好ましくは保存的アミノ酸残基）で置換されているポリペプチドを含んでもよく、さらにそのような置換アミノ酸残基は遺伝コードでコードされたものでもそうでなくてもよい。典型的な前記の置換は、Ala、Val、LeuおよびIle間で；SerとThr間で；酸性残基AspとGlu間で；AsnとGln間で、塩基性残基LysとArg間で；または芳香族残基PheとTyr間で生じる。特に好ましいものは、いくつ（すなわち5から10、1から5、1から3、1から2、または単に1つ）のアミノ酸が任意の組合せで置換されたまたは欠失または付加された変種である。特に好ましいものは、タンパク質の特性および活性を変化させないサイレント置換、付加および欠失である。さらにこれに関して特に好ましいものは保存的置換である。

30

前記変異体にはまた、1つまたは2つ以上のアミノ酸残基が置換基を含むポリペプチドが含まれる。

## 【0025】

典型的には、2つのポリペプチド間で80%を越える同一性（好ましくは特定の領域で）が機能的等価物を示すと考えられる。好ましくは、本発明の第一の特徴の機能的に等価なポリペプチドは、LBDG1ポリペプチドもしくはLBDG4ポリペプチドに関して、またはその活性なフラグメントに関して80%を越える配列同一性を有する。より好ましいポリペプチドは、LBDG1もしくはLBDG4ポリペプチドまたはその活性なフラグメントに関してそれぞれ85%、90%、95%、98%または99%を越える同一性を有する。

40

本明細書で言及される同一性のパーセンテージは、BLASTバージョン2.1.3でNCBI(the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>)によって特定されたデフォルトパラメーター{Blosum62マトリックス; ギャップ開放(open)ペナルティー=11およびギャップ伸長(extension)ペナルティー=1}を用いて決定され

50

るとおりである。

本件では、LBDG1またはLBDG4ポリペプチドの好ましい活性フラグメントは、LBDG1またはLBDG4の核内ホルモンレセプターリガンド結合ドメイン領域を含むもの、および残基LEU878、ASP885、GLN886、LEU889およびLEU890、または等価な残基を有する“LBDモチーフ”を保有するものである。“等価な残基”とは、“LBDモチーフ”残基と等価である残基を意味するが、ただし核内ホルモンレセプターリガンド結合ドメイン領域が核内ホルモンレセプターリガンド結合ドメインとしての活性を保持していることを条件とする。例えば、LEU878はILE、ALA、VAL、MET、PHE、TYRまたはTRPによって置換することができる。例えば、ASP885はGLUと置換することができる。例えば、GLN886は、ASN、LYS、HIS、ARG、SERまたはTHRと置換できる。例えば、LEU889は、ILE、ALA、VAL、MET、PHE、TYRまたはTRPと置換することができる。例えば、LEU890は、ILE、ALA、VAL、MET、PHE、TYRまたはTRPと置換することができる。LBDG4(活性な残基はILE863、ASP870、GLN871およびLEU875)についても、同様な態様で残基を置換することができる。したがって本発明のこの特徴は、LBDG1またはLBDG4ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域に関して80%を越える同一性、好ましくは85%、90%、95%、98%または99%を越える同一性をそれぞれ有するポリペプチド、および、LEU878、ASP885、GLN886、LEU889およびLEU890、または等価な残基(LBDG4ポリペプチドではILE863、ASP870、GLN871およびLEU875)をもつ“LBDモチーフ”を保有するものを含む。上記で考察したように、LBDG1の核内ホルモンレセプターリガンド結合ドメイン領域は、LBDG1ポリペプチド配列の残基822と残基1020の間に広がっていると考えられ、一方、LBDG4核内ホルモンレセプターリガンド結合ドメイン領域は、LBDG4ポリペプチド配列の残基805と残基1005の間に広がっていると考えられる。

#### 【0026】

本発明の第一の特徴の機能的に等価なポリペプチドはまた、1つまたは2つ以上の構造的アラインメント技術を用いて同定されたポリペプチドであろう。例えば、バイオペンジウム検索データベースを作製するために用いられた検索ツールの1つの特徴を形成するインファーマチカゲノムスレッダー(商標)(Inpharmatica Genome Treader)技術を用いて(同時係属国際特許出願PCT/GB01/01105を参照されたい)、LBDG1またはLBDG4ポリペプチドと比較したとき、低い配列同一性を有するが、LBDG1またはLBDG4ポリペプチド配列と顕著な構造的相同性を共有するがゆえに、核内ホルモンレセプターリガンド結合ドメイン活性をもつと予測される、現在は未知の機能を有するポリペプチドを同定することができる。

“顕著な構造的相同性”とは、インファーマチカゲノムスレッダー(商標)が、2つのタンパク質またはタンパク質領域が、少なくとも10%、より好ましくは少なくとも20%、30%、40%、50%、60%、70%、80%、90%および前記を越える確実性で構造的相同性を共有することを予測することを意味する。前記インファーマチカゲノムスレッダー(登録商標)の前記確実性の値は以下のように計算される。既知の構造を有する配列をもつばらを使用して、初めに一組の比較をインファーマチカゲノムスレッダー(登録商標)を用いて行った。いくつかの比較は(構造を基準にして)関連することが判明しているタンパク質間で行った。続いてニューラルネットワークを、CATH構造分類([www.biochem.ucl.ac.uk/bsm/cath](http://www.biochem.ucl.ac.uk/bsm/cath))から得られる既知の関係と既知の非関係とを最も良く識別することを必要とすることに基づいて調整(train)した。これによって0と1の間のニューラルネットワークスコアが得られた。しかしながら、一方で関連するタンパク質の数および無関係のタンパク質の数は既知であるので、前記ニューラルネットワークの結果を小群に分配し、正確な結果のパーセンテージを経験的に計算することが可能であった。このようにして、バイオペンジウム検索データベースにおける全ての真正の予測はニューラルネットワークスコアが付随しており、信頼百分率は、インファーマチカゲノムスレッダー(登録商標)がいかに良好なトレーニング/テストセットであるかを反映したものである。

#### 【0027】

10

20

30

40

50

LBDG1の構造的相同体は、LBDG1の核内ホルモンレセプターリガンド結合ドメイン領域と構造的相同性を共有し、“LBDモチーフ”残基LEU878、ASP885、GLN886、LEU889およびLEU890、または等価な残基を保有するはずである。そのような構造的相同体は、前記ポリペプチド配列と顕著な構造的相同性を共有し、さらに“LBDモチーフ”残基を保有するがゆえに、核内ホルモンレセプターリガンド結合ドメイン活性を有すると予測される。

LBDG4の構造的相同体は、LBDG4の核内ホルモンレセプターリガンド結合ドメイン領域と構造的相同性を共有し、“LBDモチーフ”残基ILE863、ASP870、GLN871およびLEU875、または等価な残基を保有するはずである。そのような構造的相同体は、前記ポリペプチド配列と顕著な構造的相同性を共有し、さらに“LBDモチーフ”残基を保有するがゆえに、核内ホルモンレセプターリガンド結合ドメイン活性を有すると予想される。

本発明の第一の特徴のポリペプチドにはまた、LBDG1またはLBDG4ポリペプチドのフラグメント、LBDG1またはLBDG4ポリペプチドのフラグメントの機能的等価物、およびLBDG1またはLBDG4ポリペプチドの機能的等価物のフラグメントが含まれるが、ただし前記機能的等価物およびフラグメントは核内ホルモンレセプターリガンド結合ドメイン活性を保持するか、またはLBDG1またはLBDG4ポリペプチドと共通の抗原決定基を有することを条件とする。

#### 【0028】

本明細書において用いる、“フラグメント”という用語は、LBDG1またはLBDG4ポリペプチド、またはその機能的等価物のアミノ酸配列の一部（全部ではなく）と同じアミノ酸配列を有するポリペプチドを指す。フラグメントは、前記配列に由来する少なくともn個の連続するアミノ酸を含むべきである。さらに、個々の配列に応じて、nは7またはそれより大きい（例えば8、10、12、14、16、18、20またはそれより大きい）。小さなフラグメントは抗原決定基を構成することができる。

本発明のこの特徴の好ましいポリペプチドフラグメントは、LBDG1またはLBDG4ポリペプチドのそれぞれLBDG1またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域と本明細書で明確にされる領域を含むフラグメントである。これらの領域は、核内ホルモンレセプターリガンド結合ドメインと注釈付けされた領域である。

LBDG1ポリペプチドの場合、前記領域は残基822と残基1020の間に広がっていると考えられる。LBDG4ポリペプチドの場合、この領域は残基805と残基1005の間に広がっていると考えられる。

前記フラグメントの変種は、本発明のこの特徴の態様として含まれるが、ただしこれら変種は核内ホルモンレセプターリガンド結合ドメインとしての活性を保有することを条件とする。

ある観点では、“変種”という用語は前記ポリペプチドフラグメントの伸長型または短縮型を含むことが意図される。

#### 【0029】

伸長型変種の場合、LBDG1またはLBDG4ポリペプチド配列内のこれら境界のC末端および/またはN末端にさらに付加された残基が前記ポリペプチドフラグメントに含まれるとき、LBDG1またはLBDG4ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域が正確に折り畳まれ、核内ホルモンレセプターリガンド結合ドメインの活性を示すであろうということは想像に難くない。例えば、LBDG1またはLBDG4ポリペプチド配列、または相同な配列に由来する5、10、20、30、40または50またはそれより多い付加アミノ酸残基は、LBDG1またはLBDG4ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域の境界のC末端および/またはN末端のいずれか一方または両方に含まれ、前記ポリペプチドフラグメントは、正確に折り畳まれる能力を損なうことなく核内ホルモンレセプターリガンド結合ドメイン活性を示し得る。

LBDG1またはLBDG4ポリペプチドの短縮型変種の場合、1つまたは2つ以上のアミノ酸残基をLBDG1ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域のC末端またはN末端の一方または両方で欠失させることができるが、ただし“LBDモチーフ”残基（LBDG1の場合、LEU878、ASP885、GLN886、LEU889およびLEU890；LBDG4の場合、前記残基

はILE863、ASP870、GLN871およびLEU875)または等価な残基は無傷のまま維持され、欠失は、前記残基のいずれかが欠失するほど前記ポリペプチド配列内に深く伸長することはない。

#### 【0030】

第二の観点では、“変種”という用語は、LBDG1またはLBDG4ポリペプチドの核内ホルモンレセプターリガンド結合ドメイン領域と顕著な配列相同性を保有し、かつ“LBDモチーフ”残基(LEU878、ASP885、GLN886、LEU889およびLEU890; LBDG4の場合、前記残基はILE863、ASP870、GLN871およびLEU875)または等価な残基を保有する上記で述べたポリペプチドフラグメントの相同体を含むが、ただし前記変種は核内ホルモンレセプターリガンド結合ドメインとしての活性を保持することを条件とする。

10

相同体には、LBDG1またはLBDG4ポリペプチドのうち、それぞれLBDG1またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域と80%を超える同一性を保有するポリペプチド分子が含まれる。同一性パーセンテージは、BLASTバージョン2.1.3でNCBI(the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>)によって特定されるデフォルトパラメーター{Blosum62マトリックス; ギャップ開放ペナルティー=11およびギャップ伸長ペナルティー=1}を用いて決定されるとおりである。好ましくは、本発明のこの特徴のポリペプチドフラグメントの変種相同体は、LBDG1またはLBDG4ポリペプチドのLBDG1またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域に関してそれぞれ80%を超える同一性を有する。より好ましくは、変種ポリペプチドは、LBDG1またはLBDG4ポリペプチドのLBDG1またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域に関してそれぞれ85%、90%、95%、98%または99%を超える同一性をそれぞれ有するが、ただし前記変種は核内ホルモンレセプターリガンド結合ドメインとしての活性を保持することを条件とする。変種ポリペプチドはまた、上記で考察したポリペプチドフラグメントの短縮型相同体も含むが、ただし前記変種は核内ホルモンレセプターリガンド結合ドメインとしての活性を保持することを条件とする。

20

#### 【0031】

本発明の第一の特徴のポリペプチドフラグメントは、例えばインファーマチカゲノムスレッダー(商標)によって同定される、LBDG1またはLBDG4ポリペプチド配列のLBDG1またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域によって規定されるポリペプチドフラグメントの構造と顕著な構造的相同性を示すポリペプチドフラグメントであり得る。したがって、LBDG1またはLBDG4ポリペプチド配列のLBDG1またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域によって規定されるポリペプチドフラグメントの構造的相同体であるポリペプチドフラグメントは、上記で定義されたフォールドのような、LBDG1またはLBDG4ポリペプチドフラグメントによって採用されるフォールドと同じフォールドを採用するはずである。

30

LBDG1核内ホルモンレセプターリガンド結合ドメイン領域によって規定されるポリペプチドフラグメントの構造的相同体はまた、“LBDモチーフ”残基LEU878、ASP885、GLN886、LEU889およびLEU890、または等価な残基も保持するはずである。LBDG4核内ホルモンレセプターリガンド結合ドメイン領域によって規定されるポリペプチドフラグメントの構造的相同体もまた、“LBDモチーフ”残基ILE863、ASP870、GLN871およびLEU875、または等価な残基を保持するはずである。

40

#### 【0032】

そのようなフラグメントは、“独立的存在(free-standing)”すなわち、他のアミノ酸もしくはポリペプチドの部分でなくてもよく、他のアミノ酸もしくはポリペプチドに融合されていなくてもよく、それらフラグメントがその部分または領域を構成するより大きなポリペプチドの内部に含まれていてもよい。より大きなポリペプチドの内部に含まれている場合は、本発明のフラグメントは最も好ましくは連続するただ1つの領域を形成する。例えばある種の好ましい態様は、前記フラグメントのアミノ末端に融合したプレ-および/またはプロ-ポリペプチド領域、および/または前記フラグメントのカルボキシ末端に融合した付加的領域を有するフラグメントに関する。しかしながら、いくつかのフラグメン

50

トが単一のより大きなポリペプチドの内部に含まれてもよい。

本発明のポリペプチドまたはその免疫原性フラグメント（少なくとも1つの抗原決定基を含む）を用いて、例えばポリクローナルまたはモノクローナル抗体といった、前記ポリペプチドに免疫特異的なリガンドを作製することができる。そのような抗体は、本発明のポリペプチドを発現しているクローンを単離することもしくは同定すること、またはアフィニティークロマトグラフィーによって前記ポリペプチドを精製することに用いることができる。前記抗体はまた、当業者には明らかなように他の利用の中で特に診断的または治療的補助としても用いられ得る。

#### 【0033】

“免疫特異的”という用語は、前記抗体が、従来技術において他の関連ポリペプチドに対する親和性よりも本発明のポリペプチドに対して実質的に強い親和性を有することを意味する。本明細書で用いる“抗体”という用語は、完全な分子だけでなく問題の抗原決定基と結合することができるそのフラグメント、例えばFab、F(ab')<sub>2</sub>およびFvも意味する。したがって、そのような抗体は本発明の第一の特徴のポリペプチドと結合する。

ポリクローナル抗体が所望される場合は、選択される哺乳類（例えばマウス、ウサギ、ヤギまたはウマ）は、本発明の第一の特徴のポリペプチドで免疫することができる。動物を免疫するために用いられるポリペプチドは、リコンビナントDNA技術によって誘導するか、または化学的に合成することができる。所望する場合には、前記ポリペプチドは担体タンパク質と結合させることができる。前記ポリペプチドと化学的に結合させることができる一般的に用いられる担体には、ウシ血清アルブミン、サイログロブリンおよびキーホールリンペットヘモシアニンが含まれる。続いて前記担体結合ポリペプチドを用いて動物を免疫することができる。血清は免疫した動物から採集され、既知の方法（例えばイムノアフィニティークロマトグラフィー）にしたがって処理される。

#### 【0034】

本発明の第一の特徴のポリペプチドに対するモノクローナル抗体も、当業者は容易に生成できる。ハイブリドーマ技術を用いてモノクローナル抗体を作製する一般的な方法論は周知である（例えば以下を参照されたい：G. Kohler & C. Milstein, *Nature* 256: 495-497(1975); Kozbor et al., *Immunology Today* 4: 72(1983); Cole et al., 77-96 “*Monoclonal Antibodies and Cancer Therapy*”, Alan R. Liss, Inc. (1985)）。

本発明の第一の特徴のポリペプチドに対して生成されたモノクローナル抗体のパネル(panel)を種々の特性、すなわちアイソタイプ、エピトープ、親和性などについてスクリーニングすることができる。モノクローナル抗体は、それらを作らせた個々のポリペプチドの精製に特に有用である。あるいは、対象のモノクローナル抗体をコードする遺伝子を、例えば当技術分野で知られるPCR技術によってハイブリドーマから単離し、さらにクローニングし適切なベクターで発現させることができる。

非ヒト可変領域がヒト定常領域と結合または融合されているキメラ抗体（例えば以下を参照されたい：Proc. Natl. Acad. Sci. USA, 84: 3439(1987)）もまた有用であろう。

#### 【0035】

抗体は、例えばヒト化によって改変して各個体での免疫原性を減少させることができる（例えば以下を参照されたい：Jones et al., *Nature*, 321: 522(1986); Verhoeyen et al., *Science*, 239: 1534(1988); Kabat et al., *J. Immunol.*, 147: 1709(1991); Queen et al., Proc. Natl. Acad. Sci. USA, 86: 10029(1989); Gorman et al., Proc. Natl. Acad. Sci. USA, 88: 34181(1991); Hodgson et al., *Bio/Technology* 9: 421(1991)）。

本明細書で用いられる“ヒト化抗体”という用語は、非ヒトドナー抗体の重鎖および/または軽鎖の可変ドメイン中のCDRアミノ酸および選択した他のアミノ酸がヒト抗体の等価なアミノ酸に代えて置換されている抗体分子を指す。したがって、ヒト化抗体はヒトの抗体と密接に類似するがドナー抗体の結合能力を有する。

また別の選択肢では、前記抗体は、2つの異なる抗原結合ドメインを有し、各ドメインは異なるエピトープに向けられている“二重特異性”抗体であってもよい。

ファージディスプレイ技術を用いて、本発明のポリペプチドに対する結合活性をもつ抗体

10

20

30

40

50

をコードする遺伝子を、関連抗体の保有についてスクリーニングされたヒト由来のリンパ球のPCR増幅V-遺伝子のレパトリー、または未感作ライブラリーのいずれかから選択することができる(J. McCafferty et al., (1990) Nature 348: 552-554; J. Marks et al., (1992) Biotechnology 10: 779-783)。前記抗体の親和性は、鎖のシャッフリングによって改善することもできる(T. Clackson et al., (1991) Nature 352: 624-628)。

上記の技術によって作製された抗体は(ポリクローナルであれモノクローナルであれ)、免疫アッセイ、ラジオイムノアッセイ(RIA)または酵素結合免疫吸着アッセイ(ELISA)で試薬として用いることができるというさらに別の有用性を有する。前記の利用では、これら抗体は分析的に検出可能な試薬(例えば放射性同位元素、蛍光分子または酵素)で標識することができる。

10

#### 【0036】

本発明の第二および第三の特徴の好ましい核酸分子は、配列番号: 2または配列番号: 4に記載のポリペプチド配列および機能的に等価なポリペプチドをコードするものである。前記機能的に等価なポリペプチドには、LBDG1またはLBDG4ポリペプチドの活性なフラグメント、例えばLBDG1またはLBDG4ポリペプチド配列のLBDG1またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域を含むフラグメントフラグメント、またはその相同体が含まれる。

これら配列の一続きを包含する核酸分子は本発明のこの特徴の好ましい態様を形成する。これらの核酸分子は、本明細書に記載する方法および応用で用いることができる。好ましくは本発明の核酸分子は、本明細書に開示する配列に由来する少なくともn個の連続するヌクレオチドを含む。nは、個々の配列に応じて10またはそれより大きい(例えば12、14、15、18、20、25、30、35、40またはそれより大きい)。

20

本発明の核酸分子は、上記で述べた核酸分子に相補的な配列も含む(例えばアンチセンスまたはプローブとしての目的のために)。

本発明の核酸分子は、RNA(例えばmRNA)、またはDNA(例えばcDNA、合成DNAまたはゲノムDNAを含む)の形態をとることができる。そのような核酸分子は、クローニングによって、化学合成によって、またはそれらを併用して得ることができる。前記核酸分子は、固相ホスホルアミダイト化学合成のような技術を用いるゲノムまたはcDNAライブラリーからの化学合成によって、または生物体から分離することによって調製することができる。RNA分子は一般的にはDNA配列のin vitroまたはin vivo転写によって作製することができる。

30

#### 【0037】

核酸分子は二本鎖でも一本鎖でもよい。一本鎖DNAはコード鎖(センス鎖としても知られる)でも、非コード鎖(アンチセンス鎖とも称される)でもよい。

“核酸分子”という用語には、DNAおよびRNAのアナログ(例えば改変骨格を含むもの)、並びにペプチド核酸(PNA)も含まれる。本明細書で用いられる“PNA”という用語はアンチセンス分子または抗遺伝子(anti-gene)作用因子を指し、長さが少なくとも5ヌクレオチドであってアミノ酸残基のペプチド骨格と結合したオリゴヌクレオチドを含む。前記ペプチド骨格は好ましくはリジンで終わり、前記末端リジンは当該組成物に可溶性を付与する。PNAはPEG化(pegylated)されて細胞内での寿命が延長されてもよい{細胞内では、PNAは優先的に相補性一本鎖DNAおよびRNAと結合して転写物伸長を停止させる(P.E. Nielsen et al. (1993) Anticancer Drug Des. 8: 53-63)}。

40

配列番号: 2のポリペプチドまたはその活性フラグメントをコードする核酸分子は、配列番号: 1に示した核酸分子のコード配列と同一でもよい。これらの分子は、遺伝コードの縮退の結果として、配列番号: 2のポリペプチドまたはLBDG1ポリペプチドの活性フラグメント(例えばLBDG1核内ホルモンレセプターリガンド結合ドメイン領域を含むフラグメント)、またはその相同体をコードする多様な配列を有することもできる。LBDG1核内ホルモンレセプターリガンド結合ドメイン領域は、LBDG1ポリペプチド配列の残基822と残基1020の間に広がっていると考えられる。したがって配列番号: 1では、LBDG1核内ホルモ

50

ンレセプターリガンド結合ドメイン領域は、ヌクレオチド2732から3328を含む核酸分子によってコードされる。前記配列の一続きを包含する核酸分子、および前記配列の相同体は、本発明のこの特徴の好ましい態様を形成する。

【0038】

配列番号：4またはその活性フラグメントをコードする核酸分子は、配列番号：3に示した核酸分子のコード配列と同一であろう。これらの分子はまた、遺伝コードの縮退の結果として、配列番号：4またはLBDG4ポリペプチドの活性フラグメント（例えばLBDG4の核内ホルモンレセプターリガンド結合ドメイン領域を含むフラグメント）、またはその相同体をコードする多様な配列を有することができる。LBDG4核内ホルモンレセプターリガンド結合ドメイン領域は、LBDG4ポリペプチド配列の残基805と残基1005の間に広がっていると  
10  
考えられる。配列番号：3では、LBDG4核内ホルモンレセプターリガンド結合ドメイン領域は、したがってヌクレオチド2913から3515を含む核酸分子によってコードされる。前記配列の一続きを包含する核酸分子、および前記配列の相同体は、本発明のこの特徴の好ましい態様を形成する。

配列番号：2または配列番号：4のポリペプチドをコードする前記核酸分子には、それ自体で成熟なポリペプチドのコード配列；成熟ポリペプチドおよび付加コード配列（例えばリーダー配列または分泌配列、例えばプロ-、プリ-またはプレプロ-ポリペプチド配列をコードするもの）のためのコード配列；前述の付加的コード配列を伴う、または伴わないが、さらに付加的な非コード配列（非コード5'および3'配列を含む）を伴う成熟ポリペプチドのコード配列が含まれるが、ただしこれらに限定されない。前記の非コード  
20  
コード5'および3'配列は、例えば転写される非翻訳配列で、転写（終止シグナルを含む）、リボソーム結合およびmRNA安定性で役割を果たすものである。前記核酸分子は、更なる官能性を提供するアミノ酸のような付加アミノ酸をコードする付加配列を含むこともできる。

【0039】

本発明の第二および第三の特徴の核酸分子は、本発明の第一の特徴のポリペプチドのフラグメントまたは機能的等価物およびそのフラグメントもコードし得る。

上記で考察したように、LBDG1ポリペプチドの好ましいフラグメントは、LBDG1核内ホルモンレセプターリガンド結合ドメイン領域を含むフラグメント、またはその相同体である。前記核内ホルモンレセプターリガンド結合ドメイン領域は、配列番号：1のヌクレオチド  
30  
2732から3328を含む核酸分子によってコードされる。

LBDG4ポリペプチドの好ましいフラグメントは、LBDG4核内ホルモンレセプターリガンド結合ドメイン領域を含むフラグメント、またはその相同体である。前記核内ホルモンレセプターリガンド結合ドメイン領域は、配列番号：3のヌクレオチド2913から3515を含む核酸分子によってコードされる。

本発明の機能的に等価な核酸分子は、天然に存在する変種（例えば天然に存在する対立形質変種）であっても、または前記分子は天然に存在することが知られていない変種であってもよい。前記のような天然に存在しない核酸分子の変種は、突然変異誘発技術（核酸分子、細胞または生物に対して適用される技術が含まれる）によって達成できる。

このような変種の中では、特にヌクレオチドの置換、欠失または挿入によって前述の核酸分子と異なる変種が挙げられる。置換、欠失または挿入は1つまたは2つ以上のヌクレオチドを含むことができる。変種はコード領域または非コード領域またはその両方が変化していてもよい。コード領域における変化は、保存的または非保存的なアミノ酸置換、欠失または挿入をもたらし得る。  
40

【0040】

本発明の核酸分子はまた、多様な理由で、遺伝子生成物（ポリペプチド）のクローニング、プロセッシングおよび/または発現の改変を含む当技術分野で一般的に知られている方法を用いて操作され得る。ランダムフラグメント化によるDNAシャッフリングおよび遺伝子フラグメントおよび合成オリゴヌクレオチドのPCRリアッセムブリーは、ヌクレオチド配列の操作に用いられ得る技術に含まれる。位置特異的突然変異誘発を用いて、新規な  
50

制限部位の挿入、グリコシル化パターンの変更、コドンの優先性の変化、スプライシング変種の生成、変異の導入など、その他を行うことができる。

本発明の第一の特徴のポリペプチドをコードする核酸分子は異種配列に連結され、それによって結合核酸分子が融合タンパク質をコードすることができるようにしてもよい。前記のような結合核酸分子は本発明の第二または第三の特徴に包含される。例えば、本発明のポリペプチドの阻害物質のためのペプチドライブラリーをスクリーニングするために、前記のような結合核酸分子を用いて、市販の抗体によって認識される融合タンパク質を発現させることは有用であろう。融合タンパク質はまた、本発明のポリペプチド配列と異種タンパク質配列との間に位置する切断部位を含むように操作し、それによって前記ポリペプチドを異種タンパク質から切り離して精製することができるようにしてもよい。

10

#### 【0041】

本発明の核酸分子にはまた本発明のポリペプチドをコードする核酸分子と部分的に相補的であり、したがってコード核酸分子とハイブリダイズする（ハイブリダイゼーション）アンチセンス分子が含まれる。そのようなアンチセンス分子（例えばオリゴヌクレオチド）は、当業者にはよく知られるように、本発明のポリペプチドをコードする標的核酸を認識し、その標的核酸と特異的に結合してその転写を妨げるようにデザインすることができる（例えば以下の文献を参照されたい：J.S. Cohen, Trends in Pharm. Sci., 10: 435(1989); J. Okano, Neurochem. 56: 560(1991); J. O' Connor, Neurochem. 56: 560(1991); Lee et al., Nucleic Acids Res. 6: 3073(1979); Cooney et al., Science 241: 456(1988); Dervan et al., Science 251: 1360(1991)）。

20

本明細書で用いられる“ハイブリダイゼーション”という用語は、2つの核酸分子が水素結合によって互いに結合することを指す。典型的には、1つの分子が固相支持体に固定され、他方は溶液中で遊離しているであろう。続いて2つの分子を水素結合に適した条件下で互いに接触させる。前記結合に影響する因子には以下が含まれる：溶媒の種類および体積；反応温度；ハイブリダイゼーションの時間；攪拌；液相分子の固相支持体への非特異的結合を妨害する薬剤（デンハルト試薬、またはB L O T T O）；分子の濃度；分子の結合速度を増加させる化合物の使用（硫酸デキストランまたはポリエチレングリコール）；およびハイブリダイゼーションに続く洗浄条件のストリンジェンシー（Sambrook et al.（上掲書）を参照されたい）。

#### 【0042】

完全に相補的な分子と標的分子とのハイブリダイゼーションの阻害は、当業者に知られるハイブリダイゼーションアッセイを用いて調べることができる（例えばSambrook et al.（上掲書）を参照されたい）。したがって、実質的に相同な分子は、文献（G.M. Wahl and S.L. Berger, 1987, Methods Enzymol. 152: 399 - 407; A.R. Kimmel, 1987, Methods Enzymol. 152: 507 - 511）に開示されたように完全に相同な分子と標的分子との結合を種々のストリンジェンシー条件下で競合させ阻害するであろう。

30

“ストリンジェンシー”とは、異なる分子の結合よりも非常に類似した分子の結合に適したハイブリダイゼーション反応の条件を指す。高ストリンジェンシーのハイブリダイゼーション条件は、以下を含む溶液（50%のホルムアミド、5倍のSSC（150mM NaCl、15mM クエン酸三ナトリウム）、50mMリン酸ナトリウム（pH 7.6）、5倍のデンハルト溶液、10%の硫酸デキストラン、および20 $\mu$ g/mLの変性せん断サケ精子DNA）中で42

40

で一晩インキュベーションし、続いてフィルターを0.1倍のSSCで約65 で洗浄すると定義される。低ストリンジェンシー条件は、ハイブリダイゼーション反応が35 で実施されることを含む（Sambrook et al.（上掲書）を参照されたい）。好ましくは、ハイブリダイゼーションに用いられる条件は高ストリンジェンシーを構成するものである。

#### 【0043】

本発明のこの特徴の好ましい態様は、LBDG1ポリペプチド（配列番号：2）またはLBDG4ポリペプチド（配列番号：4）をコードする核酸分子の全長にわたって少なくとも80%同一である核酸分子、および前記のような核酸分子に対して実質的に相補的な核酸分子である。好ましい活性フラグメントは、LBDG1またはLBDG4ポリペプチド配列のそれぞれLBDG1

50

またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域を含むフラグメントである。したがって、好ましい核酸分子は、LBDG1またはLBDG4ポリペプチド配列の核内ホルモンレセプターリガンド結合ドメイン領域をコードする核酸分子の全長にわたって少なくとも80%同一であるものを含む。

本明細書で言及される同一性パーセンテージは、BLASTバージョン2.1.3でNCBI ( the National Center for Biotechnology Information ; <http://www.ncbi.nlm.nih.gov/> ) によって特定されたデフォルトパラメーターを用いて決定されるとおりである。

好ましくは、本発明のこの特徴の核酸分子は、配列番号：1に示された配列を有する核酸分子、この配列のヌクレオチド2732 - 3328を含む領域の全長にわたって少なくとも80%同一の領域を含む。本発明のこの特徴の他の好ましい核酸分子は、配列番号：3に示された配列を有する核酸分子、この配列のヌクレオチド2913 - 3515を含む領域の全長にわたって少なくとも80%同一の領域を含むか、またはこれら核酸領域のいずれかと相補的である核酸分子を含む。この場合、前記の全長にわたって少なくとも90%、好ましくは少なくとも95%、より好ましくは少なくとも98%または99%同一の核酸分子が特に好ましい。これに関して、好ましい態様は、LBDG1またはLBDG4ポリペプチドと実質的に同じ生物学的機能または活性を保持するポリペプチドをコードする核酸分子である。

10

#### 【0044】

本発明はまた、以下の工程を含む、本発明の核酸分子を検出する方法を提供する：(a) 二重鎖を形成するハイブリダイゼーション条件下で本発明の核酸プローブを生物学的サンプルと接触させる工程；および(b) 形成された前記の全ての二重鎖を検出する工程。

20

本発明にしたがって利用することができるアッセイに関連して下記でさらに考察するように、上記で述べた核酸分子をRNA、cDNAまたはゲノムDNAのためのハイブリダイゼーションプローブとして用い、LBDG1またはLBDG4ポリペプチドをコードする完全長のcDNAおよびゲノムクローンを単離し、さらに前記ポリペプチドをコードする遺伝子と高い配列類似性を有する相同遺伝子またはオーソログ遺伝子のcDNAまたはゲノムクローンを単離することができる。

これに関しては、当技術分野で公知の他の技術の中で特に以下の技術を利用することができる。これらの技術は例示として下記で考察される。DNAのシーケンシングおよび解析のための方法は周知で、当技術分野では一般的に利用可能であり、本明細書で考察される本発明の態様の多くを実施するために実際に用いることができる。そのような方法では、DNAポリメラーゼIのKlenowフラグメント、シークエナーゼ(US Biochemical Corp., Cleveland, OH)、Taqポリメラーゼ(Perkin Elmer)、耐熱性T7ポリメラーゼ(Amersham, Chicago, IL)、またはポリメラーゼと校正エキソヌクレアーゼの組合せ(例えば市販(Gibco/BRL, Gaithersburg, MD)のELONGASE増幅キットで見出されるようなもの)のような酵素を利用することができる。好ましくは、シーケンシングプロセスは、例えばハミルトンマイクロラボ(Hamilton Micro Lab) 2200(Hamilton, Reno, NV)、ペルティエサーマルサイクラー(Peltier Thermal Cycler) PTC200(MJ Research, Watertown, MA)、ABIカタリスト並びに373および377DNAシークエンサー(Perkin Elmer)のような機器を用いて自動化することができる。

30

#### 【0045】

LBDG1またはLBDG4ポリペプチドの機能と等価な機能を有するポリペプチド、特にLBDG1またはLBDG4ポリペプチドのLBDG1またはLBDG4核内ホルモンレセプターリガンド結合ドメイン領域と同等の機能をもつポリペプチドをコードする核酸分子を単離する方法の1つは、当技術分野で知られている標準的な方法を用いる、天然のプローブまたは人工的にデザインしたプローブによるゲノムライブラリーまたはcDNAライブラリーの探索である(例えば以下の文献を参照されたい：“Current Protocols in Molecular Biology”, Ausubel et al.(eds). Greene Publishing Association and John Wiley Interscience, New York, 1989, 1992)。特に有用なプローブは、適切なコード遺伝子(配列番号：1または配列番号：3)、特に配列番号：1のヌクレオチド2732 - 3328または配列番号：3のヌクレオチド2913 - 3515の領域に由来する核酸配列に対応する、または前記配列と相補的である、少

40

50

なくとも15、好ましくは少なくとも30、さらに好ましくは少なくとも50の連続塩基を含むプローブである。

前記のようなプローブは分析的な検出が可能な試薬で標識して前記プローブの識別を容易にすることができる。有用な試薬には、放射性同位元素、蛍光色素、および検出可能な生成物の形成を触媒し得る酵素が含まれるが、ただしこれらに限定されない。これらのプローブを用いて、当業者は、ヒト、哺乳類または他の動物供給源から対象のタンパク質をコードするゲノムDNA、cDNAまたはRNAポリヌクレオチドの相補的なコピーを単離し、近縁配列、例えば前記のファミリー、タイプおよび/またはサブタイプに属するまた別のメンバーについて、前記の供給源をスクリーニングすることができるであろう。

#### 【0046】

多くの場合、単離されるcDNA配列は不完全で、ポリペプチドをコードする領域は短く（通常は5'末端で）切断されているであろう。完全長cDNAを得るために、または短いcDNAを伸長させるために、いくつかの方法が利用可能である。そのような配列は、部分的なヌクレオチド配列を用い、上流の配列（例えばプロモーターおよび調節エレメント）を検出するための当技術分野で公知の種々の方法を用いて伸長させることができる。例えば、使用され得るある方法は、cDNA末端迅速増幅法（RACE；例えば以下を参照されたい：Frohman et al., Proc. Natl. Acad. Sci. USA(1988) 85: 8998 - 9002)に基づく。前記技術の最近の改変（例えばマラソン（Marathon）（商標）技術（Clontech Laboratories Inc.）により例示される）は、より長いcDNAの検索を顕著に単純化した。わずかに異なる技術（“制限部位”PCRと称される）では、普遍的プライマーを用いて既知の遺伝子座に近接する未知の核酸配列が検索される（G. Sarkar (1993) PCR Methods Applic. 2: 318 - 322）。逆PCRもまた、既知の領域に基づく多様なプライマーを用いた配列の増幅または伸長に用いることができる（T. Triglia et al. (1988) Nucleic Acids Res. 16: 8186）。使用できる別の方法は捕捉PCRで、前記は、ヒトおよび酵母の人工染色体DNAで既知配列に近接するDNAフラグメントのPCR増幅を含む（M. Lagerstrom et al. (1991) PCR Methods Applic. 1: 111 - 119）。未知の配列を検索するために利用可能な別の方法はパーカーの方法である（J.D. Parker et al. (1991) Nucleic Acids Res. 19: 3055 - 3060）。さらに、ゲノムDNAを歩行させるためにPCR、入れ子（nested）プライマーおよびプロモーターファインダー（PromoterFinder）（商標）ライブラリー（Clontech, Palo Alto, CA）を用いてもよい。この方法ではライブラリーのスクリーニングが不要で、イントロン/エクソン結合部を見出すことに有用である。

#### 【0047】

完全長cDNAをスクリーニングする場合、より大きなcDNAを包含するためにサイズ選択を実施したライブラリーを用いることが好ましい。さらにまた、遺伝子の5'領域を含む配列をより多く含むという点でランダムプライミングした(random-primed)ライブラリーが好ましい。ランダムプライミングしたライブラリーの使用は、オリゴd(T)ライブラリーが完全長cDNAを生成できない状況で特に好ましいであろう。ゲノムライブラリーは、5'非転写調節領域に配列を伸長させるために有用であろう。

本発明のある態様では、染色体局在化のために本発明の核酸分子を用いることができる。この技術では、核酸分子は個々のヒト染色体上の特定の位置に対して特異的に標的化され、個々のヒト染色体上の特定の位置にハイブリダイズさせることができる。本発明の関連配列の染色体上へのマッピングは、遺伝子関連疾患に関する配列の相関性確認において重要な工程である。いったん染色体の正確な位置に配列がマッピングされたら、前記配列の染色体上の物理的な位置を遺伝子地図データと相関させることができる。そのようなデータは、例えば以下で見出すことができる：V. McKusick, Mendelian Inheritance in Man（ジョーンズホプキンス大学、ウェルチ医学図書館を通じてオンラインで入手可能である）。同じ染色体領域にマッピングされた遺伝子と疾患との関係を、次に連鎖解析（物理的に近接する遺伝子の同時遺伝(coinheritance)）によって同定する。これにより、ポジショナルクローニングまたは他の遺伝子発見技術を用いて疾患遺伝子を検索する研究者に貴重な情報が提供される。いったん疾患または症候群の位置が遺伝連鎖によって特定のゲノ

10

20

30

40

50

ム領域で大まかに限局されたら、前記領域にマッピングされるいずれの配列も、更なる解析のための関連または調節遺伝子となることができる。前記核酸分子はまた、正常な個体、キャリア個体または罹患個体間で転座、逆位などによる染色体位置上の相違を検出するために用いることができる。

#### 【0048】

本発明の核酸分子はまた組織分布同定(tissue localisation)のために貴重である。そのような技術は、ポリペプチドをコードするmRNAの検出によって組織中の前記ポリペプチドの発現パターンの決定を可能にする。これらの技術にはin situハイブリダイゼーション技術およびヌクレオチド増幅技術(例えばPCR)が含まれる。これらの研究から得られる結果は、生物内での前記ポリペプチドの正常な機能を示唆する。さらに、変異遺伝子によってコードされるmRNAの発現パターンと正常mRNA発現パターンとの比較研究によって、変異ポリペプチドの疾患における役割に対する貴重な洞察が提供される。そのような不適切な発現は時間的、位置的または量的性質を有する場合もある。

本発明のベクターは本発明の核酸分子を含み、クローニングベクターでも発現ベクターでもよい。本発明のベクターで形質転換、トランスフェクトまたは形質導入され得る本発明の宿主細胞は、原核細胞でも真核細胞でもよい。

本発明のポリペプチドは、宿主細胞内に含まれるベクター中の前記ポリペプチドをコードする核酸分子の発現によってリコンビナント形態で調製することができる。前記のような発現方法は当業者によく知られており、以下の文献により詳細に記述され得る：Sambrook et al. (上掲書) および Fernandez & Hoeffler (1998, eds. "Gene expression systems. Using nature for the art of expression", Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto)。

#### 【0049】

一般的には、要求される宿主でポリペプチドを生成させるために核酸分子の維持、増殖または発現に適したいずれの系またはベクターも用いることができる。周知であり日常的である種々の技術のいずれによっても(例えば前掲書(Sambrook et al.)に記載されたようなもの)、適切なヌクレオチド配列を発現系に挿入することができる。一般的には、コード遺伝子は制御エレメント(例えばプロモーター、リボソーム結合部位(細菌での発現の場合)、および場合によってオペレーター)の制御下に置かれ、それによって所望のポリペプチドをコードするDNA配列を形質転換宿主細胞でRNAに転写させることができる。

適切な発現系の例には、例えば染色体系、エピソーム系およびウイルス由来系、例えば以下に由来するベクターが含まれる：細菌プラスミド、バクテリオファージ、トランスポゾン、酵母エピソーム、挿入エレメント、酵母染色体エレメント、ウイルス、例えばバキュロウイルス、パポバウイルス(例えばSV40)、ワクシニアウイルス、アデノウイルス、鶏痘ウイルス、仮性狂犬病ウイルスおよびレトロウイルス、または上記の組合せ、例えばプラスミドとバクテリオファージの遺伝子エレメントに由来するもの(例えばコスミドおよびファージミドを含む)。ヒト人工染色体(HAC)もまた、プラスミドに包含させ発現させるには大きいDNAフラグメントを搬送するために用いることができる。

#### 【0050】

特に適切な発現系には、リコンビナントバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換された微生物(例えば細菌)；酵母発現ベクターで形質転換された酵母；ウイルス発現ベクター(例えばバキュロウイルス)を感染させた昆虫細胞系；ウイルス発現ベクター(例えばカリフラワーモザイクウイルス、CaMV；タバコモザイクウイルス、TMV)または細菌発現ベクター(例えばTiまたはpBR322プラスミド)で形質転換した植物細胞系；または動物細胞系が含まれる。無細胞翻訳系もまた本発明のポリペプチドの生成に用いることができる。

本発明のポリペプチドをコードする核酸分子の宿主細胞への導入は、多くの標準的な実験室マニュアル(例えば、Davis et al., Basic Methods in Molecular Biology (1986)および上掲書(Sambrook et al.))に記載された方法によって達成できる。特に適切な方法には、リン酸カルシウムトランスフェクション、DEAEデキストラン仲介トランスフ

10

20

30

40

50

エクシオン、トランスベクション(transvection)、マイクロインジェクション、陽イオン脂質仲介トランスフェクション、エレクトロポレーション、トランスダクション、擦過ローディング(scrape loading)、弾道導入または感染が含まれる(以下を参照されたい: Sambrook et al. (1989) 上掲書; Ausbel et al. (1991) 上掲書; Spector, Goldman & L einward, (1998) )。

#### 【0051】

コード核酸分子は、所望であれば、例えばシグナルペプチドまたはリーダー配列のような制御配列をコードする配列を(例えば翻訳ポリペプチドの小胞体内、細胞周辺腔または細胞外環境への分泌のために)含んでいても含んでいなくてもよい。このシグナルは前記ポリペプチドにとって内因性であっても異種シグナルであってもよい。リーダー配列は、翻訳後プロセッシングで細菌宿主によって取り除くことができる。

コントロール配列の他に、宿主細胞の増殖に関連して前記ポリペプチドの発現の調節を可能にする調節配列を付加することが望ましい場合がある。調節配列の例は、化学的または物理的刺激(調節化合物の存在を含む)または多様な温度もしくは代謝条件に应答して遺伝子の発現を増加させたり低下させたりする配列である。調節配列は、ベクターの非翻訳領域、例えばエンハンサー、プロモーター並びに5'および3'非翻訳領域である。これらは、宿主細胞タンパク質と相互作用して、転写および翻訳を実行する。そのような調節配列は、その強さおよび特異性を変化させることができる。用いられるベクター系および宿主に依存して、多くの適切な転写および翻訳エレメント(構成性および誘発性プロモーターを含む)を用いることができる。例えば、細菌系でクローニングするときは、誘発性プロモーター、例えばBluescriptファージミド(Stratagene, La Jolla, CA)またはpSport1(商標)プラスミド(Gibco BRL)などのハイブリッドlacZプロモーターを用いることができる。パキウウイルスポリヘドリン(polyhedrin)プロモーターは昆虫細胞で用いることができる。植物細胞ゲノムに由来するプロモーターまたはエンハンサー(例えば熱ショック、RUBISCOおよび貯蔵タンパク質遺伝子)または植物ウイルスに由来するプロモーターまたはエンハンサー(例えばウイルスプロモーターまたはリーダー配列)は、ベクターへクローニングすることができる。哺乳類細胞系では、哺乳類遺伝子由来または哺乳類ウイルス由来のプロモーターが好ましい。配列の多数コピーを含む細胞株の作製が必要な場合、SV40またはEBVをベースにしたベクターを適切な選択マーカーとともに用いることができる。

#### 【0052】

発現ベクターは、特定の核酸コード配列を適切な調節配列とともにベクター内に配置させることができるように構築される。前記コード配列の調節配列に関する位置および向きは、前記コード配列が調節配列の“制御下”で転写されるような位置および向きである(すなわちコントロール配列にてDNA分子と結合するRNAポリメラーゼは前記コード配列を転写する)。いくつかの事例では、前記配列を適切な向きで制御配列に付属させることができるように(すなわちリーディングフレームを維持するために)、前記配列を改変する必要があるであろう。

コントロール配列および他の調節配列は、ベクターへの挿入の前に核酸コード配列に連結させることができる。あるいは、コード配列は、コントロール配列および適切な制限部位を既に含む発現ベクターへ直接クローニングすることができる。

長期的なりコンビナントポリペプチドの高収量の生成のためには、安定な発現が好ましい。例えば、対象のポリペプチドを安定に発現する細胞株は、ウイルスの複製起点および/または内因性発現エレメント並びに選択マーカー(同じベクターまたは別個のベクターに存在する)を含む発現ベクターを用いて形質転換させることができる。ベクターの導入に続き、選択培地に切り替える前に細胞を栄養(enriched)培地で1-2日間増殖させることができる。選択マーカーの目的は、選択に対する耐性を付与することで、選択マーカーの存在によって、導入された配列をうまく発現する細胞の増殖および回収が可能になる。安定に形質転換された細胞の耐性クローンは、細胞の種類に適した組織培養技術を用いて増殖させることができる。

10

20

30

40

50

## 【0053】

発現のための宿主として利用可能な哺乳類細胞株は当技術分野で公知であり、米国菌培養収集所 (American Type Culture Collection, ATCC) から入手可能な多くの不朽化細胞株が含まれる。前記には、例えばチャイニーズハムスター卵巣細胞 (CHO)、HeLa、ベイビーハムスター腎 (BHK)、サル腎 (COS)、C127、3T3、BHK、HEK293、ボウズ (Bowes) メラノーマおよびヒト肝細胞癌 (例えばHepG2) 細胞および多数の細胞株が含まれるが、これだけに限られない。

バキュロウイルス系では、バキュロウイルス / 昆虫細胞発現系のための材料は、特にインビトロジェン (Invitrogen, San Diego, CA) からキットの形態で ("MaxBac" キット) 商業的に入手可能である。前記の技術は一般的に当業者に知られており、文献には完全に記載されている (Summers & Smith, Texas Agricultural Experiment Station Bulletin No.1555(1987))。この系での使用に特に適切な宿主細胞には昆虫細胞、例えばドロソフィラ (*Drosophila*) S 2 およびスポドプテラ (*Spodoptera*) Sf9細胞が含まれる。

当技術分野で公知である多くの植物細胞培養および植物体 (whole plant) 遺伝子発現系が存在する。適切な植物細胞遺伝子発現系の例には米国特許第5,693,506号、5,659,122号および5,608,143号に記載されたものが含まれる。植物細胞培養における遺伝子発現の別の例は、文献に記載されている (Zenk (1991) *Phytochemistry* 30: 3861 - 3863)。

特に、プロトプラストを単離し、これを培養して完全な再生植物を形成することが可能な植物は全て利用することができ、それによって移入遺伝子を含む完全な植物が回収される。特に、サトウキビ、サトウダイコン、綿花、果実および他の樹木、マメ類および野菜の主要な種の全てを含む (ただしこれらに限定されない) 全ての植物は、培養細胞または組織から再生させることができる。

## 【0054】

特に好ましい細菌宿主細胞の例には連鎖球菌、ブドウ球菌、大腸菌 (*E. coli*)、ストレプトマイセスおよびバチルス・ズブチリス (*Bacillus subtilis*) 細胞が含まれる。

真菌での発現に特に適切な宿主細胞の例には酵母細胞 (例えば *S. cerevisiae*) およびアスペルギルス細胞が含まれる。

形質転換細胞株の回収に用いることができる多くの選択系は、当技術分野で公知である。例えば、単純ヘルペスウイルスチミジンキナーゼ (*M. Wigler et al.*(1977) *Cell* 11: 223 - 32) およびアデニンホスホリボシルトランスフェラーゼ (*I. Lowy et al.*(1980) *Cell* 22: 817 - 23) の遺伝子が含まれ、前記はそれぞれ tk - または *aprt* ± 細胞で用いることができる。

さらにまた、抗代謝物質耐性、抗生物質耐性または除草剤耐性を選択基準として用いてもよい。例えばジヒドロ葉酸レダクターゼ (DHFR) はメトトレキセートに対する耐性を付与し (*M. Wigler et al.*(1980) *Proc. Natl. Acad. Sci.* 77: 3567 - 70)、*npt* はアミノグリコシド系ネオマイシンおよび G - 418 に耐性を付与し (*F. Colbere - Garapin et al.*(1981) *J. Mol. Biol.* 150: 1 - 14)、さらに *als* または *pat* はそれぞれクロロスルフロン (*chlorsulfuron*) およびホスフィントリシン (*phosphinotricin*) アセチルトランスフェラーゼに対する耐性を付与する。さらに別の選択可能な遺伝子が報告されており、当業者にはそれらの例は明白であろう。

## 【0055】

マーカー遺伝子の発現の有無は対象の遺伝子も存在することを示唆するが、対象の遺伝子の存在および発現を確認する必要がある。例えば、関連する配列がマーカー遺伝子配列内に挿入されている場合、適切な配列を含む形質転換細胞はマーカー遺伝子機能が存在しないことによって識別することができる。あるいは、マーカー遺伝子は、ただ1つのプロモーターの制御下で本発明のポリペプチドをコードする配列とともに直列に配置することができる。通常、誘発または選択に応答するマーカー遺伝子の発現は、直列遺伝子の発現も示している。

あるいは、本発明のポリペプチドをコードする核酸配列を含み、前記ポリペプチドを発現する宿主細胞は、当業者に知られている多様な方法で同定することができる。前記方法に

10

20

30

40

50

は、DNA - DNAまたはDNA - RNAハイブリダイゼーションおよびタンパク質バイオアッセイ、例えば蛍光活性化細胞ソーティング (FACS) またはイムノアッセイ技術 (例えば酵素結合免疫吸着アッセイ (ELISA) および放射性イムノアッセイ (RIA) ) が含まれ (ただしこれらに限定されない)、核酸またはタンパク質の検出および/または定量的ためのメンブレン、溶液またはチップをベースとする技術が含まれる (例えば以下を参照されたい: R. Hampton et al. (1990) Serological Methods, A Laboratory Manual, APS Press, St Paul, MN; および D.E. Maddox et al. (1983) J. Exp. Med. 158: 1211 - 1216)。

**【0056】**

多様な標識および結合技術が当業者に知られており、種々の核酸およびアミノ酸アッセイで用いることができる。本発明のポリペプチドをコードする核酸分子に関連する配列を検出するための標識ハイブリダイゼーションプローブまたはPCRプローブの作製手段には、標識したポリヌクレオチドを用いるオリゴ標識、ニックトランスレーション、末端標識またはPCR増幅が含まれる。あるいは、本発明のポリペプチドをコードする配列をベクターにクローニングしてmRNAプローブを作製することができる。そのようなベクターは当技術分野で公知で、商業的に入手可能であり、適切なRNAポリメラーゼ (例えばT7、T3またはSP6) および標識ヌクレオチドを添加することにより *in vitro* でRNAプローブを合成することに用いることができる。これらの方法は、種々の市販キット (Pharmacia & Upjohn (Kalamazoo, MI); Promega (Madison, WI); U.S. Biochemical Corp., (Cleveland, OH)) を用いて実施することができる。

適切なレポーター分子または標識 (前記は検出を容易にするために用いることができる) には、放射性核種、酵素および蛍光、化学発光または色素生産性物質の他に基質、コファクター、阻害剤、磁性粒子などが含まれる。

**【0057】**

本発明の核酸分子はまたトランスジェニック動物 (特にげっ歯類動物) の作製に用いることができる。そのようなトランスジェニック動物は本発明の別の特徴を構成する。これは、体細胞の改変によって局部的に、または遺伝性改変を導入する生殖細胞系療法によって実施することができる。前記のようなトランスジェニック動物は、本発明のポリペプチドのモジュレーターとして有効な薬剤分子のための動物モデルを作製するために特に有用であろう。

ポリペプチドは、周知の方法によってリコンビナント細胞培養物から回収し精製することができる。前記周知の方法には、硫酸またはエタノール沈澱、酸性抽出、陰イオンまたは陽イオン交換クロマトグラフィー、リン酸化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が含まれる。高性能液体クロマトグラフィーは、精製に特に有用である。単離および精製の間ポリペプチドが変性した場合には、タンパク質のリフォールディングのためによく知られている技術を用いて活性な構造を再生することができる。

**【0058】**

所望の場合には、可溶性タンパク質の精製を容易にするポリペプチドドメインをコードするヌクレオチド配列に本発明のポリペプチドをコードする配列を連結させることにより特殊化したベクター構築物も、タンパク質の精製を容易にするために用いることができる。そのような精製促進ドメインの例には、金属キレートペプチド (例えば固定化金属上での精製を可能にするヒスチジン - トリプトファンモジュール、固定化免疫グロブリン上での精製を可能にするプロテインAドメイン、およびFLAG伸長/アフィニティー精製システム (Immunex Corp., Seattle, WA) で用いられるドメイン) が含まれる。切断可能なリンカー配列 (例えばXA因子またはエンテロキナーゼ (Invitrogen, San Diego, CA) に特異的なもの) を精製ドメインと本発明のポリペプチドとの間に包含させて、精製を容易にすることに用いてもよい。そのような発現ベクターの1つは、チオレドキシニンまたはエンテロキナーゼ切断部位に先行するいくつかのヒスチジン残基と融合させた本発明のポリペプチドを含む融合タンパク質の発現を提供する。ヒスチジン残基は、IMAC (固定

10

20

30

40

50

金属イオンアフィニティークロマトグラフィー；J. Porath et al.(1992) Prot. Exp. Purif. 3: 263 - 281) により精製を容易にし、一方、チオレドキシンまたはエンテロキナーゼ切断部位は融合タンパク質からポリペプチドを精製するための手段を提供する。融合タンパク質を含むベクターについての考察は以下で提供される：D.J. Kroll et al.(1993) DNA Cell Biol. 12: 441 - 453)。

#### 【0059】

スクリーニングアッセイで使用するためにポリペプチドを発現させる場合は、前記ポリペプチドを発現する宿主細胞の表面で前記ポリペプチドを生成させることが一般には好ましい。この場合、宿主細胞はスクリーニングアッセイで使用する前に、例えば蛍光活性化細胞ソーティング (FACS) またはイムノアフィニティー技術のような技術を用いて収穫することができる。ポリペプチドが培養液中に分泌される場合は、前記培養液を回収して発現されたポリペプチドを回収および精製することができる。ポリペプチドが細胞内で生成される場合、ポリペプチドを回収する前に、先ず初めに細胞を溶解させねばならない。本発明のポリペプチドを用いて、種々の薬剤スクリーニング技術のいずれかで化合物ライブラリーをスクリーニングすることができる。そのような化合物は、本発明のポリペプチドの遺伝子発現レベルまたは活性レベルを活性化 (作働) させ、または阻害 (拮抗) することができる。それらは本発明のさらなる特徴を形成する。好ましい化合物は、本発明の第一の特徴のポリペプチドをコードする天然の遺伝子の発現を変化させることに有効であるか、または本発明の第一の特徴のポリペプチドの活性を調節することに有効である。アゴニスト化合物またはアンタゴニスト化合物は、例えば細胞、無細胞調製物、化学物質ライブラリーまたは天然生成物の混合物から単離することができる。これらのアゴニストまたはアンタゴニストは天然または改変された基質、リガンド、酵素、レセプターまたは構造的もしくは機能的模倣物質であってよい。前記のスクリーニング技術の適切な概論のためには以下を参照されたい：Coligan et al.(1991) Current Protocols in Immunology 1(2): Chapter 5。

#### 【0060】

おそらく良好なアンタゴニストと考えられる化合物は、本発明のポリペプチドと結合し、結合したときに前記ポリペプチドの生物学的作用を誘発しない分子である。潜在的なアンタゴニストには、本発明のポリペプチドと結合し、それによって本発明のポリペプチドの活性を阻害または消滅させる小型有機分子、ペプチド、ポリペプチドおよび抗体が含まれる。そのようなやり方で、前記ポリペプチドと正常な細胞の結合分子との結合が阻害され、その結果前記ポリペプチドの正常な生物学的活性が阻害され得る。このようなスクリーニング技術で用いられる本発明のポリペプチドは溶液中で遊離していても、固相支持体に固定されていても、細胞表面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。一般に、このようなスクリーニングの方法は、前記のポリペプチドを発現している適切な細胞または細胞膜を用いることを含み、前記細胞または細胞膜をテスト化合物と接触させて、結合または機能的な応答の刺激もしくは阻害を観察する。続いて前記テスト化合物と接触させた細胞の機能的応答を、前記テスト化合物と接触させなかったコントロール細胞と比較する。前記のようなアッセイによって、テスト化合物が前記ポリペプチドの活性化によって発生するシグナルをもたらすか否かを、適切な検出系を用いて評価する。活性化の阻害剤は、一般的には既知のアゴニストの存在下でアッセイされ、テスト化合物の存在下でアゴニストによる活性化の影響が観察される。

#### 【0061】

或いは、単純な結合アッセイを用いてもよい。この場合、テスト化合物のポリペプチド保持表面への付着が直接または間接的にテスト化合物と結合させた標識手段によって検出されるか、または標識競合物質を用いた競合を含むアッセイで検出される。別の態様では、競合薬剤のスクリーニングアッセイを用いることができる。この場合、ポリペプチドと特異的に結合することができる中和抗体がテスト化合物と結合について競合する。このようにして、前記抗体を用いて前記ポリペプチドに対して特異的な結合親和性を保有する一切のテスト化合物の存在を検出することができる。

10

20

30

40

50

前記ポリペプチドをコードするmRNAの細胞内の生成に対する添加テスト化合物の影響を検出するアッセイをデザインすることもできる。例えば、当技術分野で公知の標準的な方法によりモノクローナルまたはポリクローナル抗体を用いてポリペプチドの分泌レベルまたは細胞結合レベルを測定するELISAを構築することができ、それを用いて適切に操作した細胞または組織からのポリペプチド生成を阻害または増強し得る化合物について検索することができる。続いて、前記ポリペプチドとテストされる化合物との結合複合体の生成を測定することができる。

#### 【0062】

使用され得る薬剤スクリーニングのための別の技術は、対象のポリペプチドに対して適切な結合親和性を有する化合物の高速大量処理スクリーニングを提供する（国際特許出願WO 84/03564を参照されたい）。前記方法では、多数の異なる小型のテスト化合物が固相基質上で合成され、次に本発明のポリペプチドと反応させられ洗浄され得る。ポリペプチドを固定する方法の1つは非中和抗体を使用することである。続いて、当技術分野で周知の方法を用いて結合ポリペプチドを検出することができる。精製ポリペプチドはまた、前述の薬剤スクリーニング技術で使用するためにプレートに直接被覆させることができる。本発明のポリペプチドを用いて、膜結合レセプターまたは可溶性レセプターを当技術分野で公知の標準的なレセプター結合技術により同定することができる。前記標準的な技術は、例えばリガンド結合アッセイおよび架橋アッセイであり、そのようなアッセイでは、ポリペプチドは放射性同位体で標識されているか、化学的に改変されているか、またはその検出もしくは精製を容易にするペプチド配列と融合されており、推定上のレセプター供給源（例えば細胞の組成物、細胞膜、細胞上清、組織抽出物または体液）とインキュベートされる。結合有効性は、生物物理的技術、例えば表面プラズモン共鳴および分光法を用いて測定することができる。結合アッセイは、レセプターの精製およびクローニングのために用いることができるが、ポリペプチドとそのレセプターとの結合に競合する前記ポリペプチドのアゴニストおよびアンタゴニストを同定するためにも用いることができる。スクリーニングアッセイを実施する標準的方法は当技術分野ではよく理解されている。

#### 【0063】

本発明はまた、上記で述べたアゴニスト、アンタゴニスト、リガンド、レセプター、基質、酵素を同定する方法で有用なスクリーニングキットを含む。

本発明は、上記アゴニスト、アンタゴニスト、リガンド、レセプター、基質および酵素、並びに上記で述べた方法によって発見される、本発明のポリペプチドの活性または抗原性を調節する他の化合物を含む。

本発明はまた、本発明のポリペプチド、核酸、リガンドまたは化合物とともに適切な医薬担体を含む医薬組成物を提供する。これらの組成物は、下記で詳細に説明するように、治療用もしくは診断用試薬として、ワクチンとして、または他の免疫原性組成物として適切であり得る。

本明細書で用いられる専門用語にしたがえば、ポリペプチド、核酸、リガンドまたは化合物{X}を含む組成物は、組成物中のX+Yの合計の少なくとも85質量%がXである場合は不純物（本明細書中ではY）を“実質的に含まない”。好ましくはXは、組成物中のX+Yの合計の少なくとも約90質量%を、より好ましくは少なくとも約95質量%、98質量%または99質量%を構成する。

#### 【0064】

医薬組成物は、好ましくは治療的に有効な量の本発明のポリペプチド、核酸分子、リガンドまたは化合物を含むべきである。本明細書で用いられる“治療的に有効な量”という用語は、標的疾患または症状を治療、緩和もしくは予防するために、または検出可能な治療効果もしくは予防効果を示すために必要な治療薬剤の量を指す。いずれの化合物についても、治療的に有効な投与量は、最初に細胞培養アッセイ（例えば新生物細胞培養アッセイ）または動物モデル（通常はマウス、ウサギ、イヌまたはブタ）のいずれかで見積もることができる。動物モデルは、適切な濃度範囲および投与経路の決定にも用いることができる。次にそのような情報を用いて、ヒトで有用な投与量および投与経路を決定することが

できる。

ヒト対象者に対する正確な有効量は、疾患状態の重篤度、対象者の全身の健康状態、対象者の年齢、体重および性別、食事、投与時間および投与回数、併用薬剤、反応感受性および治療に対する許容性/応答性に依存するであろう。前記の量は日常的検査により決定でき、それは臨床医の判断の範囲内であろう。一般には有効用量は0.01mg/kgから50mg/kg、好ましくは0.05mg/kgから10mg/kgであろう。本組成物は個別にまたは他の薬剤、医薬品またはホルモンと一緒に患者に投与することができる。

#### 【0065】

医薬組成物はまた、治療薬の投与のために医薬的に許容できる担体を含むことができる。そのような担体には、抗体および他のポリペプチド、遺伝子並びに他の治療薬剤（例えばリポソーム）が含まれるが、ただし担体はそれ自体で前記組成物を投与される個体に有害な抗体の産生を誘発せず、さらに不都合な毒性をもたらすことなく投与されることを条件とする。適切な担体は大型でゆっくりと代謝される巨大分子、例えばタンパク質、多糖類、ポリ乳酸、ポリグリコール酸、重合アミノ酸、アミノ酸コポリマーおよび不活性ウイルス粒子であり得る。

それらの中には医薬的に許容できる塩、例えば鉍酸塩（塩酸塩、臭化水素酸塩、リン酸塩、硫酸塩などのような）；および有機酸の塩（酢酸塩、プロピオン酸塩、マロン酸塩、安息香酸塩などのような）を用いることができる。医薬的に許容できる担体についての綿密な考察は以下のテキストで入手可能である：Remington's Pharmaceutical Sciences (Marck Pub. Co., N.J. 1991)。

治療用組成物中の医薬的に許容できる担体は、さらに液体、例えば水、生理食塩水、グリセロールおよびエタノールを含むことができる。さらに、助剤（例えば湿潤剤または乳化剤、pH緩衝物質など）が前記組成物中に存在していてもよい。そのような担体は、患者が摂取できるように前記医薬組成物を錠剤、ピル、糖衣錠、カプセル、液剤、ゲル、シロップ、スラリー、懸濁剤などとして製剤化することを可能にする。

#### 【0066】

いったん製剤化されたら、本発明の組成物は直接対象者に投与することができる。治療される対象者は動物で、特にヒト対象者が治療され得る。

本発明で用いられる医薬組成物は、多数の経路（経口、静脈内、筋肉内、動脈内、骨髄内、硬膜下腔内、心室内、経皮適用（例えばW098/20734を参照）、皮下、腹腔内、鼻内、腸内、局所、舌下、腔内または直腸的手段が含まれるが、ただしこれらに限定されない）によって投与できる。遺伝子銃またはハイポスプレーもまた、本発明の医薬組成物の投与に用いることができる。典型的には、本治療用組成物は、注射用物質（液体溶液または懸濁剤のいずれか）として調製できる。注射前に液体賦形剤中で溶液または懸濁剤とするために適した固体を調製することもできる。

本組成物の直接的デリバリーは一般に、皮下、腹腔内、静脈内または筋肉内に注射によって達成されるか、または組織の間隙腔に搬送されるであろう。前記組成物はまた、病巣に投与してもよい。投薬治療は単回投与スケジュールでも複数回投与スケジュールでもよい。

本発明のポリペプチドの活性が特定の疾患状態で過剰である場合には、いくつかのアプローチが利用可能である。あるアプローチは、医薬的に許容できる担体とともに上記のような阻害化合物（アンタゴニスト）を、前記ポリペプチドの機能阻害に有効な量で対象者に投与することを含む（前記化合物による前記機能阻害は、例えばリガンド、基質、酵素、レセプターの結合を遮断するか、または第二のシグナルを阻害し、それによって異常な症状を緩和することによって達成される）。好ましくは、前記アンタゴニストは抗体である。最も好ましくは、そのような抗体は、先に記載したようなその免疫原性を最少にするキメラ抗体および/またはヒト化抗体である。

#### 【0067】

別のアプローチでは、問題のリガンド、基質、酵素、レセプターに対する結合親和性を保持するポリペプチドの可溶形を投与することができる。典型的には、前記ポリペプチドは

、関連部分を保持するフラグメントの形態で投与することができる。

また別のアプローチでは、前記ポリペプチドをコードする遺伝子の発現は、発現遮断技術を用いて、例えば内部で生成されるかまたは別々に投与されるアンチセンス核酸分子を使用して（上記で述べたように）阻害することができる。遺伝子発現の改変は、ポリペプチドをコードする遺伝子の制御領域、5'領域または調節領域（シグナル配列、プロモーター、エンハンサーおよびイントロン）に対して相補的な配列またはアンチセンス分子（DNA、RNAまたはPNA）をデザインすることによって達成できる。同様に、阻害は“三重らせん”塩基対方法論を用いて達成することができる。三重らせん対形成は、ポリメラーゼ、転写因子または調節分子の結合のために二重らせんが充分に開く能力を阻害することから有用である。三重DNAを用いる最近の治療の進歩は文献に記載されている（J.E. Gee et al. (1994) In: B.E. Huber & B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY）。相補的配列またはアンチセンス分子をデザインし、リボソームに対する結合を妨げて転写を妨害することによってmRNAの翻訳を遮断することもできる。そのようなオリゴヌクレオチドは投与されてもよいし、またin vivoでの発現によりin situで生成させてもよい。

#### 【0068】

さらに、本発明のポリペプチドの発現は、そのコードmRNA配列に特異的なリボザイムを用いることによって妨げることができる。リボザイムは、天然または合成であり得る触媒的活性型のRNAである（例えば以下を参照されたい：N. Usman et al., Curr. Opin. Struct. Biol. (1996) 6(4) : 527 - 533）。合成リボザイムをデザインして、選択した位置でmRNAを特異的に切断し、それによってmRNAの機能的ポリペプチドへの翻訳を妨げることができる。リボザイムは、通常RNA分子で見出されるような天然のリン酸リボース骨格および天然の塩基から合成することができる。或いは、リボザイムは、非天然の骨格（例えば2'-0-メチルRNA）を用いて合成して、リボヌクレアーゼ分解から保護することができる。これらは改変塩基を含んでいてもよい。

RNA分子は細胞内安定性および半減期を増加させるために改変することができる。可能な改変には、RNA分子の5'および/または3'末端へのフランキング配列の付加、または分子の骨格内でホスホジエステル結合に代わるホスホロチオエートまたは2'-0-メチルの使用が含まれるが、ただしこれらに限定されない。この概念はPNAの生成にも受け継がれ、さらに非慣用塩基（例えばイノシン、ケオシン (queosine) およびブトシン (butosine)（アセチル-、メチル-、チオ-も同様に）、および同様に改変された形態のアデニン、シチジン、グアニン、チミンおよびウリジンで、これらは内因性エンドヌクレアーゼによって容易に認識されないようなものである）の包含によってPNA分子の全てに前記概念を広げることができる。

#### 【0069】

本発明のポリペプチドおよびその活性の過小発現に関連する異常な状態を治療するためには、いくつかのアプローチも利用可能である。あるアプローチは、前記ポリペプチドを活性化する化合物（すなわち上記で述べたアゴニスト）の治療的に有効な量を対象者に投与し、異常な状態を緩和することを含む。あるいは、本ポリペプチドの治療量を適切な医薬担体と組合せて投与し、ポリペプチドの相対的な生理学的バランスを回復させてもよい。遺伝子治療を用い、対象者の関連する細胞による本ポリペプチドの内因性生成を行わせることができる。遺伝子治療は、不完全な遺伝子を修正した治療用遺伝子と置き換えることによって、前記ポリペプチドの不適切な生成を永久的に治療するために用いられる。

本発明の遺伝子治療はin vivoまたはex vivoで実施できる。Ex vivo遺伝子治療は、患者の細胞の単離および精製、治療用遺伝子の導入、および遺伝的に改変した細胞を患者に戻して導入することを必要とする。対照的に、in vivo遺伝子治療は、患者の細胞の単離および精製を必要としない。

治療用遺伝子は患者に投与するために典型的には“封入されて”いる。遺伝子デリバリー賦形剤はリボソームのような非ウイルス、または、例えばK.L. Berkner (1992) Curr. Top. Microbiol. Immunol., 158 : 39 - 66に記載されたアデノウイルスのような複製欠損ウ

10

20

30

40

50

イルスまたはN. Muzyczka (1992) Curr. Top. Microbiol. Immunol., 158: 97 - 129および米国特許第5,252,479号に記載されたアデノ付随ウイルス(AAV)ベクターであり得る。例えば、本発明のポリペプチドをコードする核酸分子は、複製欠損レトロウイルスベクターで発現させるために操作を施すことができる。次にこの発現構築物を単離し、前記ポリペプチドをコードするRNAを含有するレトロウイルスプラスミドベクターで形質導入したパッケージ細胞に導入することができる。その結果、前記パッケージ細胞は対象の遺伝子を含有する感染性ウイルス粒子を産生することができるようになる。これらのプロデューサー細胞はin vivoで細胞を操作するために、さらにin vivoでポリペプチドを発現させるために対象者に投与することができる(以下を参照されたい: Gene Therapy and Other Molecular Genetic-based Therapeutic Approaches, Chapter 20(およびその中に引用された文献), "Human Molecular Genetics" (1996) T. Strachan & A.P. Read, BIOS Scientific Publishers Ltd.)。 10

#### 【0070】

別のアプローチは“裸のDNA”の投与で、この場合、治療用遺伝子は血流または筋肉組織に直接注射される。

本発明のポリペプチドまたは核酸分子が疾患をひき起こす原因物質である場合には、本発明は、前記疾患をひき起こす原因物質に対する抗体を生成するワクチンとして用いることができる。

本発明のワクチンは予防的(すなわち感染を防ぐ)または治療的(すなわち感染後の疾患を治療する)であろう。そのようなワクチンは、免疫性を付与する抗原、免疫原、ポリペプチド、タンパク質または核酸を、通常は上記で述べた医薬的に許容できる担体と組合せて含む。前記担体には、それ自体で組成物を投与される個体に対して有害な抗体の産生を誘発しない担体のいずれもが含まれる。さらに、これらの担体は免疫刺激剤(“アジュバント”)として機能してもよい。さらにまた、前記抗原または免疫原は、細菌の類毒素(例えばジフテリア、破傷風、コレラ、H.ピロリ菌(pyroli))および他の病原体と結合させることができる。 20

ポリペプチドは胃で分解されるので、ポリペプチドを含むワクチンは好ましくは非経口的に(例えば皮下、筋肉内、静脈内または皮内注射)投与される。非経口投与に適した製剤には、水性および非水性滅菌注射溶液(前記は抗酸化剤、緩衝剤、抗菌剤および製剤を受容者の血液に対し等張にする溶質を含むことができる)、並びに水性および非水性滅菌懸濁剤(前記は分散剤または粘稠剤を含むことができる)が含まれる。 30

#### 【0071】

本発明のワクチン製剤は単位用量または複数単位用量容器で提供されてもよい。例えば封入したアンプルおよびバイアルは、使用直前に滅菌された液状担体を添加することだけを必要とする凍結乾燥状態で保存することができる。投与量はワクチンの比活性に依存し、日常的な検査で容易に決定することができる。

本発明はまた、診断薬としての本発明の核酸分子の使用に関する。本発明の核酸分子により特徴付けられ、機能不全に付随する遺伝子の変異型の検出は、前記遺伝子の過小発現、過剰発現または位置的もしくは時間的発現の変化から生じる疾患の診断、または疾患に対する感受性の診断に付け加えることができるか、またはそれらを明確にすることができる診断ツールを提供する。前記遺伝子に変異を保有する個体は、種々の技術によってDNAレベルで検出することができる。 40

診断のための核酸分子は対象者の細胞、例えば血液、尿、唾液、組織生検または剖検材料から入手できる。ゲノムDNAを直接検出に用いてもよいし、PCR、リガーゼ連鎖反応(LCR)、鎖置換増幅(SDA)または他の増幅技術を分析前に用いることによって、ゲノムDNAを酵素的に増幅してもよい(以下の文献を参照されたい: Saiki et al., Nature 324: 163 - 166(1986); Bej et al., Crit. Rev. Biochem. Molec. Biol., 26: 301 - 334(1991); Birkenmeyer et al., J. Virol. Meth., 35: 117 - 126(1991); Van Brunt, J., Bio/Technology, 8: 291 - 294(1990))。

#### 【0072】

ある態様では、本発明のこの特徴は、本発明のポリペプチドをコードする天然の遺伝子の発現レベルを評価し、さらに前記発現レベルをコントロールのレベルと比較することを含む、患者における疾患を診断する方法を提供する（ここで前記コントロールレベルと異なるレベルは疾患を示唆する）。前記方法は以下の工程を含む：

- a) 本発明の核酸分子と核酸プローブとの間でハイブリッド複合体の形成を可能にするストリンジェント条件下で、患者由来の組織サンプルを前記核酸プローブと接触させる工程；
- b) コントロールサンプルを工程 a) で用いた条件と同じ条件下で前記プローブと接触させる工程；および、
- c) 前記サンプル中のハイブリッド複合体の存在を検出する工程；

この場合、コントロールサンプル中のハイブリッド複合体レベルと異なるハイブリッド複合体レベルが患者サンプルで検出されることは、疾患を示唆する。

本発明のさらなる特徴は、以下の工程を含む診断方法を含む：

- a) 疾患について検査される患者から組織サンプルを入手する工程；
- b) 前記組織サンプルから本発明の核酸分子を単離する工程；および、
- c) 前記核酸分子内の疾患に付随する変異の存在を検出することによって患者を疾患について診断する工程。

#### 【0073】

上記に記載した方法における核酸分子の検出を補助するために、増幅工程、例えばPCRの使用を含むことができる。

正常な遺伝子型と比較すると、増幅生成物におけるサイズの変化によって、欠失および挿入が検出される。点変異は、増幅DNAを本発明の標識RNAとハイブリダイズさせるか、あるいは本発明の標識アンチセンスDNA配列とハイブリダイズさせることによって同定することができる。完全にマッチした配列は、RNase消化によって、または溶融温度における差異を評価することによって、ミスマッチを有する二重鎖と区別することができる。DNAをストリンジェントな条件下で前記DNAとハイブリダイズする核酸プローブと接触させてハイブリッド二重鎖分子を形成させ（前記ハイブリッド二重鎖は、疾患に付随する変異に対応するいずれかの部分で前記核酸プローブ鎖のハイブリダイズしていない部分を有する）、DNA鎖の対応する部分における疾患付随変異の有無を示すものとして、前記プローブ鎖のハイブリダイズしていない部分の有無を検出することによって、患者における変異の有無を検出することができる。

前記のような診断は特に出生前検査で有用であり、新生児検査においてすら有用である。

#### 【0074】

点変異、および参照遺伝子と“変異”遺伝子間で異なる他の配列は、他の周知の技術、例えば直接DNAシーケンシングまたは一本鎖構造多型性（Orita et al., Genomics, 5: 874 - 879 (1989)）によって同定できる。例えば、シーケンシングプライマーは、二本鎖PCR生成物または改変PCRによって作製された一本鎖テンプレート分子とともに用いることができる。配列決定は、放射能標識ヌクレオチドを用いる通常の方法によって、または蛍光タグを用いる自動シーケンシング法によって実施される。クローン化DNAセグメントを、特異的DNAセグメントを検出するためのプローブとして用いることもできる。この方法の感受性は、PCRと併用したとき極めて増強される。さらに、点変異および他の配列の変動（例えば多型性）は、例えば対立遺伝子特異的オリゴヌクレオチドをただ1つのヌクレオチドが異なる配列のPCR増幅に用いることによって、上記のように検出することができる。

DNA配列の相違はまた、変性剤の存在下または非存在下でのゲル内のDNAフラグメントの電気泳動移動度における変化によって、または直接DNAシーケンシング（例えば、Myers et al., Science (1985) 230: 1242）によって検出することができる。特定の位置における配列の変化はまた、ヌクレアーゼ保護アッセイ（例えばRNaseおよびS1保護）または化学切断法（Cotton et al., Proc. Natl. Acad. Sci. USA (1985) 85: 4397 - 4401を参照）によって明らかにすることができる。

10

20

30

40

50

## 【0075】

通常のゲル電気泳動およびDNAシーケンシングの他に、ミクロ欠損、異数性、転座、逆位のような変異は、*in situ*分析によっても検出できる（例えば以下を参照されたい：Keller et al., DNA Probes, 2nd Ed., Stockton Press, New York, N.Y., USA(1993)）。すなわち、細胞内のDNAまたはRNA配列は、それらを単離および/またはメンブレン上に固定する必要なしに変異について分析することができる。蛍光*in situ*ハイブリダイゼーション（FISH）は、現在のところ最も一般的に用いられている方法で、FISHに関する多数の概論が存在する（例えば以下を参照されたい：Trachuck et al., Science, 250: 559 - 562(1990)；およびTrask et al., Trends Genet. 7: 149 - 154(1991)）。

本発明の別の態様では、本発明の核酸分子を含むオリゴヌクレオチドプローブのアレイを構築し、遺伝的変種、変異および多型性の効率的スクリーニングを実施することができる。アレイ技術方法はよく知られており汎用的に応用することができ、遺伝子発現、遺伝連鎖および遺伝的多様性を含む分子遺伝学における種々の疑問に取り組むために用いることができる（例えば以下を参照されたい：M. Chee et al., Science (1996) 274: 610 - 613）。

10

## 【0076】

ある態様では、前記アレイは以下の文献に記載された方法にしたがって調製され使用される（PCT出願W095/11995（Chee et al.）；D.J. Lockhart et al.(1996) Nat. Biotech. 14: 1675 - 1680；M. Schena et al.(1996) Proc. Natl. Acad. Sci. 93: 10614 - 10619）。オリゴヌクレオチド対は2つから100万個を越える範囲で変動し得る。前記オリゴマーは、光誘導化学方法を用いて基板上の指定領域で合成される。基板は紙、ナイロンまたは他のタイプのメンブレン、フィルター、チップ、ガラススライドもしくは他の適切な任意の固相支持体でよい。別の特徴では、オリゴヌクレオチドは、PCT特許出願（W095/251116, Baldeschweiler et al.）に記載されたように化学的結合方法およびインクジェット適用装置を用いて基板表面で合成することができる。別の特徴では、ドット（またはスロット）プロットに類似する“格子化（gridded）”アレイを用い、真空系、熱結合方法、UV結合方法、機械的または化学的結合方法を用いて基質表面にcDNAフラグメントまたはオリゴヌクレオチドを配置し、連結させることができる。アレイ（例えば上記で述べたようなもの）は手動で、または利用可能な装置（スロットプロットまたはドットプロット装置）、材料（適切な固相支持体すべて）および機械（ロボット機器を含む）を用いて作製することができ、8、24、96、384、1536または6144個のオリゴヌクレオチド、または2つから100万個を越える範囲の他のいずれの数をも含むことができる（このことはアレイ自体を商業的に入手可能な計測器の有効利用に向くものとしている）。

20

30

## 【0077】

上記で考察した方法の他に、対象者に由来するサンプルから、ポリペプチドまたはmRNAの異常な増加または低下レベルを決定することを含む方法によって、疾患を診断することができる。発現低下または発現増加は、ポリヌクレオチドの定量のために当技術分野で周知の方法のいずれか、例えば核酸増幅（例を挙げるとPCR、RT-PCR、RNase保護、ノザンプロット法）および他のハイブリダイゼーション方法を用いてRNAレベルで測定することができる。

40

宿主に由来するサンプルで本発明のポリペプチドレベルを決定することに用いることができるアッセイ技術は当業者によく知られており、さらに上記でいくらか詳細に考察されている（ラジオイムノアッセイ、競合結合アッセイ、ウェスタンプロット分析およびELISAアッセイを含む）。本発明のこの特徴では以下の工程を含む診断方法が提供される：（a）上記に記載したリガンドを、リガンド-ポリペプチド複合体の形成に適した条件下で生物学的サンプルと接触させる工程；および（b）前記複合体を検出する工程。

例えばELISA、RIAおよびFACSのようなポリペプチドレベルを測定するためのプロトコルは、さらにポリペプチド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ことができる。ポリペプチド発現の正常値または標準値は、正常な哺乳類対象体（好ま

50

しくはヒト)から得られた体液または細胞抽出物を、複合体形成に適した条件下で前記ポリペプチドに対する抗体と混合することによって確立される。標準的な複合体形成量は種々の方法、例えば分光測定方法によって定量することができる。

#### 【0078】

本発明のポリペプチドと特異的に結合する抗体は、前記ポリペプチドの発現によって特徴付けられる症状または疾患の診断のために、または本発明のポリペプチド、核酸分子、リガンドおよび他の化合物を用いて治療されている患者をモニターするアッセイにおいて、用いることができる。診断目的に有用な抗体は、治療薬として上記で述べたのと同じ様式で調製することができる。前記ポリペプチドについての診断アッセイは、前記抗体および標識を用いてヒトの体液または細胞もしくは組織の抽出物中のポリペプチドを検出する方法を含む。前記抗体は改変して、または改変せずに用いることができ、さらにそれらをレポーター分子と共有結合または非共有結合によって結合させることによって標識することができる。当技術分野で公知の多様なレポーター分子を用いることができる(それらのいくつかは上記に記載されている)。

生検組織由来の対象者、コントロールおよび疾患サンプルで発現されたポリペプチドの量は、標準値と比較される。標準値と対象者の値との間の偏差は疾患診断のためのパラメータを確立する。診断アッセイを用いて、ポリペプチド発現の有無および過剰を識別し、治療的処置の間のポリペプチドレベルの調節をモニターすることができる。そのようなアッセイはまた、動物実験、臨床試験または個々の患者の治療モニタリングにおける個々の治療的処置方法の有効性を評価することに用いることができる。

#### 【0079】

本発明の診断キットは以下を含むことができる：

- (a) 本発明の核酸分子；
- (b) 本発明のポリペプチド；または
- (c) 本発明のリガンド。

本発明のある特徴では、診断キットは、ストリンジェントな条件下で本発明の核酸分子とハイブリダイズする核酸プローブを含む第一の容器；核酸分子を増幅させるために有用なプライマーを含む第二の容器；および疾患の診断を容易にするために前記プローブおよびプライマーの使用についての指示書を含むことができる。前記キットはさらに、ハイブリダイズしていないRNAを消化するための試薬を保持する第三の容器を含むことができる。

本発明の別の特徴では、診断キットは核酸分子のアレイを含み、前記核酸分子の1つが本発明の核酸分子であってもよい。

本発明のポリペプチドを検出するために、診断キットは以下を含むことができる：本発明のポリペプチドと結合する1つまたは2つ以上の抗体；および前記抗体とポリペプチドとの間の結合反応の検出に有用な試薬。

#### 【0080】

そのようなキットは、以下のような疾患または疾患に対する感受性を診断することに有用であろう：特に細胞増殖疾患(新生物、メラノーマ、肺、結腸直腸、乳房、膵、頭部および頸部の腫瘍並びに他の固形腫瘍を含む)、骨髄増殖性疾患(例えば白血病、非ホジキンリンパ腫、白血球減少症、血小板減少症、血管形成疾患、カポジ肉腫)、自己免疫/炎症性疾患(アレルギー、炎症性腸疾患、関節炎、乾癬および気道の炎症、喘息および器官の移植拒絶を含む)、心脈管系疾患(高血圧、浮腫、アングナ、アテローム性動脈硬化症、血栓症、敗血症、ショック、再灌流障害、心不整脈および虚血を含む)、神経学的疾患(中枢神経系疾患、アルツハイマー病、脳損傷、脳卒中、筋萎縮性側索硬化症、不安、抑うつおよび痛みを含む)、発達障害、代謝性障害(真性糖尿病、骨粗しょう症、脂質代謝障害、甲状腺機能亢進、上皮小体機能亢進、高カルシウム血症、高コレステロール血症、高脂血症および肥満を含む)、腎疾患(糸球体腎炎、腎血管性高血圧を含む)、皮膚疾患(アクネ、湿疹および創傷治癒を含む)、加齢の負の作用、エイズ、感染(ウイルス感染、細菌感染、真菌感染および寄生虫感染を含む)、並びに他の病的状態(特に核内ホルモンレセプターが関与するもの)。

10

20

30

40

50

本発明の種々の特徴および態様は、特にLBDG1ポリペプチドへの言及を有する実施例を介してこれからより詳細に説明されるであろう。

細部の改変は本発明の範囲を逸脱することなく実施できることは理解されよう。

#### 【0081】

##### 実施例

##### 実施例 1

遠縁の新規な核内ホルモンレセプターリガンド結合ドメインの検索を開始するために、原型となるファミリーメンバー、ホモサピエンスエストロゲンレセプターリガンド結合ドメインを選択する。より具体的には、前記検索は、タンパク質データバンク(PDB)(Research Collaboratory for Structural Bioinformaticsによって運用されている)から得られた構造を用いて開始する。

10

選択した構造は、ホモサピエンスエストロゲンレセプターのリガンド結合ドメインである(PDBコード1ERR:A; 図1参照)。1ERR:Aの類縁体についてバイオペンジウムの検索を実施し(ターゲットマイニングインターフェース(Target Mining Interface)を用いる)、3614のゲノムスレッダーの結果が返される。この3614のゲノムスレッダーの結果には、典型的な核内ホルモンレセプターリガンド結合ドメイン(例えばホモサピエンスエストロゲンレセプターの残基307-547の間で見出されるもの)(図2Aの矢印参照)の例が含まれる。

核内ホルモンレセプターリガンド結合ドメインを含むことが判明しているタンパク質の中で、核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けされていないタンパク質、075385(LBDG1; 図2Bの矢印参照)が現れる。インファーマチカゲノムスレッダーは、配列075385(LBDG1)の領域(残基822-1020の間)をホモサピエンスエストロゲンレセプターのリガンド結合ドメインと類似の構造を有すると同定した。ホモサピエンスエストロゲンレセプターのリガンド結合ドメインと類似する構造を保有することは、075385(LBDG1)の残基822-1020は核内ホルモンレセプターリガンド結合ドメインとして機能することを示唆する。ゲノムスレッダーはこれを84%の信頼度で同定する。

20

#### 【0082】

1ERR:Aの類縁体についてのバイオペンジウムの検索(ターゲットマイニングインターフェースを用いる)によって、841のフォワード(forward)PSI-BLASTの結果も返される。フォワードPSI-BLAST(図2Cを参照されたい)は前記関係を同定できず、インファーマチカゲノムスレッダーだけが075385(LBDG1)は核内ホルモンレセプターリガンド結合ドメインを含むと同定できる。

30

075385(LBDG1)についてパブリックドメインデータベースで何が判明しているかを評価するために、重複配列ディスプレイページ(図3)を調査する。075385(LBDG1)が核内ホルモンレセプターリガンド結合ドメインを含むと同定するPROSITEヒットもPRINTSヒットも得られない。PROSITEおよびPRINTSは、類似ファミリーのタンパク質を記述することに役立つデータベースである。両データベースから核内ホルモンレセプターリガンド結合ドメインのヒットが返ってこないことは、PROSITEまたはPRINTSを用いた場合、075385(LBDG1)が核内ホルモンレセプターリガンド結合ドメインを含むものとして同定できないことを意味する。

40

#### 【0083】

他のいずれかのパブリックドメイン注釈付け手段が、075385(LBDG1)は核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けすることができるか否かを確認するために、075385(LBDG1)タンパク質配列をインタープロ(InterPro)ウェブサイト上でPFAMデータベース(アラインメントおよび隠れマルコフモデルのタンパク質ファミリーデータベース)に対して検索する(図4A矢印1参照)。1つのPFAM-AマッチングがPF00069/IPR000719に対して見出され、それは真核細胞タンパク質キナーゼに特徴的である。このタンパク質キナーゼドメインは、075385(LBDG1)の残基16-278の間に位置すると注釈付けされる。075385(LBDG1)が核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けするPFAM-Aマッチングは存在しない。したがってPFAMでは、075385(LBDG1)が核内ホ

50

ルモンレセプターリガンド結合ドメインを含むことを同定しない。

興味深いことには、PROSITEのPFスキャン(図4 A 矢印2 参照)によって、075385(LBDG1)の残基501-688でプロリンに富む領域が同定される。プロリンに富む領域はDNA結合ドメインとして機能することができる(図4 B 参照)。DNA結合ドメインの位置を予測核内ホルモンレセプターリガンド結合ドメイン(残基822-1020)のN-末端に位置づけることは、原型の核内ホルモンレセプターのドメイン機構を想起させる(原型核内ホルモンレセプターはN-末端ジンクフィンガー-DNA結合ドメインおよびC-末端リガンド結合ドメインを有する)。

#### 【0084】

続いてスイスバイオインフォマティクス研究所、SIB(Swiss Institute of Bioinformatics)のSWISS-PROTタンパク質データベースを、075385(LBDG1)配列に関連するパブリックドメインでさらなる既知情報があるか否かを調べた。これは、タンパク質および遺伝子配列寄託に関するスイスのパブリックドメインデータベースである(図5)。075385(LBDG1)はホモサピエンスの配列で、そのSWISS-PROTタンパク質IDは075385であり、長さは1050アミノ酸である。075385はULK1(UNC51様キナーゼ1)と呼ばれる。ULK1はKuroyanagiら(東京都老人医学研究所、分子遺伝学部門;東京都板橋区栄町2-35(173-0015))によってクローン化された。この遺伝子について前記パブリックドメインの情報は、前記遺伝子を核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けしていない。

したがって、全てのパブリックドメインの注釈付けツールを用いて、075385(LBDG1)は核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けされないであろうと結論することができる。インファーマチカゲノムスレッダーのみが前記タンパク質を核内ホルモンレセプターリガンド結合ドメインを含むと注釈付けすることができる。

#### 【0085】

ここで逆検索を実施する。075385(LBDG1)を今度はクエリー配列としてバイオペンジウムで用いる(図6 A 参照)。インファーマチカゲノムスレッダーは、075385(LBDG1)の残基822-1020はホモサピエンスエストロゲンレセプターのリガンド結合ドメインと同じ構造を有すると信頼度84%で同定する(図6 B の矢印参照)。ホモサピエンスエストロゲンレセプターのリガンド結合ドメイン(1ERR:A)は、元々のクエリー配列であった。PSI-BLASTの正の繰返しはこの結果を返さない(図6 C)。この関係を同定することができるのはインファーマチカゲノムスレッダーだけである。

ホモサピエンスエストロゲンレセプターリガンド結合ドメインの配列アラインメントを観察するために前記に対して選択される。類似の構造を有すると確認されたタンパク質に対するクエリータンパク質のAIEyeアラインメント(図7)を観察することは、相同領域の可視化に役立つ。

ホモサピエンスエストロゲンレセプターリガンド結合ドメインは“LBDモチーフ”を含み、前記モチーフは、今日まで注釈付けされた全ての核内ホルモンレセプターリガンド結合ドメインで見出されている。“LBDモチーフ”は核内ホルモンレセプターのコアクチベーターおよびコリプレッサーを補充することに関与する。6つの残基;PHE367、LEU370、ASP374、GLN375、LEU378およびLEU379は、ホモサピエンスエストロゲンレセプターリガンド結合ドメインにおいてこのモチーフを構成する(図7の四角枠を参照)。075385(LBDG1)中の4残基(ASP885、GLN886、LEU889およびLEU890)は、ホモサピエンスエストロゲンレセプターリガンド結合ドメイン中の6つの“LBDモチーフ”残基のうちの4つ(ASP374、GLN375、LEU378およびLEU379)と正確にマッチする。075385(LBDG1)の5番目の残基(LEU878)は、ホモサピエンスエストロゲンレセプターリガンド結合ドメインの“LBDモチーフ”ではPHE367に保存的置換されている。このことは、075385(LBDG1)はホモサピエンスエストロゲンレセプターリガンド結合ドメインと類似する核内ホルモンレセプターリガンド結合ドメインを含むことを示している。

#### 【0086】

同定されたタンパク質が実際にクエリータンパク質と関連することを保証するために、可

10

20

30

40

50

視化プログラム“LigEye”(図8A)および“iRasmol”(図8B)を用いる。これらの可視化ツールは、既知の小型阻害分子が既知タンパク質構造の活性部位で相互作用するアミノ酸を表示することによって、前記既知タンパク質構造の活性部位を同定する。これらの相互作用は、直接水素結合を介するか、または疎水性相互作用を介する。このようにして、活性部位のフォールド/構造が、同定された相同体と選択した既知構造のタンパク質との間で保存されているか否かを理解することができる。ホモサピエンスエストロゲンレセプターリガンド結合ドメインのLigEyeによる調査は、ラロキシフェン(Raloxifene)と結合する15残基(図7の丸付き文字)を明らかにする。しかしながら、これら15残基のうち11残基のみ(THR347、ALA350、LEU354、TRP383、LEU387、MET388、LEU391、ARG394、PHE404、MET421およびLEU428)が、ゲノムスレッダーアラインメントの範囲内に存在する。したがって、これら11残基のみが、核内ホルモンレセプターリガンド結合ドメインを含むという075385(LBDG1)のゲノムスレッダーの注釈付けを強固にするために用いることができる。これら11残基の中に9個の疎水性残基が存在し、これらはホモサピエンスエストロゲンレセプターリガンド結合ドメインのポケットの輪郭を形成する; ALA350、LEU354、TRP383、LEU387、MET388、LEU391、PHE404、MET421およびLEU428。これら9つのホモサピエンスエストロゲンレセプターリガンド結合ドメインの疎水性残基のうち5つは、075385(LBDG1)で完全に保存されている: ALA861、LEU863、TRP894、LEU900およびLEU904(図7の丸付き文字)。残りの4つのホモサピエンスエストロゲンレセプターリガンド結合ドメインの疎水性残基のうち3つは、075385(LBDG1)では別の疎水性残基に保存的に置換されている: VAL901、LEU910およびVAL932(図7の穴あき丸付き文字)。結合ポケットの輪郭を形成する9つの疎水性残基のうち(ゲノムスレッダーアラインメント領域内で)8つで疎水性が保存されていることは、075385(LBDG1)は疎水性ステロイド様リガンドと結合するであろうことを示している。

10

20

30

40

50

#### 【0087】

このことは、実際インファーマチカゲノムスレッダーによって予測されたように、075385(LBDG1)はホモサピエンスエストロゲンレセプターリガンド結合ドメインと類似の様式で折り畳まれることを示し、したがって核内ホルモンレセプターリガンド結合ドメインを含むと同定される。

075385(LBDG1)の逆最大化PSI-BLASTによって075385(LBDG1)の1つのホモサピエンス相同体および2つのマス・マスキュラス(*Mus musculus*)相同体が同定される。070405(マス・マスキュラスのULK1、図6C矢印1参照)は075385(LBDG1)に最も近縁な配列で、88%の配列同一性を共有する。したがって070405(マス・マスキュラスのULK1)は、075385(LBDG1; ホモサピエンスULK1)のマス・マスキュラスオーソログ(orthologue)を表す。

BAA31598.1(ホモサピエンスULK2、図6C矢印2参照)は075385(LBDG1; ホモサピエンスULK1)と51%の配列同一性を共有する。したがってBAA31598.1(ホモサピエンスULK2)は、075385(LBDG1; ホモサピエンスULK1)のオーソログを表す。

#### 【0088】

BAA77341.1(マス・マスキュラスULK2、図6C矢印3参照)は075385(LBDG1; ホモサピエンスULK1)と50%の配列同一性を共有する。しかしながら、BAA77341.1(マス・マスキュラスULK2)はBAA31598.1(ホモサピエンスULK2)とはるかに近縁であって91%の配列同一性を共有し、これはそれらが図6Cでクラスターを形成する理由である。BAA77341.1(マス・マスキュラスULK2)はしたがって070405(LBDG1; ホモサピエンスULK2)のマス・マスキュラスオーソログを表す。

075385(LBDG1; ホモサピエンスULK1)、070405(マス・マスキュラスULK1)、BAA31598.1(ホモサピエンスULK2)およびBAA77341.1(マス・マスキュラスULK2)で整列し、AlEyeで調査する(図9)。AlEyeは、前記リガンド結合ポケットの輪郭を形成すると予測される075385(LBDG1)の8つの疎水性残基(ゲノムスレッダーアラインメント範囲内; ALA861、LEU863、TRP894、LEU900、VAL901、LEU904、LEU910およびVAL932)は全て070405(マス・マスキュラスULK1; 図9を参照)で正確に保存されていることを明らかにして

いる。さらにまた、075385 (LBDG1) の5つの予測される“LBDモチーフ”残基のうちの4つ (LEU878、ASP885、GLN886およびLEU890) は正確に070405 (マス・マスキュラスULK1) で保存されている。

#### 【0089】

A1Ey eはまた、リガンド結合ポケットの輪郭を形成すると予測される075385 (LBDG1) の(ゲノムスレッダーアラインメント範囲内の) 8つの疎水性残基のうち5つ (TRP894、LEU900、VAL901、LEU910およびVAL932) は全て、BAA31598.1 (ホモサピエンスULK2) およびBAA77341.1 (マス・マスキュラスULK2) で正確に保存されていることを明らかにする。さらにまた、075385 (LBDG1) の5つの予測される“LBDモチーフ”残基のうち3つ (ASP885、GLN886、LEU890) は、BAA31598.1 (ホモサピエンスULK2) およびBAA77341.1 (マス・マスキュラスULK2) で正確に保存されている。

タンパク質の機能にとって必須の残基は前記タンパク質の相同体において保存されるであろう。したがって、相同体070405 (マス・マスキュラスULK1)、BAA31598.1 (ホモサピエンスULK2) およびBAA77341.1 (マス・マスキュラスULK2) における予測される075385 (LBDG1) 核内ホルモンレセプターリガンド結合ドメインの機能に必須であろう残基の保存は、075385 (LBDG1) が核内ホルモンレセプターリガンド結合ドメインを含むとの注釈付けを強く支持する。

#### 【0090】

図10はNCBIのユニジェン(UniGene)データベースから作製されたレポートである。このデータベースは、種々のヒトの組織由来の発現配列タグ(EST)の収集であり、タンパク質に対しての一般的な組織分布を与えることに用いることができる(ただし前記タンパク質配列が前記データベースに存在することを条件とする)。075385 (LBDG1) は前記データベースに存在し、血液、脳、乳房、CNS、結腸、生殖細胞、心臓、腎臓、肺臓、リンパ、筋肉、卵巣、膵臓、上皮小体、プールされたもの(pooled)、前立腺、精巣、扁桃腺、子宮、全胚、副腎、脳、脳\_\_正常、子宮頸部、結腸、結腸\_\_e s t、結腸\_\_i n s、頭部\_\_頸部、腎臓、平滑筋、肺臓、神経\_\_正常、卵巣、胎盤、前立腺\_\_正常、前立腺\_\_腫瘍および子宮として参照されている組織で発現することが示された。075385 (LBDG1) のmRNAもまたライブラリー3761NN1112 (成人神経組織ライブラリー) で高度に(3.6%) 提示されている(図11参照)。

ユニジェンデータベースは組織分布について大まかな見解を提供するが、かずさcDNAプロジェクトにより解析されたヒト未同定遺伝子によってコードされる巨大タンパク質; Human Unidentified Gene - Encoded Large Proteins Analyzed by Kazusa cDNA Project (HUG E) データベースは、RT-PCR ELISAによるcDNAレベルの直接的な実験測定値を提供する(図12を参照)。075385 (LBDG1) は識別記号KIAA0722の下で前記データベースに存在し、脳、心臓および卵巣で高度に発現されている。

#### 【0091】

075385 (LBDG1) に関連する論文が存在する; Kuroyanagi et al., Genomics (1998) 51: 76 - 85 (図13参照)。この論文は、075385 (LBDG1) をコードする遺伝子が、細胞遺伝子座12q24.3の遠位にマッピングされることを実証している。前記の位置は、II型遠位遺伝性運動神経障害(遠位性HMNII)の原因遺伝子、および肩甲骨棘筋萎縮(SPSMA)の原因遺伝子の近傍である。075385 (LBDG1) をコードする遺伝子はこれら遺伝子座の遠位にマッピングされるが、それにもかかわらず075385 (LBDG1) が前記疾患の一方または両方の原因遺伝子である可能性が存在する。したがって、075385 (LBDG1) の遺伝子はこれらの症状に關与すると推論できる。前記症状の診断および/または治療のために075385 (LBDG1) の遺伝子またはそのコードタンパク質を標的とする薬剤の開発のための道を作ることは重用であろう。特に、前記遺伝子が核内ホルモンレセプターリガンド結合ドメインを含むという同定は、(例えば核内ホルモンレセプターリガンド結合ドメインアゴニストまたはアンタゴニストの使用を通じて) 075385 (LBDG1) タンパク質に特異的な薬剤の開発を容易にする。

#### 【0092】

10

20

30

40

50

## 実施例 2

ホモサピエンスエストロゲンレセプター リガンド結合ドメインは“LBDモチーフ”を含み、前記モチーフは、今日までに注釈された全ての核内ホルモンレセプターリガンド結合ドメインで見出されている。“LBDモチーフ”は核内ホルモンレセプターのコアクチベーターおよびコリプレッサーを補充することに関与する。6つの残基PHE367、LEU370、ASP374、GLN375、LEU378およびLEU379は、ホモサピエンスエストロゲンレセプター リガンド結合ドメインにおいて前記モチーフを構成する(図7:1ERR:A参照)。上記で考察したように、これら“LBDモチーフ”残基のうちの4つは075385(LBDG1)で正確に保存されている。1ERR:AのASP374は075385(LBDG1)ではASP885として保存されている。1ERR:AのGLN375は075385(LBDG1)ではGLN886として保存されている。1ERR:AのLEU378は075385(LBDG1)ではLEU889として保存されている。1ERR:AのLEU379は075385(LBDG1)ではLEU890として保存されている。1ERR:A“LBDモチーフ”の5番目の残基、PHE367は075385(LBDG1)ではLEU878によって保存的に置換されている。

10

## 【0093】

“LBDモチーフ”残基の他に、既知の核内ホルモンレセプターリガンド結合ドメインで良く保存されている他の残基がある。075385(LBDG1)が核内ホルモンレセプターリガンド結合ドメインを含むとの注釈付けについての更なるテストは、“LBDモチーフ”の外側の重要な残基が075385(LBDG1)および5つの075385(LBDG1)相同体で保存されているか否かを解析することである。

リガンド結合およびコアクチベーター/コリプレッサーの補充の他に、核内ホルモンレセプターリガンド結合ドメインのもう1つの固有性は、他の核内ホルモンレセプターリガンド結合ドメインとのダイマー化能力である。例えば、エストロゲンレセプター(ER)リガンド結合ドメインはそれ自体でホモダイマーを形成することができる。

20

また別の例はペルオキシソーム増殖因子活性化レセプター(PPAR)リガンド結合ドメインはレチノイドXレセプター(RXR)リガンド結合ドメインとヘテロダイマーを形成することができるということである。リガンド結合ドメインがホモダイマーまたはヘテロダイマーを形成するか否かにかかわらず、同じ普遍的な相互作用の様式が結合の根底にある。結晶学的実験および突然変異誘発実験によって、ヘリックス“10”のN-末端側の半分が2つのダイマー化リガンド結合ドメイン間の主要な相互作用部位と同定された(R.T. Gampe, V.G. Montana, M.H. Lambert, A.B. Miller, R.K. Bledsoe, M.V. Milburn, S.A. Kliewer, T.M. Willson, H.E. Xu, Mol. Cell. (2000) 5(3): 545 - 55; A.J.M. Brzozowski, A.C. Pike, Z. Dauter, R.E. Hubbard, T. Bonn, O. Engstrom, L. Ohman, G.L. Greene, J.A. Gustafsson, M. Carlquist, Nature (1997) 389(6652): 753 - 8)。ヘリックス“10”のN-末端側半分の機能的な重要性はこの領域の特定の残基の強い保存性に反映されている。図14は、ホモサピエンスER、AR、RAR、RXR、PPARおよびRORに対して、ヘリックス“10”配列のN-末端側半分で整列化した075385(LBDG1)およびその3つの相同体を示している。ダイマーを形成するために、各モノマーの“10”ヘリックスは折り合わされて堅い骨格を形成し、これは極性反応および疎水性相互作用の組み合わせによって駆動される。ヘリックス“10”上にはダイマー化相互作用の極性成分において重要な役割を果たす2つの残基の位置がある(これらは図14で枠“b”および枠“g”として示されている)。ヘリックス“10”上の位置“b”は通常、極性残基(例えばERではGLN506)が占有するか、または正荷電残基(例えばRXRではLYS417)が占有する(これらは適切な側鎖特性を有し、ダイマー化パートナーと極性/電荷相互作用を生じる)。075385(LBDG1)とホモサピエンスエストロゲンレセプターリガンド結合ドメイン(1ERR:A)とのゲノムスレッダーアラインメントによれば、075385(LBDG1)のLYS1009は、ヘリックス“10”上の位置“b”を占有するであろう。明らかにLYS1009は、ダイマー化パートナーと極性/電荷相互作用を生じるのに適切な側鎖特性を有する。このことは、075385(LBDG1)は核内ホルモンレセプターリガンド結合ドメインを含むという注釈を支持する。ヘリックス“10”の位置“b”を占有する075385(LBDG1)のLYS1009の重要性は、LYS1009は、3つの075385(LBDG1)相同体(

30

40

50

070405、BAA31598.1およびBAA77341.1；図14枠“b”参照）で保存されているという知見によって強化される。

【0094】

ヘリックス“10”上の位置“g”は通常は正荷電残基（例えばERのARG516）または極性残基（例えばARのGLN867）によって占有されている（これらは適切な側鎖特性を有し、ダイマー化パートナーと極性/電荷相互作用を生じる）。075385（LBDG1）とホモサピエンスエストロゲンレセプターリガンド結合ドメイン（1ERR:A）とのゲノムスレグダーアラインメントによれば、075385（LBDG1）のGLN1018は、ヘリックス“10”上の位置“g”を占有するであろう。明らかにGLN1018は、ダイマー化パートナーと極性/電荷相互作用を生じるのに適切な側鎖特性を有する。このことは、075385（LBDG1）は核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けを支持する。ヘリックス“10”上の位置“g”を占有する075385（LBDG1）のGLN1018の重要性は、GLN1018が075385（LBDG1）のマス・マスキュラス オートログ070405で保存されているという知見（図14枠“g”参照）によって強化される。他の2つの075385（LBDG1）相同体（BAA31598.1およびBAA77341.1）は、位置“g”にSERを有すると予測される。明らかにSERもまたダイマー化パートナーと極性/電荷相互作用を生じるのに適切な側鎖特性を有する。

10

【0095】

リガンド結合ドメインのダイマー化はまた、多数の疎水性相互作用によって駆動される。例えば、ヘリックス“10”上の位置“a”は通常は芳香族残基（例えばPPARのPHE432）または大型の疎水性残基（例えばERのLEU504）によって占有され、これらはダイマー化パートナーと疎水性相互作用を生じるのに適切な側鎖を有する。075385（LBDG1）とホモサピエンスエストロゲンレセプターリガンド結合ドメイン（1ERR:A）とのゲノムスレグダーアラインメントによれば、075385（LBDG1）のTYR1008はヘリックス“10”上の位置“a”を占有しているであろう。明らかにTYR1008は、ダイマー化パートナーと疎水性相互反応を生じるのに適切な側鎖特性を有する。このことは、075385（LBDG1）は核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けを支持する。ヘリックス“10”上の位置“a”を占有する075385（LBDG1）のTYR1008の重要性は、TYR1008が3つの075385（LBDG1）相同体で保存されているという知見（070405、BAA31598.1およびBAA77341.1、図14枠“a”参照）によって強化される。

20

【0096】

ヘリックス“10”のN-末端側の半分はまた、既知のリガンド結合ドメインで保存された疎水性残基によって占有されている多数の他の位置（図14の枠c、d、e、fおよびh）を含む。これら疎水性残基は、溶媒露出してダイマー化パートナーとの疎水性相互作用に寄与するか、または埋め込まれてリガンド結合ドメインの本体に対してヘリックス“10”のフォールディングに関与する（例えばERのLEU507およびLEU514）。075385（LBDG1）とホモサピエンスエストロゲンレセプターリガンド結合ドメイン（1ERR:A）とのゲノムスレグダーアラインメントによれば、疎水性残基は、075385（LBDG1）でこれらの位置の全てを占有するであろう（位置“c”にALA1010、位置“d”にLEU1011、位置“e”にLEU1014、位置“f”にLEU1017、位置“h”にMET1020、図14参照）。このことは、075385（LBDG1）は核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けを支持する。075385（LBDG1）の位置c、d、e、fおよびhを占有する疎水性残基の重要性は、疎水性残基が、3つの075385（LBDG1）相同体でもこれらの位置を占有するという知見によってさらに強化される（070405、BAA31598.1およびBAA77341.1、図14参照）。

30

40

まとめると、前記残基解析は、075385（LBDG1）は核内ホルモンリガンド結合ドメインとホモダイマー化またはヘテロダイマー化するのに適切な補完残基を（ヘリックス“10”のN-末端に）保有していることを示している。このことは、075385（LBDG1）は核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けを支持する。

【0097】

075385（LBDG1）が核内ホルモンレセプターリガンド結合ドメインを含むというゲノムス

50

レッダーの注釈付けの更なるテストは、LBDフォールドを採用した075385(LBDG1)の相同モデルの構築を試みることである。075385(LBDG1)の相同モデルは、075385(LBDG1)とホモサピエンスレチノイン酸レセプター 構造1FCXとのゲノムスレッダーアライメントをモデラー(Modeller)バージョン5(A. Sali & T.L. Blundell, "Comparative protein modeling by satisfaction of spatial restraints", J.Mol. Biol. 234: 779 - 815, 1993)にかけることによって構築した。

前記相同モデルの全体図は図16に提示されている。075385(LBDG1)は1FCX構造に合わせてうまくモデリングされ、LBDフォールドを形成する交差格子ヘリックスの典型的な機構を示す。図17は、2つの主要ヘリックス“3”および“5”(これらは既知の核内ホルモンレセプターリガンド結合ドメイン構造の3および5に対応する)に焦点を合わせている。図18は予測されるコアクチベーター結合部位で拡大している。既知のLBD構造では、コアクチベーター結合部位はリガンド結合ドメインの表面で3と5との間に形成された溝である。図19は同じ図を示しているが、溝に結合しているコアクチベーターの予測される形式(mode)を説明するために相互作用しているコアクチベーターのアニメーションが加えられている。075385(LBDG1)の予測される“LBDモチーフ”の5つの残基、LEU878、ASP885、GLN886、LEU889およびLEU890は灰色で示されており、これらは全て適切な向きで存在し、どのような衝突または立体障害も無しにそれらの役割を果たしている。例えば、GLN886は、コアクチベーターヘリックスのC-末端と相互作用するのに適切な位置に有り、このことは、ちょうどエストロゲンレセプターがSRC-1のコアクチベーターヘリックスと結合したとき、エストロゲンレセプターの等価な残基GLN375について観察されるとおりである(A.K. Shiau, D.Barstad, P.M. Loria, Lin Cheng, P.J. Kushner, D.A. Agard & G.L. Greene, "The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen", Cell 1998, 95: 927)。同様に、LEU890は前記溝の中に突出し、LXXLLコアクチベーターヘリックスに存在するロイシンのいくつかと疎水的に接触することができよう。それは、ちょうどエストロゲンレセプターがSRC-1のコアクチベーターヘリックスと結合したときに、エストロゲンレセプターの等価な残基LEU379について観察されるとおりである(A.K. Shiau, D.Barstad, P.M. Loria, Lin Cheng, P.J. Kushner, D.A. Agard & G.L. Greene, "The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen", Cell 1989, 95: 927)。

#### 【0098】

“LBDモチーフ”の外側の残基もまた適切な向きで配置され、それらの役割を果たす。例えば、“ダイマー化”ヘリックス“10”のN-末端に配置されると予測される残基は、ダイマー化パートナーとの相互作用に加わることができるように適切な向きで存在する。図20は、ホモサピエンスエストロゲンレセプター ホモダイマーの構造の全体図を示している。前記ホモダイマーの境界の主要な特徴は、各モノマのが挟じりあわされた、ヘリックス“10”である。図21は、エストロゲンレセプター モノマー(図21A)とLBDフォールドを採用する075385(LBDG1)の相同モデル(図21B)とを比較している。ヘリックス“10”に配置されるとゲノムスレッダーによって予測される075385(LBDG1)の領域は、どのような衝突または立体障害も無しに長い連続したヘリックスとしてうまくモデリングされていることは明白である。このことは、075385(LBDG1)が核内ホルモンレセプターリガンド結合ドメインを含むというゲノムスレッダーの注釈付けを支持する。さらにまた、主要な電荷/極性相互作用を生じると予測される残基は、ダイマー化パートナーとの相互作用に適切な向きで存在する。例えば、075385(LBDG1)のLYS1009はリガンド結合ドメインの本体からダイマー化パートナーが位置すると思われる方向に向かって突出し、これは、ちょうどエストロゲンレセプターの等価な残基GLN506について観察されるとおりである。例えば、075385(LBDG1)のGLN1018は、リガンド結合ドメインの本体からダイマー化パートナーが位置すると思われる方向に向かって突出し、ちょうどエストロゲンレセプターの等価な残基ARG515について観察されるとおりである。

10

20

30

40

50

さらにまた、主要な疎水性相互作用を生じると予想される残基はダイマー化パートナーとの相互作用に適した向きで存在する。例えば、075385 (LBDG1) のTYR1007はリガンド結合ドメインの本体からダイマー化パートナーが位置すると思われる方向に向かって突出し、ちょうどエストロゲンレセプター の等価な残基LEU504について観察されるとおりである。図 2 1 は、ER (図21A)およびLBDフォールドを採用する075385 (LBDG1) の相同モデル (図 2 1 B) の両方についての、ヘリックス “ 1 0 ” のクローズアップを示している。図 2 1 はまた、リガンド結合ドメインの本体に対してヘリックス “ 1 0 ” フォールディングにおいて機能するとゲノムスレッダーにより予測される残基は、どのような衝突または立体障害も無しにその役割を果たすのに適切な向きで存在することを明らかにしている。例えば、075385 (LBDG1) のALA1010はリガンド結合ドメインの本体に突出し、その位置で前記構造のコアと疎水的に接触することができ、ちょうどエストロゲンレセプター の等価な残基LEU507について観察されるとおりである。例えば、075385 (LBDG1) のLEU1017はリガンド結合ドメインの本体に突出し、その位置で前記構造のコアと疎水的に接触することができ、ちょうどエストロゲンレセプター の等価な残基ILE514について観察されるとおりである。

10

リガンド結合ドメインとしての075385 (LBDG1) の相同モデリングは、075385 (LBDG1) がリガンド結合ドメインを含むというゲノムスレッダーの注釈付けを支持する。

【 0 0 9 9 】

### 実施例 3

本実施例は、本明細書中で核内ホルモンレセプターリガンド結合ドメインとして同定された、BAA31598.1と称される新規なタンパク質、並びに前記タンパク質およびコード遺伝子に由来する核酸配列の疾患の診断、予防および治療における使用に関する。

20

本タンパク質は075385 (LBDG1) のホモサピエンスパラログであり、本明細書においてLBDG4 (BAA31598.1) と称される。前記タンパク質は075385 (LBDG1) と51%の配列同一性を有し、さらに075385 (LBDG1) のリガンド結合ドメインフォールドで主要な役割を果たすと予測される残基は、BAA31598.1で保存されている (LBDG4、図 2 2 参照)。075385 (LBDG1) に対する高い相同性および予測される主要な残基の保存に基づいて、我々は、BAA31598.1 (LBDG4) もまた核内ホルモンレセプターリガンド結合ドメインを含むと注釈付ける。

ホモサピエンスエストロゲンレセプター リガンド結合ドメインは、今日まで注釈付けされている核内ホルモンレセプターリガンド結合ドメインの全てで見出されている “ L B D モチーフ ” を含む。 “ L B D モチーフ ” は、核内ホルモンレセプターのコアクチベーターおよびコリプレッサーを補充することに関与する。6つの残基PHE367、LEU370、ASP374、GLN375、LEU378およびLEU379は、ホモサピエンスエストロゲンレセプター リガンド結合ドメインで前記モチーフを構成する (図 2 2 : 1ERR : A参照)。これら “ L B D モチーフ ” 残基のうち3つはBAA31598.1 (LBDG4) で正確に保存されている。1ERR : AのASP374はBAA31598.1 (LBDG4) ではASP870として保存されている。1ERR : AのGLN375はBAA31598.1 (LBDG4) ではGLN871として保存されている。1ERR : AのLEU379はBAA31598.1 (LBDG4) ではLEU875として保存されている。1ERR : A “ L B D モチーフ ” の4番目の残基、PHE367はBAA31598.1 (LBDG4) ではILE863として保存的に置換されている。このことは、BAA31598.1 (LBDG4) は核内ホルモンレセプターリガンド結合ドメインを含むとの注釈付けを支持する。

30

40

【 0 1 0 0 】

“ L B D モチーフ ” の他に、既知の核内ホルモンレセプターリガンド結合ドメインで良く保存されている他の残基が存在する。BAA31598.1 (LBDG4) が核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けの更なるテストは、“ L B D モチーフ ” の外側の重要な残基がBAA31598.1 (LBDG4) で保存されているか否かを解析することである。

リガンド結合およびコアクチベーター/コリプレッサーの補充の他に、核内ホルモンレセプターリガンド結合ドメインのもう1つの固有性は他の核内ホルモンレセプターリガンド結合ドメインとのダイマー化能力である。例えば、エストロゲンレセプター (ER) リガンド結合ドメインはそれ自身とホモダイマーを形成することができる。

50

別の例は、ペルオキシソーム増殖因子活性化レセプター (PPAR) リガンド結合ドメインはレチノイドXレセプター (RXR) リガンド結合ドメインとヘテロダイマー化することができるということである。リガンド結合ドメインがホモダイマーまたはヘテロダイマーを形成するか否かにかかわらず、同じ普遍的な相互作用の様式が結合の根底にある。結晶学的実験および突然変異誘発実験によって、ヘリックス“ 10 ”のN - 末端側の半分が2つのダイマー化リガンド結合ドメイン間の主要な相互作用部位として同定された (R.T. Gampe, V.G. Montana, M.H. Lambert, A.B. Miller, R.K. Bledsoe, M.V. Milburn, S.A. Kliewer, T.M. Willson, H.E. Xu, Mol. Cell. (2000) 5(3): 545 - 55; A.M. Brzozowski, A.C. Pike, Z. Dauter, R.E. Hubbard, T. Bonn, O. Engstrom, L. Ohman, G.L. Greene, J.A. Gustafsson, M. Carlquist, Nature (1997) 389(6652): 753 - 8)。ヘリックス“ 10 ”のN - 末端側半分の機能的な重要性はこの領域に存在する特定の残基の強い保存性に反映されている。図14は、ホモサピエンスER、AR、RAR、RXR、PPARおよびRORのヘリックス“ 10 ”配列のN - 末端側半分で整列化したBAA31598.1 (LBDG4) およびその3つの相同体を示している。ダイマーを形成するために、各モノマーの“ 10 ”ヘリックスは折り合わされて堅い骨格を形成し、これは極性相互作用および疎水性相互作用の組み合わせによって駆動される。ヘリックス“ 10 ”上には2つの残基の位置があり、それら残基はダイマー化相互作用の極性成分で重要な役割を果たす (これらは図14で棒“ b ”および棒“ g ”として示されている)。ヘリックス“ 10 ”上の位置“ b ”は通常は、極性残基 (例えばERのGLN506) により占有されるか、または正荷電残基 (例えばRXRのLYS417) により占有される (これらは適切な側鎖特性を有し、ダイマー化パートナーと極性/電荷相互作用を生じる)。BAA31598.1 (LBDG4) とホモサピエンスエストロゲンレセプター リガンド結合ドメイン (1ERR: A) とのアラインメントによれば、BAA31598.1 (LBDG4) のLYS994はヘリックス“ 10 ”上の位置“ b ”を占有するであろう。明らかにLYS994は、ダイマー化パートナーと極性/電荷相互作用を生じるのに適切な側鎖特性を有する。このことは、BAA31598.1 (LBDG4) は核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けを支持する。

#### 【0101】

ヘリックス“ 10 ”上の位置“ g ”は通常は正荷電残基 (例えばERのARG516) または極性残基 (例えばRORのTHR417) によって占有されている (これらは適切な側鎖の特性を有し、ダイマー化パートナーと極性/電荷相互作用を生じる)。BAA31598.1 (LBDG4) とホモサピエンスエストロゲンレセプター リガンド結合ドメイン (1ERR: A) とのアラインメントによれば、BAA31598.1 (LBDG4) のSER1003はヘリックス“ 10 ”上の位置“ g ”を占有するであろう。明らかにSER1003は、ダイマー化パートナーと極性/電荷相互作用を生じるのに適切な側鎖特性を有する。このことは、BAA31598.1 (LBDG4) は核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けを支持する。

リガンド結合ドメインのダイマー化はまた、多数の疎水性相互作用によって駆動される。例えば、ヘリックス“ 10 ”上の位置“ a ”は通常は芳香族残基 (例えばPPARのPHE432) または大型の疎水性残基 (例えばERのLEU504) によって占有され、これらはダイマー化パートナーと疎水性相互作用を生じるのに適切な側鎖を有する。BAA31598.1 (LBDG4) とホモサピエンスエストロゲンレセプター リガンド結合ドメイン (1ERR: A) とのアラインメントによれば、BAA31598.1 (LBDG4) のTYR992は、ヘリックス“ 10 ”上の位置“ a ”を占有しているであろう。明らかにTYR992は、ダイマー化パートナーと疎水性相互作用を生じるのに適切な側鎖特性を有する。このことは、BAA31598.1 (LBDG4) は核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けを支持する。

#### 【0102】

ヘリックス“ 10 ”のN - 末端側の半分はまた、既知のリガンド結合ドメインで保存された疎水性残基によって占有されている多数の他の位置 (図14の棒c、d、e、fおよびh) を含む。これら疎水性残基は、溶媒露出してダイマー化パートナーとの疎水性相互作用に寄与するか、または埋め込まれてリガンド結合ドメインの本体に対してヘリックス“ 10 ”のフォールディングに関与する (例えばERのLEU507およびLEU514)。BAA315

10

20

30

40

50

98.1 (LBDG4) とホモサピエンスエストロゲンレセプター リガンド結合ドメイン (1ERR : A) とのアラインメントによれば、疎水性残基は、BAA31598.1 (LBDG4) で前記の位置の全てを占有するであろう (位置 “ c ” に ALA995、位置 “ d ” に ALA996、位置 “ e ” に LEU99、位置 “ f ” に LEU1002、位置 “ h ” に ILE1005、図 1 4 参照)。このことは、BAA31598.1 (LBD4) は核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けを支持する。まとめると、前記残基解析は、BAA31598.1 (LBDG4) は (ヘリックス “ 1 0 ” の N - 末端で) 核ホルモンリガンド結合ドメインとホモダイマー化またはヘテロダイマー化するための適切な補完残基を保有することを示している。このことは、BAA31598.1 (LBDG4) は核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けを支持する。

#### 【 0 1 0 3 】

BAA31598.1 (LBDG4) が核内ホルモンレセプターリガンド結合ドメインを含むという注釈付けの更なるテストは、L B D フォールドを採用する BAA31598.1 (LBDG4) の相同モデルの構築を試みることである。BAA31598.1 (LBDG4) の相同モデルは、図 2 2 に示したアラインメントを基にして構築した。相同モデリングはモデラーバージョン 5 (A. Sali & T. L. Blundell, “ Comparative protein modeling by satisfaction of spatial restraints ”, J.Mol. Biol. 234 : 779 - 815, 1993) を用いて実施した。

前記相同モデルの全体図は図 2 3 に提示されている。BAA31598.1 (LBDG4) は L B D フォールドに合わせてうまくモデリングされ、L B D フォールドを形成する交差格子ヘリックスの典型的な機構を示す。BAA31598.1 (LBDG4) の予測される “ L B D モチーフ ” の 4 つの残基、ILE863、ASP870、GLN871 および LEU875 は黒色で示されており、全て適切な向きで存在し、どのような衝突または立体障害も無しにそれらの役割を果たす。

#### 【 0 1 0 4 】

“ L B D モチーフ ” の外側の残基もまた適切な向きで配置され、それらの役割を果たす。例えば、“ダイマー化”ヘリックス “ 1 0 ” の N - 末端に配置されると予測される残基は、ダイマー化パートナーとの相互作用に加わることができるように適切な向きで存在する (図 2 4)。ヘリックス “ 1 0 ” の N - 末端に配置されると予測される BAA31598.1 (LBDG4) の領域は、どのような衝突または立体障害も無しにヘリックスとしてうまくモデリングされていることは明白である。このことは、BAA31598.1 (LBDG4) が核内ホルモンレセプターリガンド結合ドメインを含むという前記ゲノム注釈付けを支持する。さらにまた、主要な電荷 / 極性相互作用を生じると予測される残基は、ダイマー化パートナーとの相互作用に適切な向きで存在する。例えば、BAA31598.1 (LBDG4) の LYS994 はリガンド結合ドメインの本体からダイマー化パートナーが位置すると思われる方向に向かって突出し、これは、ちょうどエストロゲンレセプターの等価な残基 GLN506 について観察されるとおりである。例えば、BAA31598.1 (LBDG4) の SER1003 は、リガンド結合ドメインの本体からダイマー化パートナーが位置すると思われる方向に向かって突出し、ちょうどエストロゲンレセプターの等価な残基 ARG515 について観察されるとおりである。さらにまた、主要な疎水性相互作用を生じると予測される残基はダイマー化パートナーとの相互作用に適切な向きで存在する。例えば、BAA31598.1 (LBDG4) の TYR992 はリガンド結合ドメインの本体からダイマー化パートナーが位置すると思われる方向に向かって突出し、ちょうどエストロゲンレセプターの等価な残基 LEU504 について観察されるとおりである。図 2 4 B および 2 4 C は、L B D フォールドを採用する BAA31598.1 (LBDG4) の相同モデルについての、ヘリックス “ 1 0 ” のクローズアップを示している。図 2 4 B および 2 4 C はまた、リガンド結合ドメインの本体に対してヘリックス “ 1 0 ” のフォールディングにおいて機能すると予想される残基は、どのような衝突または立体障害も無しにその役割を果たすのに適切な向きで存在することを明らかにしている。例えば、BAA31598.1 (LBDG4) の ALA995 はリガンド結合ドメインの本体に突出し、その位置で前記構造のコアと疎水的に接触することができ、ちょうどエストロゲンレセプターの等価な残基 LEU507 について観察されるとおりである。例えば、BAA31598.1 (LBDG4) の LEU1002 はリガンド結合ドメインの本体に突出し、その位置で前記構造のコアと疎水的に接触することができ、ちょうどエストロゲンレセプターの等価な残基 ILE514 について観察されるとおりである。

10

20

30

40

50

リガンド結合ドメインとしてのBAA31598.1 (LBDG4) の相同モデリングは、BAA31598.1 (LBDG4) がリガンド結合ドメインを含むというゲノムスレッダーの注釈付けを支持する。

【図面の簡単な説明】

【0105】

【図1】バイオペンジウムターゲットマイニングインターフェースのフロントエンドを示す。データベースの検索はPDBコード“1ERR:A”を用いて開始する。

【図2A】1ERR:Aを用いた検索に対するインファーマチカゲノムスレッダーの結果の抜粋が示されている。矢印は、ホモサピエンスエストロゲンレセプター（典型的な核内ホルモンレセプターリガンド結合ドメインを有する）を示している。

【図2B】1ERR:Aを用いた検索に対するインファーマチカゲノムスレッダーの結果の抜粋が示されている。矢印は、075385 (LBDG1) を示している。 10

【図2C】1ERR:Aを用いた検索に対するフォワードPSI-BLASTの結果の完全なリストが示されている。075385 (LBDG1) は同定されない。

【図3】075385 (LBDG1) に対する重複配列表示の結果のページである。

【図4A】075385 (LBDG1) に対するインタープロPFAM検索結果を示す。矢印1を参照されたい。

【図4B】インタープロによるプロリンに富む領域についてのレポートで、前記領域がDNA結合ドメインとして機能し得ることを示している。

【図5】075385 (LBDG1) に対するSWISS-PROTのエントリーである。

【図6A】バイオペンジウムデータベースのフロントエンドである。データベースの検索はクエリー配列として075385 (LBDG1) を用いて開始する。 20

【図6B】クエリー配列として075385 (LBDG1) を用いた検索のインファーマチカゲノムスレッダーの結果の抜粋である。矢印は1ERR:Aを指している。

【図6C】クエリー配列として075385 (LBDG1) を用いて得られた逆最大化PSI-BLASTの結果の抜粋である。1、2、3の番号を付与された矢印は075385 (LBDG1) の相同体を指している。

【図7】075385 (LBDG1) および1ERR:AのAIEye配列アラインメントである。

【図8A】1ERR:Aに関するLigEyeで、ラロキシフェン (Ral600(A)) とホモサピエンスエストロゲンレセプターのリガンド結合ドメイン、1ERR:Aとの相互作用部位を図示している。 30

【図8B】1ERR:A、ホモサピエンスエストロゲンレセプターのリガンド結合ドメインのiRasMol図である。

【図9】075385 (LBDG1; ホモサピエンスULK1) と相同体070405 (マス・マスキュラスULK1)、BAA31598.1 (ホモサピエンスULK2) およびBAA77341.1 (マス・マスキュラスULK2) とのAIEye配列アラインメントである。

【図10】075385 (LBDG1) に対応するcDNAおよびESTの供給源を要約するNCBIユニジェンのレポートである。

【図11】NN1112は成人神経組織ライブラリーである。

【図12】KIAA0722 (075385 (LBDG1) についてのHUGE識別記号) に対するHUGEデータベースのレポートである。 40

【図13】Kuroyanagiらの論文 (Genomics (1998) 51: 76 - 85) の84ページである。075385 (LBDG1) の遺伝子および近傍の疾患に関する遺伝子座の細胞遺伝学的マップ位置を詳述する。

【図14】075385 (LBDG1) とホモサピエンスエストロゲンレセプター (1ERR:A) とのゲノムスレッダーアラインメントである。ダイマー化ヘリックス“10”のN-末端側の半分のみを示している。3つの075385 (LBDG1) 相同体が (075385 (LBDG1) 配列の参照によって) アラインメントに含まれている。他の5つのヒトリガンド結合ドメイン配列もまた、1ERR:A配列の参照によりアラインメントに含まれている。前記は、アンドロゲンレセプター (AR)、レチノイン酸レセプター (RAR)、レチノイドXレセプター (RXR)、ペルオキシソーム増殖因子活性化レセプター (PPAR) およびレチノ 50

イドオーファンレセプター (ROR) に由来するLBDである。保存残基は、aからhの標識を付した枠によって示されている。極性/荷電残基は灰色の枠内に、疎水性残基は黒色の枠内に示されている。

【図15】リガンド結合ドメインフォールドを採用する075385(LBDG1)の相同モデルの全体図である。特に注目される残基は黒色で示されている。

【図16】リガンド結合ドメインフォールドを採用する075385(LBDG1)の相同モデルの図で、予測されるヘリックス“3”および“5”を包含する領域のみが示されている。灰色の矢印はN-末端からC-末端へ向かうポリペプチド鎖の方向を示している。

【図17】リガンド結合ドメインフォールドを採用する075385(LBDG1)の相同モデルの予測されるコアクチベーター結合部位のクローズアップである。特に注目される残基は黒色で示されている。 10

【図18】リガンド結合ドメインフォールドを採用する075385(LBDG1)の相同モデルの予測されるコアクチベーター結合部位のクローズアップである。特に注目される残基は黒色で示されている。相同モデルへの予測される結合様式を図示するため、コアクチベーターヘリックスのアニメーションが付け加えられている。

【図19】ホモサピエンスエストロゲンレセプター ホモダイマー (1ERR:Aおよび1ERR:B)の構造の概観である。黒い線は2つのモノマーを分割している。ダイマー化ヘリックスは標識されている。

【図20】特にダイマー化ヘリックス“10”に関連する、ER構造(1ERR:A、図20A)とリガンド結合ドメインフォールドを採用する075385(LBDG1、図20B)の相同モデルとの比較である。ダイマー相互作用に重要な残基は黒色で示されている。 20

【図21】ER構造(1ERR:A、図21A)とリガンド結合ドメインフォールドを採用する075385(LBDG1、図21B)の相同モデルとの比較である。ダイマー化ヘリックス“10”のみを示してある。図21AおよびBは各々ダイマー化ヘリックス“10”の2つの垂直図である。ヘリックスの端を前に見た図では、ポリペプチド鎖はC-末端に向かって進み、ページから出て観察者の方に向かう。重要な残基は黒色で示されている。

【図22】BAA31598.1(LBDG4)とホモサピエンスエストロゲンレセプター (1ERR:A)とのゲノムスレッダーアラインメントである (BAA31598.1(LBDG4)のヒトパラログを含む)。

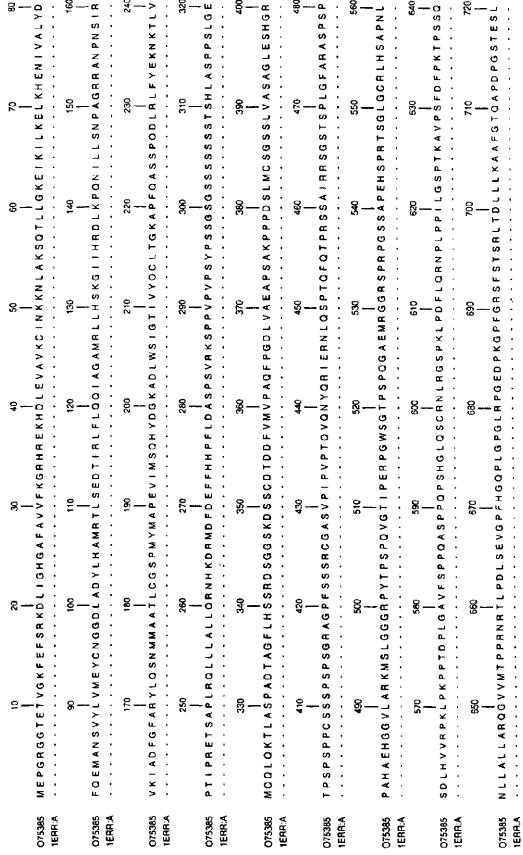
【図23】リガンド結合ドメインフォールドを採用するBAA31598.1(LBDG4)の相同モデルの全体図である。特に注目される残基は黒色で示されている。 30

【図24】特にダイマー化ヘリックス“10”に関連する、リガンド結合ドメインフォールドを採用するBAA31598.1(LBDG4)の相同モデルの図である。ダイマー相互作用に重要な残基は黒色で示されている。図24のBおよびCはダイマー化ヘリックス“10”の2つの垂直図を示す。ヘリックスの端を前に見た図では、ポリペプチド鎖はC-末端に向かって進み、ページから出て観察者の方に向かう。重要な残基は黒色で示されている。

【 図 7 - 1 】

AEI<sub>0</sub> 出力 (1月15日, 2001 5:02 PM)

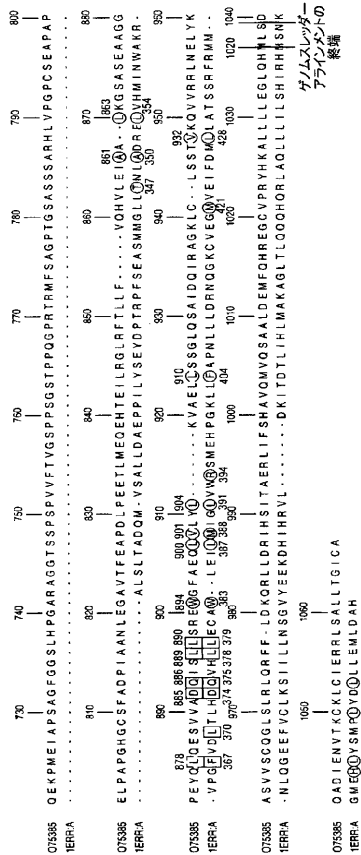
FIG. 7-1



【 図 7 - 2 】

AEI<sub>0</sub> 出力 (2月22日, 2001 12:07 PM)

FIG. 7-2



- IERR-A 中の "LBDモチーフ" 残基
- 075385I において正確に保存されている "LBDモチーフ" 残基
- 075385I において "LBDモチーフ" 残基中の残置できる変化
- IERR-A 中のリガンド結合残基
- 075385I 中で正確に保存されているリガンド結合残基
- 075385I 中で疎水性が保存されているリガンド結合残基

【 図 8 A 】

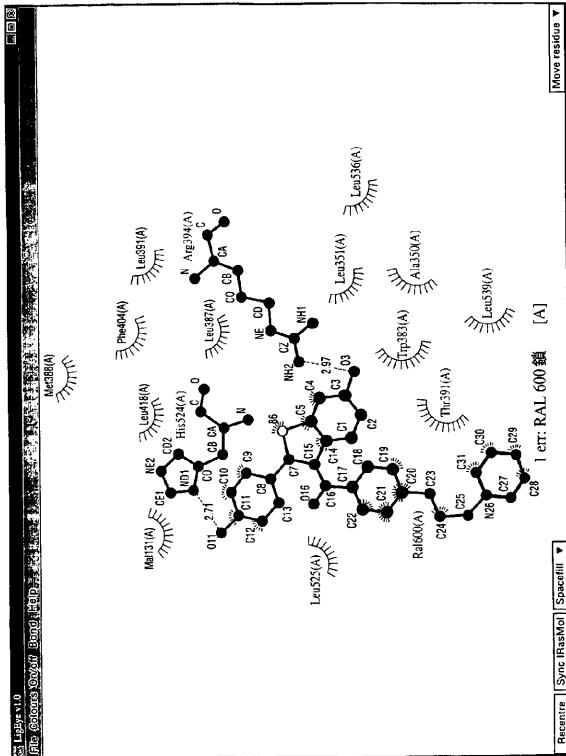


FIG. 8A

【 図 8 B 】

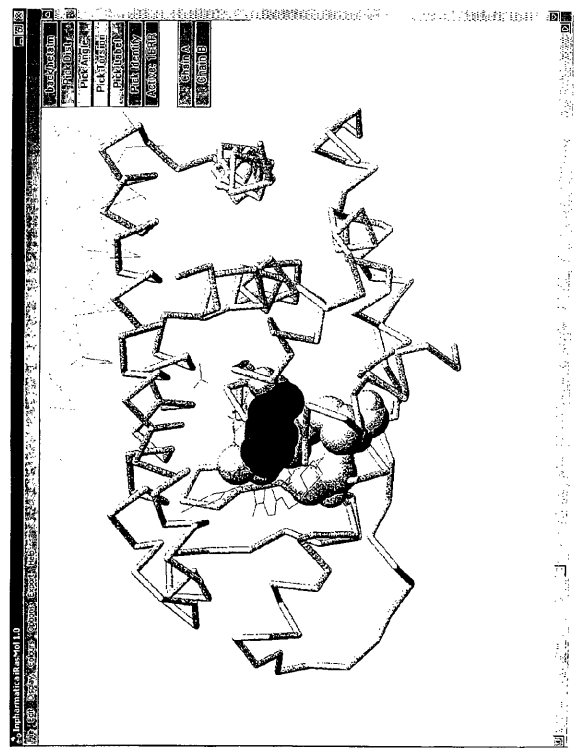
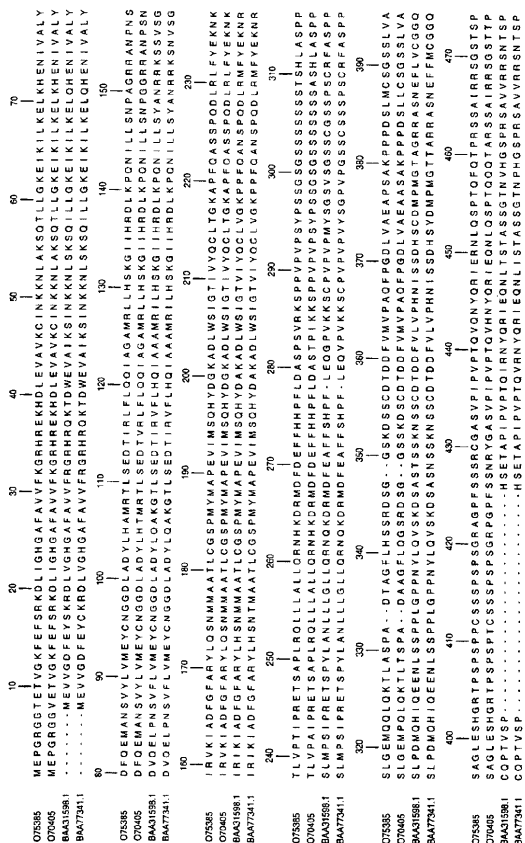


FIG. 8B

【 図 9 - 1 】

AEIeye 出力 ( 2月22日 , 2001 2:30 PM)

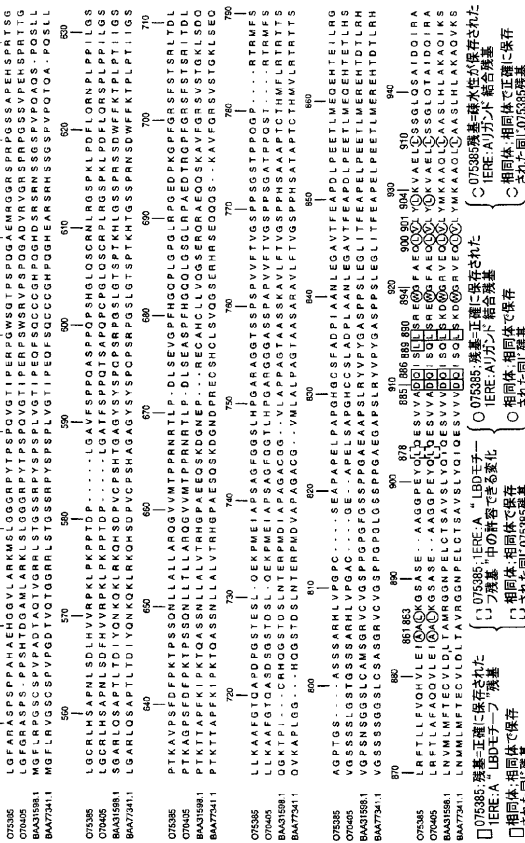
FIG. 9-1



【 図 9 - 2 】

AEIeye 出力 ( 2月22日 , 2001 2:30 PM)

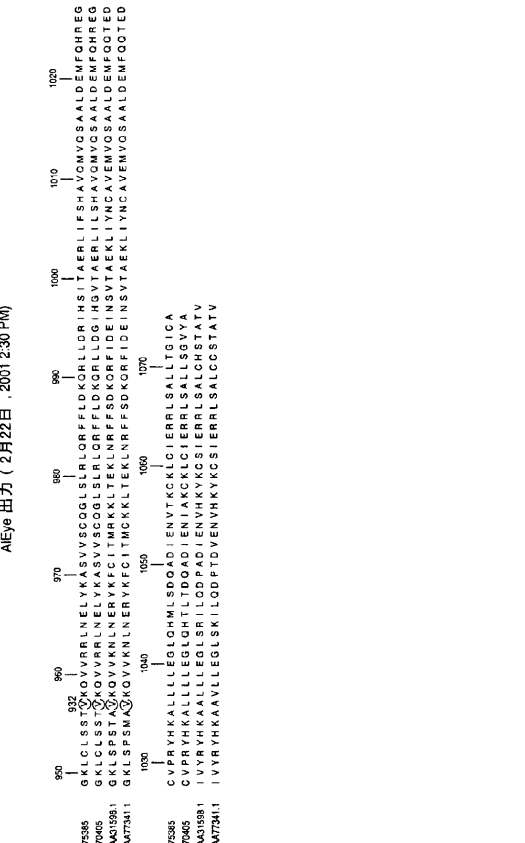
FIG. 9-2



【 図 9 - 3 】

AEIeye 出力 ( 2月22日 , 2001 2:30 PM)

FIG. 9-3



【 図 1 1 】

FIG. 11

NCBI UniGene Homo sapiens NN1112

PubMed Entrez BLAST OMI Taxonomy Structure

Search Human for [ ] display as html [G]

Lib.3761

NN1112

生物学的起源

ユニジーン ドナー: 成人  
組織: 神経 正常

ライブラリー構成の詳細

ベクター: puc18  
クローニング部位: SmaI : SmaI

提出者のコメント

ORESTES PCR (U.S. Letters Patent application No.196,716 - Ludwig 癌研究所) プロファイル由来の産物をpUC18ベクターにクローニングすることによって、cDNAライブラリーを構築した。組織mRNAの逆転写およびcDNA増幅は、低ストリンジエンシー条件下で行った。

ライブラリーディスカバリー及び多様性統計値

選ごとのライブラリー レポート ライブラリー 多様性グラフ ライブラリーディスカバリー

配列

ESTs: 418個の配列  
遺伝子 364個のESTは55個の遺伝子セットに分類された

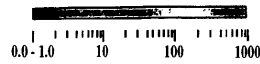
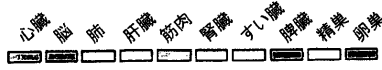
ユニジーン オルガニスム 対象遺伝子

【 図 1 2 】

FIG. 12

発現プロファイル

RT-PCR-ELISA



実験条件

プライマー<sub>f</sub> : TTTGTTCACAGGCTTCTCTGG  
 プライマー<sub>r</sub> : CCTGTGCCATCTGCTCTAAC  
 PCR条件 : 95°C 30 sec 55°C 30 sec 72°C 60 sec 30 cycles

RH マッピング情報

第12番染色体

実験条件

パネル名 : GeneBridge 4  
 プライマー<sub>f</sub> : TTTGTTCACAGGCTTCTCTGG  
 プライマー<sub>r</sub> : CCTGTGCCATCTGCTCTAAC  
 PCR産物の長さ : 209 bp  
 PCR条件 : 95°C 15 sec 64°C 60 sec 30 cycles

KIAAクローンの取得方法

Back to the [HUGE Protein Database homepage](#)  
 Send a message to [huge@kazusa.or.jp](mailto:huge@kazusa.or.jp)

説明

【 図 1 3 - 1 】

成人ヒト組織における ULK1 遺伝子の発現

ULK1 の組織発現をノーザンブロット分析により調べた (Fig. 5)。単一の 6.0kb の転写物が調べた全ての成人組織において種々の程度で検出された。ULK1 も最も高い遺伝子発現は骨格筋および心臓において検出され、中程度の発現は脾臓、脳、胎盤及び肝臓で検出され、最も低い発現は腎臓および肺に見られた。C. elegans において UNC-51 は神経に特異的に発現しており、そのヒト homolog ULK1 は普遍的な発現を示し、このことはこのクラスのセリン/スレオニンキナーゼが分子進化において転写特異性を獲得したかもしれないことを示唆する。遺伝子増幅は脊椎動物ゲノムの進化における機能的多様性に寄与するかも知れない。

ヒト ULK1 遺伝子の染色体マッピング

ヒト ULK1 遺伝子の染色体上の位置付けは FISH およびライダイエーションハイブリッドマッピングによって決定した。ヒト ULK1 cDNA プローブを用いた FISH 分析によりヒト ULK1 遺伝子が 12q24.3 に割り当てられることが明らかになった (Fig. 6)。体細胞ハイブリッドパネルの解析によりヒト第 12 番染色体を含むサンプルが予測される大きさの DNA 断片を特異的に生じさせることが示され、このことはヒト ULK1 遺伝子が第 12 番染色体上に位置することを示す。更なる測定のため、ライダイエーションハイブリッドを基礎とするマッピングを行った。ヒト ULK1 から得たデータベクトルは 00001000010000111 1100000110 001101000

説明

0 0000110000 100000000 0000001010  
 1011110001 0100011011 010 であった。

このデータベクトルの統計分析により ULK1 遺伝子座は D12S367、染色体 12q24.3 上の遺伝的マーカー、の遠位にマッピングされることが示された。ヒト遺伝子マッピング (Kucherlapati ら、1995) および OHIM を参照すると、2つの遺伝病がこの領域に割り当てられている：II 型遠位運動神経障害 (遠位 HMN II) の原因遺伝子は最近染色体 12q24 上の D12S86 と D12S340 の間の 13-cM に位置づけられており (Timmerman ら、1996)、肩甲骨骨筋萎縮症 (scapuloperoneal spinal muscular atrophy) (SPSMA) の原因遺伝子は染色体 12q24.1-q24.31 上の D12S338 と D12S366 の間の 19-cM に位置づけられている (Isozumi ら、1996)。マーカー-D12S367 はこれらの区間の遠位に位置づけられているので、ULK1 がこれらの遺伝病と関連することはありなさそうである。ヒトまたはマウスにおける ULK1 キナーゼを欠損する疾病を明らかにするためにさらなる遺伝的ストラテジーまたはジェネターゲッティングストラテジーが将来的に必要であろう。

謝辞

Koizumi Ken-ichi 博士、Sakai Tsuyoshi 氏 および Tomoda Toshifumi 氏の技術的援助に感謝する。H.K. は日本学術振興会の若手研究者促進のための特別研究員奨励学に援助されている。

【 図 1 3 - 2 】

FIG. 13

KUROYANAGI ET AL.

参考文献

Aigner, L. and Caroni, P. (1995) Absence of persistent spreading, branching, and adhesion in GAP-43-depleted growth cones. *J. Cell Biol.* 128: 647-650.  
 Appley, J. L., Parkinson, J. S., and Bourne, R. B. (1996) Signal transduction via the multi-step phosphatase: Not necessarily a road less traveled. *Cell* 86: 845-848.  
 Bonner, T. I., Oppermann, H., Seeburg, P., Kerby, S. B., Gannell, M. A., Young, A. C., and Rapp, U. R. (1986) The complete coding sequence of the human *ret* oncogene and the corresponding structure of the *c-myc* gene. *Nucleic Acids Res.* 14: 1009-1013.  
 Charest, D. L., Mordret, G., Harder, K. W., Jirik, F., and Pelech, S. L. (1993) Molecular cloning, expression, and characterization of the human mitogen-activated protein kinase *pak4*. *Mol. Cell Biol.* 13: 4679-4690.  
 Desai, C., Garriga, C., McIntire, S. L., and Horvitz, H. R. (1988) A genetic pathway for the development of the *Caenorhabditis elegans* HSN motor neuron. *Nature* 336: 638-646.  
 Dilleig, S. J. (1996) Cell cycle checkpoints: Preventing an identity crisis. *Science* 274: 1664-1672.  
 Faux, M. C., and Scott, J. D. (1996) Molecular glue: Kinase anchoring and scaffolding proteins. *Cell* 86: 9-12.  
 Finkenzeller, G., Marne, D., and Hug, H. (1990) Sequence of human protein kinase C  $\alpha$ . *Nucleic Acids Res.* 18: 2183.  
 Hanks, S. K., and Hunter, T. (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: Kinase (catalytic) domain structure and classification. *FASEB J.* 9: 576-596.  
 Harris, W. A., and Holt, C. E. (1995) From tags to RAGS: Chemoattractant receptors and ligands. *Neuron* 15: 241-244.  
 Hedgecock, E. M., Culotti, J. G., Thomson, J. N., and Perkins, L. A. (1984) Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescent dyes. *Dev. Biol.* 111: 158-170.  
 Higgins, D. G., and Sharp, P. M. (1988) CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* 73: 237-244.  
 Ito, M., K. DeLong, R., Kaplan, J., Deng, H. X., Iqbal, Z., Hung, W. T., Watanabe, K. C., Henshi, A., Percek, V. M., and Siddiquie, T. (1996) Linkage of scapuloperoneal spinal muscular atrophy to chromosome 12q24.1-q24.31. *Hum. Mol. Genet.* 5: 1377-1382.  
 Juan, M., Hidaka, H., and Schmidt, J. T. (1994) Kinase requirement for retinal growth cone motility. *J. Neurobiol.* 25: 1310-1325.  
 Johnson, L. N., Noble, M. E., and Owen, D. J. (1996) Active and inactive protein kinases: Structural basis for regulation. *Cell* 85: 149-158.  
 Keynes, R., and Cook, G. M. (1986) Axons turn to neurons find their receptor. *Neuron* 17: 1031-1034.  
 Koszik, M. (1986) Point mutations define a sequence flanking the AUC initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44: 265-272.  
 Koszik, M. (1987) At least six nucleotides preceding the AUC initiator codon enhance translation in mammalian cells. *J. Mol. Biol.* 196: 947-950.  
 Kucherlapati, R. S., Craig, I. W., and Marynen, P. (1995) Report of the Committee on the Genetic Constitution of Chromosome 12. In "Human Gene Mapping: A Compendium" (A. J. Cantichia, M. A. Chaperfield, and P. A. Foster, Eds., pp. 761-800, Johns Hopkins Univ. Press, Baltimore, MD).  
 Lee, N. G., and Nurse, P. (1987) Complementation used to clone a human homologue of the fission yeast cell cycle control gene *cdc2*. *Nature* 327: 31-35.  
 Letwin, K., Mizzen, L., Mose, B., Bero, D. Y., Bernstein, A., and Pawson, T. (1992) A mammalian dual specificity protein kinase *Nek1*, is related to the NIMA cell cycle regulator and highly expressed in meiotic germ cells. *EMBO J.* 11: 3521-3531.

【 図 1 5 - 1 】

FIG. 15-1

		10	20	30	40	50
11bd ( 225 )		saneDmPverllcAElavepk-----				-----te
21bd ( 182 )		lspqlcglitkVSLAHqgTFP-slcqlg-				-----kyttn--
1prga ( 207 )		esadrlaklklydgyIKSPtkkAqzsl				-----ltgktldkspFvlydnms
3erta ( 306 )		lalgltdqmvslaldqkapp-----				-----lylseygdp-
1a28a ( 682 )		qlpp1Inlmsjepd-----				-----vlyAgHdn-
		aaaaaaaa				
		60	70	80	90	100
11bd ( 248 )		LVVgsmwlmqspn-----				-----dptvniqgAedkqlfLveWakrIP
21bd ( 214 )		-----ssadhrvq-----				-----LdlglvdkfgeLatkcikiveFakrLp
1prga ( 255 )		lmmGedkikfkhitlqegskvAikifgqg				-----frsveavqelTeYakgIp
3erta ( 334 )		-----trpfs-----				-----easmngllLnLqgrelvHMinWakrLp
1a28a ( 706 )		-----tkpdt-----				-----ssGLLglnlqgqGLLsvvkvWkSLp
		aaaaaaaaaaaaaaaaaaaaaa				
		110	120	130	140	150
11bd ( 288 )		hFseLpLdDgVllrAgwneLIAsFshr				-----gia-vk-----dGllLatG-lhv
21bd ( 250 )		gFtGlsIaDgYlLlKsKClDlIMRlCTR				-----ytpeqdtMTfSDG-LtL
1prga ( 305 )		gFvDlLhdGvLlKvYWHlITLmAsl---				-----Ymk-dGvLlSeGGfM
3erta ( 366 )		gFvDlLhdGvLlKvYWHlITLmAsl---				-----hp-gKlFA---pplLl
1a28a ( 738 )		gFvDlLhdGvLlKvYWHlITLmAsl---				-----pdLlL
		333	aaaaaaaaaaaaaaaaaaaaaa			-----bbb
		160	170	180	190	200
11bd ( 333 )		hrcnsAhsa-gvgalFgrVlTeLvsKmz				-----MqWdkTELGCRAIVLfnDpd
21bd ( 295 )		nctQmNA---GfCplTdlVfAqgl1-p				-----leMdtETGLLSAICLTeGdc
1prga ( 349 )		tgeFLkrlrkPfgdfmepkFfAvkFh-				-----aleLdDlAlFIAlIIGdGR
3erta ( 411 )		drnqGkCv-egMveIFdmLatSsrFr-				-----mmulqgeEFVCLKSILLNLSV
1a28a ( 785 )		neqkMk-e-ssFyslQlTmVlPqefv-				-----klqVsqeEFLCkVLLLNtIp
		baaaa	aaaaaaaaaaaa	aaaaaaaaaaaa		
		210	220	230	240	250
11bd ( 381 )		k-----glsmpagVslrckVyaqLeay				-----Ckhkypeq---pgfakl1
21bd ( 342 )		m-----dlspekdldlqepLleLlyA				-----rEKP---sqpMfppM1
1prga ( 398 )		p-----glnvklpedlgnhLlqLleLl				-----CKLkhhp---essglfakl1
3erta ( 459 )		ytFlsStlksckdhHhVldkIdtDlth				-----Makag1lqqahqlaqL1
1a28a ( 832 )		l-----egLrntqgFsemSSyRElkL				-----lgLrqkvvssgrfyqlT
		aaaaaaaaaaaaaaaaaaaaaa				-----aaaaaa
		260	270	280	290	300
11bd ( 420 )		lrlpLdLrSgIkLqehlFff-----				-----KLiCstpidtflmlemla
21bd ( 381 )		mkItDlrglStgkgeaaiLLmeIp---				-----gpMpliremlenp
1prga ( 437 )		gkmtDlrglVtehvqllqvIkktetd---				-----mslhplleiykd1
3erta ( 509 )		l1lshIGHsKgmghlymkc-----				-----kxvv---plydllemlD-----
1a28a ( 875 )		klLdLhdLlVkgLHlVclntfgrsrls				-----VeFpenmgevvaagLphlIagm
		a	aaaaaaaaaaaaaaaaaaaaaa			-----aaaaaa
		11bd ( 458 )				ph---gnt
		21bd ( )				
		1prga ( )				
		3erta ( 546 )				--ahr1ha
		1a28a ( 925 )				Vkl1lfhk

【 図 1 5 - 2 】

FIG. 15-2

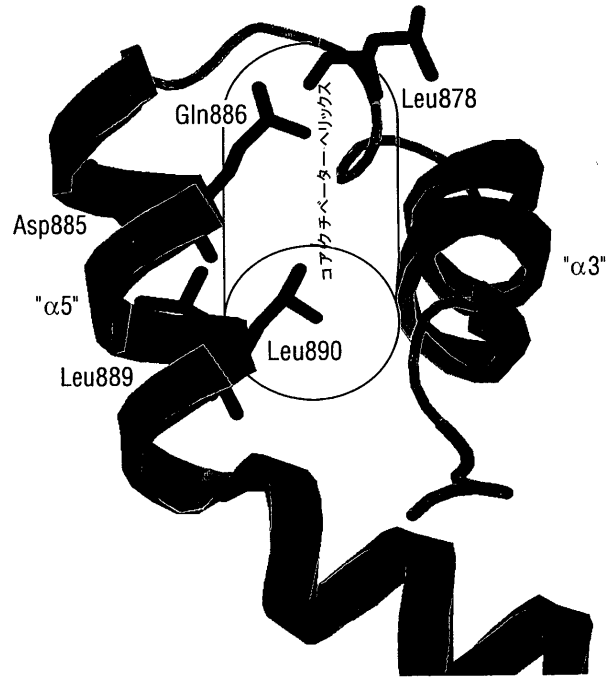
Key to Homstrad

aaa	$\alpha$ ヘリックス
bbb	$\beta$ 鎖
333	$3_{10}$ ヘリックス

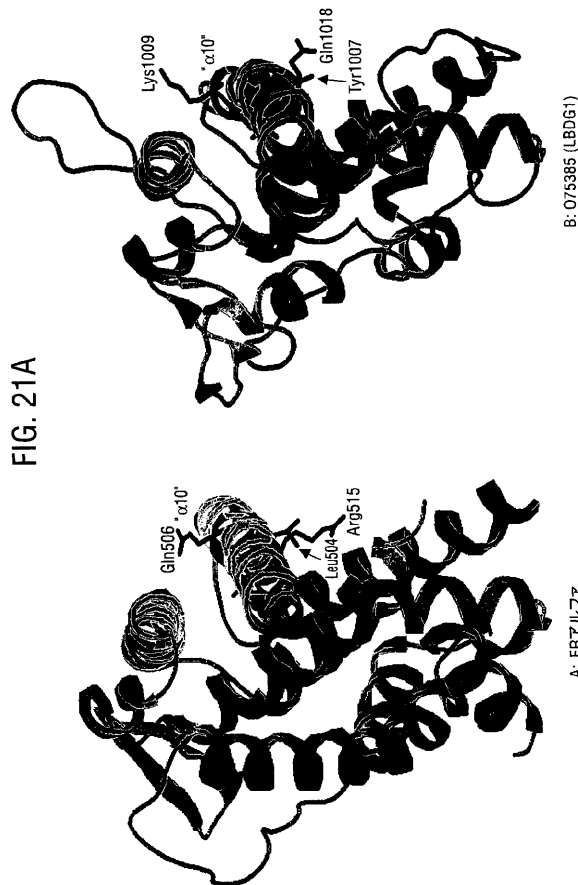
小文字	溶媒接近性
大文字	溶媒非接近性
太字	主鎖アミドに対する水素結合
下線	主鎖カルボニルに対する水素結合
イタリック	正のゆねじれ角

【 図 1 9 】

FIG. 19

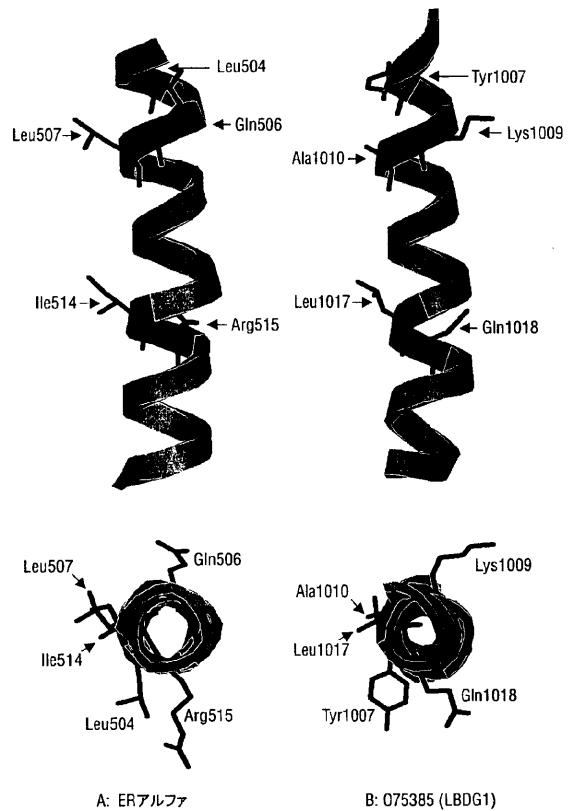


【 図 2 1 A 】



【 図 2 1 B 】

FIG. 21B





## 【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  
12 September 2002 (12.09.2002)

PCT

(10) International Publication Number  
WO 02/070559 A2

- (51) International Patent Classification: C07K 14/705 2NU (GB). POTTER, Sarah, Jane [GB/GB]; 39 Kneller Road, Breckley, London SE4 2AR (GB).
- (21) International Application Number: PCT/GB02/00948
- (22) International Filing Date: 5 March 2002 (05.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 0105402.2 5 March 2001 (05.03.2001) GB
- (71) Applicant (for all designated States except US): INPHARMATICA LIMITED [GB/GB]; 60 Charlotte Street, London W1T 2NU (GB).
- (72) Inventors; and  
(75) Inventors/Applicants (for US only): FAGAN, Richard, Joseph [US/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PHELPS, Christopher, Benjamin [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PHILLIPS, Tom [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PIERRON, Valerie, Nathalie [FR/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). ALLEN, Kathryn, Elizabeth [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). ALLEN, Janet, Marjorie [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB).
- (74) Agents: MERCER, Christopher, Paul et al.; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).
- (81) Designated States (national): AF, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BI, CI, CG, CL, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 02/070559 A2

(54) Title: NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAINS

(57) Abstract: This invention relates to the novel proteins, termed LBDG1 and LBDG4, herein identified as Nuclear Hormone Receptor Ligand Binding Domains and to the use of these proteins and nucleic acid sequences from the encoding genes in the diagnosis, prevention and treatment of disease.

WO 02/070559

PCT/GB02/00948

**NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAINS**

This invention relates to the novel proteins, termed O75385 and BAA31598.1 herein identified as Nuclear Hormone Receptor Ligand Binding Domains and to the use of these proteins and nucleic acid sequences from the encoding genes in the diagnosis, prevention and treatment of disease.

All publications, patents and patent applications cited herein are incorporated in full by reference.

**BACKGROUND**

The process of drug discovery is presently undergoing a fundamental revolution as the era of functional genomics comes of age. The term "functional genomics" applies to an approach utilising bioinformatics tools to ascribe function to protein sequences of interest. Such tools are becoming increasingly necessary as the speed of generation of sequence data is rapidly outpacing the ability of research laboratories to assign functions to these protein sequences.

As bioinformatics tools increase in potency and in accuracy, these tools are rapidly replacing the conventional techniques of biochemical characterisation. Indeed, the advanced bioinformatics tools used in identifying the present invention are now capable of outputting results in which a high degree of confidence can be placed.

Various institutions and commercial organisations are examining sequence data as they become available and significant discoveries are being made on an on-going basis. However, there remains a continuing need to identify and characterise further genes and the polypeptides that they encode, as targets for research and for drug discovery.

Recently, a remarkable tool for the evaluation of sequences of unknown function has been developed by the Applicant for the present invention. This tool is a database system, termed the Biopendium search database, that is the subject of co-pending International Patent Application No. PCT/GB01/01105. This database system consists of an integrated data resource created using proprietary technology and containing information generated from an all-by-all comparison of all available protein or nucleic acid sequences.

The aim behind the integration of these sequence data from separate data resources is to

WO 02/070559

PCT/GB02/00948

2

combine as much data as possible, relating both to the sequences themselves and to information relevant to each sequence, into one integrated resource. All the available data relating to each sequence, including data on the three-dimensional structure of the encoded protein, if this is available, are integrated together to make best use of the information that is known about each sequence and thus to allow the most educated predictions to be made from comparisons of these sequences. The annotation that is generated in the database and which accompanies each sequence entry imparts a biologically relevant context to the sequence information.

This data resource has made possible the accurate prediction of protein function from sequence alone. Using conventional technology, this is only possible for proteins that exhibit a high degree of sequence identity (above about 20%-30% identity) to other proteins in the same functional family. Accurate predictions are not possible for proteins that exhibit a very low degree of sequence homology to other related proteins of known function.

In the present case, a protein whose sequence is recorded in a publicly available database as O75385 (NCBI Genebank nucleotide accession number AF045458 and a Genebank protein accession number O75385), is implicated as a novel member of the Nuclear Hormone Receptor Ligand Binding Domain family.

#### **Introduction to Nuclear Hormone Receptor Ligand Binding Domains**

The Nuclear Hormone Receptor gene superfamily (see Table 1) encodes structurally related proteins that regulate the transcription of target genes. These proteins include receptors for steroid and thyroid hormones, vitamins, and other proteins for which no ligands have been found. Nuclear Receptors are composed of two key domains, a DNA-Binding Domain (DBD) and a Ligand Binding Domain (LBD). The DBD directs the receptors to bind specific DNA sequences as monomers, homodimers, or heterodimers. The DBD is a particular type of zinc-finger, found only in Nuclear Receptors. Nuclear Receptors with DBDs can be readily identified at the sequence level by searching for matches to the PROSITE consensus sequence (PS00031).

WO 02/070559

PCT/GB02/00948

3

The Ligand Binding Domain (LBD) binds and responds to the cognate hormone. Ligand binding to the LBD triggers a conformational change which expels a bound "Nuclear Receptor Co-Repressor". The site previously occupied by the Co-Repressor is then free to recruit a "Nuclear Receptor Co-Activator". This Ligand-triggered swap of a Co-Repressor for a Co-Activator is the mechanism by which Ligand binding leads to the transcriptional activation of target genes. All ligand binding domains contain a consensus sequence, the "LBD motif" (see Table 2) which mediates Co-Repressor and Co-Activator binding. The LBD is the binding site for all Nuclear Hormone Receptor targeted drugs to date and it is thus desirable to identify novel Ligand Binding Domains since these will be attractive drug targets. Ligand Binding Domains share low sequence identity (~15%) but have very similar structures and so present ideal targets for a structure-based relationship tool such as Genome Threader.

Many protein sequences have already been annotated in the public domain as Nuclear Hormone Receptors by their possession of DBDs using basic search tools like PROSITE, and their LBDs inferred on the basis of this. Because of this it is anticipated that any novel LBDs identified by Genome Threader *which are not annotated as nuclear receptors* will lack the DBD entirely. A precedent for a protein which has an LBD but lacks a DBD is provided by DAX1. Thus we annotate these DBD-less hits not as "Nuclear Hormone Receptors" but rather as containing a "Nuclear Hormone Receptor Ligand Binding Domain".

WO 02/070559

PCT/GB02/00948

4

Table 1: Nuclear hormone Receptor Superfamily

<i>Family: Steroid Hormone Receptors</i>	
Subfamilies	Glucocorticoid Receptors
	Progesterone Receptors
	Androgen Receptors
	Estrogen Receptors
<i>Family: Thyroid Hormone Receptor-like Factors</i>	
Subfamilies	Retinoic Acid Receptors (RARs)
	Retinoid X Receptors (RXRs)
	Thyroid Hormone Receptors
	Vitamin D Receptor
	NGFI-B
	FTZ-F1
	Peroxisome Proliferator Activated Receptors (PPARs)
	Ecdysone Receptors
	Retinoid Orphan Receptors (RORs)
	Tailless/COUP
	HNF-4
	CF1
	Knirps
<i>Family: DAX1</i>	
Subfamilies	DAX1

WO 02/070559

PCT/GB02/00948

5

Table 2: The "LBD motif". Numbers along the top row refer to residue position within the motif. Letters refer to amino acids by the 1-letter code. Letters within one column are all acceptable for that position within the motif. For example L, I, A, V, M, F, Y or W can occupy the first position of the "LBD motif". Note that there is observed variation in the number of residues found between position 4 and 8, and position 9 and 12. The "LBD motif" was constructed by aligning 681 sequences of Nuclear Hormone Receptor Ligand Binding Domains, and identifying conserved patterns of residues.

1	2	3	4	5	6	7	8	9	10	11	12	13
L	Any 2 residues		L	Any 3 residues (or 2 residues or 4 residues)			D	Q	Any 2 residues (or 1 or 3 residues)		L	L
I			I				E	N			I	I
A			A					R			A	A
V			V					H			V	V
M			M					K			M	M
F			F					S			F	F
Y			Y					T			Y	Y
W			W								W	W

WO 02/070559

PCT/GB02/00948

6

**II. Nuclear Hormone Receptors and Disease**

Nuclear Hormone Receptors have been shown to play a role in diverse physiological functions, many of which can play a role in disease processes (see Table 3).

5

Table 3. Nuclear Hormone Receptors and disease.

Nuclear Hormone Receptor	Disease
Androgen Receptor	Androgen Insensitivity Syndrome (Lubahn <i>et al.</i> 1989 Proc. Natl. Acad. Sci. USA 86, 9534-9538).
	Reifenstein syndrome (Wooster <i>et al.</i> 1992 Nat. Genet. 2, 132-134).
	X-linked recessive spinal and bulbar muscular atrophy (MacLean <i>et al.</i> 1995 Mol. Cell. Endocrinol. 112,133-141).
	Male breast cancer ((Wooster <i>et al.</i> 1992 Nat. Genet. 2, 132-134).
Glucocorticoid Receptor	Nelson's syndrome (Karl <i>et al.</i> 1996 J. Clin. Endocrinol. Metab. 81, 124-129).
	Glucocorticoid resistant acute T-cell leukemia (Hala <i>et al.</i> 1996 Int. J. Cancer 68, 663-668).
Mineralocorticoid Receptor	Pseudohypoaldosteronism (Chung <i>et al.</i> 1995 J. Clin. Endocrinol. Metab. 80, 3341-3345).
Estrogen Receptor alpha	ER alpha expression is elevated in a subset of human breast cancers. The application of Tamoxifen is the major therapy to prevent breast tumour progression. Unfortunately 35% of ER alpha positive breast cancers are Tamoxifen resistant (Petrangeli <i>et al.</i> 1994 J. Steroid Biochem. Mol. Biol. 49, 327-331).
Vitamin D3 Receptor	Mutations in the Vitamin D3 receptor produce a hereditary disorder similar in phenotype to Vitamin D3 deficiency (Rickets) (Hughes <i>et al.</i> 1988 Science 242, 1702-1725).
Retinoic Acid Receptor alpha	Acute Myeloid Leukemia (Lavau and Dejean 1994 Leukemia 8, 9-15).
Thyroid Hormone Receptor beta	"Generalised Resistance to Thyroid Hormones" (GRTH) (Refetoff 1994 Thyroid 4, 345-349).
DAX1	X-linked Adrenal Hypoplasia Congenita (AHC) and Hypogonadism (Ito <i>et al.</i> 1997 Mol. Cell. Biol. 17, 1476-1483).

WO 02/070559

PCT/GB02/00948

7

Alteration of Nuclear Hormone Receptors by ligands which bind to their LBD thus provides a means to alter the disease phenotype. There is thus a great need for the identification of novel Nuclear Hormone Receptor Ligand Binding Domains, as these  
5 proteins may play a role in the diseases identified above, as well as in other disease states. The identification of novel Nuclear Hormone Receptor Ligand Binding Domains is thus highly relevant for the treatment and diagnosis of disease, particularly those identified in Table 3.

#### THE INVENTION

10 The invention is based on the discovery that the O75385 and BAA31598.1 proteins function as Nuclear Hormone Receptor Ligand Binding Domains.

For the O75385 protein, it has been found that a region including residues 822-1020 of this protein sequence adopts an equivalent fold to residues 1 (ALA307) to 201 (MET517) of the Human Estrogen Receptor alpha (PDB code 1ERR:A). Human Estrogen Receptor  
15 alpha is known to function as a Nuclear Hormone Receptor Ligand Binding Domain. Furthermore, the "LBD motif" residues PHE367, ASP374, GLN375, LEU378 and LEU379 of the Human Estrogen Receptor alpha are conserved as LEU878, ASP885, GLN886, LEU889 and LEU890 in O75385, respectively. This relationship is not just to Human Estrogen Receptor alpha, but rather to the Nuclear Hormone Receptor Ligand  
20 Binding Domain family as a whole. Thus, by reference to the Genome Threader™ alignment of O75385 with the Human Estrogen Receptor alpha (1ERR:A) LEU878, ASP885, GLN886, LEU889 and LEU890 of O75385 are predicted to form the "LBD motif" residues.

The combination of equivalent fold and conservation of "LBD motif" residues allows the  
25 functional annotation of this region of O75385, and therefore proteins that include this region, as possessing Nuclear Hormone Receptor Ligand Binding Domain activity.

In one embodiment of the first aspect of the invention, there is provided a polypeptide, which polypeptide:

- (i) comprises the amino acid sequence as recited in SEQ ID NO:2;

WO 02/070559

PCT/GB02/00948

8

(ii) is a fragment thereof having Nuclear Hormone Receptor Ligand Binding Domain activity or having an antigenic determinant in common with the polypeptides of (i); or

(iii) is a functional equivalent of (i) or (ii).

5 Preferably, the polypeptide:

(i) consists of the amino acid sequence as recited in SEQ ID NO:2;

(ii) is a fragment thereof having Nuclear Hormone Receptor Ligand Binding Domain activity or having an antigenic determinant in common with the polypeptides of (i); or

10 (iii) is a functional equivalent of (i) or (ii).

The polypeptide having the sequence recited in SEQ ID NO:2 is referred to hereafter as "the LBDG1 polypeptide".

According to this aspect of the invention, a preferred polypeptide fragment according to part ii) above includes the region of the LBDG1 polypeptide that is predicted as that responsible for Nuclear Hormone Receptor Ligand Binding Domain activity (hereafter, the "LBDG1 Nuclear Hormone Receptor Ligand Binding Domain region"), or is a variant thereof that possesses the "LBD motif" (LEU878, ASP885, GLN886, LEU889 and LEU890, or equivalent residues). As defined herein, the LBDG1 Nuclear Hormone Receptor Ligand Binding Domain region is considered to extend between residue 822 and residue 1020 of the LBDG1 polypeptide sequence.

This aspect of the invention also includes fusion proteins that incorporate polypeptide fragments and variants of these polypeptide fragments as defined above, provided that said fusion proteins possess activity as a Nuclear Hormone Receptor Ligand Binding Domain.

25 A *Homo sapiens* paralogue of O75385 (LBDG1), has also been identified, and will be referred to herein as LBDG4. This polypeptide has the accession code BAA31598.1. BAA31598.1 exhibits 51% sequence identity to O75385 (LBDG1), and furthermore, residues predicted to play key roles in the Ligand Binding Domain fold of O75385

WO 02/070559

PCT/GB02/00948

9

(LBDG1) are conserved in BAA31598.1 (LBDG4, see Figure 22). On the basis of the high homology to O75385 (LBDG1), and the conservation of key predicted residues, BAA31598.1 (LBDG4) is also annotated as containing a Nuclear Hormone Receptor Ligand Binding Domain.

- 5 The combination of equivalent fold and conservation of "LBD motif" residues allows the functional annotation of this region of BAA31598.1, and therefore proteins that include this region, as possessing Nuclear Hormone Receptor Ligand Binding Domain activity.

In a second embodiment of the first aspect of the invention, there is thus provided a polypeptide, which polypeptide:

- 10 (i) comprises the amino acid sequence as recited in SEQ ID NO:4;
- (ii) is a fragment thereof having Nuclear Hormone Receptor Ligand Binding Domain activity or having an antigenic determinant in common with the polypeptides of (i); or
- (iii) is a functional equivalent of (i) or (ii).

- 15 Preferably, the polypeptide:

- (i) consists of the amino acid sequence as recited in SEQ ID NO:4;
- (ii) is a fragment thereof having Nuclear Hormone Receptor Ligand Binding Domain activity or having an antigenic determinant in common with the polypeptides of (i); or
- 20 (iii) is a functional equivalent of (i) or (ii).

The polypeptide having the sequence recited in SEQ ID NO:4 is referred to hereafter as "the LBDG4 polypeptide".

- According to this aspect of the invention, a preferred polypeptide fragment according to part ii) above includes the region of the LBDG4 polypeptide that is predicted as that responsible for Nuclear Hormone Receptor Ligand Binding Domain activity (hereafter, the "LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region"), or is a variant thereof that possesses the "LBD motif" (ILE863, ASP870, GLN871 and LEU875, or equivalent residues). As defined herein, the LBDG4 Nuclear Hormone Receptor
- 25

WO 02/070559

PCT/GB02/00948

10

Ligand Binding Domain region is considered to extend between residue 805 and residue 1005 of the LBDG4 polypeptide sequence.

This aspect of the invention also includes fusion proteins that incorporate polypeptide fragments and variants of these polypeptide fragments as defined above, provided that said fusion proteins possess activity as a Nuclear Hormone Receptor Ligand Binding Domain.

In a second aspect, the invention provides a purified nucleic acid molecule that encodes a polypeptide according to the first aspect of the invention. Preferably, the purified nucleic acid molecule has the nucleic acid sequence as recited in SEQ ID NO:1 (encoding the LBDG1 polypeptide) or SEQ ID NO:3 (encoding the LBDG4 polypeptide), or is a redundant equivalent or fragment of these sequences. A preferred nucleic acid fragment is one that encodes a polypeptide fragment according to part ii) above, preferably a polypeptide fragment that includes the LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain regions, or that encodes a variant of these fragments as this term is defined above.

In a third aspect, the invention provides a purified nucleic acid molecule which hybridizes under high stringency conditions with a nucleic acid molecule of the second aspect of the invention.

In a fourth aspect, the invention provides a vector, such as an expression vector, that contains a nucleic acid molecule of the second or third aspect of the invention.

In a fifth aspect, the invention provides a host cell transformed with a vector of the fourth aspect of the invention.

In a sixth aspect, the invention provides a ligand which binds specifically to, and which preferably inhibits the Nuclear Hormone Receptor Ligand Binding Domain activity of, a polypeptide of the first aspect of the invention.

In a seventh aspect, the invention provides a compound that is effective to alter the expression of a natural gene which encodes a polypeptide of the first aspect of the invention or to regulate the activity of a polypeptide of the first aspect of the invention.

A compound of the seventh aspect of the invention may either increase (agonise) or

WO 02/070559

PCT/GB02/00948

11

decrease (antagonise) the level of expression of the gene or the activity of the polypeptide. Importantly, the identification of the function of the region defined herein as the LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 or LBDG4 polypeptide, respectively, allows for the design of screening methods capable of identifying compounds that are effective in the treatment and/or diagnosis of diseases in which Nuclear Hormone Receptor Ligand Binding Domains are implicated.

In an eighth aspect, the invention provides a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a ligand of the fifth aspect of the invention, or a compound of the sixth aspect of the invention, for use in therapy or diagnosis. These molecules may also be used in the manufacture of a medicament for the treatment of cell proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and other solid tumours, myeloproliferative disorders, such as leukemia, non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposi's sarcoma, autoimmune/inflammatory disorders, including allergy, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation, asthma, and organ transplant rejection, cardiovascular disorders, including hypertension, oedema, angina, atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia, neurological disorders including, central nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain, developmental disorders, metabolic disorders including diabetes mellitus, osteoporosis, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, and obesity, renal disorders, including glomerulonephritis, renovascular hypertension, dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging, AIDS, infections including viral infection, bacterial infection, fungal infection and parasitic infection and other pathological conditions, particularly those in which nuclear hormone receptors are implicated.

In a ninth aspect, the invention provides a method of diagnosing a disease in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide of the first aspect of the invention or the activity of a polypeptide of the first aspect of the

WO 02/070559

PCT/GB02/00948

12

invention in tissue from said patient and comparing said level of expression or activity to a control level, wherein a level that is different to said control level is indicative of disease. Such a method will preferably be carried out *in vitro*. Similar methods may be used for monitoring the therapeutic treatment of disease in a patient, wherein altering the  
5 level of expression or activity of a polypeptide or nucleic acid molecule over the period of time towards a control level is indicative of regression of disease.

A preferred method for detecting polypeptides of the first aspect of the invention comprises the steps of: (a) contacting a ligand, such as an antibody, of the sixth aspect of the invention with a biological sample under conditions suitable for the formation of a  
10 ligand-polypeptide complex; and (b) detecting said complex.

A number of different such methods according to the ninth aspect of the invention exist, as the skilled reader will be aware, such as methods of nucleic acid hybridization with short probes, point mutation analysis, polymerase chain reaction (PCR) amplification and methods using antibodies to detect aberrant protein levels. Similar methods may be used  
15 on a short or long term basis to allow therapeutic treatment of a disease to be monitored in a patient. The invention also provides kits that are useful in these methods for diagnosing disease.

In a tenth aspect, the invention provides for the use of a polypeptide of the first aspect of the invention as a Nuclear Hormone Receptor Ligand Binding Domain. The invention  
20 also provides for the use of a nucleic acid molecule according to the second or third aspects of the invention to express a protein that possesses Nuclear Hormone Receptor Ligand Binding Domain activity. The invention also provides a method for effecting Nuclear Hormone Receptor Ligand Binding Domain activity, said method utilising a polypeptide of the first aspect of the invention.

In an eleventh aspect, the invention provides a pharmaceutical composition comprising a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a ligand of the sixth aspect of the invention, or a compound of the seventh aspect of the invention, in conjunction with a pharmaceutically-acceptable carrier.

30 In a twelfth aspect, the present invention provides a polypeptide of the first aspect of the

WO 02/070559

PCT/GB02/00948

13

invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a ligand of the sixth aspect of the invention, or a compound of the seventh aspect of the invention, for use in the manufacture of a medicament for the diagnosis or treatment of a disease, such as cell

5 proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and other solid tumours, myeloproliferative disorders, such as leukemia, non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposis' sarcoma, autoimmune/inflammatory disorders, including allergy, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation, asthma, and organ

10 transplant rejection, cardiovascular disorders, including hypertension, oedema, angina, atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia, neurological disorders including, central nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain, developmental disorders, metabolic disorders including diabetes mellitus, osteoporosis,

15 lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, and obesity, renal disorders, including glomerulonephritis, renovascular hypertension, dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging, AIDS, infections including viral infection, bacterial infection, fungal infection and parasitic infection and other

20 pathological conditions, particularly those in which nuclear hormone receptors are implicated.

In a thirteenth aspect, the invention provides a method of treating a disease in a patient comprising administering to the patient a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the

25 fourth aspect of the invention, or a ligand of the sixth aspect of the invention, or a compound of the seventh aspect of the invention.

For diseases in which the expression of a natural gene encoding a polypeptide of the first aspect of the invention, or in which the activity of a polypeptide of the first aspect of the invention, is lower in a diseased patient when compared to the level of expression or

30 activity in a healthy patient, the polypeptide, nucleic acid molecule, ligand or compound administered to the patient should be an agonist. Conversely, for diseases in which the

WO 02/070559

PCT/GB02/00948

14

expression of the natural gene or activity of the polypeptide is higher in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, ligand or compound administered to the patient should be an antagonist. Examples of such antagonists include antisense nucleic acid molecules, ribozymes and ligands, such as antibodies.

In a fourteenth aspect, the invention provides transgenic or knockout non-human animals that have been transformed to express higher, lower or absent levels of a polypeptide of the first aspect of the invention. Such transgenic animals are very useful models for the study of disease and may also be using in screening regimes for the identification of compounds that are effective in the treatment or diagnosis of such a disease.

A summary of standard techniques and procedures which may be employed in order to utilise the invention is given below. It will be understood that this invention is not limited to the particular methodology, protocols, cell lines, vectors and reagents described. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and it is not intended that this terminology should limit the scope of the present invention. The extent of the invention is limited only by the terms of the appended claims.

Standard abbreviations for nucleotides and amino acids are used in this specification.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA technology and immunology, which are within the skill of the those working in the art.

Such techniques are explained fully in the literature. Examples of particularly suitable texts for consultation include the following: Sambrook Molecular Cloning: A Laboratory Manual, Second Edition (1989); DNA Cloning, Volumes I and II (D.N. Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames & S.J. Higgins eds. 1984); Animal Cell Culture (R.I. Freshney ed. 1986); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide to Molecular Cloning (1984); the Methods in Enzymology series (Academic Press, Inc.), especially volumes 154 & 155; Gene Transfer Vectors for Mammalian Cells (J.H. Miller and M.P. Calos eds. 1987,

WO 02/070559

PCT/GB02/00948

15

Cold Spring Harbor Laboratory); Immunochemical Methods in Cell and Molecular Biology (Mayer and Walker, eds. 1987, Academic Press, London); Scopes, (1987) Protein Purification: Principles and Practice, Second Edition (Springer Verlag, N.Y.); and Handbook of Experimental Immunology, Volumes I-IV (D.M. Weir and C. C. Blackwell eds. 1986).

5 As used herein, the term "polypeptide" includes any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e. peptide isosteres. This term refers both to short chains (peptides and oligopeptides) and to longer chains (proteins).

10 The polypeptide of the present invention may be in the form of a mature protein or may be a pre-, pro- or prepro- protein that can be activated by cleavage of the pre-, pro- or prepro- portion to produce an active mature polypeptide. In such polypeptides, the pre-, pro- or prepro- sequence may be a leader or secretory sequence or may be a sequence that is employed for purification of the mature polypeptide sequence.

15 The polypeptide of the first aspect of the invention may form part of a fusion protein. For example, it is often advantageous to include one or more additional amino acid sequences which may contain secretory or leader sequences, pro-sequences, sequences which aid in purification, or sequences that confer higher protein stability, for example during recombinant production. Alternatively or additionally, the mature polypeptide may be  
20 fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol).

Polypeptides may contain amino acids other than the 20 gene-encoded amino acids, modified either by natural processes, such as by post-translational processing or by chemical modification techniques which are well known in the art. Among the known  
25 modifications which may commonly be present in polypeptides of the present invention are glycosylation, lipid attachment, sulphation, gamma-carboxylation, for instance of glutamic acid residues, hydroxylation and ADP-ribosylation. Other potential modifications include acetylation, acylation, amidation, covalent attachment of flavin, covalent attachment of a haeme moiety, covalent attachment of a nucleotide or nucleotide  
30 derivative, covalent attachment of a lipid derivative, covalent attachment of

WO 02/070559

PCT/GB02/00948

16

phosphatidylinositol, cross-linking, cyclization, disulphide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, GPI anchor formation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, 5 racemization, selenoylation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl terminus in a polypeptide, or both, by a covalent modification is common in 10 naturally-occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention.

The modifications that occur in a polypeptide often will be a function of how the polypeptide is made. For polypeptides that are made recombinantly, the nature and extent of the modifications in large part will be determined by the post-translational 15 modification capacity of the particular host cell and the modification signals that are present in the amino acid sequence of the polypeptide in question. For instance, glycosylation patterns vary between different types of host cell.

The polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally-occurring polypeptides (for example purified from 20 cell culture), recombinantly-produced polypeptides (including fusion proteins), synthetically-produced polypeptides or polypeptides that are produced by a combination of these methods.

The functionally-equivalent polypeptides of the first aspect of the invention may be polypeptides that are homologous to the LBDG1 or LBDG4 polypeptides. Two 25 polypeptides are said to be "homologous", as the term is used herein, if the sequence of one of the polypeptides has a high enough degree of identity or similarity to the sequence of the other polypeptide. "Identity" indicates that at any particular position in the aligned sequences, the amino acid residue is identical between the sequences. "Similarity" indicates that, at any particular position in the aligned sequences, the amino acid residue 30 is of a similar type between the sequences. Degrees of identity and similarity can be

WO 02/070559

PCT/GB02/00948

17

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

Homologous polypeptides therefore include natural biological variants (for example, allelic variants or geographical variations within the species from which the polypeptides 10 are derived) and mutants (such as mutants containing amino acid substitutions, insertions or deletions) of the LBDG1 polypeptide or the LBDG4 polypeptide. Such mutants may include polypeptides in which one or more of the amino acid residues are 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 15 genetic code. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; among the basic residues Lys and Arg; or among the aromatic residues Phe and Tyr. Particularly preferred are variants in which several, i.e. between 5 and 10, 1 and 5, 1 and 3, 1 and 2 or just 1 amino acids are substituted, deleted or added in any combination. Especially preferred are 20 silent substitutions, additions and deletions, which do not alter the properties and activities of the protein. Also especially preferred in this regard are conservative substitutions.

Such mutants also include polypeptides in which one or more of the amino acid residues includes a substituent group.

25 Typically, greater than 80% identity between two polypeptides (preferably, over a specified region) is considered to be an indication of functional equivalence. Preferably, functionally equivalent polypeptides of the first aspect of the invention have a degree of sequence identity with the LBDG1 polypeptide or the LBDG4 polypeptide, or with active fragments thereof, of greater than 80%. More preferred polypeptides have degrees of 30 identity of greater than 85%, 90%, 95%, 98% or 99%, respectively with the LBDG1 or

WO 02/070559

PCT/GB02/00948

18

LBDG4 polypeptides, or with active fragments thereof.

Percentage identity, as referred to herein, is as determined using BLAST version 2.1.3 using the default parameters specified by the NCBI (the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>) [Blosom 62 matrix; gap open penalty=11 and gap extension penalty=1].

In the present case, preferred active fragments of the LBDG1 or LBDG4 polypeptides are those that include the LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region and which possess the "LBD motif" of residues LEU878, ASP885, GLN886, LEU889 and LEU890, or equivalent residues. By "equivalent residues" is meant residues that are equivalent to the "LBD motif" residues, provided that the Nuclear Hormone Receptor Ligand Binding Domain region retains activity as a Nuclear Hormone Receptor Ligand Binding Domain. For example LEU878 may be replaced by ILE, ALA, VAL, MET, PHE, TYR or TRP. For example ASP885 may be replaced by GLU. For example GLN886 may be replaced by ASN, LYS, HIS, ARG, SER or THR. For example LEU889 may be replaced by ILE, ALA, VAL, MET, PHE, TYR or TRP. For example LEU890 may be replaced by ILE, ALA, VAL, MET, PHE, TYR or TRP. Residues may be replaced in a similar manner for LBDG4 (the active residues are ILE863, ASP870, GLN871 and LEU875). Accordingly, this aspect of the invention includes polypeptides that have degrees of identity of greater than 80%, preferably, greater than 85%, 90%, 95%, 98% or 99%, respectively, with the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 or LBDG4 polypeptide and which possess the "LBD motif" of LEU878, ASP885, GLN886, LEU889 and LEU890, or equivalent residues (ILE863, ASP870, GLN871 and LEU875 in the LBDG4 polypeptide). As discussed above, the LBDG1 Nuclear Hormone Receptor Ligand Binding Domain region is considered to extend between residue 822 and residue 1020 of the LBDG1 polypeptide sequence, while the LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region is considered to extend between residue 805 and residue 1005 of the LBDG4 polypeptide sequence.

The functionally-equivalent polypeptides of the first aspect of the invention may also be polypeptides which have been identified using one or more techniques of structural

WO 02/070559

PCT/GB02/00948

19

alignment. For example, the Inpharmatica Genome Threader™ technology that forms one aspect of the search tools used to generate the Biopendium search database may be used (see co-pending International patent application PCT/GB01/01105) to identify polypeptides of presently-unknown function which, while having low sequence identity as compared to the LBDG1 or LBDG4 polypeptide, are predicted to have Nuclear Hormone Receptor Ligand Binding Domain activity, by virtue of sharing significant structural homology with the LBDG1 or LBDG4 polypeptide sequence.

By "significant structural homology" is meant that the Inpharmatica Genome Threader™ predicts two proteins, or protein regions, to share structural homology with a certainty of at least 10% more preferably, at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% and above. The certainty value of the Inpharmatica Genome Threader™ is calculated as follows. A set of comparisons was initially performed using the Inpharmatica Genome Threader™ exclusively using sequences of known structure. Some of the comparisons were between proteins that were known to be related (on the basis of structure). A neural network was then trained on the basis that it needed to best distinguish between the known relationships and known not-relationships taken from the CATH structure classification ([www.biochem.ucl.ac.uk/bsm/cath](http://www.biochem.ucl.ac.uk/bsm/cath)). This resulted in a neural network score between 0 and 1. However, again as the number of proteins that are related and the number that are unrelated were known, it was possible to partition the neural network results into packets and calculate empirically the percentage of the results that were correct. In this manner, any genuine prediction in the Biopendium search database has an attached neural network score and the percentage confidence is a reflection of how successful the Inpharmatica Genome Threader™ was in the training/testing set.

Structural homologues of LBDG1 should share structural homology with the LBDG1 Nuclear Hormone Receptor Ligand Binding Domain region and possess the "LBD motif" residues LEU878, ASP885, GLN886, LEU889 and LEU890, or equivalent residues. Such structural homologues are predicted to have Nuclear Hormone Receptor Ligand Binding Domain activity by virtue of sharing significant structural homology with this polypeptide sequence and possessing the "LBD motif" residues.

Structural homologues of LBDG4 should share structural homology with the LBDG4

WO 02/070559

PCT/GB02/00948

20

Nuclear Hormone Receptor Ligand Binding Domain region and possess the "LBD motif" residues ILE863, ASP870, GLN871 and LEU875, or equivalent residues. Such structural homologues are predicted to have Nuclear Hormone Receptor Ligand Binding Domain activity by virtue of sharing significant structural homology with this polypeptide sequence and possessing the "LBD motif" residues.

The polypeptides of the first aspect of the invention also include fragments of the LBDG1 or LBDG4 polypeptides, functional equivalents of the fragments of the LBDG1 or LBDG4 polypeptide, and fragments of the functional equivalents of the LBDG1 or LBDG4 polypeptides, provided that those functional equivalents and fragments retain Nuclear Hormone Receptor Ligand Binding Domain activity or have an antigenic determinant in common with the LBDG1 or LBDG4 polypeptide.

As used herein, the term "fragment" refers to a polypeptide having an amino acid sequence that is the same as part, but not all, of the amino acid sequence of the LBDG1 or LBDG4 polypeptides or one of its functional equivalents. The fragments should comprise at least n consecutive amino acids from the sequence and, depending on the particular sequence, n preferably is 7 or more (for example, 8, 10, 12, 14, 16, 18, 20 or more). Small fragments may form an antigenic determinant.

Preferred polypeptide fragments according to this aspect of the invention are fragments that include a region defined herein as the LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 or LBDG4 polypeptides, respectively. These regions are the regions that have been annotated as Nuclear Hormone Receptor Ligand Binding Domain.

For the LBDG1 polypeptide, this region is considered to extend between residue 822 and residue 1020. For LBDG4 polypeptide, this region is considered to extend between residue 805 and residue 1005.

Variants of this fragment are included as embodiments of this aspect of the invention, provided that these variants possess activity as a Nuclear Hormone Receptor Ligand Binding Domain.

WO 02/070559

PCT/GB02/00948

21

In one respect, the term "variant" is meant to include extended or truncated versions of this polypeptide fragment.

For extended variants, it is considered highly likely that the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 or LBDG4 polypeptide will fold correctly and show Nuclear Hormone Receptor Ligand Binding Domain activity if additional residues C terminal and/or N terminal of these boundaries in the LBDG1 or LBDG4 polypeptide sequence are included in the polypeptide fragment. For example, an additional 5, 10, 20, 30, 40 or even 50 or more amino acid residues from the LBDG1 or LBDG4 polypeptide sequence, or from a homologous sequence, may be included at either or both the C terminal and/or N terminal of the boundaries of the Nuclear Hormone Receptor Ligand Binding Domain regions of the LBDG1 or LBDG4 polypeptide, without prejudicing the ability of the polypeptide fragment to fold correctly and exhibit Nuclear Hormone Receptor Ligand Binding Domain activity.

For truncated variants of the LBDG1 or LBDG4 polypeptide, one or more amino acid residues may be deleted at either or both the C terminus or the N terminus of the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 polypeptide, although the "LBD motif" residues (LEU878, ASP885, GLN886, LEU889 and LEU890 for LBDG1; these residues are ILE863, ASP870, GLN871 and LEU875 for LBDG4), or equivalent residues should be maintained intact; deletions should not extend so far into the polypeptide sequence that any of these residues are deleted.

In a second respect, the term "variant" includes homologues of the polypeptide fragments described above, that possess significant sequence homology with the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 or LBDG4 polypeptide and which possess the "LBD motif" residues (LEU878, ASP885, GLN886, LEU889 and LEU890; for the LBDG1 polypeptide, these residues are ILE863, ASP870, GLN871 and LEU875), or equivalent residues, provided that said variants retain activity as an Nuclear Hormone Receptor Ligand Binding Domain.

Homologues include those polypeptide molecules that possess greater than 80% identity with the LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain regions, of the LBDG1 or LBDG4 polypeptides, respectively. Percentage identity is as

WO 02/070559

PCT/GB02/00948

22

determined using BLAST version 2.1.3 using the default parameters specified by the NCBI (the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>) [Blosum 62 matrix; gap open penalty=11 and gap extension penalty=1]. Preferably, variant homologues of polypeptide fragments of this aspect of the invention have a degree of sequence identity with the LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain regions, of the LBDG1 or LBDG4 polypeptides, respectively, of greater than 80%. More preferred variant polypeptides have degrees of identity of greater than 85%, 90%, 95%, 98% or 99%, respectively with the LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain regions of the LBDG1 or LBDG4 polypeptides, provided that said variants retain activity as a Nuclear Hormone Receptor Ligand Binding Domain. Variant polypeptides also include homologues of the truncated forms of the polypeptide fragments discussed above, provided that said variants retain activity as a Nuclear Hormone Receptor Ligand Binding Domain.

15 The polypeptide fragments of the first aspect of the invention may be polypeptide fragments that exhibit significant structural homology with the structure of the polypeptide fragment defined by the LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain regions, of the LBDG1 or LBDG4 polypeptide sequence, for example, as identified by the Inpharmatica Genome Threader™. Accordingly, 20 polypeptide fragments that are structural homologues of the polypeptide fragments defined by the LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain regions of the LBDG1 or LBDG4 polypeptide sequence should adopt the same fold as that adopted by this polypeptide fragment, as this fold is defined above.

Structural homologues of the polypeptide fragment defined by the LBDG1 Nuclear Hormone Receptor Ligand Binding Domain region should also retain the "LBD motif" residues LEU878, ASP885, GLN886, LEU889 and LEU890, or equivalent residues. 25

Structural homologues of the polypeptide fragment defined by the LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region should also retain the "LBD motif" residues ILE863, ASP870, GLN871 and LEU875, or equivalent residues.

WO 02/070559

PCT/GB02/00948

23

Such fragments may be "free-standing", i.e. not part of or fused to other amino acids or polypeptides, or they may be comprised within a larger polypeptide of which they form a part or region. When comprised within a larger polypeptide, the fragment of the invention most preferably forms a single continuous region. For instance, certain preferred  
5 embodiments relate to a fragment having a pre- and/or pro- polypeptide region fused to the amino terminus of the fragment and/or an additional region fused to the carboxyl terminus of the fragment. However, several fragments may be comprised within a single larger polypeptide.

The polypeptides of the present invention or their immunogenic fragments (comprising at least one antigenic determinant) can be used to generate ligands, such as polyclonal or monoclonal antibodies, that are immunospecific for the polypeptides. Such antibodies may be employed to isolate or to identify clones expressing the polypeptides of the invention or to purify the polypeptides by affinity chromatography. The antibodies may also be employed as diagnostic or therapeutic aids, amongst other applications, as will be  
15 apparent to the skilled reader.

The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides of the prior art. As used herein, the term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, F(ab')<sub>2</sub> and Fv, which are capable of binding to the antigenic determinant in question. Such antibodies thus bind to the polypeptides of the  
20 first aspect of the invention.

If polyclonal antibodies are desired, a selected mammal, such as a mouse, rabbit, goat or horse, may be immunised with a polypeptide of the first aspect of the invention. The polypeptide used to immunise the animal can be derived by recombinant DNA technology or can be synthesized chemically. If desired, the polypeptide can be  
25 conjugated to a carrier protein. Commonly used carriers to which the polypeptides may be chemically coupled include bovine serum albumin, thyroglobulin and keyhole limpet haemocyanin. The coupled polypeptide is then used to immunise the animal. Serum from the immunised animal is collected and treated according to known procedures, for example by immunoaffinity chromatography.  
30

WO 02/070559

PCT/GB02/00948

24

Monoclonal antibodies to the polypeptides of the first aspect of the invention can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies using hybridoma technology is well known (see, for example, Kohler, G. and Milstein, C., *Nature* 256: 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, 77-96 in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985).

Panels of monoclonal antibodies produced against the polypeptides of the first aspect of the invention can be screened for various properties, i.e., for isotype, epitope, affinity, etc. Monoclonal antibodies are particularly useful in purification of the individual polypeptides against which they are directed. Alternatively, genes encoding the monoclonal antibodies of interest may be isolated from hybridomas, for instance by PCR techniques known in the art, and cloned and expressed in appropriate vectors.

Chimeric antibodies, in which non-human variable regions are joined or fused to human constant regions (see, for example, Liu *et al.*, *Proc. Natl. Acad. Sci. USA*, 84, 3439 (1987)), may also be of use.

The antibody may be modified to make it less immunogenic in an individual, for example by humanisation (see Jones *et al.*, *Nature*, 321, 522 (1986); Verhoeven *et al.*, *Science*, 239: 1534 (1988); Kabat *et al.*, *J. Immunol.*, 147: 1709 (1991); Queen *et al.*, *Proc. Natl. Acad. Sci. USA*, 86, 10029 (1989); Gorman *et al.*, *Proc. Natl. Acad. Sci. USA*, 88: 34181 (1991); and Hodgson *et al.*, *Bio/Technology* 9: 421 (1991)). The term "humanised antibody", as used herein, refers to antibody molecules in which the CDR amino acids and selected other amino acids in the variable domains of the heavy and/or light chains of a non-human donor antibody have been substituted in place of the equivalent amino acids in a human antibody. The humanised antibody thus closely resembles a human antibody but has the binding ability of the donor antibody.

In a further alternative, the antibody may be a "bispecific" antibody, that is an antibody having two different antigen binding domains, each domain being directed against a different epitope.

Phage display technology may be utilised to select genes which encode antibodies with binding activities towards the polypeptides of the invention either from repertoires of

WO 02/070559

PCT/GB02/00948

25

PCR amplified V-genes of lymphocytes from humans screened for possessing the relevant antibodies, or from naive libraries (McCafferty, J. *et al.*, (1990), Nature 348, 552-554; Marks, J. *et al.*, (1992) Biotechnology 10, 779-783). The affinity of these antibodies can also be improved by chain shuffling (Clackson, T. *et al.*, (1991) Nature  
5 352, 624-628).

Antibodies generated by the above techniques, whether polyclonal or monoclonal, have additional utility in that they may be employed as reagents in immunoassays, radioimmunoassays (RIA) or enzyme-linked immunosorbent assays (ELISA). In these applications, the antibodies can be labelled with an analytically-detectable reagent such  
10 as a radioisotope, a fluorescent molecule or an enzyme.

Preferred nucleic acid molecules of the second and third aspects of the invention are those which encode the polypeptide sequences recited in SEQ ID NO:2 or SEQ ID NO:4, and functionally equivalent polypeptides, including active fragments of the LBDG1 or LBDG4 polypeptide, such as a fragment including the LBDG1 or LBDG4 Nuclear  
15 Hormone Receptor Ligand Binding Domain region of the LBDG1 or LBDG4 polypeptide sequence, or a homologue thereof.

Nucleic acid molecules encompassing these stretches of sequence form a preferred embodiment of this aspect of the invention.

These nucleic acid molecules may be used in the methods and applications described  
20 herein. The nucleic acid molecules of the invention preferably comprise at least n consecutive nucleotides from the sequences disclosed herein where, depending on the particular sequence, n is 10 or more (for example, 12, 14, 15, 18, 20, 25, 30, 35, 40 or more).

The nucleic acid molecules of the invention also include sequences that are  
25 complementary to nucleic acid molecules described above (for example, for antisense or probing purposes).

Nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance cDNA, synthetic DNA or genomic DNA. Such nucleic acid molecules may be obtained by cloning, by chemical

WO 02/070559

PCT/GB02/00948

26

synthetic techniques or by a combination thereof. The nucleic acid molecules can be prepared, for example, by chemical synthesis using techniques such as solid phase phosphoramidite chemical synthesis, from genomic or cDNA libraries or by separation from an organism. RNA molecules may generally be generated by the *in vitro* or *in vivo* transcription of DNA sequences.

The nucleic acid molecules may be double-stranded or single-stranded. Single-stranded DNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

The term "nucleic acid molecule" also includes analogues of DNA and RNA, such as those containing modified backbones, and peptide nucleic acids (PNA). The term "PNA", as used herein, refers to an antisense molecule or an anti-gene agent which comprises an oligonucleotide of at least five nucleotides in length linked to a peptide backbone of amino acid residues, which preferably ends in lysine. The terminal lysine confers solubility to the composition. PNAs may be pegylated to extend their lifespan in a cell, where they preferentially bind complementary single stranded DNA and RNA and stop transcript elongation (Nielsen, P.E. *et al.* (1993) *Anticancer Drug Des.* 8:53-63).

A nucleic acid molecule which encodes the polypeptide of SEQ ID NO:2, or an active fragment thereof, may be identical to the coding sequence of the nucleic acid molecule shown in SEQ ID NO:1. These molecules also may have a different sequence which, as a result of the degeneracy of the genetic code, encodes the polypeptide SEQ ID NO:2, or an active fragment of the LBDG1 polypeptide, such as a fragment including the LBDG1 Nuclear Hormone Receptor Ligand Binding Domain region, or a homologue thereof. The LBDG1 Nuclear Hormone Receptor Ligand Binding Domain region is considered to extend between residue 822 and residue 1020 of the LBDG1 polypeptide sequence. In SEQ ID NO:1 the LBDG1 Nuclear Hormone Receptor Ligand Binding Domain region is thus encoded by a nucleic acid molecule including nucleotide 2732 to 3328. Nucleic acid molecules encompassing this stretch of sequence, and homologues of this sequence, form a preferred embodiment of this aspect of the invention.

A nucleic acid molecule which encodes the polypeptide of SEQ ID NO:4, or an active fragment thereof, may be identical to the coding sequence of the nucleic acid molecule

WO 02/070559

PCT/GB02/00948

27

shown in SEQ ID NO:3. These molecules also may have a different sequence which, as a result of the degeneracy of the genetic code, encodes the polypeptide SEQ ID NO:4, or an active fragment of the LBDG4 polypeptide, such as a fragment including the LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region, or a homologue thereof. The LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region is considered to extend between residue 805 and residue 1005 of the LBDG4 polypeptide sequence. In SEQ ID NO:3 the LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region is thus encoded by a nucleic acid molecule including nucleotide 2913 to 3515. Nucleic acid molecules encompassing this stretch of sequence, and homologues of this sequence, form a preferred embodiment of this aspect of the invention.

Such nucleic acid molecules that encode the polypeptide of SEQ ID NO:2 or SEQ ID NO:4 may include, but are not limited to, the coding sequence for the mature polypeptide by itself; the coding sequence for the mature polypeptide and additional coding sequences, such as those encoding a leader or secretory sequence, such as a pro-, pre- or prepro- polypeptide sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences, together with further additional, non-coding sequences, including non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription (including termination signals), ribosome binding and mRNA stability. The nucleic acid molecules may also include additional sequences which encode additional amino acids, such as those which provide additional functionalities.

The nucleic acid molecules of the second and third aspects of the invention may also encode the fragments or the functional equivalents of the polypeptides and fragments of the first aspect of the invention.

As discussed above, a preferred fragment of the LBDG1 polypeptide is a fragment including the LBDG1 Nuclear Hormone Receptor Ligand Binding Domain region, or a homologue thereof. The Nuclear Hormone Receptor Ligand Binding Domain region is encoded by a nucleic acid molecule including nucleotide 2732 to 3328 of SEQ ID NO:1.

A preferred fragment of the LBDG4 polypeptide is a fragment including the LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region, or a homologue thereof. The

WO 02/070559

PCT/GB02/00948

28

Nuclear Hormone Receptor Ligand Binding Domain region is encoded by a nucleic acid molecule including nucleotide 2913 to 3515 of SEQ ID NO:3.

Functionally equivalent nucleic acid molecules according to the invention may be naturally-occurring variants such as a naturally-occurring allelic variant, or the molecules  
5 may be a variant that is not known to occur naturally. Such non-naturally occurring variants of the nucleic acid molecule may be made by mutagenesis techniques, including those applied to nucleic acid molecules, cells or organisms.

Among variants in this regard are variants that differ from the aforementioned nucleic acid molecules by nucleotide substitutions, deletions or insertions. The substitutions,  
10 deletions or insertions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or insertions.

The nucleic acid molecules of the invention can also be engineered, using methods generally known in the art, for a variety of reasons, including modifying the cloning,  
15 processing, and/or expression of the gene product (the polypeptide). DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides are included as techniques which may be used to engineer the nucleotide sequences. Site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce  
20 mutations and so forth.

Nucleic acid molecules which encode a polypeptide of the first aspect of the invention may be ligated to a heterologous sequence so that the combined nucleic acid molecule encodes a fusion protein. Such combined nucleic acid molecules are included within the  
25 second or third aspects of the invention. For example, to screen peptide libraries for inhibitors of the activity of the polypeptide, it may be useful to express, using such a combined nucleic acid molecule, a fusion protein that can be recognised by a commercially-available antibody. A fusion protein may also be engineered to contain a cleavage site located between the sequence of the polypeptide of the invention and the sequence of a heterologous protein so that the polypeptide may be cleaved and purified  
30 away from the heterologous protein.

WO 02/070559

PCT/GB02/00948

29

The nucleic acid molecules of the invention also include antisense molecules that are partially complementary to nucleic acid molecules encoding polypeptides of the present invention and that therefore hybridize to the encoding nucleic acid molecules (hybridization). Such antisense molecules, such as oligonucleotides, can be designed to recognise, specifically bind to and prevent transcription of a target nucleic acid encoding a polypeptide of the invention, as will be known by those of ordinary skill in the art (see, for example, Cohen, J.S., Trends in Pharm. Sci., 10, 435 (1989), Okano, J. Neurochem. 56, 560 (1991); O'Connor, J. Neurochem 56, 560 (1991); Lee *et al.*, Nucleic Acids Res 6, 3073 (1979); Cooney *et al.*, Science 241, 456 (1988); Dervan *et al.*, Science 251, 1360 (1991).

The term "hybridization" as used here refers to the association of two nucleic acid molecules with one another by hydrogen bonding. Typically, one molecule will be fixed to a solid support and the other will be free in solution. Then, the two molecules may be placed in contact with one another under conditions that favour hydrogen bonding. Factors that affect this bonding include: the type and volume of solvent; reaction temperature; time of hybridization; agitation; agents to block the non-specific attachment of the liquid phase molecule to the solid support (Denhardt's reagent or BLOTTO); the concentration of the molecules; use of compounds to increase the rate of association of molecules (dextran sulphate or polyethylene glycol); and the stringency of the washing conditions following hybridization (see Sambrook *et al.* [*supra*]).

The inhibition of hybridization of a completely complementary molecule to a target molecule may be examined using a hybridization assay, as known in the art (see, for example, Sambrook *et al.* [*supra*]). A substantially homologous molecule will then compete for and inhibit the binding of a completely homologous molecule to the target molecule under various conditions of stringency, as taught in Wahl, G.M. and S.L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A.R. (1987; Methods Enzymol. 152:507-511).

"Stringency" refers to conditions in a hybridization reaction that favour the association of very similar molecules over association of molecules that differ. High stringency hybridisation conditions are defined as overnight incubation at 42°C in a solution

WO 02/070559

PCT/GB02/00948

30

comprising 50% formamide, 5XSSC (150mM NaCl, 15mM trisodium citrate), 50mM sodium phosphate (pH7.6), 5x Denhardts solution, 10% dextran sulphate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1X SSC at approximately 65°C. Low stringency conditions involve the hybridisation  
5 reaction being carried out at 35°C (see Sambrook *et al.* [*supra*]). Preferably, the conditions used for hybridization are those of high stringency.

Preferred embodiments of this aspect of the invention are nucleic acid molecules that are at least 80% identical over their entire length to a nucleic acid molecule encoding the LBDG1 polypeptide (SEQ ID NO:2) or LBDG4 polypeptide (SEQ ID NO:4), and nucleic  
10 acid molecules that are substantially complementary to such nucleic acid molecules. A preferred active fragment is a fragment that includes an LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 or LBDG4 polypeptide sequences, respectively. Accordingly, preferred nucleic acid molecules include those that are at least 80% identical over their entire length to a nucleic acid  
15 molecule encoding the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 or LBDG4 polypeptide sequence.

Percentage identity, as referred to herein, is as determined using BLAST version 2.1.3 using the default parameters specified by the NCBI (the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>).

20 Preferably, a nucleic acid molecule according to this aspect of the invention comprises a region that is at least 80% identical over its entire length to the nucleic acid molecule having the sequence given in SEQ ID NO:1, to a region including nucleotides 2732-3328 of this sequence. Other preferred nucleic acid molecules according to this aspect of the invention are those that comprise a region that is at least 80% identical over its entire  
25 length to the nucleic acid molecule having the sequence given in SEQ ID NO:3, to a region including nucleotides 2913-3515 of this sequence, or a nucleic acid molecule that is complementary to any one of these regions of nucleic acid. In this regard, nucleic acid molecules at least 90%, preferably at least 95%, more preferably at least 98% or 99% identical over their entire length to the same are particularly preferred. Preferred  
30 embodiments in this respect are nucleic acid molecules that encode polypeptides which

WO 02/070559

PCT/GB02/00948

31

retain substantially the same biological function or activity as the LBDG1 or LBDG4 polypeptide.

The invention also provides a process for detecting a nucleic acid molecule of the invention, comprising the steps of: (a) contacting a nucleic probe according to the invention with a biological sample under hybridizing conditions to form duplexes; and  
5 (b) detecting any such duplexes that are formed.

As discussed additionally below in connection with assays that may be utilised according to the invention, a nucleic acid molecule as described above may be used as a hybridization probe for RNA, cDNA or genomic DNA, in order to isolate full-length  
10 cDNAs and genomic clones encoding the LBDG1 or LBDG4 polypeptide and to isolate cDNA and genomic clones of homologous or orthologous genes that have a high sequence similarity to the gene encoding this polypeptide.

In this regard, the following techniques, among others known in the art, may be utilised and are discussed below for purposes of illustration. Methods for DNA sequencing and  
15 analysis are well known and are generally available in the art and may, indeed, be used to practice many of the embodiments of the invention discussed herein. Such methods may employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase (US Biochemical Corp, Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham, Chicago, IL), or combinations of polymerases and proof-reading  
20 exonucleases such as those found in the ELONGASE Amplification System marketed by Gibco/BRL (Gaithersburg, MD). Preferably, the sequencing process may be automated using machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, NV), the Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA) and the ABI Catalyst and 373 and 377 DNA Sequencers (Perkin Elmer).

25 One method for isolating a nucleic acid molecule encoding a polypeptide with an equivalent function to that of the LBDG1 or LBDG4 polypeptide, particularly with an equivalent function to the LBDG1 or LBDG4 Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 or LBDG4 polypeptide, is to probe a genomic or cDNA library with a natural or artificially-designed probe using standard procedures that  
30 are recognised in the art (see, for example, "Current Protocols in Molecular Biology",

WO 02/070559

PCT/GB02/00948

32

Ausubel *et al.* (eds). Greene Publishing Association and John Wiley Interscience, New York, 1989,1992). Probes comprising at least 15, preferably at least 30, and more preferably at least 50, contiguous bases that correspond to, or are complementary to, nucleic acid sequences from the appropriate encoding gene (SEQ ID NO:1 or SEQ ID NO:3), particularly a region from nucleotides 2732-3328 of SEQ ID NO:1, or nucleotides 2913-3515 of SEQ ID NO:3, are particularly useful probes.

Such probes may be labelled with an analytically-detectable reagent to facilitate their identification. Useful reagents include, but are not limited to, radioisotopes, fluorescent dyes and enzymes that are capable of catalysing the formation of a detectable product. Using these probes, the ordinarily skilled artisan will be capable of isolating complementary copies of genomic DNA, cDNA or RNA polynucleotides encoding proteins of interest from human, mammalian or other animal sources and screening such sources for related sequences, for example, for additional members of the family, type and/or subtype.

In many cases, isolated cDNA sequences will be incomplete, in that the region encoding the polypeptide will be cut short, normally at the 5' end. Several methods are available to obtain full length cDNAs, or to extend short cDNAs. Such sequences may be extended utilising a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed is based on the method of Rapid Amplification of cDNA Ends (RACE; see, for example, Frohman *et al.*, Proc. Natl. Acad. Sci. USA (1988) 85: 8998-9002). Recent modifications of this technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.), for example, have significantly simplified the search for longer cDNAs. A slightly different technique, termed "restriction-site" PCR, uses universal primers to retrieve unknown nucleic acid sequence adjacent a known locus (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). Inverse PCR may also be used to amplify or to extend sequences using divergent primers based on a known region (Triglia, T., *et al.* (1988) Nucleic Acids Res. 16:8186). Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent a known sequence in human and yeast artificial chromosome DNA (Lagerstrom, M. *et al.* (1991) PCR Methods Applic. 1: 111-119). Another method which may be used to retrieve

WO 02/070559

PCT/GB02/00948

33

unknown sequences is that of Parker, J.D. *et al.* (1991); Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PromoterFinder™ libraries to walk genomic DNA (Clontech, Palo Alto, CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

- 5 When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they will contain more sequences that contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for  
10 extension of sequence into 5' non-transcribed regulatory regions.

In one embodiment of the invention, the nucleic acid molecules of the present invention may be used for chromosome localisation. In this technique, a nucleic acid molecule is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to  
15 the present invention is an important step in the confirmatory correlation of those sequences with the gene-associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins  
20 University Welch Medical Library). The relationships between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes). This provides valuable information to investigators searching for disease genes using positional cloning or other gene  
25 discovery techniques. Once the disease or syndrome has been crudely localised by genetic linkage to a particular genomic region, any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleic acid molecule may also be used to detect differences in the chromosomal location due to  
translocation, inversion, etc. among normal, carrier, or affected individuals.

The nucleic acid molecules of the present invention are also valuable for tissue  
30 localisation. Such techniques allow the determination of expression patterns of the

WO 02/070559

PCT/GB02/00948

34

polypeptide in tissues by detection of the mRNAs that encode them. These techniques include in situ hybridization techniques and nucleotide amplification techniques, such as PCR. Results from these studies provide an indication of the normal functions of the polypeptide in the organism. In addition, comparative studies of the normal expression  
5 pattern of mRNAs with that of mRNAs encoded by a mutant gene provide valuable insights into the role of mutant polypeptides in disease. Such inappropriate expression may be of a temporal, spatial or quantitative nature.

The vectors of the present invention comprise nucleic acid molecules of the invention and may be cloning or expression vectors. The host cells of the invention, which may be  
10 transformed, transfected or transduced with the vectors of the invention may be prokaryotic or eukaryotic.

The polypeptides of the invention may be prepared in recombinant form by expression of their encoding nucleic acid molecules in vectors contained within a host cell. Such expression methods are well known to those of skill in the art and many are described in  
15 detail by Sambrook *et al.* (*supra*) and Fernandez & Hoeffler (1998, eds. "Gene expression systems. Using nature for the art of expression". Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto).

Generally, any system or vector that is suitable to maintain, propagate or express nucleic acid molecules to produce a polypeptide in the required host may be used. The  
20 appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those described in Sambrook *et al.*, (*supra*). Generally, the encoding gene can be placed under the control of a control element such as a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator, so that the DNA sequence encoding the desired polypeptide  
25 is transcribed into RNA in the transformed host cell.

Examples of suitable expression systems include, for example, chromosomal, episomal and virus-derived systems, including, for example, vectors derived from: bacterial plasmids, bacteriophage, transposons, yeast episomes, insertion elements, yeast  
30 chromosomal elements, viruses such as baculoviruses, papova viruses such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses,

WO 02/070559

PCT/GB02/00948

35

or combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, including cosmids and phagemids. Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained and expressed in a plasmid.

- 5 Particularly suitable expression systems include microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (for example, baculovirus); plant cell systems transformed with virus expression vectors (for example, cauliflower mosaic virus, CaMV; tobacco
- 10 mosaic virus, TMV) or with bacterial expression vectors (for example, Ti or pBR322 plasmids); or animal cell systems. Cell-free translation systems can also be employed to produce the polypeptides of the invention.

Introduction of nucleic acid molecules encoding a polypeptide of the present invention into host cells can be effected by methods described in many standard laboratory

15 manuals, such as Davis *et al.*, *Basic Methods in Molecular Biology* (1986) and Sambrook *et al.*, (*supra*). Particularly suitable methods include calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection (see Sambrook *et al.*, 1989 [*supra*]; Ausubel *et al.*, 1991 [*supra*]; Spector,

20 Goldman & Leinwald, 1998). In eukaryotic cells, expression systems may either be transient (for example, episomal) or permanent (chromosomal integration) according to the needs of the system.

The encoding nucleic acid molecule may or may not include a sequence encoding a control sequence, such as a signal peptide or leader sequence, as desired, for example, for

25 secretion of the translated polypeptide into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment. These signals may be endogenous to the polypeptide or they may be heterologous signals. Leader sequences can be removed by the bacterial host in post-translational processing.

In addition to control sequences, it may be desirable to add regulatory sequences that

30 allow for regulation of the expression of the polypeptide relative to the growth of the host

WO 02/070559

PCT/GB02/00948

36

cell. Examples of regulatory sequences are those which cause the expression of a gene to be increased or decreased in response to a chemical or physical stimulus, including the presence of a regulatory compound or to various temperature or metabolic conditions. Regulatory sequences are those non-translated regions of the vector, such as enhancers, promoters and 5' and 3' untranslated regions. These interact with host cellular proteins to carry out transcription and translation. Such regulatory sequences may vary in their strength and specificity. Depending on the vector system and host utilised, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript phagemid (Stratagene, LaJolla, CA) or pSport1<sup>TM</sup> plasmid (Gibco BRL) and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (for example, heat shock, RUBISCO and storage protein genes) or from plant viruses (for example, viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of the sequence, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

An expression vector is constructed so that the particular nucleic acid coding sequence is located in the vector with the appropriate regulatory sequences, the positioning and orientation of the coding sequence with respect to the regulatory sequences being such that the coding sequence is transcribed under the "control" of the regulatory sequences, i.e., RNA polymerase which binds to the DNA molecule at the control sequences transcribes the coding sequence. In some cases it may be necessary to modify the sequence so that it may be attached to the control sequences with the appropriate orientation; i.e., to maintain the reading frame.

The control sequences and other regulatory sequences may be ligated to the nucleic acid coding sequence prior to insertion into a vector. Alternatively, the coding sequence can be cloned directly into an expression vector that already contains the control sequences and an appropriate restriction site.

WO 02/070559

PCT/GB02/00948

37

For long-term, high-yield production of a recombinant polypeptide, stable expression is preferred. For example, cell lines which stably express the polypeptide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells that successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalised cell lines available from the American Type Culture Collection (ATCC) including, but not limited to, Chinese hamster ovary (CHO), HeLa, baby hamster kidney (BHK), monkey kidney (COS), C127, 3T3, BHK, HEK 293, Bowes melanoma and human hepatocellular carcinoma (for example Hep G2) cells and a number of other cell lines.

In the baculovirus system, the materials for baculovirus/insect cell expression systems are commercially available in kit form from, inter alia, Invitrogen, San Diego CA (the "MaxBac" kit). These techniques are generally known to those skilled in the art and are described fully in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987). Particularly suitable host cells for use in this system include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells.

There are many plant cell culture and whole plant genetic expression systems known in the art. Examples of suitable plant cellular genetic expression systems include those described in US 5,693,506; US 5,659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture has been described by Zenk, (1991) *Phytochemistry* 30, 3861-3863.

In particular, all plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be utilised, so that whole plants are recovered which contain the transferred gene. Practically all plants can be regenerated from cultured cells or tissues,

WO 02/070559

PCT/GB02/00948

38

including but not limited to all major species of sugar cane, sugar beet, cotton, fruit and other trees, legumes and vegetables.

Examples of particularly preferred bacterial host cells include streptococci, staphylococci, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells.

- 5 Examples of particularly suitable host cells for fungal expression include yeast cells (for example, *S. cerevisiae*) and *Aspergillus* cells.

Any number of selection systems are known in the art that may be used to recover transformed cell lines. Examples include the herpes simplex virus thymidine kinase (Wigler, M. *et al.* (1977) Cell 11:223-32) and adenine phosphoribosyltransferase (Lowy, 10 I. *et al.* (1980) Cell 22:817-23) genes that can be employed in tk- or aprt± cells, respectively.

- Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dihydrofolate reductase (DHFR) that confers resistance to methotrexate (Wigler, M. *et al.* (1980) Proc. Natl. Acad. Sci. 77:3567-70); npt, which 15 confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. *et al.* (1981) J. Mol. Biol. 150:1-14) and als or pat, which confer resistance to chloresulfuron and phosphinotricin acetyltransferase, respectively. Additional selectable genes have been described, examples of which will be clear to those of skill in the art.

- Although the presence or absence of marker gene expression suggests that the gene of 20 interest is also present, its presence and expression may need to be confirmed. For example, if the relevant sequence is inserted within a marker gene sequence, transformed cells containing the appropriate sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a polypeptide of the invention under the control of a single promoter. 25 Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells that contain a nucleic acid sequence encoding a polypeptide of the invention and which express said polypeptide may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not

WO 02/070559

PCT/GB02/00948

39

limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassays, for example, fluorescence activated cell sorting (FACS) or immunoassay techniques (such as the enzyme-linked immunosorbent assay [ELISA] and radioimmunoassay [RIA]), that include membrane, solution, or chip based technologies for the detection and/or  
5 quantification of nucleic acid or protein (see Hampton, R. *et al.* (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN) and Maddox, D.E. *et al.* (1983) J. Exp. Med, 158, 1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing  
10 labelled hybridization or PCR probes for detecting sequences related to nucleic acid molecules encoding polypeptides of the present invention include oligolabelling, nick translation, end-labelling or PCR amplification using a labelled polynucleotide. Alternatively, the sequences encoding the polypeptide of the invention may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art,  
15 are commercially available, and may be used to synthesise RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3 or SP6 and labelled nucleotides. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, MI); Promega (Madison WI); and U.S. Biochemical Corp., Cleveland, OH)).

20 Suitable reporter molecules or labels, which may be used for ease of detection, include radionuclides, enzymes and fluorescent, chemiluminescent or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Nucleic acid molecules according to the present invention may also be used to create transgenic animals, particularly rodent animals. Such transgenic animals form a further  
25 aspect of the present invention. This may be done locally by modification of somatic cells, or by germ line therapy to incorporate heritable modifications. Such transgenic animals may be particularly useful in the generation of animal models for drug molecules effective as modulators of the polypeptides of the present invention.

The polypeptide can be recovered and purified from recombinant cell cultures by well-  
30 known methods including ammonium sulphate or ethanol precipitation, acid extraction,

WO 02/070559

PCT/GB02/00948

40

anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. High performance liquid chromatography is particularly useful for purification. Well known techniques for refolding proteins may be employed to regenerate an active conformation when the polypeptide is denatured during isolation and or purification.

Specialised vector constructions may also be used to facilitate purification of proteins, as desired, by joining sequences encoding the polypeptides of the invention to a nucleotide sequence encoding a polypeptide domain that will facilitate purification of soluble proteins. Examples of such purification-facilitating domains include metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilised metals, protein A domains that allow purification on immobilised immunoglobulin, and the domain utilised in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, WA). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the polypeptide of the invention may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing the polypeptide of the invention fused to several histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilised metal ion affinity chromatography as described in Porath, J. *et al.* (1992) *Prot. Exp. Purif.* 3: 263-281) while the thioredoxin or enterokinase cleavage site provides a means for purifying the polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D.J. *et al.* (*DNA Cell Biol.* 199312:441-453).

If the polypeptide is to be expressed for use in screening assays, generally it is preferred that it be produced at the surface of the host cell in which it is expressed. In this event, the host cells may be harvested prior to use in the screening assay, for example using techniques such as fluorescence activated cell sorting (FACS) or immunoaffinity techniques. If the polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the expressed polypeptide. If polypeptide is produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

WO 02/070559

PCT/GB02/00948

41

The polypeptide of the invention can be used to screen libraries of compounds in any of a variety of drug screening techniques. Such compounds may activate (agonise) or inhibit (antagonise) the level of expression of the gene or the activity of the polypeptide of the invention and form a further aspect of the present invention. Preferred compounds are effective to alter the expression of a natural gene which encodes a polypeptide of the first aspect of the invention or to regulate the activity of a polypeptide of the first aspect of the invention.

Agonist or antagonist compounds may be isolated from, for example, cells, cell-free preparations, chemical libraries or natural product mixtures. These agonists or antagonists may be natural or modified substrates, ligands, enzymes, receptors or structural or functional mimetics. For a suitable review of such screening techniques, see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991).

Compounds that are most likely to be good antagonists are molecules that bind to the polypeptide of the invention without inducing the biological effects of the polypeptide upon binding to it. Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to the polypeptide of the invention and thereby inhibit or extinguish its activity. In this fashion, binding of the polypeptide to normal cellular binding molecules may be inhibited, such that the normal biological activity of the polypeptide is prevented.

The polypeptide of the invention that is employed in such a screening technique may be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. In general, such screening procedures may involve using appropriate cells or cell membranes that express the polypeptide that are contacted with a test compound to observe binding, or stimulation or inhibition of a functional response. The functional response of the cells contacted with the test compound is then compared with control cells that were not contacted with the test compound. Such an assay may assess whether the test compound results in a signal generated by activation of the polypeptide, using an appropriate detection system. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist in the presence of the test compound is observed.

WO 02/070559

PCT/GB02/00948

42

Alternatively, simple binding assays may be used, in which the adherence of a test compound to a surface bearing the polypeptide is detected by means of a label directly or indirectly associated with the test compound or in an assay involving competition with a labelled competitor. In another embodiment, competitive drug screening assays may be used, in which neutralising antibodies that are capable of binding the polypeptide specifically compete with a test compound for binding. In this manner, the antibodies can be used to detect the presence of any test compound that possesses specific binding affinity for the polypeptide.

Assays may also be designed to detect the effect of added test compounds on the production of mRNA encoding the polypeptide in cells. For example, an ELISA may be constructed that measures secreted or cell-associated levels of polypeptide using monoclonal or polyclonal antibodies by standard methods known in the art, and this can be used to search for compounds that may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues. The formation of binding complexes between the polypeptide and the compound being tested may then be measured.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the polypeptide of interest (see International patent application WO84/03564). In this method, large numbers of different small test compounds are synthesised on a solid substrate, which may then be reacted with the polypeptide of the invention and washed. One way of immobilising the polypeptide is to use non-neutralising antibodies. Bound polypeptide may then be detected using methods that are well known in the art. Purified polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques.

The polypeptide of the invention may be used to identify membrane-bound or soluble receptors, through standard receptor binding techniques that are known in the art, such as ligand binding and crosslinking assays in which the polypeptide is labelled with a radioactive isotope, is chemically modified, or is fused to a peptide sequence that facilitates its detection or purification, and incubated with a source of the putative receptor (for example, a composition of cells, cell membranes, cell supernatants, tissue

WO 02/070559

PCT/GB02/00948

43

extracts, or bodily fluids). The efficacy of binding may be measured using biophysical techniques such as surface plasmon resonance and spectroscopy. Binding assays may be used for the purification and cloning of the receptor, but may also identify agonists and antagonists of the polypeptide, that compete with the binding of the polypeptide to its  
5 receptor. Standard methods for conducting screening assays are well understood in the art.

The invention also includes a screening kit useful in the methods for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, that are described above.

The invention includes the agonists, antagonists, ligands, receptors, substrates and  
10 enzymes, and other compounds which modulate the activity or antigenicity of the polypeptide of the invention discovered by the methods that are described above.

The invention also provides pharmaceutical compositions comprising a polypeptide, nucleic acid, ligand or compound of the invention in combination with a suitable pharmaceutical carrier. These compositions may be suitable as therapeutic or diagnostic reagents, as vaccines, or as other immunogenic compositions, as outlined in detail below.  
15

According to the terminology used herein, a composition containing a polypeptide, nucleic acid, ligand or compound [X] is "substantially free of" impurities [herein, Y] when at least 85% by weight of the total X+Y in the composition is X. Preferably, X comprises at least about 90% by weight of the total of X+Y in the composition, more preferably at least about 95%, 98% or even 99% by weight.  
20

The pharmaceutical compositions should preferably comprise a therapeutically effective amount of the polypeptide, nucleic acid molecule, ligand, or compound of the invention. The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent needed to treat, ameliorate, or prevent a targeted disease or condition,  
25 or to exhibit a detectable therapeutic or preventative effect. For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, for example, of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and  
30 routes for administration in humans.

WO 02/070559

PCT/GB02/00948

44

The precise effective amount for a human subject will depend upon the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. This amount can be determined by routine experimentation and is within the judgement of the clinician. Generally, an effective dose will be from 0.01 mg/kg to 50 mg/kg, preferably 0.05 mg/kg to 10 mg/kg. Compositions may be administered individually to a patient or may be administered in combination with other agents, drugs or hormones.

A pharmaceutical composition may also contain a pharmaceutically acceptable carrier, for administration of a therapeutic agent. Such carriers include antibodies and other polypeptides, genes and other therapeutic agents such as liposomes, provided that the carrier does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Suitable carriers may be large, slowly metabolised macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers and inactive virus particles.

Pharmaceutically acceptable salts can be used therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulphates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable carriers is available in Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Pharmaceutically acceptable carriers in therapeutic compositions may additionally contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such compositions. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Once formulated, the compositions of the invention can be administered directly to the subject. The subjects to be treated can be animals; in particular, human subjects can be treated.

WO 02/070559

PCT/GB02/00948

45

The pharmaceutical compositions utilised in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal or transcutaneous applications (for example, see WO98/20734), subcutaneous, intraperitoneal, intranasal, 5 enteral, topical, sublingual, intravaginal or rectal means. Gene guns or hyposprays may also be used to administer the pharmaceutical compositions of the invention. Typically, the therapeutic compositions may be prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared.

10 Direct delivery of the compositions will generally be accomplished by injection, subcutaneously, intraperitoneally, intravenously or intramuscularly, or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Dosage treatment may be a single dose schedule or a multiple dose schedule.

If the activity of the polypeptide of the invention is in excess in a particular disease state, 15 several approaches are available. One approach comprises administering to a subject an inhibitor compound (antagonist) as described above, along with a pharmaceutically acceptable carrier in an amount effective to inhibit the function of the polypeptide, such as by blocking the binding of ligands, substrates, enzymes, receptors, or by inhibiting a second signal, and thereby alleviating the abnormal condition. Preferably, such 20 antagonists are antibodies. Most preferably, such antibodies are chimeric and/or humanised to minimise their immunogenicity, as described previously.

In another approach, soluble forms of the polypeptide that retain binding affinity for the ligand, substrate, enzyme, receptor, in question, may be administered. Typically, the polypeptide may be administered in the form of fragments that retain the relevant 25 portions.

In an alternative approach, expression of the gene encoding the polypeptide can be inhibited using expression blocking techniques, such as the use of antisense nucleic acid molecules (as described above), either internally generated or separately administered. Modifications of gene expression can be obtained by designing complementary 30 sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5' or regulatory

WO 02/070559

PCT/GB02/00948

46

regions (signal sequence, promoters, enhancers and introns) of the gene encoding the polypeptide. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. *et al.* (1994) In: Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY). The complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes. Such oligonucleotides may be administered or may be generated *in situ* from expression *in vivo*.

In addition, expression of the polypeptide of the invention may be prevented by using ribozymes specific to its encoding mRNA sequence. Ribozymes are catalytically active RNAs that can be natural or synthetic (see for example Usman, N, *et al.*, Curr. Opin. Struct. Biol (1996) 6(4), 527-33). Synthetic ribozymes can be designed to specifically cleave mRNAs at selected positions thereby preventing translation of the mRNAs into functional polypeptide. Ribozymes may be synthesised with a natural ribose phosphate backbone and natural bases, as normally found in RNA molecules. Alternatively the ribozymes may be synthesised with non-natural backbones, for example, 2'-O-methyl RNA, to provide protection from ribonuclease degradation and may contain modified bases.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of non-traditional bases such as inosine, queosine and butosine, as well as acetyl-, methyl-, thio- and similarly modified forms of adenine, cytidine, guanine, thymine and uridine which are not as easily recognised by endogenous endonucleases.

For treating abnormal conditions related to an under-expression of the polypeptide of the invention and its activity, several approaches are also available. One approach comprises

WO 02/070559

PCT/GB02/00948

47

administering to a subject a therapeutically effective amount of a compound that activates the polypeptide, i.e., an agonist as described above, to alleviate the abnormal condition. Alternatively, a therapeutic amount of the polypeptide in combination with a suitable pharmaceutical carrier may be administered to restore the relevant physiological balance of polypeptide.

Gene therapy may be employed to effect the endogenous production of the polypeptide by the relevant cells in the subject. Gene therapy is used to treat permanently the inappropriate production of the polypeptide by replacing a defective gene with a corrected therapeutic gene.

Gene therapy of the present invention can occur *in vivo* or *ex vivo*. *Ex vivo* gene therapy requires the isolation and purification of patient cells, the introduction of a therapeutic gene and introduction of the genetically altered cells back into the patient. In contrast, *in vivo* gene therapy does not require isolation and purification of a patient's cells.

The therapeutic gene is typically "packaged" for administration to a patient. Gene delivery vehicles may be non-viral, such as liposomes, or replication-deficient viruses, such as adenovirus as described by Berkner, K.L., in *Curr. Top. Microbiol. Immunol.*, 158, 39-66 (1992) or adeno-associated virus (AAV) vectors as described by Muzyczka, N., in *Curr. Top. Microbiol. Immunol.*, 158, 97-129 (1992) and U.S. Patent No. 5,252,479. For example, a nucleic acid molecule encoding a polypeptide of the invention may be engineered for expression in a replication-defective retroviral vector. This expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding the polypeptide, such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo* (see Chapter 20, *Gene Therapy and other Molecular Genetic-based Therapeutic Approaches*, (and references cited therein) in *Human Molecular Genetics* (1996), T Strachan and A P Read, BIOS Scientific Publishers Ltd).

Another approach is the administration of "naked DNA" in which the therapeutic gene is directly injected into the bloodstream or muscle tissue.

WO 02/070559

PCT/GB02/00948

48

In situations in which the polypeptides or nucleic acid molecules of the invention are disease-causing agents, the invention provides that they can be used in vaccines to raise antibodies against the disease causing agent.

Vaccines according to the invention may either be prophylactic (ie. to prevent infection) or therapeutic (ie. to treat disease after infection). Such vaccines comprise immunising antigen(s), immunogen(s), polypeptide(s), protein(s) or nucleic acid, usually in combination with pharmaceutically-acceptable carriers as described above, which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the antigen or immunogen may be conjugated to a bacterial toxoid, such as a toxoid from diphtheria, tetanus, cholera, *H. pylori*, and other pathogens.

Since polypeptides may be broken down in the stomach, vaccines comprising polypeptides are preferably administered parenterally (for instance, subcutaneous, intramuscular, intravenous, or intradermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient, and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents.

The vaccine formulations of the invention may be presented in unit-dose or multi-dose containers. For example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

This invention also relates to the use of nucleic acid molecules according to the present invention as diagnostic reagents. Detection of a mutated form of the gene characterised by the nucleic acid molecules of the invention which is associated with a dysfunction will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals carrying mutations in the

WO 02/070559

PCT/GB02/00948

49

gene may be detected at the DNA level by a variety of techniques.

Nucleic acid molecules for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR, ligase chain reaction (LCR), strand displacement amplification (SDA), or other amplification techniques (see Saiki *et al.*, *Nature*, 324, 163-166 (1986); Bej, *et al.*, *Crit. Rev. Biochem. Molec. Biol.*, 26, 301-334 (1991); Birkenmeyer *et al.*, *J. Virol. Meth.*, 35, 117-126 (1991); Van Brunt, J., *Bio/Technology*, 8, 291-294 (1990)) prior to analysis.

In one embodiment, this aspect of the invention provides a method of diagnosing a disease in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide according to the invention and comparing said level of expression to a control level, wherein a level that is different to said control level is indicative of disease. The method may comprise the steps of:

a) contacting a sample of tissue from the patient with a nucleic acid probe under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule of the invention and the probe;

b) contacting a control sample with said probe under the same conditions used in step a);

c) and detecting the presence of hybrid complexes in said samples;

wherein detection of levels of the hybrid complex in the patient sample that differ from levels of the hybrid complex in the control sample is indicative of disease.

A further aspect of the invention comprises a diagnostic method comprising the steps of:

a) obtaining a tissue sample from a patient being tested for disease;

b) isolating a nucleic acid molecule according to the invention from said tissue sample; and,

c) diagnosing the patient for disease by detecting the presence of a mutation in the nucleic acid molecule which is associated with disease.

To aid the detection of nucleic acid molecules in the above-described methods, an amplification step, for example using PCR, may be included.

WO 02/070559

PCT/GB02/00948

50

Deletions and insertions can be detected by a change in the size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labelled RNA of the invention or alternatively, labelled antisense DNA sequences of the invention. Perfectly-matched sequences can be distinguished from mismatched duplexes by RNase digestion or by assessing differences in melting temperatures. The presence or absence of the mutation in the patient may be detected by contacting DNA with a nucleic acid probe that hybridises to the DNA under stringent conditions to form a hybrid double-stranded molecule, the hybrid double-stranded molecule having an unhybridised portion of the nucleic acid probe strand at any portion corresponding to a mutation associated with disease; and detecting the presence or absence of an unhybridised portion of the probe strand as an indication of the presence or absence of a disease-associated mutation in the corresponding portion of the DNA strand.

Such diagnostics are particularly useful for prenatal and even neonatal testing.

Point mutations and other sequence differences between the reference gene and "mutant" genes can be identified by other well-known techniques, such as direct DNA sequencing or single-strand conformational polymorphism, (see Orita *et al.*, *Genomics*, 5, 874-879 (1989)). For example, a sequencing primer may be used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabelled nucleotides or by automatic sequencing procedures with fluorescent-tags. Cloned DNA segments may also be used as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. Further, point mutations and other sequence variations, such as polymorphisms, can be detected as described above, for example, through the use of allele-specific oligonucleotides for PCR amplification of sequences that differ by single nucleotides.

DNA sequence differences may also be detected by alterations in the electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (for example, Myers *et al.*, *Science* (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton *et al.*, *Proc. Natl. Acad. Sci.*

WO 02/070559

PCT/GB02/00948

51

USA (1985) 85: 4397-4401).

In addition to conventional gel electrophoresis and DNA sequencing, mutations such as microdeletions, aneuploidies, translocations, inversions, can also be detected by in situ analysis (see, for example, Keller *et al.*, DNA Probes, 2nd Ed., Stockton Press, New York, N.Y., USA (1993)), that is, DNA or RNA sequences in cells can be analysed for mutations without need for their isolation and/or immobilisation onto a membrane. Fluorescence in situ hybridization (FISH) is presently the most commonly applied method and numerous reviews of FISH have appeared (see, for example, Trachuck *et al.*, Science, 250: 559-562 (1990), and Trask *et al.*, Trends, Genet. 7:149-154 (1991)).

In another embodiment of the invention, an array of oligonucleotide probes comprising a nucleic acid molecule according to the invention can be constructed to conduct efficient screening of genetic variants, mutations and polymorphisms. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M.Chee *et al.*, Science (1996) 274: 610-613).

In one embodiment, the array is prepared and used according to the methods described in PCT application WO95/11995 (Chee *et al.*); Lockhart, D. J. *et al.* (1996) Nat. Biotech. 14: 1675-1680; and Schena, M. *et al.* (1996) Proc. Natl. Acad. Sci. 93: 10614-10619). Oligonucleotide pairs may range from two to over one million. The oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass slide or any other suitable solid support. In another aspect, an oligonucleotide may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application WO95/251116 (Baldeschweiler *et al.*). In another aspect, a "gridded" array analogous to a dot (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. An array, such as those described above, may be produced by hand or by using available devices (slot blot or dot blot apparatus), materials (any suitable solid support), and machines (including robotic instruments), and may contain 8, 24, 96, 384, 1536 or 6144

WO 02/070559

PCT/GB02/00948

52

oligonucleotides, or any other number between two and over one million which lends itself to the efficient use of commercially-available instrumentation.

In addition to the methods discussed above, diseases may be diagnosed by methods comprising determining, from a sample derived from a subject, an abnormally decreased  
5 or increased level of polypeptide or mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods.

10 Assay techniques that can be used to determine levels of a polypeptide of the present invention in a sample derived from a host are well-known to those of skill in the art and are discussed in some detail above (including radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays). This aspect of the invention provides a diagnostic method which comprises the steps of: (a) contacting a ligand as described  
15 above with a biological sample under conditions suitable for the formation of a ligand-polypeptide complex; and (b) detecting said complex.

Protocols such as ELISA, RIA, and FACS for measuring polypeptide levels may additionally provide a basis for diagnosing altered or abnormal levels of polypeptide expression. Normal or standard values for polypeptide expression are established by  
20 combining body fluids or cell extracts taken from normal mammalian subjects, preferably humans, with antibody to the polypeptide under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, such as by photometric means.

Antibodies which specifically bind to a polypeptide of the invention may be used for the  
25 diagnosis of conditions or diseases characterised by expression of the polypeptide, or in assays to monitor patients being treated with the polypeptides, nucleic acid molecules, ligands and other compounds of the invention. Antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for the polypeptide include methods that utilise the antibody and a  
30 label to detect the polypeptide in human body fluids or extracts of cells or tissues. The

WO 02/070559

PCT/GB02/00948

53

antibodies may be used with or without modification, and may be labelled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules known in the art may be used, several of which are described above.

Quantities of polypeptide 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. Diagnostic assays may be used to distinguish between absence, presence, and excess expression of polypeptide and to monitor regulation of polypeptide levels during therapeutic intervention. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

A diagnostic kit of the present invention may comprise:

- (a) a nucleic acid molecule of the present invention;
- (b) a polypeptide of the present invention; or
- (c) a ligand of the present invention.

In one aspect of the invention, a diagnostic kit may comprise a first container containing a nucleic acid probe that hybridises under stringent conditions with a nucleic acid molecule according to the invention; a second container containing primers useful for amplifying the nucleic acid molecule; and instructions for using the probe and primers for facilitating the diagnosis of disease. The kit may further comprise a third container holding an agent for digesting unhybridised RNA.

In an alternative aspect of the invention, a diagnostic kit may comprise an array of nucleic acid molecules, at least one of which may be a nucleic acid molecule according to the invention.

To detect polypeptide according to the invention, a diagnostic kit may comprise one or more antibodies that bind to a polypeptide according to the invention; and a reagent useful for the detection of a binding reaction between the antibody and the polypeptide.

Such kits will be of use in diagnosing a disease or susceptibility to disease, particularly cell proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and other solid tumours, myeloproliferative disorders, such as

WO 02/070559

PCT/GB02/00948

54

leukemia, non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposi's sarcoma, autoimmune/inflammatory disorders, including allergy, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation, asthma, and organ transplant rejection, cardiovascular disorders, including hypertension, oedema, angina, atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia, neurological disorders including, central nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain, developmental disorders, metabolic disorders including diabetes mellitus, osteoporosis, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, and obesity, renal disorders, including glomerulonephritis, renovascular hypertension, dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging, AIDS, infections including viral infection, bacterial infection, fungal infection and parasitic infection and other pathological conditions, particularly those in which nuclear hormone receptors are implicated.

Various aspects and embodiments of the present invention will now be described in more detail by way of example, with particular reference to the LBDG1 polypeptide.

It will be appreciated that modification of detail may be made without departing from the scope of the invention.

#### 20 Brief description of the Figures

Figure 1: This is the front end of the Biopendium Target Mining Interface. A search of the database is initiated using the PDB code "1ERR:A".

Figure 2A: A selection is shown of the Inpharmatica Genome Threader results for the search using 1ERR:A. The arrow indicates *Homo sapiens* Estrogen Receptor, which has a typical Nuclear Hormone Receptor Ligand Binding Domain.

Figure 2B: A selection is shown of the Inpharmatica Genome Threader results for the search using 1ERR:A. The arrow indicates O75385 (LBDG1).

Figure 2C: Full list of forward PSI-BLAST results for the search using 1ERR:A. O75385 (LBDG1) is not identified.

WO 02/070559

PCT/GB02/00948

55

- Figure 3: The Redundant Sequence Display results page for O75385 (LBDG1).
- Figure 4A: InterPro PFAM search results for O75385 (LBDG1), see arrow ①.
- Figure 4B: InterPro report on Proline rich regions, indicating they may function as DNA binding domains.
- 5 Figure 5: SWISS-PROT entry for O75385 (LBDG1).
- Figure 6A: This is the front end of the Biopendium database. A search of the database is initiated using O75385 (LBDG1), as the query sequence.
- Figure 6B: A selection of the Inpharmatica Genome Threader results of search using O75385 (LBDG1), as the query sequence. The arrow points to 1ERR:A.
- 10 Figure 6C: A selection of the reverse-maximised PSI-BLAST results obtained using O75385 (LBDG1), as the query sequence. The arrows numbered ①, ② and ③ point to homologues of O75385 (LBDG1).
- Figure 7: ALEye sequence alignment of O75385 (LBDG1) and 1ERR:A.
- Figure 8A: LigEye for 1ERR:A that illustrates the sites of interaction of Raloxifene (Ral600(A)) with the Ligand Binding Domain of *Homo sapiens* Estrogen Receptor, 1ERR:A.
- 15 Figure 8B: iRasMol view of 1ERR:A, the Ligand Binding Domain of *Homo sapiens* Estrogen Receptor.
- Figure 9: ALEye sequence alignment of O75385 (LBDG1; *Homo sapiens* ULK1) with the homologues O70405 (*Mus musculus* ULK1), BAA31598.1 (*Homo sapiens* ULK2) and BAA77341.1 (*Mus musculus* ULK2).
- 20 Figure 10: NCBI UniGene report summarising sources of cDNAs and ESTs which correspond to O75385 (LBDG1).
- Figure 11: NN1112 is an adult nervous tissue library.
- 25 Figure 12: HUGE Database report for KIAA0722 (the HUGE identifier for O75385 (LBDG1)).

WO 02/070559

PCT/GB02/00948

56

Figure 13: Page 84 of the article Kuroyanagi *et al*, Genomics (1998) vol.51:76-85, which details the cytogenetic map position of the O75385 (LBDG1) gene and nearby disease loci.

Figure 14: Genome Threader alignment of O75385 (LBDG1) with Homo sapiens Estrogen Receptor alpha (1ERR:A), showing only the N-terminal half of the dimerisation helix " $\alpha 10$ ". The three O75385 (LBDG1) homologues are included in the alignment (by reference to the O75385 (LBDG1) sequence). Five other human Ligand binding domain sequences are also included in the alignment by reference to the 1ERR:A sequence. These are LBDs from the Androgen receptor (AR), Retinoic Acid Receptor gamma (RARgamma), Retinoid X Receptor alpha (RXRalpha), Peroxisome Proliferator Activated Receptor gamma (PPARGgamma) and Retinoid Orphan Receptor beta (RORbeta). Conserved residues are marked by boxes labeled a to h. Polar/charged residues are in grey boxes, hydrophobic residues are in black boxes.

Figure 15: Overall view of the homology model of O75385 (LBDG1) adopting the Ligand Binding Domain fold. Residues of particular interest are marked in black.

Figure 16: View of the homology model of O75385 (LBDG1) adopting the Ligand Binding Domain fold, showing only the region encompassing the predicted helices " $\alpha 3$ " and " $\alpha 5$ ". Grey arrows mark the direction of the polypeptide chain running N-terminal to C-terminal.

Figure 17: Close-up of the predicted Co-activator binding site of the homology model of O75385 (LBDG1) adopting the Ligand Binding Domain fold. Residues of particular interest are shown in black.

Figure 18: Close-up of the predicted Co-activator binding site of the homology model of O75385 (LBDG1) adopting the Ligand Binding Domain fold. Residues of particular interest are shown in black. A cartoon of a Co-activator helix has been added to illustrate the predicted mode of binding to the homology model.

Figure 19: Overview of the structure of the *Homo sapiens* Estrogen Receptor alpha homodimer (1ERR:A and 1ERR:B). The black line divides the two monomers. The dimerisation helix " $\alpha 10$ " is labeled.

WO 02/070559

PCT/GB02/00948

57

Figure 20: Comparison of the ERalpha structure (1ERR:A, Figure 20A) with the homology model of O75385 (LBDG1, Figure 20B) adopting the Ligand Binding Domain fold, with particular reference to the dimerisation helix "α10". Residues which are critical to the dimer interaction are marked in black.

- 5 Figure 21: Comparison of the ERalpha structure (1ERR:A, Figure 21A) with the homology model of O75385 (LBDG1, Figure 21B) adopting the Ligand Binding Domain fold, showing only the dimerisation helix "α10". Figure 21A and B each two perpendicular views of the dimerisation helix "α10". In the edge-on view of the helix, the polypeptide chain is running towards the C-terminus, out of the page towards the viewer.
- 10 Important residues are marked in black.

Figure 22: Genome Threader alignment of O75385 (LBDG1) with Homo sapiens Estrogen Receptor alpha (1ERR:A), including the human paralog BAA31598.1 (LBDG4).

- Figure 23: Overall view of the homology model of BAA31598.1 (LBDG4) adopting the  
15 Ligand Binding Domain fold. Residues of particular interest are marked in black.

- Figure 24: View of the homology model of BAA31598.1 (LBDG4) adopting the Ligand Binding Domain fold, with particular reference to the dimerisation helix "α10". Residues which are critical to the dimer interaction are marked in black. Figure 24B and C show two perpendicular views of the dimerisation helix "α10". In the edge-on view of the  
20 helix, The polypeptide chain is running towards the C-terminus, out of the page towards the viewer. Important residues are marked in black.

#### Example 1

- In order to initiate a search for novel, distantly related Nuclear Hormone Receptor Ligand Binding Domains, an archetypal family member is chosen, the Ligand Binding Domain  
25 of *Homo sapiens* Estrogen Receptor. More specifically, the search is initiated using a structure from the Protein Data Bank (PDB) which is operated by the Research Collaboratory for Structural Bioinformatics.

The structure chosen is the Ligand Binding Domain of *Homo sapiens* Estrogen Receptor (PDB code 1ERR:A; see Figure 1).

WO 02/070559

PCT/GB02/00948

58

A search of the Biopendium (using the Target Mining Interface) for relatives of IERR:A takes place and returns 3614 Genome Threader results. The 3614 Genome Threader results include examples of typical Nuclear Hormone Receptor Ligand Binding Domains, such as that found between residues 307-547 of the *Homo sapiens* Estrogen Receptor (see arrow in Figure 2A).

Among the proteins known to contain a Nuclear Hormone Receptor Ligand Binding Domain appears a protein which is not annotated as containing a Nuclear Hormone Receptor Ligand Binding Domain, O75385 (LBDG1; see arrow in Figure 2B). The Inpharmatica Genome Threader has identified a region of the sequence O75385 (LBDG1), between residues 822-1020, as having a structure similar to the Ligand Binding Domain of *Homo sapiens* Estrogen Receptor. The possession of a structure similar to the Ligand Binding Domain of *Homo sapiens* Estrogen Receptor suggests that residues 822-1020 of O75385 (LBDG1) function as a Nuclear Hormone Receptor Ligand Binding Domain. The Genome Threader identifies this with 84% confidence.

The search of the Biopendium (using the Target Mining Interface) for relatives of IERR:A also returns 841 Forward PSI-Blast results. Forward PSI-Blast (see Figure 2C) is unable to identify this relationship; only the Inpharmatica Genome Threader is able to identify O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

In order to assess what is known in the public domain databases about O75385 (LBDG1) the Redundant Sequence Display Page (Figure 3) is viewed. There are no PROSITE or PRINTS hits which identify O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain. PROSITE and PRINTS are databases that help to describe proteins of similar families. Returning no Nuclear Hormone Receptor Ligand Binding Domain hits from both databases means that O75385 (LBDG1) is unidentifiable as containing a Nuclear Hormone Receptor Ligand Binding Domain using PROSITE or PRINTS.

In order to identify if any other public domain annotation vehicle is able to annotate O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain, the O75385 (LBDG1) protein sequence is searched against the PFAM database (Protein

WO 02/070559

PCT/GB02/00948

59

Family Database of Alignment and hidden Markov models) at the InterPro website (see Figure 4A arrow ①). A PFAM-A match is found to PF00069/IPR000719, which is diagnostic of a Eukaryotic protein kinase domain. This protein kinase domain is annotated as being located between residues 16-278 of O75385 (LBDG1). There are no  
5 PFAM-A matches annotating O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain. Thus PFAM does not identify O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

Interestingly, PROSITE PFscan (see Figure 4A arrow ②) identifies a Proline rich region in O75385 (LBDG1) at residues 501-688. Proline rich regions may act as DNA binding domains (see Figure 4B). The positioning of a DNA binding domain to the N-terminus of  
10 the predicted Nuclear Hormone Receptor Ligand Binding Domain (residues 822-1020) is reminiscent of the domain organisation of archetypal Nuclear Hormone Receptors (which have an N-terminal Zinc Finger DNA Binding Domain and C-terminal Ligand Binding Domain).

15 The Swiss Institute of Bioinformatics (SIB) SWISS-PROT protein database is then viewed to examine if there is any further information that is known in the public domain relating to the O75385 (LBDG1) sequence. This is the Swiss public domain database for protein and gene sequence deposition (Figure 5). O75385 (LBDG1) is a *Homo Sapiens* sequence, its SWISS-PROT protein ID is O75385 and it is 1050 amino acids in length.  
20 O75385 is called ULK1 (UNC51-Like Kinase 1). ULK1 was cloned by Kuroyanagi *et al.* At the Department of Molecular Genetics, Tokyo Metropolitan Institute of Gerontology, 35-2 Sakaecho, Itabashi-ku, Tokyo, 173-0015, Japan. The public domain information for this gene does not annotate it as containing a Nuclear Hormone Receptor Ligand Binding Domain.

25 Therefore, it can be concluded that using all public domain annotation tools, O75385 (LBDG1) may not be annotated as containing a Nuclear Hormone Receptor Ligand Binding Domain. Only the Inpharmatica Genome Threader is able to annotate this protein as containing a Nuclear Hormone Receptor Ligand Binding Domain.

The reverse search is now carried out. O75385 (LBDG1) is now used as the query  
30 sequence in the Biopendium (see Figure 6A). The Inpharmatica Genome Threader

WO 02/070559

PCT/GB02/00948

60

identifies residues 822-1020 of O75385 (LBDG1) as having a structure that is the same as the Ligand Binding Domain of *Homo sapiens* Estrogen Receptor with 84% confidence (see arrow in Figure 6B). The Ligand Binding Domain of *Homo sapiens* Estrogen Receptor (1JERR:A) was the original query sequence. Positive iterations of PSI-Blast do not return this result (Figure 6C). It is only the Inpharmatica Genome Threader that is able to identify this relationship.

The sequence of the *Homo sapiens* Estrogen Receptor Ligand Binding Domain is chosen against which to view the sequence alignment of O75385 (LBDG1). Viewing the AIEye alignment (Figure 7) of the query protein against the protein identified as being of a similar structure helps to visualize the areas of homology.

The *Homo sapiens* Estrogen Receptor Ligand Binding Domain contains an "LBD motif" which has been found in all annotated Nuclear Hormone Receptor Ligand Binding Domains to date. The "LBD motif" is involved in recruiting Nuclear Hormone Receptor Co-Activators and Co-Repressors. The 6 residues; PHE367, LEU370, ASP374, GLN375, LEU378 and LEU379 constitute this motif in the *Homo sapiens* Estrogen Receptor Ligand Binding Domain (see square boxes Figure 7). 4 residues (ASP885, GLN886, LEU889 and LEU890) in O75385 (LBDG1) precisely match 4 (ASP374, GLN375, LEU378 and LEU379) out of the 6 "LBD motif" residues in the *Homo sapiens* Estrogen Receptor Ligand Binding Domain. A fifth residue, LEU878, in O75385 (LBDG1) is conservatively substituted for the PHE367 in the "LBD motif" of *Homo sapiens* Estrogen Receptor Ligand Binding Domain. This indicates that O75385 (LBDG1) contains a Nuclear Hormone Receptor Ligand Binding Domain similar to The *Homo sapiens* Estrogen Receptor Ligand Binding Domain.

In order to ensure that the protein identified is in fact a relative of the query sequence, the visualization programs "LigEye" (Figure 8A) and "iRasmol" (Figure 8B) are used. These visualization tools identify the active site of known protein structures by indicating the amino acids with which known small molecule inhibitors interact at the active site. These interactions are either through a direct hydrogen bond or through hydrophobic interactions. In this manner, one can see if the active site fold/structure is conserved between the identified homologue and the chosen protein of known structure. The LigEye

view of the *Homo sapiens* Estrogen Receptor Ligand Binding Domain reveals 15 residues which bind Raloxifene (circled in Figure 7). However, only 11 (THR347, ALA350, LEU354, TRP383, LEU387, MET388, LEU391, ARG394, PHE404, MET421 and LEU428) of these 15 residues lie within the Genome Threader alignment. Thus only these 11 residues can be used to consolidate the Genome Threader annotation of O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain. Of these 11 residues there are 9 hydrophobic residues which line the pocket of the *Homo sapiens* Estrogen Receptor Ligand Binding Domain; ALA350, LEU354, TRP383, LEU387, MET388, LEU391, PHE404, MET421 and LEU428. 5 of these 9 *Homo sapiens* Estrogen Receptor Ligand Binding Domain hydrophobic residues are perfectly conserved in O75385 (LBDG1): ALA861, LEU863, TRP894, LEU900 and LEU904 (circled in Figure 7). 3 of the remaining 4 *Homo sapiens* Estrogen Receptor Ligand Binding Domain hydrophobic residues are conservatively substituted for another hydrophobic residue in O75385 (LBDG1): VAL901, LEU910 and VAL932 (broken circles in Figure 7). This conservation of hydrophobicity in 8 out of the 9 hydrophobic residues (within the region of Genome Threader alignment) which line the binding pocket indicates that O75385 (LBDG1) will bind a hydrophobic steroid-like ligand.

This indicates that indeed as predicted by the Inpharmatica Genome Threader, O75385 (LBDG1) folds in a similar manner to the *Homo sapiens* Estrogen Receptor Ligand Binding Domain and as such is identified as containing a Nuclear Hormone Receptor Ligand Binding Domain.

Reverse-maximised PSI-BLAST of O75385 (LBDG1) identifies 1 *Homo sapiens* homologue, and 2 *Mus musculus* homologues of O75385 (LBDG1). O70405 (*Mus musculus* ULK1, see Figure 6C arrow Ⓞ) is the most closely related sequence to O75385 (LBDG1), sharing 88% sequence identity. O70405 (*Mus musculus* ULK1) thus represents the *Mus musculus* orthologue of O75385 (LBDG1; *Homo sapiens* ULK1).

BAA31598.1 (*Homo sapiens* ULK2, see Figure 6C arrow Ⓢ) shares 51% sequence identity with O75385 (LBDG1; *Homo sapiens* ULK1). BAA31598.1 (*Homo sapiens* ULK2) thus represents a paralogue of O75385 (LBDG1; *Homo sapiens* ULK1).

WO 02/070559

PCT/GB02/00948

62

BAA77341.1 (*Mus musculus* ULK2, see Figure 6C arrow ③) shares 50% sequence identity with O75385 (LBDG1; *Homo sapiens* ULK1). However, BAA77341.1 (*Mus musculus* ULK2) is much more closely related to BAA31598.1 (*Homo sapiens* ULK2), which share 91% sequence identity and this is why they are clustered in Figure 6C).

5 BAA77341.1 (*Mus musculus* ULK2) thus represents the *Mus musculus* orthologue of O70405 (LBDG1; *Homo sapiens* ULK2).

O75385 (LBDG1; *Homo sapiens* ULK1), O70405 (*Mus musculus* ULK1), BAA31598.1 (*Homo sapiens* ULK2) and BAA77341.1 (*Mus musculus* ULK2) are aligned and viewed in AIEye (Figure 9). AIEye reveals that the 8 hydrophobic residues (within the Genome Threader alignment; ALA861, LEU863, TRP894, LEU900, VAL901, LEU904, LEU910 and VAL932) of O75385 (LBDG1) predicted to line the ligand binding pocket are all precisely conserved in O70405 (*Mus musculus* ULK1; see Figure 9). Furthermore, 4 (LEU878, ASP885, GLN886, LEU890) of the 5 predicted "LBD motif" residues in O75385 (LBDG1) are precisely conserved in O70405 (*Mus musculus* ULK1).

15 AIEye also reveals that 5 (TRP894, LEU900, VAL901, LEU910 and VAL932) of the 8 hydrophobic residues (within the Genome Threader alignment) of O75385 (LBDG1) predicted to line the ligand binding pocket are all precisely conserved in BAA31598.1 (*Homo sapiens* ULK2) and BAA77341.1 (*Mus musculus* ULK2). Furthermore, 3 (ASP885, GLN886, LEU890) of the 5 predicted "LBD motif" residues in O75385  
20 (LBDG1) are precisely conserved in BAA31598.1 (*Homo sapiens* ULK2) and BAA77341.1 (*Mus musculus* ULK2).

Residues which are essential for the function of a protein will be conserved in homologues of that protein. Thus the conservation of residues which would be essential for the function of the predicted O75385 (LBDG1) Nuclear Hormone Receptor Ligand Binding Domain in the homologues O70405 (*Mus musculus* ULK1), BAA31598.1  
25 (*Homo sapiens* ULK2) and BAA77341.1 (*Mus musculus* ULK2) strongly supports the annotation of O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

Figure 10 is a report generated from the NCBI UniGene database. This database is a  
30 collection of expressed sequence tags (ESTs) from various human tissues, it can be used to

WO 02/070559

PCT/GB02/00948

63

give a general tissue distribution for a protein provided that its sequence is present in the database. O75385 (LBDG1) is present in the database and is shown to be expressed in tissues referred to as Blood, Brain, Breast, CNS, Colon, Germ Cell, Heart, Kidney, Lung, Lymph, Muscle, Ovary, Pancreas, Parathyroid, Pooled, Prostate, Testis, Tonsil, Uterus, Whole embryo, adrenal gland, brain, breast\_normal, cervix, colon, colon\_est, colon\_ins, head\_neck, kidney, leiomyos, lung, nervous\_normal, ovary, placenta, prostate\_normal, prostate\_tumor and uterus. mRNA for O75385 (LBDG1) is also highly represented (3.6%) in library 3761 NN1112, an adult nervous tissue library (see Figure 11).

Although the UniGene database gives a rough idea of tissue distribution, the Human Unidentified Gene-Encoded Large Proteins Analyzed by Kazusa cDNA Project (HUGE) database provides a direct experimental measure of cDNA levels by RT-PCR ELISA (see Figure 12). O75385 (LBDG1) is present in the database under the identifier KIAA0722 and is highly expressed in brain, heart and ovary.

There is an article associated with O75385 (LBDG1); Kuroyanagi *et al.*, Genomics (1998) vol.51:76-85, see Figure 13. This article demonstrates that the gene encoding O75385 (LBDG1) maps distal to the cytogenetic locus 12q24.3. This position is close to the causal gene for distal hereditary motor neuropathy type II (distal HMN II) and the causal gene for scapuloperoneal spinal muscular atrophy (SPSMA). While the gene encoding O75385 (LBDG1) maps distal to these loci, there is nonetheless a possibility that O75385 (LBDG1) is the causal gene for one or both of these disorders. It may thus be inferred that the O75385 (LBDG1) gene may be implicated in these conditions, so paving the way for the development of agents that target the O75385 (LBDG1) gene or its encoded protein, to diagnose and/or treat these conditions. In particular, the identification of this gene as containing a Nuclear Hormone Receptor Ligand Binding Domain facilitates the development of agents that are specific to the O75385 (LBDG1) protein, for example, through the use of Nuclear Hormone Receptor Ligand Binding Domain agonists or antagonists.

#### Example 2

The *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain contains an "LBD motif" which has been found in all annotated Nuclear Hormone Receptor Ligand Binding

WO 02/070559

PCT/GB02/00948

64

Domains to date. The "LBD motif" is involved in recruiting Nuclear Hormone Receptor Co-Activators and Co-Repressors. The 6 residues PHE367, LEU370, ASP374, GLN375, LEU378 and LEU379 constitute this motif in the *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain (refer back to Figure 7: 1ERR:A). As discussed above, four of these "LBD motif" residues are precisely conserved in O75385 (LBDG1). 1ERR:A ASP374 is conserved as ASP885 in O75385 (LBDG1). 1ERR:A GLN375 is conserved as GLN886 in O75385 (LBDG1). 1ERR:A LEU378 is conserved as LEU889 in O75385 (LBDG1). 1ERR:A LEU379 is conserved as LEU890 in O75385 (LBDG1). A fifth residue of the 1ERR:A "LBD motif", PHE367 is conservatively substituted by LEU878 in O75385 (LBDG1).

In addition to the "LBD motif" residues, there are other residues which are well conserved in known Nuclear Hormone Receptor Ligand Binding Domains. A further test of the annotation of O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain is to analyse whether important residues outside of the "LBD motif" are conserved in O75385 (LBDG1) and the five O75385 (LBDG1) homologues.

In addition to Ligand binding and Co-Activator/Co-Repressor recruitment, another characteristic property of Nuclear Hormone Receptor Ligand Binding Domains is their ability to dimerise with other Nuclear Hormone Receptor Ligand Binding Domains. For example, the Estrogen Receptor alpha (ER $\alpha$ ) Ligand Binding Domain can homodimerise with itself.

Another example is that the Peroxisome Proliferator Activated Receptor gamma (PPAR $\gamma$ ) Ligand Binding domain can heterodimerise with the Retinoid X Receptor alpha (RXR $\alpha$ ) Ligand Binding Domain. Regardless of whether a Ligand Binding Domain is forming a homo- or hetero-dimer, the same general mode of interaction underlies the association. Crystallographic and mutagenesis experiments have identified the N-terminal half of helix "α10" as the primary site of interaction between two dimerising Ligand Binding Domains (Gampe, R. T., Montana, V G., Lambert, M H., Miller, A B., Bledsoe, R K., Milburn, M V., Kliewer, S A., Willson, T M., Xu, H E., Mol. Cell. 2000, 5 (3):545-55 and (Brzozowski, A M., Pike, A C., Dauter, Z., Hubbard, R E., Bonn, T., Engstrom, O., Ohman, L., Greene, G L., Gustafsson, J A., Carlquist, M., Nature. 1997, vol.389

WO 02/070559

PCT/GB02/00948

65

6652):753-8). The functional importance of the N-terminal half of helix "α10" is reflected in the strong conservation of particular residues in this region. Figure 14 shows O75385 (LBDG1), and its three homologues, aligned with the N-terminal half of helix "α10" sequences for *Homo sapiens* ERα, AR, RARγ, RXRα, PPARγ and RORβ. To form a dimer, the "α10" helix of each monomer intertwine to form a rigid backbone, and this is driven by a combination of polar and hydrophobic interactions. There are two residue positions on helix "α10" which play a significant role in the polar component of the dimerisation interactions (these are marked as box "b" and box "g" in Figure 14). Position "b" on helix "α10" is normally occupied by a polar residue (eg. GLN506 in ERα) or a positively charged residue (eg. LYS417 in RXRα), which have the appropriate side chain properties to make a polar/charge interaction with the dimerisation partner. According to the Genome Threader alignment of O75385 (LBDG1) with *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain (1ERR:A), LYS1009 of O75385 (LBDG1) would occupy position "b" on helix "α10". Clearly LYS1009 has the appropriate side chain properties to make a polar/charge interaction with the dimerisation partner. This supports the annotation of O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain. The significance of LYS1009 of O75385 (LBDG1) occupying position "b" on helix "α10" is enhanced by the observation that LYS1009 is conserved in the three O75385 (LBDG1) homologues (O70405, BAA31598.1 and BAA77341.1, see Figure 14 box "b").

Position "g" on helix "α10" is normally occupied by a positively charged residue (eg. ARG516 in ERα) or a polar residue (eg. GLN867 in AR), which have the appropriate side chain properties to make a polar/charge interaction with the dimerisation partner. According to the Genome Threader alignment of O75385 (LBDG1) with *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain (1ERR:A), GLN1018 of O75385 (LBDG1) would occupy position "g" on helix "α10". Clearly GLN1018 has the appropriate side chain properties to make a polar/charge interaction with the dimerisation partner. This supports the annotation of O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain. The significance of GLN1018 of O75385 (LBDG1) occupying position "g" on helix "α10" is enhanced by the observation that

WO 02/070559

PCT/GB02/00948

66

GLN1018 is conserved in the O75385 (LBDG1) *Mus musculus* ortholog O70405 (see Figure 14 box "g"). The other two O75385 (LBDG1) homologs (BAA31598.1 and BAA77341.1) are predicted to have a SER at position "g". Clearly a SER also has the appropriate side chain properties to make a polar/charge interaction with the dimerisation partner.

Dimerisation of Ligand Binding Domains is also driven by a number of hydrophobic interactions. For example, position "a" on helix " $\alpha$ 10" is normally occupied by an aromatic residue (eg. PHE432 in PPAR $\gamma$ ) or a large hydrophobic residue (eg. LEU504 in ER $\alpha$ ) which have the appropriate side chain properties to make a hydrophobic interaction with the dimerisation partner. According to the Genome Threader alignment of O75385 (LBDG1) with *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain (IERR:A), TYR1008 of O75385 (LBDG1) would occupy position "a" on helix " $\alpha$ 10". Clearly TYR1008 has the appropriate side chain properties to make a hydrophobic interaction with the dimerisation partner. This supports the annotation of O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain. The significance of TYR1008 of O75385 (LBDG1) occupying position "a" on helix " $\alpha$ 10" is enhanced by the observation that TYR1008 is conserved in the three O75385 (LBDG1) homologues (O70405, BAA31598.1 and BAA77341.1, see Figure 14 box "a").

The N-terminal half of helix " $\alpha$ 10" also contains a number of other positions (boxes c, d, e, f and h, Figure 14) which are occupied by conserved hydrophobic residues in known Ligand Binding Domains. These hydrophobic residues are either solvent exposed and contribute to the hydrophobic interaction with the dimerisation partner or are buried and are involved in folding helix " $\alpha$ 10" against the main body of the Ligand Binding Domain (eg. LEU507 and LEU514 of ER $\alpha$ ). According to the Genome Threader alignment of O75385 (LBDG1) with the *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain (IERR:A), hydrophobic residues would occupy all of these positions in O75385 (LBDG1) (ALA1010 in position "c", LEU1011 in position "d", LEU1014 in position "e", LEU1017 in position "f" and MET1020 in position "h", see Figure 14). This supports the annotation of O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain. The significance of hydrophobic residues occupying positions c, d, e, f,

WO 02/070559

PCT/GB02/00948

67

and h, in O75385 (LBDG1) is further enhanced by the observation that hydrophobic residues also occupy these positions in the three O75385 (LBDG1) homologues (O70405, BAA31598.1 and BAA77341.1, see Figure 14).

5 Taken together, this residue analysis indicates that O75385 (LBDG1) possesses the appropriate complement of residues (at the N-terminus of helix "α10") to homo- or hetero-dimerise with Nuclear Hormone Ligand Binding Domains. This supports the annotation of O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

10 A further test of the Genome Threader annotation of O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain is to attempt to construct a homology model of O75385 (LBDG1) adopting the LBD fold. A homology model of O75385 (LBDG1) was constructed by submitting the Genome Threader alignment of O75385 (LBDG1) with the *Homo sapiens* Retinoic Acid Receptor gamma structure 1FCX to Modeller version 5 (Sali, A and Blundell T.L. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234,779-815, 1993).

15 An overall view of the homology model is presented in Figure 16. It can be seen that O75385 (LBDG1) has been successfully modelled onto the 1FCX structure and exhibits the typical organization of cross-latticed α-helices which make up the LBD fold. Figure 17 focuses on the two key alpha helices "α3" and "α5" (which correspond to α3 and α5 in known Nuclear Hormone Receptor Ligand Binding Domain structures). Figure 18 zooms in on the predicted Co-Activator binding site. In known LBD structures the Co-Activator binding site is a groove formed between α3 and α5 on the Ligand Binding Domain surface. Figure 19 shows the same view, but with a cartoon of the interacting Co-Activator helix added to illustrate the predicted mode of Co-Activator binding to the groove). The five residues of the predicted "LBD motif" in O75385 (LBDG1), LEU878, ASP885, GLN886, LEU889 and LEU890 are marked in dark grey, and are all in the appropriate orientations to fulfill their roles without any clashes or steric hindrance. For example, GLN886 is in a suitable position to interact with the C-terminus of the Co-Activator helix, just as is observed for the equivalent residue GLN375 of Estrogen receptor alpha, when it binds the Co-activator helix of SRC-1 (A. K. Shiau, D. Barstad, P.

20  
25  
30

WO 02/070559

PCT/GB02/00948

68

M. Loria, Lin Cheng, P. J. Kushner, D. A. Agard, and G. L. Greene, "The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen", *Cell* 1998 95: 927). Similarly, LEU890 projects into the groove and could make hydrophobic contacts with several of the Leucines present in the LXXLL Co-Activator helix, just as is observed for the equivalent residue LEU379 of Estrogen receptor alpha, when it binds the Co-activator helix of SRC-1 (A. K. Shiau, D. Barstad, P. M. Loria, Lin Cheng, P. J. Kushner, D. A. Agard, and G. L. Greene, The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen *Cell* 1998 95: 927).

10 Residues outside of the "LBD motif" are also positioned in appropriate orientations to fulfill their roles. For example, residues predicted to be positioned at the N-terminus of the "dimerisation" helix " $\alpha 10$ " are in suitable orientations to participate in interactions with a dimerisation partner. Figure 20 shows an overview of the structure of the *Homo sapiens* Estrogen Receptor alpha homodimer. The principle feature of the homodimer interface is the intertwined helices " $\alpha 10$ " of each monomer. Figure 21 compares the Estrogen Receptor alpha monomer (Figure 21A) with the homology model of O75385 (LBDG1) adopting the LBD fold (Figure 21B). It is clear that the region of O75385 (LBDG1) predicted by Genome Threader to be positioned at helix " $\alpha 10$ " has been successfully modelled as a long continuous alpha helix without any clashes or steric hindrance. This supports the Genome Threader annotation of O75385 (LBDG1) as containing a Nuclear Hormone Receptor Ligand Binding Domain. Furthermore, residues predicted to make key charge/polar interactions are in the appropriate orientation to interact with a dimerisation partner. For example, LYS1009 of O75385 (LBDG1) projects away from the body of the Ligand Binding Domain, towards where a dimerisation partner would be located, just as is observed for the equivalent residue GLN506 of Estrogen receptor alpha. For example, GLN1018 of O75385 (LBDG1) projects away from the body of the Ligand Binding Domain, towards where a dimerisation partner would be located, just as is observed for the equivalent residue ARG515 of Estrogen receptor alpha. Furthermore, residues predicted to make key hydrophobic interactions are in the appropriate orientation to interact with a dimerisation partner. For example, TYR1007 of O75385 (LBDG1) projects away from the body of the

15

20

25

30

WO 02/070559

PCT/GB02/00948

69

Ligand Binding Domain, towards where a dimerisation partner would be located, just as is observed for the equivalent residue LEU504 of Estrogen receptor alpha. Figure 21 shows a close-up of helix "α10" for both ERalpha (Figure 21A) and the homology model of O75385 (LBDG1) adopting the LBD fold (Figure 21B). Figure 21 also reveals that residues predicted by the Genome Threader alignment to function in the folding of helix "α10" against the body of the Ligand Binding Domain are in the appropriate orientation to fulfill this role without any clashes or steric hindrance. For example, ALA1010 of O75385 (LBDG1) projects into the body of the Ligand Binding Domain, where it can make hydrophobic contacts with the core of the structure, just as is observed for the equivalent residue LEU507 of Estrogen Receptor alpha. For example, LEU1017 of O75385 (LBDG1) projects into the body of the Ligand Binding Domain, where it can make hydrophobic contacts with the core of the structure, just as is observed for the equivalent residue ILE514 of Estrogen receptor alpha.

The homology modelling of O75385 (LBDG1) as a Ligand Binding Domain supports the Genome Threader annotation of O75385 (LBDG1) as containing a Ligand Binding Domain.

### Example 3

This example relates to a novel protein, termed BAA31598.1 herein identified as a Nuclear Hormone Receptor Ligand Binding Domain and to the use of this protein and nucleic acid sequence from the encoding gene in the diagnosis, prevention and treatment of disease.

This protein is a *Homo sapiens* paralogue of O75385 (LBDG1), and will be referred to herein as LBDG4 (BAA31598.1). This protein has 51% sequence identity to O75385 (LBDG1), and furthermore, residues predicted to play key roles in the Ligand Binding Domain fold of O75385 (LBDG1) are conserved in BAA31598.1 (LBDG4, see Figure 22). On the basis of the high homology to O75385 (LBDG1), and the conservation of key predicted residues, we annotate BAA31598.1 (LBDG4) as also containing a Nuclear Hormone Receptor Ligand Binding Domain.

The *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain contains an "LBD motif" which has been found in all annotated Nuclear Hormone Receptor Ligand Binding

WO 02/070559

PCT/GB02/00948

70

Domains to date. The "LBD motif" is involved in recruiting Nuclear Hormone Receptor Co-Activators and Co-Repressors. The 6 residues PHE367, LEU370, ASP374, GLN375, LEU378 and LEU379 constitute this motif in the *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain (see Figure 22: 1ERR:A). Three of these "LBD motif" residues are precisely conserved in BAA31598.1 (LBDG4). 1ERR:A ASP374 is conserved as ASP870 in BAA31598.1 (LBDG4). 1ERR:A GLN375 is conserved as GLN871 in BAA31598.1 (LBDG4). 1ERR:A LEU379 is conserved as LEU875 in BAA31598.1 (LBDG4). A fourth residue of the 1ERR:A "LBD motif", PHE367 is conservatively substituted by ILE863 in BAA31598.1 (LBDG4). This supports the annotation of BAA31598.1 (LBDG4) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

In addition to the "LBD motif" residues, there are other residues which are well conserved in known Nuclear Hormone Receptor Ligand Binding Domains. A further test of the annotation of BAA31598.1 (LBDG4) as containing a Nuclear Hormone Receptor Ligand Binding Domain is to analyse whether important residues outside of the "LBD motif" are conserved in BAA31598.1 (LBDG4).

In addition to Ligand binding and Co-Activator/Co-Repressor recruitment, another characteristic property of Nuclear Hormone Receptor Ligand Binding Domains is their ability to dimerise with other Nuclear Hormone Receptor Ligand Binding Domains. For example, the Estrogen Receptor alpha (ER $\alpha$ ) Ligand Binding Domain can homodimerise with itself.

Another example is that the Peroxisome Proliferator Activated Receptor gamma (PPAR $\gamma$ ) Ligand Binding domain can heterodimerise with the Retinoid X Receptor alpha (RXR $\alpha$ ) Ligand Binding Domain. Regardless of whether a Ligand Binding Domain is forming a homo- or hetero-dimer, the same general mode of interaction underlies the association. Crystallographic and mutagenesis experiments have identified the N-terminal half of helix " $\alpha 10$ " as the primary site of interaction between two dimerising Ligand Binding Domains (Gampe, R T., Montana, V G., Lambert, M H., Miller, A B., Bledsoe, R K., Milburn, M V., Kliewer, S A., Willson, T M., Xu, H E., Mol. Cell. 2000, 5 (3):545-55 and (Brzozowski, A M., Pike, A C., Dauter, Z., Hubbard, R E., Bonn, T., Engstrom, O.,

WO 02/070559

PCT/GB02/00948

71

Ohman, L., Greene, G. L., Gustafsson, J. A., Carlquist, M., Nature. 1997, vol.389 (6652):753-8). The functional importance of the N-terminal half of helix "α10" is reflected in the strong conservation of particular residues in this region. Figure 14 shows BAA31598.1 (LBDG4), and its three homologues, aligned with the N-terminal half of helix "α10" sequences for *Homo sapiens* ERα, AR, RARγ, RXRα, PPARγ and RORβ. To form a dimer, the "α10" helix of each monomer intertwine to form a rigid backbone, and this is driven by a combination of polar and hydrophobic interactions. There are two residue positions on helix "α10" which play a significant role in the polar component of the dimerisation interactions (these are marked as box "b" and box "g" in Figure 14).

10 Position "b" on helix "α10" is normally occupied by a polar residue (eg. GLN506 in ERα) or a positively charged residue (eg. LYS417 in RXRα), which have the appropriate side chain properties to make a polar/charge interaction with the dimerisation partner. According to the alignment of BAA31598.1 (LBDG4) with *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain (1ERR:A), LYS994 of BAA31598.1 (LBDG4) would occupy position "b" on helix "α10". Clearly LYS994 has the appropriate side chain properties to make a polar/charge interaction with the dimerisation partner. This supports the annotation of BAA31598.1 (LBDG4) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

20 Position "g" on helix "α10" is normally occupied by a positively charged residue (eg. ARG516 in ERα) or a polar residue (eg. THR417 in RORbeta), which have the appropriate side chain properties to make a polar/charge interaction with the dimerisation partner. According to the alignment of BAA31598.1 (LBDG4) with *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain (1ERR:A), SER1003 of BAA31598.1 (LBDG4) would occupy position "g" on helix "α10". Clearly SER1003 has the appropriate side chain properties to make a polar/charge interaction with the dimerisation partner. This supports the annotation of BAA31598.1 (LBDG4) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

25 Dimerisation of Ligand Binding Domains is also driven by a number of hydrophobic interactions. For example, position "a" on helix "α10" is normally occupied by an aromatic residue (eg. PHE432 in PPARγ) or a large hydrophobic residue (eg. LEU504 in

30

WO 02/070559

PCT/GB02/00948

72

ER $\alpha$ ) which have the appropriate side chain properties to make a hydrophobic interaction with the dimerisation partner. According to the alignment of BAA31598.1 (LBDG4) with *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain (1ERR:A), TYR992 of BAA31598.1 (LBDG4) would occupy position "a" on helix " $\alpha$ 10". Clearly TYR992 has  
5 the appropriate side chain properties to make a hydrophobic interaction with the dimerisation partner. This supports the annotation of BAA31598.1 (LBDG4) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

The N-terminal half of helix " $\alpha$ 10" also contains a number of other positions (boxes c, d, e, f and h, Figure 14) which are occupied by conserved hydrophobic residues in known  
10 Ligand Binding Domains. These hydrophobic residues are either solvent exposed and contribute to the hydrophobic interaction with the dimerisation partner or are buried and are involved in folding helix " $\alpha$ 10" against the main body of the Ligand Binding Domain (eg. LEU507 and LEU514 of ER $\alpha$ ). According to the alignment of BAA31598.1 (LBDG4) with the *Homo sapiens* Estrogen Receptor alpha Ligand Binding Domain  
15 (1ERR:A), hydrophobic residues would occupy all of these positions in BAA31598.1 (LBDG4) (ALA995 in position "c", ALA996 in position "d", LEU999 in position "e", LEU1002 in position "f" and ILE1005 in position "h", see Figure 14). This supports the annotation of BAA31598.1 (LBDG4) as containing a Nuclear Hormone Receptor Ligand Binding Domain. Taken together, this residue analysis indicates that BAA31598.1  
20 (LBDG4) possesses the appropriate complement of residues (at the N-terminus of helix " $\alpha$ 10") to homo- or hetero-dimerise with Nuclear Hormone Ligand Binding Domains. This supports the annotation of BAA31598.1 (LBDG4) as containing a Nuclear Hormone Receptor Ligand Binding Domain.

A further test of the annotation of BAA31598.1 (LBDG4) as containing a Nuclear  
25 Hormone Receptor Ligand Binding Domain is to attempt to construct a homology model of BAA31598.1 (LBDG4) adopting the LBD fold. A homology model of BAA31598.1 (LBDG4) was constructed based on the alignment shown in Figure 22. The homology modeling was performed using Modeller version 5 (Sali, A and Blundell T.L. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234,779-  
30 815, 1993).

WO 02/070559

PCT/GB02/00948

73

An overall view of the homology model is presented in Figure 23. It can be seen that BAA31598.1 (LBDG4) has been successfully modelled on the LBD fold and exhibits the typical organization of cross-latticed  $\alpha$ -helices which make up the LBD fold. The four residues of the predicted "LBD motif" in BAA31598.1 (LBDG4), ILE863, ASP870, 5 GLN871 and LEU875 are marked in black, and are all in the appropriate orientations to fulfill their roles without any clashes or steric hindrance.

Residues outside of the "LBD motif" are also positioned in appropriate orientations to fulfill their roles. For example, residues predicted to be positioned at the N-terminus of the "dimerisation" helix " $\alpha$ 10" are in suitable orientations to participate in interactions with a dimerisation partner (Figure 24). It is clear that the region of BAA31598.1 10 (LBDG4) predicted to be positioned at the N-terminus of helix " $\alpha$ 10" has been successfully modelled as an alpha helix without any clashes or steric hindrance. This supports the Genome annotation of BAA31598.1 (LBDG4) as containing a Nuclear Hormone Receptor Ligand Binding Domain. Furthermore, residues predicted to make 15 key charge/polar interactions are in the appropriate orientation to interact with a dimerisation partner. For example, LYS994 of BAA31598.1 (LBDG4) projects away from the body of the Ligand Binding Domain, towards where a dimerisation partner would be located, just as is observed for the equivalent residue GLN506 of Estrogen receptor alpha. For example, SER1003 of BAA31598.1 (LBDG4) projects away from the 20 body of the Ligand Binding Domain, towards where a dimerisation partner would be located, just as is observed for the equivalent residue ARG515 of Estrogen receptor alpha. Furthermore, residues predicted to make key hydrophobic interactions are in the appropriate orientation to interact with a dimerisation partner. For example, TYR992 of BAA31598.1 (LBDG4) projects away from the body of the Ligand Binding Domain, 25 towards where a dimerisation partner would be located, just as is observed for the equivalent residue LEU504 of Estrogen receptor alpha. Figure 24B and 24C show a close-up of helix " $\alpha$ 10" for the homology model of BAA31598.1 (LBDG4) adopting the LBD fold. Figure 24B and 24C also reveals that residues predicted to function in the folding of helix " $\alpha$ 10" against the body of the Ligand Binding Domain are in the 30 appropriate orientation to fulfill this role without any clashes or steric hindrance. For example, ALA995 of BAA31598.1 (LBDG4) projects into the body of the Ligand

WO 02/070559

PCT/GB02/00948

74

Binding Domain, where it can make hydrophobic contacts with the core of the structure, just as is observed for the equivalent residue LEU507 of Estrogen receptor alpha. For example, LEU1002 of BAA31598.1 (LBDG4) projects into the body of the Ligand Binding Domain, where it can make hydrophobic contacts with the core of the structure, just as is observed for the equivalent residue ILE514 of Estrogen receptor alpha.

The homology modelling of BAA31598.1 (LBDG4) as a Ligand Binding Domain supports the annotation of BAA31598.1 (LBDG4) as containing a Ligand Binding Domain.

WO 02/070559

PCT/GB02/00948

75

## CLAIMS

1. A polypeptide, which polypeptide:
- (i) comprises or consists of the amino acid sequence as recited in SEQ ID NO:2;
  - (ii) is a fragment thereof having Nuclear Hormone Receptor Ligand Binding Domain activity or having an antigenic determinant in common with the polypeptide of (i); or
  - (iii) is a functional equivalent of (i) or (ii).
2. A polypeptide which is a fragment according to claim 1(ii), which includes the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 polypeptide, said Nuclear Hormone Receptor Ligand Binding Domain region being defined as including residues 822 to 1020 inclusive, of the amino acid sequence recited in SEQ ID NO:2, wherein said fragment possesses the "LBD motif" residues LEU878, ASP885, GLN886 LEU889 and LEU890, or equivalent residues, and possesses Nuclear Hormone Receptor Ligand Binding Domain activity.
3. A polypeptide which is a functional equivalent according to claim 1(iii), is homologous to the amino acid sequence as recited in SEQ ID NO:2, possesses the catalytic residues LEU878, ASP885, GLN886 LEU889 and LEU890, or equivalent residues, and has Nuclear Hormone Receptor Ligand Binding Domain activity.
4. A polypeptide, which polypeptide:
- (i) comprises or consists of the amino acid sequence as recited in SEQ ID NO:4;
  - (ii) is a fragment thereof having Nuclear Hormone Receptor Ligand Binding Domain activity or having an antigenic determinant in common with the polypeptide of (i); or
  - (iii) is a functional equivalent of (i) or (ii).

WO 02/070559

PCT/GB02/00948

76

5. A polypeptide which is a fragment according to claim 4(ii), which includes the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG4 polypeptide, said Nuclear Hormone Receptor Ligand Binding Domain region being defined as including residues 805 to 1005 inclusive, of the amino acid sequence recited in SEQ ID NO:4, wherein said fragment possesses the "LBD motif" residues ILE863, ASP870, GLN871 and LEU875, or equivalent residues, and possesses Nuclear Hormone Receptor Ligand Binding Domain activity.
6. A polypeptide which is a functional equivalent according to claim 4(iii), is homologous to the amino acid sequence as recited in SEQ ID NO:4, possesses the catalytic residues ILE863, ASP870, GLN871 and LEU875, or equivalent residues, and has Nuclear Hormone Receptor Ligand Binding Domain activity.
7. A polypeptide according to claim 3 or 6, wherein said functional equivalent is homologous to the Nuclear Hormone Receptor Ligand Binding Domain region of the LBDG1 polypeptide.
8. A fragment or functional equivalent according to any one of claims 1-7, which has greater than 80% sequence identity with an amino acid sequence as recited in SEQ ID NO:2 or SEQ ID NO:4, or with a fragment thereof that possesses Nuclear Hormone Receptor Ligand Binding Domain activity, preferably greater than 85%, 90%, 95%, 98% or 99% sequence identity, as determined using BLAST version 2.1.3 using the default parameters specified by the NCBI (the National Center for Biotechnology Information; <http://www.ncbi.nlm.nih.gov/>) [Blosum 62 matrix; gap open penalty=11 and gap extension penalty=1].
9. A functional equivalent according to any one of claims 1-8, which exhibits significant structural homology with a polypeptide having the amino acid sequence given in any one of SEQ ID NO:2 or SEQ ID NO:4, or with a fragment thereof that possesses Nuclear Hormone Receptor Ligand Binding Domain activity.
10. A fragment as recited in claim 1, 2, 4 or 5, having an antigenic determinant in common with the polypeptide of claim 1(i), which consists of 7 or more (for example, 8, 10, 12, 14, 16, 18, 20 or more) amino acid residues from the sequence of SEQ ID NO:2 or SEQ ID NO:4.

WO 02/070559

PCT/GB02/00948

77

11. A purified nucleic acid molecule which encodes a polypeptide according to any one of the preceding claims.
12. A purified nucleic acid molecule according to claim 11, which has the nucleic acid sequence as recited in SEQ ID NO:1 or SEQ ID NO:3, or is a redundant equivalent or  
5 fragment thereof.
13. A fragment of a purified nucleic acid molecule according to claim 11 or claim 12, which comprises nucleotides 2732 to 3328 of SEQ ID NO:1 or nucleotides 2913 to 3515 of SEQ ID NO:3, or is a redundant equivalent thereof.
14. A purified nucleic acid molecule which hybridizes under high stringency conditions  
10 with a nucleic acid molecule according to any one of claims 11-13.
15. A vector comprising a nucleic acid molecule as recited in any one of claims 11-14.
16. A host cell transformed with a vector according to claim 15.
17. A ligand which binds specifically to, and which preferably inhibits the Nuclear Hormone Receptor Ligand Binding Domain activity of, a polypeptide according to  
15 any one of claims 1-10.
18. A ligand according to claim 17, which is an antibody.
19. A compound that either increases or decreases the level of expression or activity of a polypeptide according to any one of claims 1-10.
20. A compound according to claim 19 that binds to a polypeptide according to any one  
20 of claims 1-10 without inducing any of the biological effects of the polypeptide.
21. A compound according to claim 19 or claim 20, which is a natural or modified substrate, ligand, enzyme, receptor or structural or functional mimetic.
22. A polypeptide according to any one of claims 1-10, a nucleic acid molecule according to any one of claims 11-14, a vector according to claim 15, a ligand according to  
25 claim 17 or 18, or a compound according to any one of claims 19-21, for use in therapy or diagnosis of disease.
23. A method of diagnosing a disease in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide according to any one of claim 1-

SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

78

10, or assessing the activity of a polypeptide according to any one of claim 1-10, in tissue from said patient and comparing said level of expression or activity to a control level, wherein a level that is different to said control level is indicative of disease.

24. A method according to claim 23 that is carried out *in vitro*.

5 25. A method according to claim 23 or claim 24, which comprises the steps of: (a) contacting a ligand according to claim 17 or claim 18 with a biological sample under conditions suitable for the formation of a ligand-polypeptide complex; and (b) detecting said complex.

26. A method according to claim 23 or claim 24, comprising the steps of:

10 a) contacting a sample of tissue from the patient with a nucleic acid probe under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule according to any one of claims 11-14 and the probe;

b) contacting a control sample with said probe under the same conditions used in step a); and

15 c) detecting the presence of hybrid complexes in said samples;

wherein detection of levels of the hybrid complex in the patient sample that differ from levels of the hybrid complex in the control sample is indicative of disease.

27. A method according to claim 23 or claim 24, comprising:

20 a) contacting a sample of nucleic acid from tissue of the patient with a nucleic acid primer under stringent conditions that allow the formation of a hybrid complex between a nucleic acid molecule according to any one of claims 11-14 and the primer;

b) contacting a control sample with said primer under the same conditions used in step a); and

25 c) amplifying the sampled nucleic acid; and

d) detecting the level of amplified nucleic acid from both patient and control samples;

WO 02/070559

PCT/GB02/00948

79

wherein detection of levels of the amplified nucleic acid in the patient sample that differ significantly from levels of the amplified nucleic acid in the control sample is indicative of disease.

28. A method according to claim 23 or claim 24 comprising:

- 5 a) obtaining a tissue sample from a patient being tested for disease;
- b) isolating a nucleic acid molecule according to any one of claims 11-14 from said tissue sample; and
- c) diagnosing the patient for disease by detecting the presence of a mutation which is associated with disease in the nucleic acid molecule as an indication of the disease.

10 29. The method of claim 28, further comprising amplifying the nucleic acid molecule to form an amplified product and detecting the presence or absence of a mutation in the amplified product.

15 30. The method of either claim 28 or 29, wherein the presence or absence of the mutation in the patient is detected by contacting said nucleic acid molecule with a nucleic acid probe that hybridises to said nucleic acid molecule under stringent conditions to form a hybrid double-stranded molecule, the hybrid double-stranded molecule having an unhybridised portion of the nucleic acid probe strand at any portion corresponding to a mutation associated with disease; and detecting the presence or absence of an unhybridised portion of the probe strand as an indication of the presence or absence

20 of a disease-associated mutation.

31. A method according to any one of claims 23-30, wherein said disease is selected from cell proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and other solid tumours, myeloproliferative disorders, such as leukemia, non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposi's sarcoma, autoimmune/inflammatory disorders, including allergy, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation, asthma, and organ transplant rejection, cardiovascular disorders, including hypertension, oedema, angina, atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia, neurological disorders including, central

SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

80

- nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain, developmental disorders, metabolic disorders including diabetes mellitus, osteoporosis, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, and obesity, renal disorders, including glomerulonephritis, renovascular hypertension, dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging, AIDS, infections including viral infection, bacterial infection, fungal infection and parasitic infection and other pathological conditions, particularly those in which nuclear hormone receptors are implicated.
- 10 32. Use of a polypeptide according to any one of claims 1-10 as a Nuclear Hormone Receptor Ligand Binding Domain.
33. Use of a nucleic acid molecule according to any one of claims 11-14 to express a protein that possesses Nuclear Hormone Receptor Ligand Binding Domain activity.
34. A method for effecting cell-cell adhesion, utilising a polypeptide according to any one of claims 1-10.
- 15 35. A pharmaceutical composition comprising a polypeptide according to any one of claims 1-10, a nucleic acid molecule according to any one of claims 11-14, a vector according to claim 15, a ligand according to claim 17 or 18, or a compound according to any one of claims 19-21.
- 20 36. A vaccine composition comprising a polypeptide according to any one of claims 1-10 or a nucleic acid molecule according to any one of claims 11-14.
37. A polypeptide according to any one of claims 1-10, a nucleic acid molecule according to any one of claims 11-14, a vector according to claim 15, a ligand according to claim 17 or 18, a compound according to any one of claims 19-21, or a pharmaceutical composition according to claim 35 for use in the manufacture of a medicament for the treatment of a cell proliferative disorders, including neoplasm, melanoma, lung, colorectal, breast, pancreas, head and neck and other solid tumours, myeloproliferative disorders, such as leukemia, non-Hodgkin lymphoma, leukopenia, thrombocytopenia, angiogenesis disorder, Kaposis' sarcoma,
- 25

SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

81

- autoimmune/inflammatory disorders, including allergy, inflammatory bowel disease, arthritis, psoriasis and respiratory tract inflammation, asthma, and organ transplant rejection, cardiovascular disorders, including hypertension, oedema, angina, atherosclerosis, thrombosis, sepsis, shock, reperfusion injury, heart arrhythmia, and ischemia, neurological disorders including, central nervous system disease, Alzheimer's disease, brain injury, stroke, amyotrophic lateral sclerosis, anxiety, depression, and pain, developmental disorders, metabolic disorders including diabetes mellitus, osteoporosis, lipid metabolism disorder, hyperthyroidism, hyperparathyroidism, hypercalcemia, hypercholesterolemia, hyperlipidemia, and obesity, renal disorders, including glomerulonephritis, renovascular hypertension, dermatological disorders, including, acne, eczema, and wound healing, negative effects of aging, AIDS, infections including viral infection, bacterial infection, fungal infection and parasitic infection and other pathological conditions, particularly those in which nuclear hormone receptors are implicated.
- 5
- 10
- 15 38. A method of treating a disease in a patient, comprising administering to the patient a polypeptide according to any one of claims 1-10, a nucleic acid molecule according to any one of claims 11-14, a vector according to claim 15, a ligand according to claim 17 or 18, a compound according to any one of claims 19-21, or a pharmaceutical composition according to claim 35.
- 20 39. A method according to claim 38, wherein, for diseases in which the expression of the natural gene or the activity of the polypeptide is lower in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, vector, ligand, compound or composition administered to the patient is an agonist.
- 25 40. A method according to claim 38, wherein, for diseases in which the expression of the natural gene or activity of the polypeptide is higher in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, vector, ligand, compound or composition administered to the patient is an antagonist.

WO 02/070559

PCT/GB02/00948

82

41. A method of monitoring the therapeutic treatment of disease in a patient, comprising monitoring over a period of time the level of expression or activity of a polypeptide according to any one of claims 1-10, or the level of expression of a nucleic acid molecule according to any one of claims 11-14 in tissue from said patient, wherein  
5 altering said level of expression or activity over the period of time towards a control level is indicative of regression of said disease.
42. A method for the identification of a compound that is effective in the treatment and/or diagnosis of disease, comprising contacting a polypeptide according to any one of claims 1-10, a nucleic acid molecule according to any one of claims 11-14, or a host  
10 cell according to claim 16 with one or more compounds suspected of possessing binding affinity for said polypeptide or nucleic acid molecule, and selecting a compound that binds specifically to said nucleic acid molecule or polypeptide.
43. A kit useful for diagnosing disease comprising a first container containing a nucleic acid probe that hybridises under stringent conditions with a nucleic acid molecule  
15 according to any one of claims 11-14; a second container containing primers useful for amplifying said nucleic acid molecule; and instructions for using the probe and primers for facilitating the diagnosis of disease.
44. The kit of claim 43, further comprising a third container holding an agent for digesting unhybridised RNA.
- 20 45. A kit comprising an array of nucleic acid molecules, at least one of which is a nucleic acid molecule according to any one of claims 11-14.
46. A kit comprising one or more antibodies that bind to a polypeptide as recited in any one of claims 1-10; and a reagent useful for the detection of a binding reaction between said antibody and said polypeptide.
- 25 47. A transgenic or knockout non-human animal that has been transformed to express higher, lower or absent levels of a polypeptide according to any one of claims 1-10.
48. A method for screening for a compound effective to treat disease, by contacting a non-human transgenic animal according to claim 47 with a candidate compound and determining the effect of the compound on the disease of the animal.

SUBSTITUTE SHEET (RULE 26)

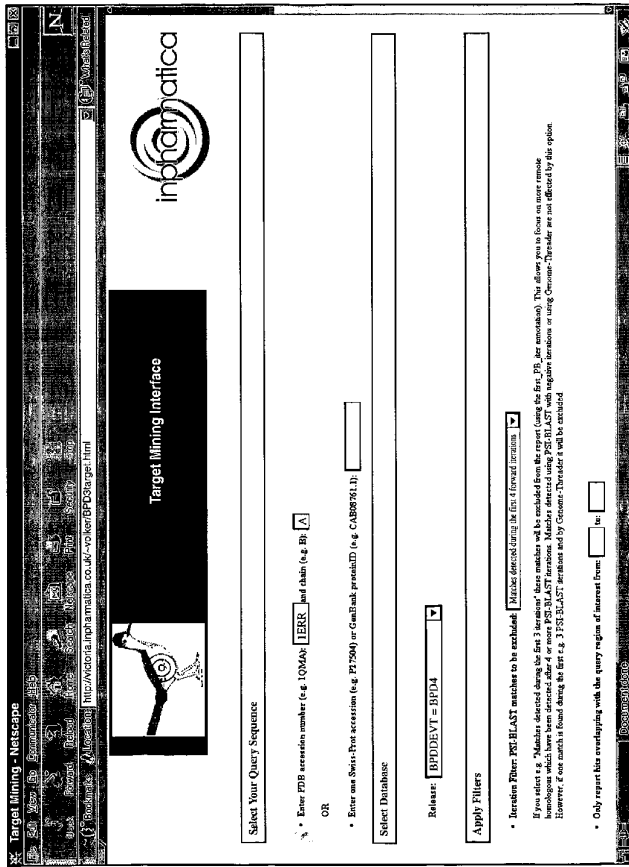


FIG. 1

SUBSTITUTE SHEET (RULE 26)









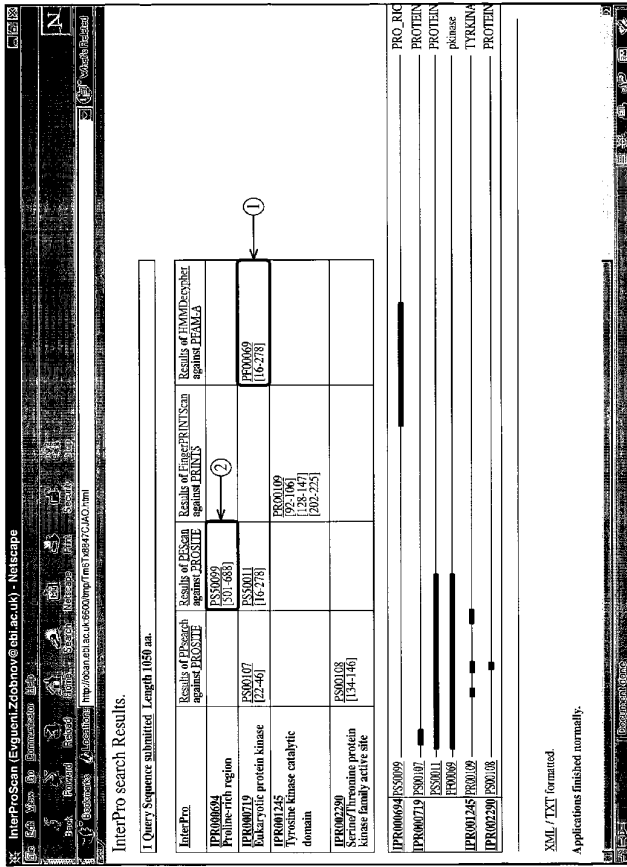


FIG. 4A

SUBSTITUTE SHEET (RULE 26)

InterPro Entry IPRO00694 - Nucleosome rich region

InterPro Entry IPRO00694

Prolin-rich region

Database	InterPro
Accession	IPRO00694 - Nucleosome rich region
Name	Prolin-rich region
Type	Region
Dates	08-OCT-1999 (created) 19-JUL-2000 (last modified)
Signatures	IPRO00694 - Nucleosome rich region
Abstract	Keratins from the skin epidermis are transcriptionally upregulated to protect against the damaging effects of external agents, including ultraviolet (UV) light. The presence of damage response genes in skin has been shown to be important for the development of skin cancer. One of the genes upregulated in skin cancer is the p16 <sup>INK4</sup> gene, which is known to mediate DNA-binding in repressor proteins. Small proline-rich proteins are also found in higher concentrations during normal keratinocyte differentiation, although the reason for this is unclear. Due to the large number of false positives found by the PROSITE signature, no protein matches are shown.
Examples	<ul style="list-style-type: none"> <li>133232 SP2B_HUMAN: Human Small proline rich protein 2B</li> <li>133233 SP2B_HUMAN: Human Small proline rich protein 2B</li> <li>543333 COBA_RAT: Rat Corrin alpha</li> <li>552527 COBB_MOUSE: Mouse Corrin B</li> </ul>
References	<ol style="list-style-type: none"> <li>1. <i>Journal of Cell Biology</i>, 148: 1011-1021 (2000)</li> <li>2. <i>Mol. Cell Biol.</i> 21: 95-100 (2001)</li> <li>3. <i>Journal of Cell Biology</i>, 152: 1011-1021 (2001)</li> <li>4. <i>Journal of Cell Biology</i>, 152: 1011-1021 (2001)</li> </ol>
Database links	PROSITE profile: 00005097
Matches	1 Match

FIG. 4B  
SUBSTITUTE SHEET (RULE 26)



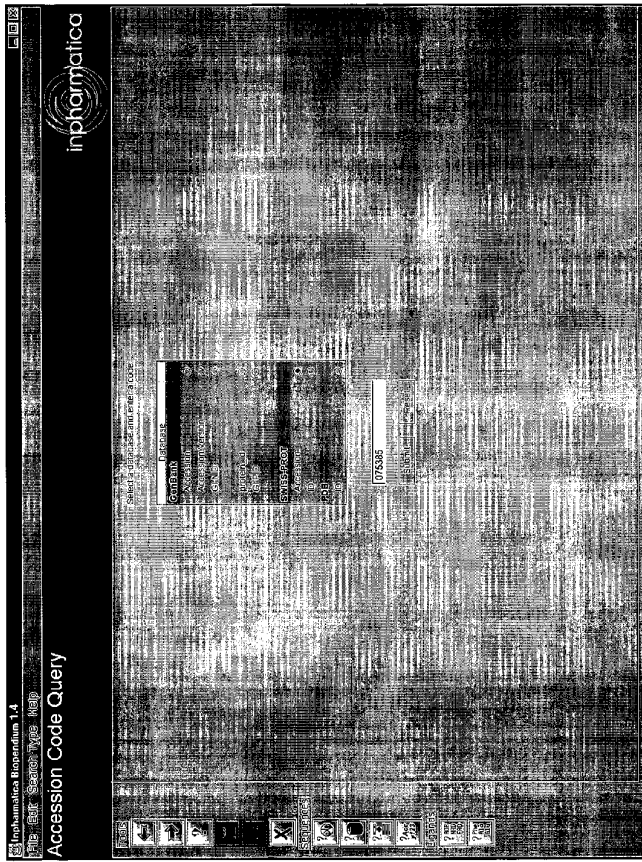


FIG. 6A

SUBSTITUTE SHEET (RULE 26)

inpharmatica **inpharmatica**  
 Query/details  
 O75385: SERINE/THREONINE-PROTEIN KINASE ULK1 (EC 2.7.1.) (UNC-51-LIKE KINASE 1).

inpharmatica Genomic Threader results

Query Detail	Link	No.	Contig	MD	Query	Label	Seq. Method	Contig	MD	Query	Label	Seq. Method
1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9
10	10	10	10	10	10	10	10	10	10	10	10	10
11	11	11	11	11	11	11	11	11	11	11	11	11
12	12	12	12	12	12	12	12	12	12	12	12	12
13	13	13	13	13	13	13	13	13	13	13	13	13
14	14	14	14	14	14	14	14	14	14	14	14	14
15	15	15	15	15	15	15	15	15	15	15	15	15
16	16	16	16	16	16	16	16	16	16	16	16	16
17	17	17	17	17	17	17	17	17	17	17	17	17
18	18	18	18	18	18	18	18	18	18	18	18	18
19	19	19	19	19	19	19	19	19	19	19	19	19
20	20	20	20	20	20	20	20	20	20	20	20	20
21	21	21	21	21	21	21	21	21	21	21	21	21
22	22	22	22	22	22	22	22	22	22	22	22	22
23	23	23	23	23	23	23	23	23	23	23	23	23
24	24	24	24	24	24	24	24	24	24	24	24	24
25	25	25	25	25	25	25	25	25	25	25	25	25
26	26	26	26	26	26	26	26	26	26	26	26	26
27	27	27	27	27	27	27	27	27	27	27	27	27
28	28	28	28	28	28	28	28	28	28	28	28	28
29	29	29	29	29	29	29	29	29	29	29	29	29
30	30	30	30	30	30	30	30	30	30	30	30	30
31	31	31	31	31	31	31	31	31	31	31	31	31
32	32	32	32	32	32	32	32	32	32	32	32	32
33	33	33	33	33	33	33	33	33	33	33	33	33
34	34	34	34	34	34	34	34	34	34	34	34	34
35	35	35	35	35	35	35	35	35	35	35	35	35
36	36	36	36	36	36	36	36	36	36	36	36	36
37	37	37	37	37	37	37	37	37	37	37	37	37
38	38	38	38	38	38	38	38	38	38	38	38	38
39	39	39	39	39	39	39	39	39	39	39	39	39
40	40	40	40	40	40	40	40	40	40	40	40	40
41	41	41	41	41	41	41	41	41	41	41	41	41
42	42	42	42	42	42	42	42	42	42	42	42	42
43	43	43	43	43	43	43	43	43	43	43	43	43
44	44	44	44	44	44	44	44	44	44	44	44	44
45	45	45	45	45	45	45	45	45	45	45	45	45
46	46	46	46	46	46	46	46	46	46	46	46	46
47	47	47	47	47	47	47	47	47	47	47	47	47
48	48	48	48	48	48	48	48	48	48	48	48	48
49	49	49	49	49	49	49	49	49	49	49	49	49
50	50	50	50	50	50	50	50	50	50	50	50	50
51	51	51	51	51	51	51	51	51	51	51	51	51
52	52	52	52	52	52	52	52	52	52	52	52	52
53	53	53	53	53	53	53	53	53	53	53	53	53
54	54	54	54	54	54	54	54	54	54	54	54	54
55	55	55	55	55	55	55	55	55	55	55	55	55
56	56	56	56	56	56	56	56	56	56	56	56	56
57	57	57	57	57	57	57	57	57	57	57	57	57
58	58	58	58	58	58	58	58	58	58	58	58	58
59	59	59	59	59	59	59	59	59	59	59	59	59
60	60	60	60	60	60	60	60	60	60	60	60	60
61	61	61	61	61	61	61	61	61	61	61	61	61
62	62	62	62	62	62	62	62	62	62	62	62	62
63	63	63	63	63	63	63	63	63	63	63	63	63
64	64	64	64	64	64	64	64	64	64	64	64	64
65	65	65	65	65	65	65	65	65	65	65	65	65
66	66	66	66	66	66	66	66	66	66	66	66	66
67	67	67	67	67	67	67	67	67	67	67	67	67
68	68	68	68	68	68	68	68	68	68	68	68	68
69	69	69	69	69	69	69	69	69	69	69	69	69
70	70	70	70	70	70	70	70	70	70	70	70	70
71	71	71	71	71	71	71	71	71	71	71	71	71
72	72	72	72	72	72	72	72	72	72	72	72	72
73	73	73	73	73	73	73	73	73	73	73	73	73
74	74	74	74	74	74	74	74	74	74	74	74	74
75	75	75	75	75	75	75	75	75	75	75	75	75
76	76	76	76	76	76	76	76	76	76	76	76	76
77	77	77	77	77	77	77	77	77	77	77	77	77
78	78	78	78	78	78	78	78	78	78	78	78	78
79	79	79	79	79	79	79	79	79	79	79	79	79
80	80	80	80	80	80	80	80	80	80	80	80	80
81	81	81	81	81	81	81	81	81	81	81	81	81
82	82	82	82	82	82	82	82	82	82	82	82	82
83	83	83	83	83	83	83	83	83	83	83	83	83
84	84	84	84	84	84	84	84	84	84	84	84	84
85	85	85	85	85	85	85	85	85	85	85	85	85
86	86	86	86	86	86	86	86	86	86	86	86	86
87	87	87	87	87	87	87	87	87	87	87	87	87
88	88	88	88	88	88	88	88	88	88	88	88	88
89	89	89	89	89	89	89	89	89	89	89	89	89
90	90	90	90	90	90	90	90	90	90	90	90	90
91	91	91	91	91	91	91	91	91	91	91	91	91
92	92	92	92	92	92	92	92	92	92	92	92	92
93	93	93	93	93	93	93	93	93	93	93	93	93
94	94	94	94	94	94	94	94	94	94	94	94	94
95	95	95	95	95	95	95	95	95	95	95	95	95
96	96	96	96	96	96	96	96	96	96	96	96	96
97	97	97	97	97	97	97	97	97	97	97	97	97
98	98	98	98	98	98	98	98	98	98	98	98	98
99	99	99	99	99	99	99	99	99	99	99	99	99
100	100	100	100	100	100	100	100	100	100	100	100	100

inpharmatica > SWISS-PROT > Job

FIG. 6B

SUBSTITUTE SHEET (RULE 26)

inpharmatica Biostrudium 1.4  
 File Edit Search Help  
 inpharmatica  
 Query details: 075588: SERINE/THREONINE-PROTEIN KINASE ULK1 (EC 2.1.1) (UNG-51-LIKE KINASE 1).  
 Reverse Maximised PSI-BLAST results  
 Page 1 of 273  
 Results: 11827  
 Select: 0  
 View alignment

Query Local	Query	Accession	Length	Identical	Similar	Expect	Low Comp. B	High Comp. B	Low Comp. E	High Comp. E
1	Q10405	SERINE/THREONINE-PROTEIN KINASE ULK1 (EC 2.1.1)	481	100	100	1.0E-106	1	1	1	1
2	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
3	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
4	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
5	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
6	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
7	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
8	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
9	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
10	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
11	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
12	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
13	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
14	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
15	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
16	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
17	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
18	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
19	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
20	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
21	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
22	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
23	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
24	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
25	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
26	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
27	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
28	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
29	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
30	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
31	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
32	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
33	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
34	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
35	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
36	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
37	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
38	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
39	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
40	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
41	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
42	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
43	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
44	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
45	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
46	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
47	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
48	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
49	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
50	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
51	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
52	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
53	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
54	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
55	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
56	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
57	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
58	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
59	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
60	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
61	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
62	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
63	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
64	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
65	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
66	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
67	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
68	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
69	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
70	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
71	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
72	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
73	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
74	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
75	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
76	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
77	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
78	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
79	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
80	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
81	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
82	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
83	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
84	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
85	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
86	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
87	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
88	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
89	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
90	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
91	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
92	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
93	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
94	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
95	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
96	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
97	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
98	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
99	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1
100	AA43871	homo sapiens	481	100	100	1.0E-106	1	1	1	1

FIG. 6C Aligned Sequence Display



SUBSTITUTE SHEET (RULE 26)

AIEye output (January 15, 2001 5:02 PM)

FIG. 7-1

```

075385 10 20 30 40 50 60 70 80
1ERRA 1 M E P G R G G T E T V G K F F S R K D L I G H G A F A V F K G R R R E K H D L E V A V K I N K N L A K S O T L L K E I K I L K E L K H E N I V A L Y D
.....
075385 90 100 110 120 130 140 150 160
1ERRA 2 F O E M A N S V Y L V M E Y C N G D L A D Y L H A M R T L S E D T I R L F L O G I A G A M R L L H S K G I I H R D L K P C N I L L S N P A G R R A N P N S I R
.....
075385 170 180 190 200 210 220 230 240
1ERRA 3 V K I A D F G A R Y L O S N M M A A T L C G S P M Y M A P E V I M S Q H Y D G K A D L W S I G T I V Y Q C L T G K A P F O A S S P O D L R L F Y E K M K T L V
.....
075385 250 260 270 280 290 300 310 320
1ERRA 4 P T I P R E T S A P L R Q L L L L O R N H K D R M D F D E F F H P F L D A S P S V R K S P P V P S Y P S G S G S S S T S H L A S P S L G E
.....
075385 330 340 350 360 370 380 390 400
1ERRA 5 M Q L O K T L A S P A D T A G F L H S R D S G G S K D S S C D T D E F V M V P A Q F P G D L V A E A P S A K P P P D S L M C G S S L V A S A G L E S H G R
.....
075385 410 420 430 440 450 460 470 480
1ERRA 6 T P S P P C S S P S P S G R A G F S S S R C G A S V P I P V P T O V N Y O R I E R N L O S P T O F O T P H S S A I R R S G S T S P L G F A R A S P S P
.....
075385 490 500 510 520 530 540 550 560
1ERRA 7 P A N A H G G V L A R K M S L G G G R P Y T P S P Q V G T I P E R P G W S G T P S P G A E M R G G R S P R P G S S A P E H S P R T S G L G C R L H S A P N L
.....
075385 570 580 590 600 610 620 630 640
1ERRA 8 S D L H V Y R P K L P K P P T D P L G A V F S P P O A S P P S H G L S C R N L R S P K L P D L Q R N P L P P I L G S P T K A V P S F D F P K T P S S O
.....
075385 650 660 670 680 690 700 710 720
1ERRA 9 N L L A L L A R Q G V V M T P P R N R T L P D L S E V G P F H G Q P L G P G L R P G E D K P G P F F G R S F S T S R L T D L L L K A A F G T C A P D P G S T E S L
.....

```

FIG. 7-2

AlEye output (February 22, 2001 12:07 PM)

```

075385 730 740 750 760 770 780 790 800
TERRA O E K P M E I A P S A G F G G S L H P G A R A G G T S S P S P V F T V G S P P S G S T P P O G P R T R M F S A G P T G S A S S A R H L V P G P C S E A P A P
      810 820 830 840 850 860 870 880
075385 E L P A P G H G C S F A D P I A A N L E G A V T F E A P D L P E E T L M E Q E H T E I L R G L R F T L L F . . . . . V Q H V L E (A) A . . . . . K K S A S F A A G S
TERRA . . . . . A L S L T A D Q M . . . . . V S A L L D A E P P I L Y S E Y D P T R P F F S E A S M M G L L (Q) N L (Q) D R (C) Q (S) F M I N W A K R . .
      890 900 910 920 930 940 950 960
075385 P E Y (L) D E S V V A (D) I S (I) S R E (G) G F A E C C (L) V (Q) . . . . . K V A E L (C) S G G L O S A I D Q I R A G K L C . . . . . L S S T (Q) V V R R L N E L Y K
TERRA - V P G (F) V D (I) T L H (D) V H L L L E C A (D) . . . . . L E I (D) (D) I G D V W (G) S M E H P G K L (C) A P N L L L D R N O G K C V E G (G) V E I F D M (L) A T S S R F R M M . .
      970 980 990 1000 1010 1020 1030 1040
075385 A S Y V S C G G L S R L R O R F F . L D K O R L L D R I H S I T A E R L I F S H A V Q M V Q S A L D E M F Q H R E G C V P R Y H K A L L L L E G L O M L S D
TERRA - N L G E E F V C L K S I I L L N S G V Y E E K D H I H R V L . . . . . D K I T D T L I H L M A K A G L T Q Q Q H Q R L A Q L L L I L S H I R H W S N K
      1050 1060
075385 Q A D I E N Y T K C K L C I E R R L S A L L T G I C A
TERRA G M E (Q) Q Y S M P Q Y D (L) L E M L D A H
      524 525 536 539

```

SUBSTITUTE SHEET (RULE 26)

- "LBD motif" residues in IERR:A
- Precisely conserved "LBD motif" residues in 075385
- Tolerated change in "LBD motif" residues in 075385
- Ligand binding residues in IERR:A
- Precisely conserved Ligand binding residue in 075385
- hydrophobicity conserved Ligand binding residue in 075385

end of  
genome  
threader  
alignment

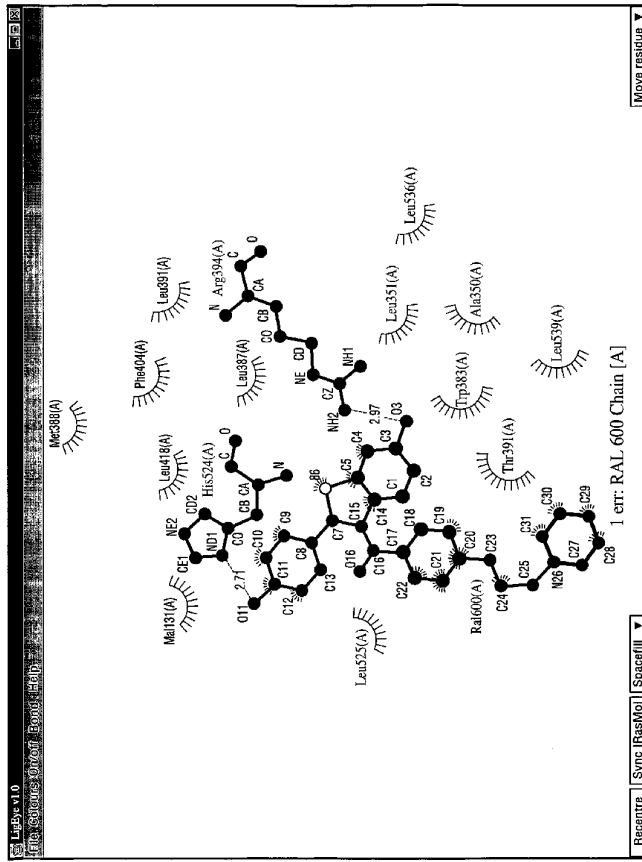


FIG. 8A

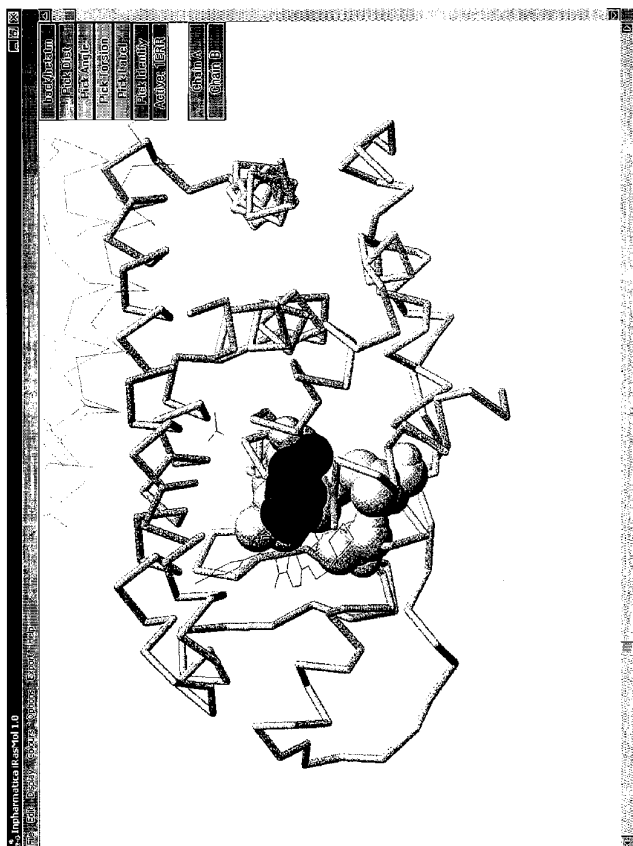


FIG. 8B

SUBSTITUTE SHEET (RULE 26)

AIEye output (February 22, 2001 2:30 PM)

FIG. 9-1

```

075385  M E P R G G T E T Y G K K F E F S R K D L I G H G A F A V F K G R H R E K H D L E V A V K I N K N L K A K S Q T L L K E I K I L K E L K H E N I V A L Y
070405  M E P R G G V E T Y G K K F E F S R K D L I G H G A F A V F K G R H R E K H D L E V A V K I N K N L K A K S Q T L L K E I K I L K E L K H E N I V A L Y
BAA31598.1  . . . . . M E V Y G D F E Y S K R D L V G H G A F A V F E R G R H R O T D E V A I K S I N K N L K S Q I L L G K E I K I L K E L Q H E N I V A L Y
BAA77341.1  . . . . . M E V Y G D F E Y C K R D L V G H G A F A V F E R G R H R O T D E V A I K S I N K N L K S Q I L L G K E I K I L K E L Q H E N I V A L Y
80      |
100     |
110     |
120     |
130     |
140     |
150     |
075385  D F Q E M A N S V Y L V M E Y C N G G D L A D Y L H A M R T L S E D T I R L F L Q I A G A M R L L H S K G I I H R D L K P O N I L L S N P A G R R A N P M S
070405  D F Q E M A N S V Y L V M E Y C N G G D L A D Y L H M R T L S E D T V R L F L Q O I A G A M R L L H S K G I I H R D L K P O N I L L S N P G R R A M P N S
BAA31598.1  D V Q E L P N S V F L V M E Y C N G G D L A D Y L Q A K G T L S E D T I R V F L H Q I A A A M R I L H S K G I I H R D L K P O N I L L S Y A N R R K S S V S G
BAA77341.1  D V Q E L P N S V F L V M E Y C N G G D L A D Y L Q A K G T L S E D T I R V F L H Q I A A A M R I L H S K G I I H R D L K P O N I L L S Y A N R R K S N V S G
160     |
170     |
180     |
190     |
200     |
210     |
220     |
075385  I R V K I A D F G F A R Y L O S N M M A A T L C G S P M Y M A P E V I M S O H Y D G K A D L W S I G T I V Y O C L T G K A P F O A S S P O D L R L F Y E K V K
070405  I R V K I A D F G F A R Y L O S N M M A A T L C G S P M Y M A P E V I M S O H Y D G K A D L W S I G T I V Y O C L T G K A P F O A S S P O D L R L F Y E K V K
BAA31598.1  I R I K I A D F G F A R Y L H S N M M A A T L C G S P M Y M A P E V I M S O H Y A K A D L W S I G T V I Y O C L V G K P P F O A N S P O D L R M F Y E K R I R
BAA77341.1  I R I K I A D F G F A R Y L H S N T M A A T L C G S P M Y M A P E V I M S O H Y A K A D L W S I G T V I Y O C L V G K P P F O A N S P O D L R M F Y E K R I R
240     |
250     |
260     |
270     |
280     |
290     |
300     |
075385  T L V P T I P R E T S A P I R O L L A I L Q R N H K D R M D F E F F H P F L D A S P S V R K S P P V P S Y P S S G S S S S S S S S S S T S U L A S P
070405  T L V P A I P R E T S A P I R O L L A L L Q R N H K D R M D F E F F H P F L D A S P T P K K S P V P V P S Y P S S G S S S S S S S S S S S S S S S S L A S P
BAA31598.1  S L M P S I P R E T S P Y L A N L L L G L L Q R N Q D R M D F E A F F S H P F - L E G P Y K K S C P V P P M Y G S V S O S S C G S S P S C R F A S P
BAA77341.1  S L M P S I P R E T S P Y L A N L L L G L L Q R N Q D R M D F E A F F S H P F - L E R Y P V K K S C P V P P V Y G G P P P S S S S S S S S S S S S P S C R F A S P
320     |
330     |
340     |
350     |
360     |
370     |
380     |
390     |
075385  S L G E M Q L O K T L A S P A . . . D T A G F L H S R D S G . . . G S K D S C D T D F V M Y P A G P P G D L V A E A P S A K P P P D S L M C S S S L V A
070405  S L G E M P L O K T L T S P A . . . D A A G F L O S R D S G . . . G S K D S C D T D F V M Y P A G P P G D L V A E A A S A K P P P D S L L C S S S L V A
BAA31598.1  S L P D M Q H I Q E E N L S P P L G P P N Y L O V S K D S A S T S K N S C D T D D F V L P P H N I S S O H S C D M P M G T A G R A S H E F L V C G G Q
BAA77341.1  S L P D M Q H I Q E E N L S P P L G P P N Y L O V S K D S A S N S K N S C D T D D F V L P P H N I S S D H S Y D M P M G T T A R R A S N E F F M C G G Q
400     |
410     |
420     |
430     |
440     |
450     |
460     |
470     |
075385  S A G L E S H G R T P S P S P P C S S S P S G R A G P F S S S R C G A S V P I P V P T Q V O N Y O R I E N L Q S P T O F Q T P R S S A I R R S G S T T P
070405  S A G L E S H G R T P S P S P S P S G R A G P F S S S R C G A S V P I P V P T Q V H N Y O R I E O N L Q S P T O Q O T P R S S A I R R S G S T T P
BAA31598.1  C O P T V S P . . . . . H S E T A P I P V P T Q I R N Y O R I E O N L T S A S S G T N V H G S P R S A V V R R S N T S P
BAA77341.1  C O P T V S P . . . . . H S E T A P I P V P T Q V R N Y O R I E O N L I S T A S S G T N P H G S P R S A V V R R S N T S P

```



FIG. 9-3

ALEye output (February 22, 2001 2:30 PM)

950 | 950 | 960 | 970 | 980 | 990 | 1000 | 1010 | 1020  
 075385 | GKCLST | KGVVRRRLNELYKASVVS | GGLSRLR | LORFFLDK | RLLDR | HSI | TAERL | IESHAV | QMYS | AALDEMF | OHREG  
 079405 | GKCLST | KGVVRRRLNELYKASVVS | GGLSRLR | LORFFLDK | RLLDR | HSI | TAERL | IESHAV | QMYS | AALDEMF | OHREG  
 BAA31598.1 | GKLSPTA | KGVVRRRLNELYKFCITMR | KKLTEKLN | RFFSDK | RFI | DEINSV | TAEKLI | YNCAVEM | YGSAALDEMF | OOTED  
 BAA77341.1 | GKLSPTA | KGVVRRRLNELYKFCITMR | KKLTEKLN | RFFSDK | RFI | DEINSV | TAEKLI | YNCAVEM | YGSAALDEMF | OOTED  
 1030 | 1040 | 1050 | 1060 | 1070  
 075385 | CVPRYH | KALLLEGL | QHMLSD | QAD | IENV | TKKLC | IERR | L | SALL | TGICA  
 079405 | CVPRYH | KALLLEGL | QHMLSD | QAD | IENV | TKKLC | IERR | L | SALL | TGICA  
 BAA31598.1 | IVRYHKA | ALLLEGL | SRIL | ODPAD | IENV | H | YKCSI | ERR | L | SAL | CH | STATV  
 BAA77341.1 | IVRYHKA | ALLLEGL | SRIL | ODPAD | IENV | H | YKCSI | ERR | L | SAL | CH | STATV





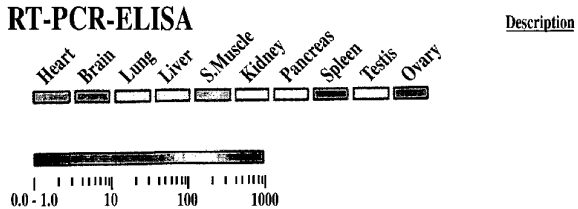
WO 02/070559

21/38

PCT/GB02/00948

FIG. 12

## Expression profile

**RT-PCR-ELISA**

## Experimental conditions

Primer\_f : TTTGTCAAGCGTTCCTCTGG

Primer\_r : CCTGTGCCATCTGCCTCTAAC

PCR conditions : 95 °C 30 sec 55 °C 30 sec 72 °C 60 sec 30 cycles

## RH mapping information

Chromosome No. : 12

## Experimental conditions

Panel name : GeneBridge 4

Primer\_f : TTTGTCAAGCGTTCCTCTGG

Primer\_r : CCTGTGCCATCTGCCTCTAAC

PCR product length : 209 bp

PCR conditions : 95 °C 15 sec 64 °C 60 sec 30 cycles

How to obtain KIAA clone(s)Back to the [HUGE Protein Database homepage](#).Send a message to [huge@kazusa.or.jp](mailto:huge@kazusa.or.jp).

## FIG. 13

KUROYANAGI ET AL.

**Expression of the *ULK1* Gene in Adult Human Tissues**

Tissue expression of *ULK1* was investigated by Northern blot analysis (Fig. 5). A single 6.0-kb transcript was detected in varying degrees in all adult tissues examined. The highest gene expression of *ULK1* was detected in the skeletal muscle and heart, moderate expression was detected in the pancreas, brain, placenta, and liver, and the lowest expression was found in the kidney and lung. While *UNC-51* in *C. elegans* is specifically expressed in neurons, its human homologue *ULK1* showed ubiquitous expression, suggesting that this class of serine/threonine kinase may have acquired transcriptional specificities in molecular evolution. Gene amplifications may contribute to functional diversity in the evolution of the vertebrate genome.

**Chromosome Mapping of the Human *ULK1* Gene**

Chromosomal localization of the human *ULK1* gene was determined by FISH and radiation hybrid mapping. FISH analysis using a human *ULK1* cDNA probe revealed that the human *ULK1* gene was assigned to 12q24.3 (Fig. 6). Analysis of a somatic cell hybrid panel showed that the samples containing human chromosome 12 specifically produced a DNA fragment of the expected size, indicating that the human *ULK1* gene is located on chromosome 12. For further determination, radiation hybrid-based mapping was performed. The data vector obtained for human *ULK1* was 0000100001 0000100001 110000110 0011010000 0000110000 1000000000 000001010 1011100001 0100011011 010. Statistical analysis of the data vector indicated that the *ULK1* locus was mapped to a location distal to D12S367, a genetic marker on chromosome 12q24.3. With reference to *Human Gene Mapping* (Kucherlapati et al., 1995) and OMIM, two hereditary diseases have been assigned to this region: the causal gene for distal hereditary motor neuropathy type II (distal HMN II) has recently been assigned to the 13-cM interval between D12S86 and D12S340 on chromosome 12q24 (Timmerman et al., 1996), and the causal gene for scapulohumeral spinal muscular atrophy (SPSMA) has also been assigned to the 19-cM interval between D12S338 and D12S366 on chromosome 12q24.1-24.31 (Isozumi et al., 1996). As the marker D12S367 is located distal to these intervals, it is less likely that *ULK1* is related to these hereditary diseases. Further genetic or gene-targeting strategies would be needed to define the disease that is defective for *ULK1* kinase in human or mouse in the future.

**ACKNOWLEDGMENTS**

We thank Drs. Ken-ichi Koizumi, Tsuyoshi Sakai, and Toshifumi Tomoda for their technical assistance. H.K. is supported by Research Fellowships from the Japan Society for the Promotion of Science for Young Students.

**REFERENCES**

- Aigner, L., and Carenzi, P. (1995). Absence of persistent spreading, branching and adhesion in GAP-43-depleted growth cones. *J. Cell Biol.* 128: 647-660.
- Appleby, J. L., Parkinson, J. S., and Bourret, R. B. (1996). Signal transduction via the multi-step phosphorelay: Not necessarily a road less traveled. *Cell* 86: 845-848.
- Bonner, T. I., Opperman, H., Seeburg, P., Kerby, S. B., Gunnell, M. A., Young, A. C., and Rapp, U. R. (1986). The complete coding sequence of the human *raf* oncogene and the corresponding structure of the *c-raf-1* gene. *Nucleic Acids Res.* 14: 1009-1015.
- Charest, D. L., Mordret, G., Harder, K. W., Jurik, F., and Pelech, S. L. (1993). Molecular cloning, expression, and characterization of the human mitogen-activated protein kinase p44<sup>MAPK</sup>. *Mol. Cell Biol.* 13: 4879-4890.
- Dessi, C., Garriga, G., McIntire, S. L., and Horvitz, H. R. (1985). A genetic pathway for the development of the *Caenorhabditis elegans* HSN motor neurons. *Nature* 338: 638-646.
- Ellledge, S. J. (1996). Cell cycle checkpoints: Preventing an identity crisis. *Science* 274: 1664-1672.
- Faux, M. C., and Scott, J. D. (1996). Molecular glue: Kinase anchoring and scaffold proteins. *Cell* 85: 9-12.
- Fitzkeweller, G., Marme, D., and Hug, H. (1990). Sequence of human protein kinase C  $\alpha$ . *Nucleic Acids Res.* 18: 2183.
- Hanks, S. K., and Hunter, T. (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: Kinase (catalytic) domain structure and classification. *FASEB J.* 9: 576-596.
- Harris, W. A., and Holt, C. E. (1995). From tags to RAGS: Chemofinity finally has receptors and ligands. *Neuron* 15: 241-244.
- Hedgecock, E. M., Culotti, J. G., Thomson, J. N., and Perkins, L. A. (1995). Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescent dyes. *Dev. Biol.* 171: 158-170.
- Higgins, D. G., and Sharp, P. M. (1988). CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* 73: 237-244.
- Isozumi, K., DeLong, R., Kaplan, J., Deng, H. X., Iqbal, Z., Hung, W. Y., Wilhelmsson, K. C., Hentati, A., Feriac, V. M., and Siddique, T. (1996). Linkage of scapulohumeral spinal muscular atrophy to chromosome 12q24.1-q24.31. *Hum. Mol. Genet.* 5: 1317-1332.
- Jian, X., Hidaka, H., and Schmidt, J. T. (1994). Kinase requirement for retinal growth cone motility. *J. Neurobiol.* 25: 1310-1328.
- Johnson, L. N., Noble, M. E., and Owen, D. J. (1996). Active and inactive protein kinases: Structural basis for regulation. *Cell* 85: 149-153.
- Kevnes, R., and Cook, G. M. (1996). Axons turn as netrins find their receptor. *Neuron* 17: 1031-1034.
- Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44: 283-292.
- Kozak, M. (1987). At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J. Mol. Biol.* 196: 947-950.
- Kucherlapati, R. S., Craig, J. W., and Marynen, P. (1995). Report of the Committee on the Genetic Constitution of Chromosome 12. In "Human Gene Mapping: A Compendium" (A. J. Cuticchia, M. A. Chipperfield, and P. A. Foster, Eds.), pp. 762-800, Johns Hopkins Univ. Press, Baltimore, MD.
- Lee, M. G., and Nurse, P. (1987). Complementation used to clone a human homologue of the fission yeast cell cycle control gene *cdc2*. *Nature* 327: 31-35.
- Letwin, K., Mizzen, L., Motro, B., Ben, D. Y., Bernstein, A., and Pawson, T. (1992). A mammalian dual specificity protein kinase, Nek1, is related to the NIMA cell cycle regulator and highly expressed in meiotic germ cells. *EMBO J.* 11: 3521-3531.

SUBSTITUTE SHEET (RULE 26)





WO 02/070559

25/38

PCT/GB02/00948

FIG. 15-1

```

11bd ( 225 )      sanedMpverIleAEIavepk-----te
21bd ( 182 )      lspqlelikvSkAHqeIFp-slcqlg-----kyttn---
1prga ( 207 )      esadlralAkHlydsYikEFltkakAraILtgktdkspfyIydmns
3erta ( 306 )      lalgltadqMvsAlgAepp-----ilySegdp-
1a28a ( 682 )      qlipplInlLMgiepd-----viyAghdn-
                    aaaaaaaaaa

11bd ( 248 )      tyvaanmglhpsspn-----dpvtnicgAadkqiftLveWakrIP
21bd ( 214 )      -----ssadhrvq---LdlgLwdkfseLatkciikIveFakrLp
1prga ( 255 )      lmmGedkikfkhitplqeqskevAikifgggfrveavqeieYakgIp
3erta ( 334 )      -----trpfs-----easmglltnLadrelvhMInWakrVp
1a28a ( 706 )      -----tkpdt-----ssgLitslngLgergLsvvkWSksLp
                    aaaaaaaaaaaaaaaaaa

11bd ( 288 )      hFseLpdgvilLragwneLlIAsFShrSia-vk---dgllLatG-lhV
21bd ( 250 )      gFtgLsiadqitLlkaACldilMLRiCLR---ytpeqdtMTfsdG-LtL
1prga ( 305 )      gFvnldndgvtLlkyGVHEIIYTmlAsl---Mnk--dGvLiseGgFm
3erta ( 366 )      gFvdLtihdqVhLLegAwlEIlMigiVvrSme--hp-gkLFA--pnLiL
1a28a ( 738 )      gFrnLhidqitLiqvgMmsLmVFgLgwrSykhvsg--qnLyfa--pdLiL
                    333 aaaaaaaaaaaaaaaaaa          bbb          bb

11bd ( 333 )      hrnsAhsa--gvgaIFdrVieLvsKugdMgMdkTELCLRAivLFnpds
21bd ( 295 )      nrtQMNA--GfgplTdiVfaFggll-pLeMidtETGLSAICLicGdr
1prga ( 349 )      treFLkgirkpFgdfmepKFeFavkFn-aleLdDsDLAIFTAVILGdR
3erta ( 411 )      drnqGkcV-egMveiFdmletSrFr--muvLggeEFVCLRSiILLnSGV
1a28a ( 785 )      neqrMk-e-ssFysiCLlmqIPqeFv-klqVsgeEFLCMVLLLNtIp
                    baaaa aaaaaaaaaa          aaaaaaaaaa

11bd ( 381 )      k-----gLsnpaeVeaLrekVyagLeaygkhkypeq---pgRfakLl
21bd ( 342 )      m-----dLeepekVdklqepLeaLrlyArrrrp---sqpyMfrMl
1prga ( 398 )      p-----gLnvkpedigdnLgALlQLlnhp---essgLfakLl
3erta ( 459 )      ytFlstlgleekdhirVldkItdTLihlMakgltlqqhqrlaqLl
1a28a ( 832 )      l-----egLrsqtgFeemrssyirELikAIglrqgvvssqrfyqLT
                    aaaaaaaaaaaaaaaaaa          aaaaaa

11bd ( 420 )      lrlpaLrsIgkClehIFff-----kLiGdtpidtflmemlea
21bd ( 381 )      mkitdLrgiStkgaegaiLLmeIp---gpMpliremlenp
1prga ( 437 )      gkmtdLrqIvtehvqllqwkketd--mslhplqeiykdl
3erta ( 509 )      lilshlhSnkgmehlygmkc----knvv---plydllemLd-----
1a28a ( 875 )      klldnLhdlVkqLHlycintfiqsralsVeFpemngevaaqLpkilgm
                    aaaaaaaaaaaaaaaaaa          aaaaaa

11bd ( 458 )      ph---qmt
21bd ( )
1prga ( )
3erta ( 546 )      --ahrlha
1a28a ( 925 )      Vkpllfhk

```

## FIG. 15-2

## Key to Homstrad

aaa	alpha helix
bbb	beta strand
333	$3_{10}$ helix
lower case	solvent accessible
UPPER CASE	solvent inaccessible
<b>Bold</b>	hydrogen bond to main-chain amide
<u>Underline</u>	hydrogen bond to main-chain carbonyl
<i>Italic</i>	positive phi torsion angle

FIG. 16

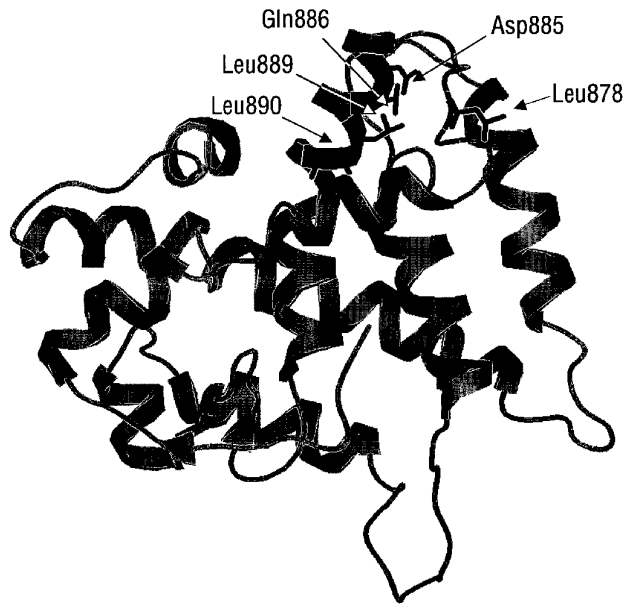


FIG. 17

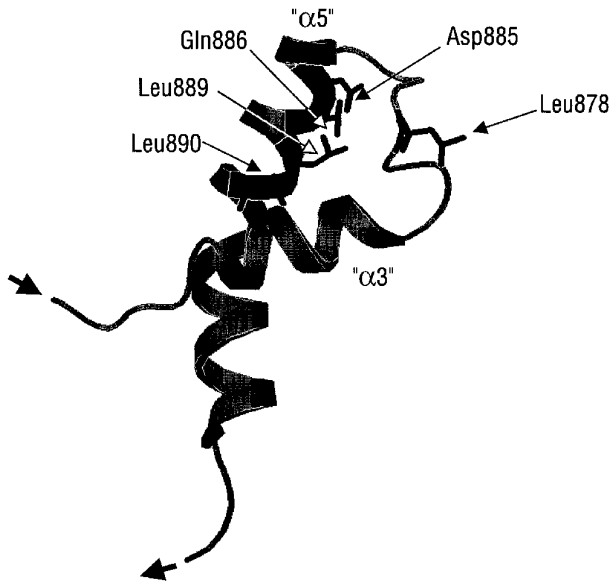


FIG. 18

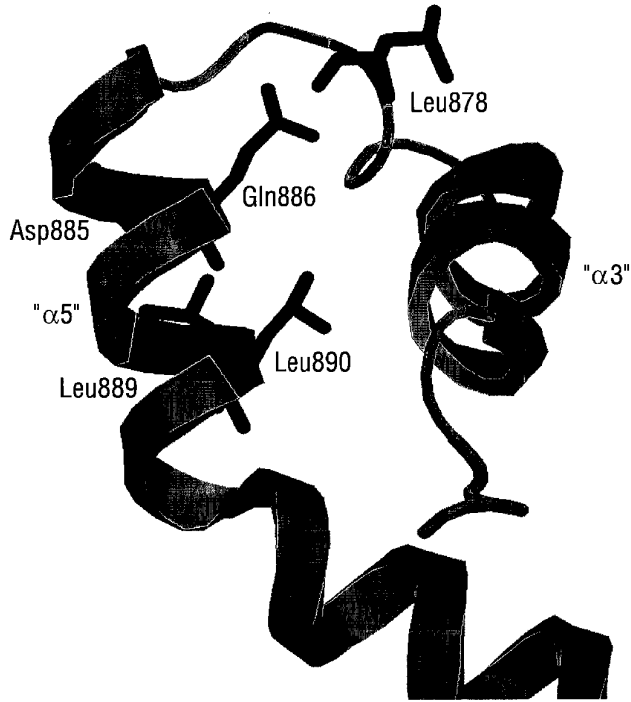
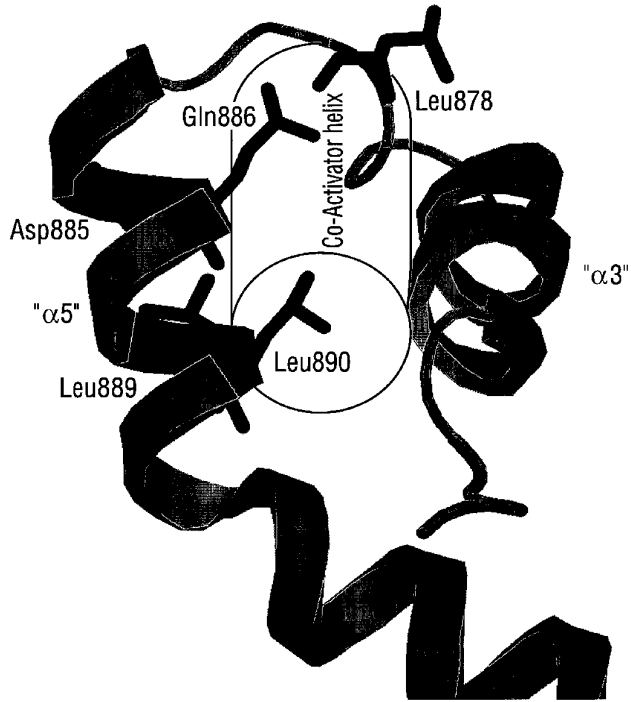


FIG. 19



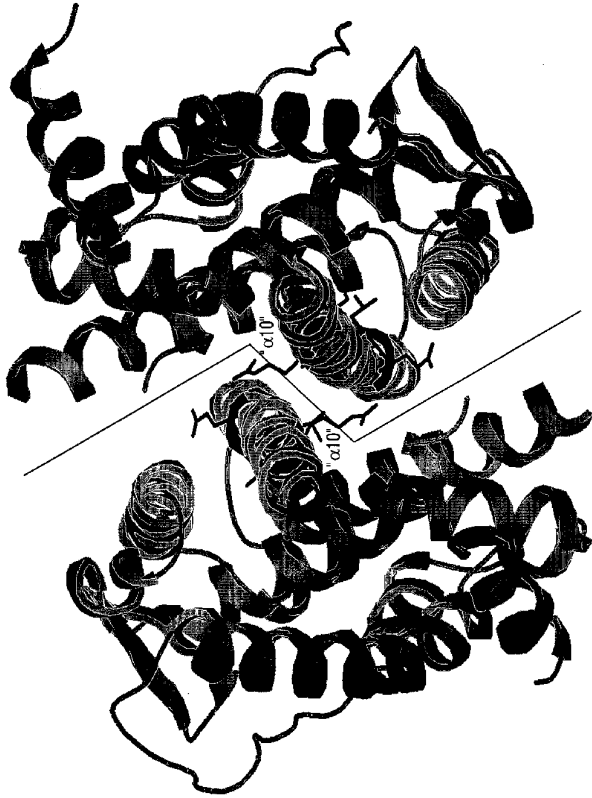


FIG. 20

SUBSTITUTE SHEET (RULE 26)

FIG. 21A

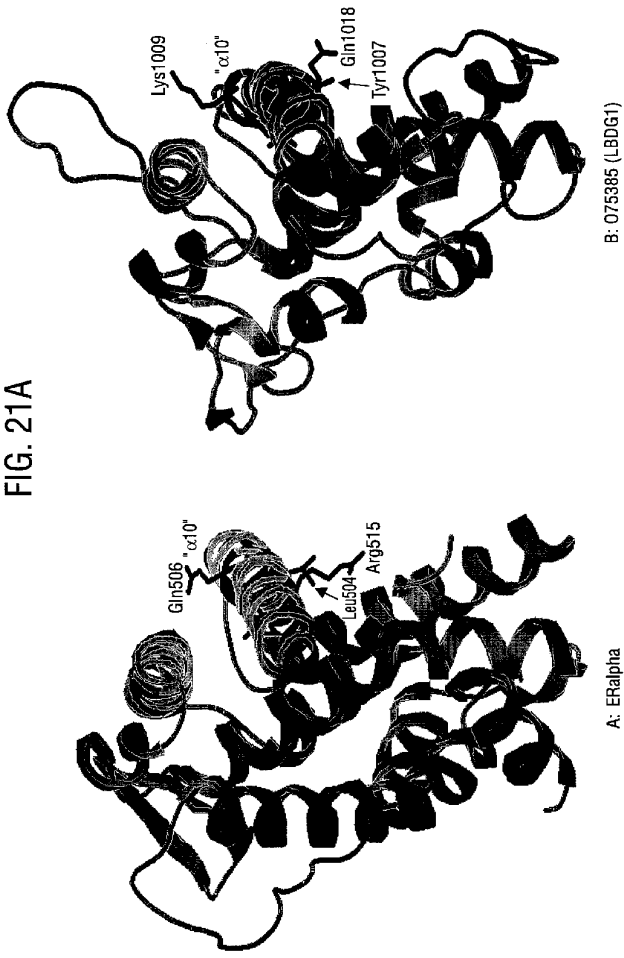
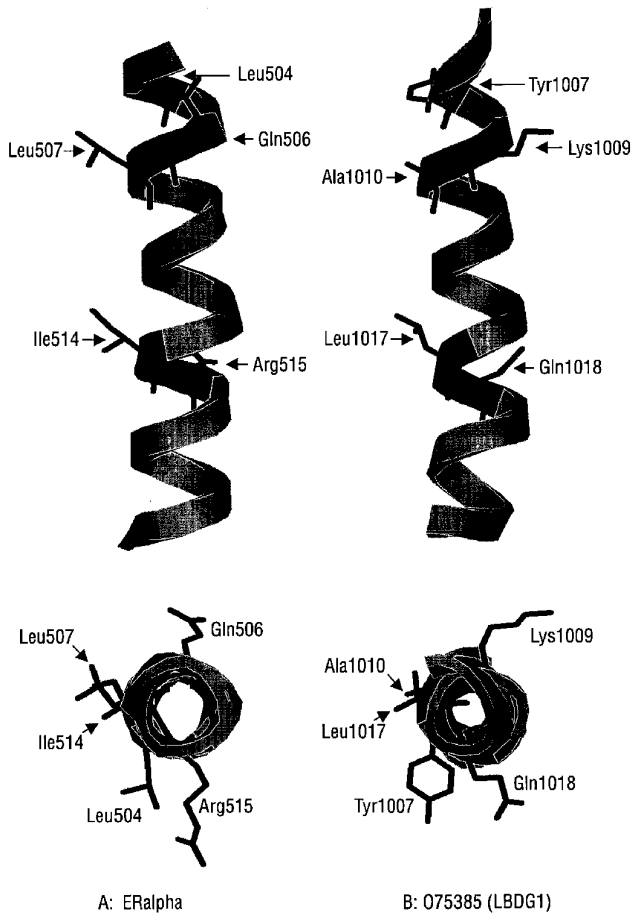


FIG. 21B



WO 02/070559

34/38

PCT/GB02/00948

FIG. 22-1

075385  
BAA31598.1  
1ERR  
MEFGRGGTETVCGKFEFSRDLIGHGAFVVFVGRHREKHDELVAVKVCINRKNLAKSQTL  
-----MEVVGDFEYSKDLVGHGAFVVFGRHROKTDWEVAIKSINKNLKSKQILL  
-----

075385  
BAA31598.1  
1ERR  
GREIKILKELKHENIVALYDFQEMANSVYLVMVYCNGDLDADYLQAKGTLSEDTIRLFLQ  
-----

075385  
BAA31598.1  
1ERR  
QIAGAMRLLHSKGIHRDLKPNILLSNPAGRANPNSIRVKIADFGFARYLQSNMMAAT  
QIAAAMRILHSKGIHRDLKPNILLSNANRRKSSVSGIRIKIADFGFARYLHENMMAAT  
-----

075385  
BAA31598.1  
1ERR  
LCGSPMYMAPEVIMSOHYDGKADLWSIGTIVVYOCILTGKAPQASSPQDLRFYEKNTLV  
LCGSPMYMAPEVIMSOHYDAKADLWSIGTIVYOCILVKGPPFQANSQDRLRMFYEKNSLIM  
-----

075385  
BAA31598.1  
1ERR  
PTIPRETSAPLRQLLALLQNRHKDRMDFEFPHHFFLDASPSVRKSPVVPVPSYSSGS  
PSTPRETSPYLANLLGLLQNRQDRMDFEAFESHPELEQGPVKKSCPVPVPMY-----

075385  
BAA31598.1  
1ERR  
GSSSSSSTSHLASPPSLGEMOOLKTLASPADTAGFLHSSRDSGSKDSSCDTDFVYV  
-----SGSVSGSSCGSSPSCR--FASPPSLPDMQHIQENLSSPLGPPNYLQVSK  
-----

075385  
BAA31598.1  
1ERR  
PAQFPGLVAEAFSAKPPPDLMCGSSSLVASAGLESHGRTPPSPPCSSSPSSGRAGP  
DSASTSSKNSSCDTDFVLPVPHNISDHSKDMGTAGRR-----ASNEFLVCGGQCP  
-----

075385  
BAA31598.1  
1ERR  
FSSSRCGASVPIPVFTQVQNYQRLEARNLQSPYTF-----QTPRSSAIRRSGSTPLGFARA  
PTVSPHSETAPIPVFTQIRNYQRLEARNLQSPYTF-----QTPRSSAIRRSGSTPLGFARA  
-----

WO 02/070559

35/38

PCT/GB02/00948

FIG. 22-2

075385 SPSPPAHAEHCGVLAHKMSLGGRRPYTPSPQVGTIPEPFGMSGTRSPQGAEMRGGRRSRP  
 BAA31598.1 GSCSPVADTAGTVGRRRLSTGSSRPYSPLVGTIPEQFSQCCCHPQGHDRSRNSSGS  
 1ERR -----  
 075385 GSSAPEHSPTSGLGCRLHSAFNI.SDLHVVRPKLPKPPDPLGAVF-----SPQASPPQ  
 BAA31598.1 PVP-QAQSPQSLLSGARLQSAFTLDIYQNKQKLRKQHSDPVCPSHTGAGYSYSPQPSRP  
 1ERR -----  
 075385 PSHGLQSCRNLRGSPKLPDFLQRNPLPPLIGSPTKAVPSDFPKTPSSQNLLALLAROGV  
 BAA31598.1 GSLGTSFTKHLGSSPRSSDWFKPLPLTIIGSPTKTTAPFKIPKTOASSNLLALVTRHG-  
 1ERR -----  
 075385 VMTPPRNRTLPLDLSEYGFPHGQPLGPGLRPGEDPKGPFGRSFTSRLTDLILLKAAFGTQA  
 BAA31598.1 -PAEQSKDGNPRECAHCLLVQSERQRAEQSKAVFGRSVSTGKLSQQGKTIPICRHQ  
 1ERR -----  
 075385 PDPGSTESLQEKPMETAFSAGFGSLHFGARAGTSSPVPVTVGSPESGSTPPQGP--  
 BAA31598.1 GSTDS--LINTERPMDIAPAGACGGVLA--PAGTAASSKAVLFTVGSPPHSAAPTCTHM  
 1ERR -----  
 075385 --RTRMFSAGFTGSASS--SAHHLVFGPC---SEAPAPLPAHGCSFADPIAANL  
 BAA31598.1 FLRTRTTSVGFNSGGSLCAMSGRVCGSPGFGSGSPGAEAAPSLRYVYVYGASPPSL  
 1ERR -----  
 075385 EGAVTFAEAPDLPEETIMQEHEHTLLRGLRFTLLF---VOHVLEIAA--LKGSASE--A  
 BAA31598.1 EGLITFEAPELPEETIMEREHTTLRHLNVMMLMF-----TECVLDLTA--MRGNPEICT  
 1ERR --ALSLTADQW-VSALLDAEPPILYSEYDPTTRPFSEASMMGLLNADRELHVHMINWAKR

SUBSTITUTE SHEET (RULE 26)



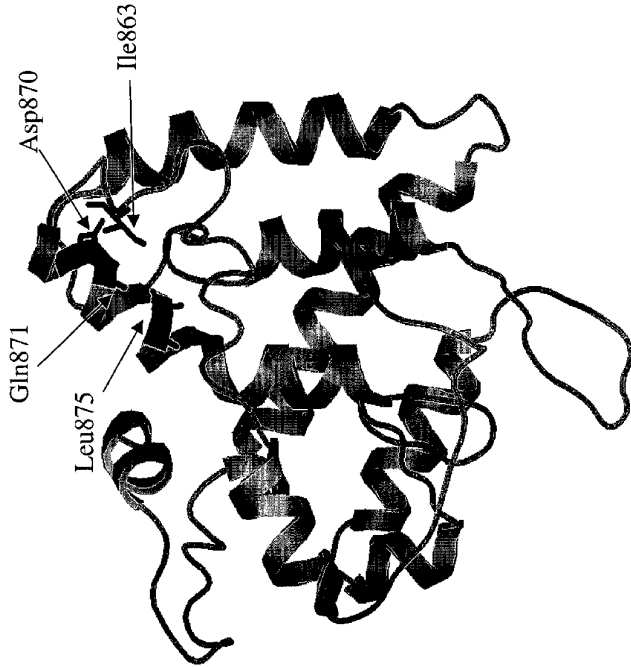
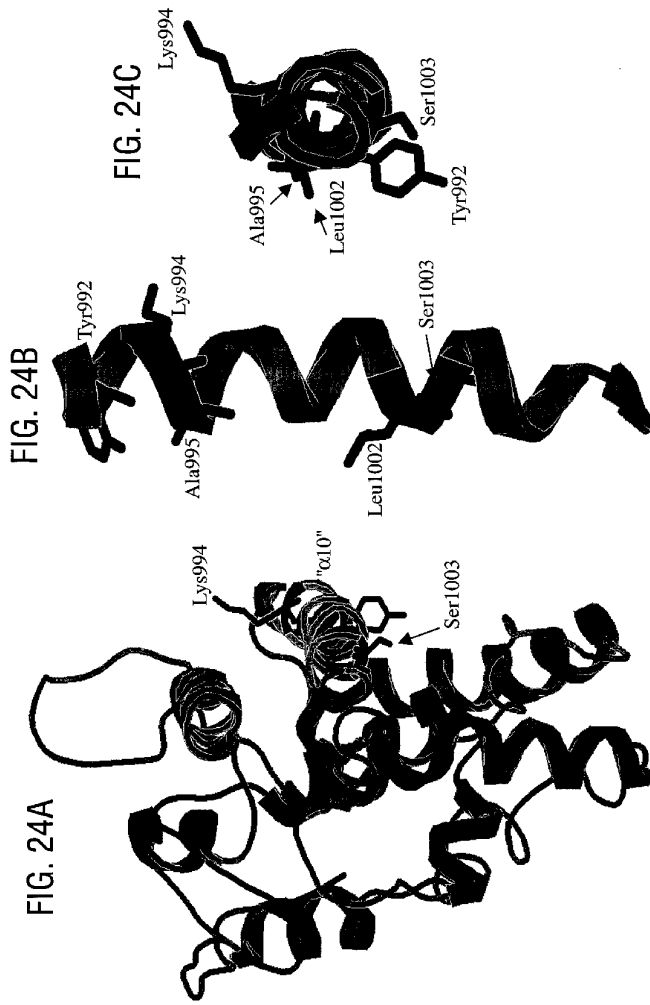


FIG. 23

WO 02/070559

38/38

PCT/GB02/00948



SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

1

**Sequence Listing**

SEQ ID NO: 1 Nucleotide coding sequence for O75385 (LBDG1) protein

```

1  ggatcgggat  tccgattagc  agcccgggaa  gagtgcctg  gcacagggcc  cggagggagc
5  61  gagacctcgc  gacctccctc  ggcccgcggg  gctgggaacc  gggcccggcc  tgcccgatgg
121  ggccgcgggc  cccggagatg  cgcctcggcc  cggcccgcgc  ccccggccc  cgcgcccgcg
181  gcccgccgc  cccggccgc  gctcgcctc  gagtccccg  cgccttggcc  cgcaccccc
241  cggcccgcgc  cccggccgc  cctgcgcct  gtagcccgcc  cggggcgcca  cagagaccgt
301  gggcaagttc  gatttctccc  gcaaggacct  gatcgccac  ggcgcttcg  cggtygttt
10  361  caaggccgc  caccgcgaga  agcacgattt  gtaggtcgcc  gtaagtgra  ttaacaagaa
421  gaacctgcc  aagtctcaga  cgtgctggg  gaagaaatc  aaaatcctga  aggaactgaa
481  acatgaaaac  atcgtggccc  tgtacgact  ccagaaatg  gtaattctg  tctacctggt
541  tatggagtac  tgcaacgggt  gggacctggc  cgaactctg  cagccatgc  gcacgtgag
601  ctagggacac  atcaggctct  tctgcagca  gatcgcggg  gccatcggc  ttctgcacg
15  661  caaaggcacc  atcccgcgc  acctgaaacc  gcagaacatc  ctgctgtcca  accccgcggg
721  ccgcccgcgc  aacccaaca  gcctccgct  caagatcgct  gacttcggct  tcgcccggta
781  cctccagagc  aacatgatgg  cggccacct  ctgggctcc  cccatgtaca  tggccccga
841  ggtcatcatg  tcccagcact  acgacgggaa  gycggacctg  tggagcatcg  gcaecatct
901  ctaccagtgc  ctgacgggga  aggcgcctt  ccaggccagc  agccccagc  acctgcgct
20  961  gttctacgag  aagaacaaga  cgttggctcc  caccatcccc  cgggagacct  cggcccgcct
1021  gggcagctg  ctctggccc  tactgcaacg  caaccacaag  gaccgatgg  acttcgatga
1081  gtttttcat  cacccttcc  tcgatgccag  cccctggtc  aggaaatccc  caccctgccc
1141  tgtgcccctg  tacccaagct  cggggtccg  cagcagctcc  agcagcagct  ccacctccca
1201  cctggcctcc  ccgcccgcc  tgggagagat  gcagcagctg  cagaagaccc  tggcctcccc
25  1261  ggcctgacac  gctggcttcc  tgcacagctc  ccgggactct  ggtggcagca  aggactcttc
1321  ctgtgacaca  gacgacttgc  tcattgctcc  cgcgcagttt  ccagtgacc  tggtygtga
1381  ggcgcccagt  gccaaaaccc  cgcagacag  cctgatgtgc  agtgggagct  cactggtggc

```

SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

2

1441 ctctgaggggc ttggagagcc agggccggac cccatctcca tccccaccct gcagcagctc  
 1501 ccccagttccc tcaggccggg ctggcccggt ctccagcagc aggtgcggcg cctctgtccc  
 1561 catcccagtc cccacgcagg tgcaagaacta ccagcgcatt gagcgaaac tcagcgcacc  
 1621 cccccagttc caaacacctc ggtcctctgc cctccgcagg tcaggcagca ccagccccct  
 5 1681 gggctttgca agggccagcc cctccgcccc tgccccagct gagcatggag gcgtcctggc  
 1741 caggaagatg tctctgggtg gaggccggcc ctacaogcca tctcctcaag ttggaacct  
 1801 cctgagcggg ccaggctgga gcgggacgcc ctccccacag ggagctgaga tcgggggtgg  
 1861 caggteccct cgtccaggct cctctgcacc cagcaactct ccccgcactt ccgggtggg  
 1921 ctgcgcctg cacagcggcc ccaacctgtc tgacttgacc gtctccgcc ccaagctgcc  
 10 1981 caaacccccc acggaccccc tgggagctgt gttcagccca ccacaggcca gccctcccca  
 2041 gccgtccacc ggcctgcagt cctgcgggaa cctgcccggg tcaccaagc tgcccagctt  
 2101 cctgcagcga aacccccctc cccccatcct gggctccccc accaaggctg tgcccctctt  
 2161 tganttcccg aagaccceca gctcccagaa cctgctggcc ctctagccc ggcaggggct  
 2221 ggtgatgacg ccccccgaa accggacgct gcccgacctc tcggaggtgg gacccttcca  
 15 2281 tggtcagcgg ttgggcccctg gctgcggcc aggcgagac cccaaggcc cctttggccg  
 2341 gctttcagc accagccgcc tcactgacct gctccttaag gcggcgtttg ggacacaagc  
 2401 cccggaccgg ggcagcaggg agagcctgca ggagaagccc atggagatcg caccctcagc  
 2461 tggctttgga gggagcctgc acccaggagc ccgtgctggg ggcaccagca gccctcccc  
 2521 ggtggtcttc accgtgggct ctccccgag cgggagcag cccccccag gcccccgcac  
 20 2581 caggatgttc tcagcgggccc ccactggctc tgccagctct tctgccgccc acctggtgcc  
 2641 tgggcccctg agcagggccc cagcccctga gctccctgct ccaggacagc gctgcagctt  
 2701 tgccgacccc attgctgoga acctggaggg ggctgtgacc ttcgaggccc ccgacctccc  
 2761 tyaggagacc ctcatggagc aagagcacac gsagatcctg cgtggcctgc gcttcaagct  
 2821 gctgttctg cagcacgtcc tggagatcgc agcccgaag ggcagcgcga gtgaggcggc  
 25 2881 gggggcccct gagtaccagc tgcaggagag tytggtggcc gaccagatca gcctgctgag  
 2941 ccgagaatgg ggcttcgagg aacagctggt gctgtacctg aaggtyggcc agctactgtc  
 3001 ctccggcctg caaagtcca tcgaccagat ccggccggc aagctctgcc tgtctccac

SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

3

3061 tgtgaagcag gtggtgagca ggctgaatga gctgtacaag gccagcgtgg tgctctgcca  
 3121 gggcctgagc ctgocgctgc agccttctt cctggacaag cagcggctcc tggaccgcat  
 3181 tcacagcacc actgocgaga ggctcatctt cagccacgct gtgcagatgg tgcagtccgc  
 3241 tgcocctggac gagatgttcc agcaacctga gggctgctc ccacgctacc acaaggccct  
 5 3301 gctgctcctg gaggggctgc agcaatgct ctcggaccag gccgacatcg agaacgtcac  
 3361 caagtgcacg ctgtgcattg agcggagact ctggcgctg ctgactggca tctgtgctg  
 3421 acctttctg cctggctggg ccccccgtcc tgcagagccc tgcagagtgg gctctgtgtg  
 3481 ctggctggac tctcgggac aagcccatgg cgtgtatgct tgggtgctgag cctgcccctg  
 3541 ggcgccacgg acagtccagc tgcggcctc cctgcagctc accggggcaga accagcacat  
 10 3601 ctggagccac acagcttggg ggggtgtctc catcttttac agtggggat cacagaattt  
 3661 ctgcccctcc agctgctgg ctcagcagge gtgggtgcca ccacctota gccccagggc  
 3721 agccccggag gacagccaag ggcctgagac cactgccgac tcaagccaa agcagctcc  
 3781 tgcctagggc aggtcagcag gcactgtgcc caggaagagc ctgcccctc ggcgtccccc  
 3841 agtctccagg agcctctccc tccagatac ccccccagct ttgtcaatca cccaagcact  
 15 3901 ttatgcatat agagacagaa cctggacctc accaggact gctgggcagc gattcctggc  
 3961 agtggcctgg tgtttgtaca taccatattg cagacatg ccagggcccc ccaagcccca  
 4021 gcaccggacc acgttctgct ccaggtctgg acctcagcgg gagaactggc tccgggggga  
 4081 gtggggccct gcgctagagg cagaggcagt tcttltgtca agcgttctc tggggaccgg  
 4141 cagcagagge accgtgttct ctcagcctg gatacgtct gtaatcttc acactttatt  
 20 4201 cctaaaaagt gtcttatttt tatgcagctc atttttctt taaaggagaa aactttagg  
 4261 tgttaagaa ttggttttgg gagggcaggg actgggccag gttagaggca gatggcacag  
 4321 gggcgtgtgg cggcgggtg agcctgcttt gcacacctgt gttgggtgct gtcccctgcc  
 4381 gccctccct gtggcagcag caggacaggt gtgtgccag caccctcct acctgggct  
 4441 ggaagcagat gaggggata ctcatgcaa agaaaaagt aacatgtgca aaagctccc  
 25 4501 gtccagcttt gacagtcagt tttgatgca gctcctcggc agggtaggct tgatgacagc  
 4561 cctgtccctc cctgctccg ccttgcccaa ggcacggag ggcgtctgca gagaggcctg  
 4621 cctccggat tccagcggg catgcoctgc aaaccccgc tggcctccc ttggtctgcc

SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

4

4681 cagccctcgg ttagccctgc ctgaatcagt agatacttga acgagtcgcc agtctgcggg  
 4741 aggcagtggt ggggccatgg acccatgocgg ggggttccag ggtcacacgc cacataacag  
 4801 acaaaaatac acacacgtgt gtttttcttt gcaatacttg aaatattgcc actgtgcttg  
 4861 gacttagaag aagaaaatoc ccgtgacttc ttccctcatca ccttgatggc ttattctca  
 5 4921 ccttgtagggg catgtttgaa ttattgctt catggccgac tggaatcctg agtctcggga  
 4981 agctggcact gcgggatct tgcccgggtg cctggctcct ttgcttcctg cgcggccgca  
 5041 tgtgctgtg tocaagcagg tcttgggcgc ctcaactgct gccctgggtt gaatgttctc  
 5101 ttgatagtgc tggacccttt gtctatttta aagcgaattt tglgtgattt cctgcccctt  
 5161 gcgttatatt gtataatacc aacgtaagga aataaacctt tggaattggt gaaaaaaaa  
 10 5221 aaaaaaaa

## SEQ ID NO: 2 Protein sequence for O75385 (LBDG1)

1 mepgrggtet vgkfefsrkd lighgafavv fkgvrhekhd levavkcink knlaksqtll  
 15 61 gkeikikel khenivalyd fgemansvyl vmeycnggd1 adylhamrtl sedtirflq  
 121 giagamrllh skgiihrdlk pgnillsnpa grranpsir vkiadfgfar ylqsnmmaat  
 181 lcgspmyap evimsqhydg kadlwsigtv vyccltqkap fqaesspdlr lfyekntlv  
 241 ptipretsap lrqlllallq rnhkdrmdfd effhnpflda spsvrksppv pvpsypssgs  
 301 gsssssssts hlaaspslge mgqlqktlas padtagflhs srdsggskds scdtddfvmv  
 20 361 pacffpgdlva eapsakpppd slmcsgeisl asagleshgr tpspsppcss spspgragp  
 421 fsssrcgasv pipvptqvm yqriernlgs ptqfqtprss airrsgstsp lgfaraespap  
 481 pahaehgvl arkmslgggr pytppqvgt iperpgwsgt pspggaemrg greprpsssa  
 541 pepsrtsagl gcrhsapnl sdhvvvrpk1 pkpdtplga vspggaapp qpsbglsqr  
 601 nlrgepklpd flqznlppi lgeptkavps fdfpktpsaq nilallargg vmtpprnr1  
 25 661 lpdlsevgpf hsqplgpglr pgedpkpfg rsfstsr1td lllkaafgtq apdpgsteel  
 721 qekmeiaps agfsgslhpg araggtssps pvvftvgsp sgstppgpr trmfagptg  
 781 ssssarhlv pgpcseapap elpapghcs fadpiaanle gavtfeapdl peetlmeqeh  
 841 teilrglrft llfvghviei aalkgsasea aggpeyqlqe svvadqisll srewgfaeql  
 901 vlylkvaell seglgsaidq iragklclss tvkqvvrln elykasvpsc gglslrlqrf  
 30 961 fldkqrlldr ihsitaerli fshavqmvgs aaldemfghr egcvpryhka lllleglqhm  
 1021 ledqadienv tkcklcierl lsalltgica

SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

5

SEQ ID NO:3 (Genbank nucleotide accession number AB014523; coding for LBDG4)

```

1  ggcgcgcgca gggcgttggg cgcgcgcgcg aggcggggaa gcgcggggcc gggcgggtgc
61  gggttctagg gcggcgcccg tcgccgtcgc agcagcgcgc cgagcgggga gggccgagga
5  121  gggccgacga gctggggatg gagagtaccg ggcctctcgc tcctcagag cgcgtgtgcg
181  gctctgggag cgcacagtga cggtgacggc accctcggcc cggcagcgcc gaggccgctt
241  cggcagacag ccagcggccg gcggcagggc gggccatgag cggcaggggc cggccggggc
301  ctgctgagcc ctggctccgc gcggcagctt cccagtttc cgtccggtc tctcggcatg
361  agagtcgcgc cgggcccggg gctgcggctg cccagaccgc gcgcacgct ggcgcgctcc
10  421  gggcccgccg agccgcggtg ctgataacct gcgccactg cgcgcgccgc ccgtccgctg
481  tgtgccccgg gggcgcggcc atggaagttg tgggtgactt cagtagacgc aagagggatc
541  tcgtgggaca cggggccttc gccgtggtct tccggggcgc gcaccgccag aaaactgatt
601  gggaggtagc tattaagaat attaataaaa agaacttgc aaaatcacia atactgcttg
661  gaaaggaat taaaatctta aaggaacttc agcatgaaaa tattgtagca ctctatgatg
15  721  ttcaggaat acccaactct gtcttttgg tgatggagt ttgcaatggt ggagacctcg
781  cagattattt gcaagcgaaa gggactctca gtgaagacac galcagagtg tttctgcatc
841  agattcctgc tgccatgcga atcctgcaca gcaaaaggaat catccacaga gatccaaac
901  cacagaacat ctgtctgtcc tatgccaat gcagaaaatc aagtgtcagt ggtattcgca
961  tcaaaatagc ggattttggt tttgctcgtt acctacatag taacatgatg gctgcaacac
20  1021  tgtgtggatc cccgatgtac atggctcctg aggttattat gtctcaacat tatgatgata
1081  aggctgactt gtggagcata ggaacagtga tataccaatg cctagttaga aaaaaccctt
1141  ttcaggccaa tagtctcaa gacttaagga tgtttatga aaaaaacagg agcttaatgc
1201  ctgtatttcc cagagaacca tcaacttatt tggtaatct ccttttgggt ttgcttcaga
1261  gaaacaaaa agatagaatg gactttgaag cgttttttag ccatcctttt cttgagcaag
25  1321  gtccagtaaa aanaatcttc ccagttccag tgcccatgta ttctggttct gtctctgaaa
1381  gctctctgag cagctctcca tcttctgctt ttgcttccc acctccctt ccagatagtc
1441  agcatattca ggaagaaaaa ttatcttccc caccattggg tcctcccaac tatctacaag
1501  tttccaaaag ttctgccagt actagtagca agaactcttc ttgtgacag gatgactttg
1561  ttttgggtgc acacaacatc tcgtcagacc actcatgtga tatgccaatg gggactgctg
30  1621  gcagacgtgc ttcaaatgaa ttcttgggtg gtggagggca gtgtcagcct actgtgtcac
1681  ctccagcga aacagcacca attccagttc ctactcaaat aaggaattat cagcgcatag
1741  agcagaatct tacatctact gccagctcag gcacaaatgt acatggttct ccaagatctg
1801  cagtggtacg aaggtccaac accagcccca tgggttctct cggccggga tcatgtctcc
1861  cagtccagc agacacagca cagacagttg gacgaaggct ctccactggg tottctagtc
35  1921  ctactcacc ttccccttgg gttggtacca ttcctgagca attcagtag tgtctgtg
1981  ggcactctca gggccatgac tccaggagta gaaactctct aggttctcca gtgccacaag
2041  ctcatctccc acagttctct ttatcgggtg ctgactgca ggcgcccc accctcactg
2101  acatctatca gaacaagcag aagctcagaa aacagcactc tgacccctg tgccatccc

```

SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

6

2161 atactggggc tgggtacagc tactgcgctc agcccagtcg gcctggcagc cttggaactt  
2221 cccccaccaa gcacttgggg tccctccacc ggagttctga ctggttcttt aaaactcctt  
2281 tgccaacaaat cattggctct cctactaaga ccacagctcc ttccaaaac cctaaaactc  
2341 aagcatcttc caacctgta gccttgggta ctctgcatgg gctctgtgaa gaacagtcga  
5 2401 aagatgggaa tgagccacgg gaatgtgccc attgctcttt agtgcaggga agtgagaggg  
2461 agcgggcccga gcagcagagc aaggcagctt ttggcagatc tgcagctacc gggaggttat  
2521 cagatcaaca aggaagact cctatatgtc gacatcaggg cagcacagac agtttaaata  
2581 cagaacgacc aatggatata gctccggcag gacccctgtg ttggtttctg gcacctcctg  
2641 caggtacagc agcaagttcc aaggctgtcc tcttcactgt aggtctctcc ccacacagtg  
10 2701 cggcagcccc cacttgtacc caatgttcc ttgaaacaag acaaacctca gtggggccca  
2761 gcaactccgg gggctctctt tgtgccatga gtggccgctg gtgctggggg tccccgctg  
2821 gccagggtt cggctcttcc cctccaggag cagaggcagc tcccagcctg agatacgtc  
2881 ctacaggtgc ttcaccccc agcctagagg gctcactcac ctttgaagcc cctgaactgc  
2941 cggaggagac gctgatggag cgggaacaca cagcaactt acgcatctg aatgtgatgc  
15 3001 tgatgttccac tgagtgtgtg ctggacctga cagccatgag gggaggaaac cctgagctgt  
3061 gcacatctgc tgtgtccttg taccagatcc aggagatgt ggtggtggac cagatcagtc  
3121 agctgagcaa agactggggg cgggtggagc agctggtgt gtacatgaaa gcagcacagc  
3181 tgcttggggc ttctctgcat ctggccaaa cccagatcaa gtccgggaaa ctgagcccat  
3241 ccacagctgt gaacaagtt gtcaagaatc tgaacgaac atataaattc tgcacacca  
20 3301 tgcgcaagaa acttacagaa aagctgaatc gattcttctc tgacaaacag aggtttattg  
3361 atgaatcaa cagtgtagct gcagagaac tcatctataa ttgtgctgta gaaatggctc  
3421 agtctgcagc cctgagtag atgtttcagc agaccgaaga tattgtttat cgcctacata  
3481 aggcagccct tcttttggaa ggcctaagta ggalctaca ggaacctga gatattgaaa  
3541 atgtgcataa atataaatgt agtattgaga gaagactgfc ggcgctctgc catagcaccg  
25 3601 caaccgtgtg agcagcagc tcacccctg gaccgggtgt gggaaactga ggtgatgctt  
3661 ttgggattac agctgagtt ctgtcaccoc atccccagga aactgtagct tcttaactgg  
3721 tgactaccaa agaacaagca gtgattgaa aaaggaaaa caatccaaa actacatatt  
3781 ttaggaaat ctgccttatt ggagaaaac accctttccc ttttctttg tagaagcagg  
3841 agcaagagtg ttggctccc agttggact tggtaataa atglacctta gaactaggat  
30 3901 aatcgggtaca gttattctta aagataata aaaaatgaaac aaagtgagt cctgctactg  
3961 ggtctcatag agcagtggt gaaattccat gtgtttgctg aggtgtaaag gtaaatgtat  
4021 tcacccctca tccaggcagt ttgatatttg gactaagttt gtttaaatc gagcatgcat  
4081 ctttaaacag ctccaggaaga aatagcttaa gaagaagtga aacatggatc ttggaaqaaa  
4141 ttttgaatc ttcaatttga tccataatg gatacatgtt aatcttcaa aatcttcat  
35 4201 attgcaactaa ttattaaaa caactgtgta ttggattttg taatttaact aaggcaaat  
4261 ggaactgtttt aaaatatttt acttgattgt atacatagac cctttccaga attcactgt  
4321 aatctccagt gaacttttaa gtggttaaaa ctgtattca tgtgaacctt tgcacatttt  
4381 tttttttta cttctttac tacacctaca gattttctca gtaigtgtt tgttagcttt

SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

7

4441 tggttccatt ttttatttg catgcagaat gtacattgat goctgtgacc ttaggtttat  
 4501 taaaggctag gtttatttgg gcagtattag aaacaaaatc atggatcaag agatactett  
 4561 gataatttga atagggccaa aacaaagttg gtgacctaaa ggcttgtag tgatgtggag  
 4621 ttctcatag cagtgagtg aaaatgaagt tctgtttctc ttaggaaaat gggcagctgt  
 5 4681 ctctgccta atgtgtattt ttcattgtaa ttctgacagt tcaccaata gctagtcag  
 4741 gagaatgcag gcagtaact taatatccct ccaggaatgg ttctacagt gtgtattatt  
 4801 tggtttcttt tacttacctg cttgaatact tgaataaacc attccaat ttaactcctt  
 4861 ttattttaat ccttttacct aaaataact gaactcttg acaaatgca cagagctctt  
 4921 tggcattaat ctaattttaa tgtactgata aaaacaaca tggttgtcct ttactttgac  
 10 4981 aaagtaagt aattttacc ttatttatct gtatgaatt ccagtagta atttgaacat  
 5041 ttatttatat gacgtttgta tttttaggtc ttaataacag tgtttctacc tctcatttgt  
 5101 aactgcagc attattcttg aaactagga aaactcactg aattgttggt taatagcctt  
 5161 ttattatttg cctgtacaaa tgtatattaa ggtaaaata aactgacaaa gtgtttctag  
 5221 ggtgtagctg ggtacatatt aagtggcttg ttgagccagg tacttctcta gtgagtttag  
 15 5281 agactggccc atgaatatcc ttgtctcgc cccaggattt agactctggc tactgtcag  
 5341 caggcttcca ggaecataga ctgttttacc tccacaacc tattgttat tagtgatact  
 5401 ttattttata taatattttt ttttcacagt gaaatttcat tcatgttctt toagttatca  
 5461 cctgtgttat ctcagttgta ggtttattct atctctctct ctctctctcc cattctcttt  
 5521 ttaacacagg atgaacacag ttcagagag ggaagtgatt ggcctaagt caggaactag  
 20 5581 gcaagtggc aagccatgct ttgtgactt caagttaatt cttctgttcc ttgtatatta  
 5641 aaggtcttgg ggtagatggt gtgtgtgaaa cagtgaagtc tcaacagcag aaaagaacaa  
 5701 aatgtaaaat catgaataat gttctggtt atacttccat tatcaaggct aattaagaga  
 5761 ttttgccttg agtatagcaa taataacaa atgcittatg tttccttg

25 SEQ ID NO:4 (Genbank protein accession BAA31598.1; LBDG4)

1 mevvgdfey krdlvghgaf avvfrgrhrq ktoweaiaks inkknlsksq illgkeikil  
 61 kelqheniva lydvqelpns vflvmeycng gdladylqak gtlstedtivr flhqiaaamr  
 121 ilhskgihr dlkqgnills yanrrkssvs girikiadfg farylhnmm aatlcgspmy  
 30 181 mapevimsqh ydakadlws i gtvlyqclvg kppfqanspq dlrmfyeknr slmipsipret  
 241 spylanlllg llqrnqkdrx dfeaffshpf leqgpvkksc pvpvpmysgs vsgsscgssp  
 301 scrfasppsl pdmqhiqeen lssplgppn ylvskdsas tssknascdt ddfvlvphni  
 361 ssdhsdmpm gtagrrasne flvcgggqcp tvsphsetap ipvptqirny grieqnltst  
 421 assgtnvhgs prsavvrrsn tspmgflrpg scspvpadta qtvgrrlstg ssrxpyspspl  
 35 481 vgtipeqfsg cocghpgghd srsrnsqsp vpaqspqsl lsgarlqsap tldiyqnkq  
 541 klrkqhadv cpshtgagys yspqsrpqs lqstptkhlq ssprsdwff ktplptiigs  
 601 ptkttapfki pktqassnll alvtrhgpa egskdgnep ecahcllvqg serqraegqs  
 661 kavfgrevst gklsdqgqkt picrhqgstd slnterpmidi apagacggvl appagtaass

SUBSTITUTE SHEET (RULE 26)

WO 02/070559

PCT/GB02/00948

8

721 kavltvgsp phsaaaptot hmflrttts vgpsnsgsl camsgrcvvg sppgpfgss  
781 ppgaaaapsl ryvpygaspp sleglitfea pelpeetlme rehtdtrhl nwnlmftec  
841 lditamrgn peltsavsl yqiqesvvvd qisqlskdwg rveqlvlymk aaqllaaslh  
901 lakaqiksgk lspstavkv vknlnerykf citmrkkite klnrffsdqk rfideinsvt  
5 961 aekliyncav emvqsaalde mfqqtedivy ryhkaalle glsrilqdpd dienvhkyk  
1021 sierrlsalc hstatv

## 【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  
12 September 2002 (12.09.2002)

PCT

(10) International Publication Number  
WO 02/070559 A3

- (51) International Patent Classification: C07K 14/705, 2NU (GB). POTTER, Sarah, Jane [GB/GB]; 39 Kneller Road, Brockley, London SE4 2AR (GB).  
14/72, A61K 38/17
- (21) International Application Number: PCT/GB02/00948 (74) Agents: MERCER, Christopher, Paul et al.; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).
- (22) International Filing Date: 5 March 2002 (05.03.2002)
- (25) Filing Language: English (81) Designated States (national): AH, 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, GR, GU, HD, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, OM, PA, PE, PG, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (26) Publication Language: English (84) Designated States (regional): ARIPO patent (GI, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BI, CH, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (30) Priority Data: 0105402.2 5 March 2001 (05.03.2001) GB
- (71) Applicant (for all designated States except US): INPHARMATICA LIMITED [GB/GB]; 60 Charlotte Street, London W1T 2NU (GB).  
(72) Inventors; and  
(75) Inventors/Applicants (for US only): FAGAN, Richard, Joseph [US/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PHELPS, Christopher, Benjamin [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PHILLIPS, Tom [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). PIERRON, Valerie, Nathalie [FR/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). ALLEN, Kathryn, Elizabeth [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB). ALLEN, Janet, Marjorie [GB/GB]; Inpharmatica Limited, 60 Charlotte Street, London W1T 2NU (GB).
- Published:  
with international search report  
(88) Date of publication of the international search report: 3 April 2003
- 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 02/070559 A3

(54) Title: NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAINS

(57) Abstract: This invention relates to the novel proteins, termed LBDG1 and LBDG4, herein identified as Nuclear Hormone Receptor Ligand Binding Domains and to the use of these proteins and nucleic acid sequences from the encoding genes in the diagnosis, prevention and treatment of disease.

## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/GB 02/00948
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C07K14/705 C07K14/72 A61K38/17		
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 C07K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the International search (name of data base and, where practical, search terms used) BIOSIS, EPO-Internal, ENBL, SEQUENCE SEARCH, WPI Data, PAJ, MEDLINE		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BURRIS THOMAS P ET AL: "The gene responsible for adrenal hypoplasia congenita, DAX-1, encodes a nuclear hormone receptor that defines a new class within the superfamily." RECENT PROGRESS IN HORMONE RESEARCH, vol. 51, 1996, pages 241-260, XP001105530 Conference; Skamania, Washington, USA, 1995, 1996 The Endocrine Society 9650 Rockville Pike, Bethesda, Maryland 20814, USA ISBN: 1-879225-22-0 abstract --- -/--	1
<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		*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*E* earlier document but published on or after the international filing date		*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
*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)		*Y* document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
*O* document relating to an oral disclosure, use, exhibition or other means		*Z* document member of the same patent family
*P* document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 24 September 2002	Date of mailing of the international search report 16.01.03	
Name and mailing address of the ISA European Patent Office, P.B. 5518 Patentlaan 2 NL - 2009 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 opo nl, Fax: (+31-70) 340-2016	Authorized officer Nichogiannopoulou, A	

Form PCT/ISA/210 (second sheet) July 1992

## INTERNATIONAL SEARCH REPORT

 International Application No  
 PCT/GB 02/00948

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SEOL WONGI ET AL: "An orphan nuclear hormone receptor that lacks a DNA binding domain and heterodimerizes with other receptors." SCIENCE (WASHINGTON D C), vol. 272, no. 5266, 1996, pages 1336-1339, XP001093740 ISSN: 0036-8075 abstract	1
X	WO 98 49276 A (HELIX RESEARCH INST; MURAMATSU MASAOKI (JP); TOKUMITSU HIROSHI (JP) 5 November 1998 (1998-11-05) page 43 -page 58	1-3, 7-16,18, 22-48
X	KUROYANAGI HIDEHITO ET AL: "Human ULK1, a novel serine/threonine kinase related to UNC-51 kinase of <i>Caenorhabditis elegans</i> : cDNA cloning, expression, and chromosomal assignment." GENOMICS, vol. 51, no. 1, 1 July 1998 (1998-07-01), pages 76-85, XP002214556 ISSN: 0888-7543 cited in the application the whole document	1-3,7-16
X	-& DATABASE EMBL 'Online! 19 August 1998 (1998-08-19) KUROYANAGI H ET AL: "Homo sapiens serine/threonine kinase ULK1 (ULK1) mRNA" retrieved from EMBL Database accession no. AF045458 XP002214557 cited in the application Sequences with 100% identity with SEQ ID Nos:1 and 2 over 528 nucleotides and 1050 amino acids respectively abstract	1-3,7-16
A	SEIELSTAD D A ET AL: "ANALYSIS OF THE STRUCTURAL CORE OF THE HUMAN ESTROGEN RECEPTOR LIGAND BINDING DOMAIN BY SELECTIVE PROTEOLYSIS/MASS SPECTROMETRIC ANALYSIS" BIOCHEMISTRY, AMERICAN CHEMICAL SOCIETY, EASTON, PA, US, vol. 34, no. 39, 1995, pages 12605-12615, XP002063749 ISSN: 0006-2960 figure 1	1

1

Form PCT/IB/A/2 (continuation of second sheet) (July 1996)

INTERNATIONAL SEARCH REPORT	International application No. PCT/GB 02/00948
<b>Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)</b>	
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:	
1.	<input checked="" type="checkbox"/> Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: see FURTHER INFORMATION sheet PCT/ISA/210
2.	<input checked="" type="checkbox"/> Claims Nos.: 17, 19-21 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
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).
<b>Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)</b>	
This International Searching Authority found multiple inventions in this international application, as follows:	
see additional sheet	
1.	<input type="checkbox"/> As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
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.
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.:
4.	<input checked="" 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.: 1-3 all completely; 7-48 partially
<b>Remark on Protest</b>	
	<input type="checkbox"/> The additional search fees were accompanied by the applicant's protest.
	<input type="checkbox"/> No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International Application No. PCT/GB 02/0948

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: 1-3 all completely and 7-48 all partially

Polypeptides, nucleic acids, vectors, host cells, ligands, agonists, antagonists, pharmaceuticals, vaccines, kits, transgenic animals and methods relating to the LBD61 peptide (Acc. No.: 075385) (SEQ ID Nos:1 and 2).

2. Claims: 4-6 all completely and 7-48 all partially

Polypeptides, nucleic acids, vectors, host cells, ligands, agonists, antagonists, pharmaceuticals, vaccines, kits, transgenic animals and methods relating to the LBD64 peptide (Acc. No.: BAA31598.1) (SEQ ID Nos:3 and 4).

## INTERNATIONAL SEARCH REPORT

International Application No. PCT/JP 02/0948

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

## Continuation of Box I.1

Although claims 38-41 and 48 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claims 23, 25, 26-31 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

## Continuation of Box I.2

Claims Nos.: 17, 19-21

Present claims 17 and 19-21 relate to compounds defined by reference to a desirable characteristic or property, namely their ability to bind to a claimed polypeptide or their ability to increase or decrease the level of expression or activity of a claimed polypeptide.

The claims cover all compounds having this characteristic or property, whereas the application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for any such compound. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, no search has been carried out for these claims.

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 SEARCH REPORT**  
Information on patent family members

International Application No  
PCT/GB 02/00948

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9849276	A	AU 6421498 A	24-11-1998
		WO 9849276 A1	05-11-1998
		US 6358720 B1	19-03-2002

## フロントページの続き

(51)Int.Cl. <sup>7</sup>	F I	テーマコード(参考)
A 6 1 K 39/00	A 6 1 K 45/00	4 C 0 8 5
A 6 1 K 45/00	A 6 1 K 48/00	4 C 0 8 6
A 6 1 K 48/00	A 6 1 P 1/00	4 C 0 8 7
A 6 1 P 1/00	A 6 1 P 3/04	
A 6 1 P 3/04	A 6 1 P 3/06	
A 6 1 P 3/06	A 6 1 P 3/10	
A 6 1 P 3/10	A 6 1 P 3/14	
A 6 1 P 3/14	A 6 1 P 5/14	
A 6 1 P 5/14	A 6 1 P 7/00	
A 6 1 P 7/00	A 6 1 P 9/00	
A 6 1 P 9/00	A 6 1 P 9/06	
A 6 1 P 9/06	A 6 1 P 9/10	
A 6 1 P 9/10	A 6 1 P 9/10	1 0 1
A 6 1 P 9/12	A 6 1 P 9/10	1 0 3
A 6 1 P 11/00	A 6 1 P 9/12	
A 6 1 P 11/06	A 6 1 P 11/00	
A 6 1 P 13/12	A 6 1 P 11/06	
A 6 1 P 17/02	A 6 1 P 13/12	
A 6 1 P 17/04	A 6 1 P 17/02	
A 6 1 P 17/06	A 6 1 P 17/04	
A 6 1 P 17/10	A 6 1 P 17/06	
A 6 1 P 19/02	A 6 1 P 17/10	
A 6 1 P 19/10	A 6 1 P 19/02	
A 6 1 P 25/00	A 6 1 P 19/10	
A 6 1 P 25/02	A 6 1 P 25/00	
A 6 1 P 25/04	A 6 1 P 25/02	
A 6 1 P 25/22	A 6 1 P 25/04	
A 6 1 P 25/24	A 6 1 P 25/22	
A 6 1 P 25/28	A 6 1 P 25/24	
A 6 1 P 29/00	A 6 1 P 25/28	
A 6 1 P 31/04	A 6 1 P 29/00	
A 6 1 P 31/10	A 6 1 P 31/04	
A 6 1 P 31/12	A 6 1 P 31/10	
A 6 1 P 31/18	A 6 1 P 31/12	
A 6 1 P 33/00	A 6 1 P 31/18	
A 6 1 P 35/00	A 6 1 P 33/00	
A 6 1 P 35/02	A 6 1 P 35/00	
A 6 1 P 37/02	A 6 1 P 35/02	
A 6 1 P 37/06	A 6 1 P 37/02	
A 6 1 P 37/08	A 6 1 P 37/06	
A 6 1 P 43/00	A 6 1 P 37/08	
C 0 7 K 14/72	A 6 1 P 43/00	1 0 5
C 1 2 N 1/15	C 0 7 K 14/72	
C 1 2 N 1/19	C 1 2 N 1/15	
C 1 2 N 1/21	C 1 2 N 1/19	
C 1 2 N 5/10	C 1 2 N 1/21	
C 1 2 Q 1/68	C 1 2 Q 1/68	A

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
	C 1 2 N 5/00	A
	A 6 1 K 37/02	

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,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,OM,PH,P L,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZM,ZW

(74)代理人 100084009  
弁理士 小川 信夫

(74)代理人 100082821  
弁理士 村社 厚夫

(74)代理人 100086771  
弁理士 西島 孝喜

(74)代理人 100084663  
弁理士 箱田 篤

(72)発明者 ファガン リチャード ジョセフ  
イギリス ロンドン ダブリュ1ティー 2エヌユー チャーロット ストリート 60 インフ  
ァーマティカ リミテッド

(72)発明者 フェルプス クリストファー ベンジャミン  
イギリス ロンドン ダブリュ1ティー 2エヌユー チャーロット ストリート 60 インフ  
ァーマティカ リミテッド

F ターム(参考) 2G045 AA34 AA35 DA13 DA36 FB02  
4B024 AA01 AA11 AA12 BA63 CA04 CA09 EA04 GA11 HA12  
4B063 QA18 QA19 QQ53 QQ79 QR32 QR38 QR55 QR60 QR62 QR80  
QS11 QS24 QS25 QS33 QS34  
4B065 AA93Y AB01 AC14 BA02 CA24 CA44 CA46  
4C084 AA02 AA03 AA07 AA13 AA17 BA01 BA22 BA23 BA35 BA44  
CA01 CA18 CA53 DB63 MA13 MA17 MA22 MA23 MA31 MA35  
MA37 MA41 MA43 MA52 MA59 MA60 MA65 MA66 NA14 ZA022  
ZA032 ZA082 ZA122 ZA152 ZA162 ZA202 ZA222 ZA362 ZA402 ZA422  
ZA452 ZA512 ZA592 ZA622 ZA662 ZA702 ZA812 ZA892 ZA962 ZB072  
ZB092 ZB112 ZB132 ZB262 ZB322 ZB332 ZB352 ZB372 ZC062 ZC212  
ZC332 ZC352 ZC412 ZC422  
4C085 AA03 BB11 BB24 CC21 CC24 DD62 EE01 EE06 GG02 GG03  
GG04 GG05  
4C086 AA01 AA02 AA03 EA16 MA01 MA04 MA13 MA17 MA22 MA23  
MA31 MA35 MA37 MA41 MA43 MA52 MA59 MA60 MA65 MA66  
NA14 ZA02 ZA05 ZA08 ZA12 ZA15 ZA16 ZA20 ZA22 ZA36  
ZA40 ZA42 ZA45 ZA51 ZA59 ZA66 ZA70 ZA81 ZA89 ZA96  
ZB05 ZB07 ZB09 ZB11 ZB13 ZB26 ZB32 ZB33 ZB35 ZB37  
ZC06 ZC21 ZC33 ZC35 ZC41 ZC42  
4C087 AA01 AA02 BC83 CA12 CA47 MA01 MA13 MA17 MA22 MA23  
MA31 MA35 MA37 MA41 MA43 MA52 MA59 MA60 MA65 MA66  
NA14 ZA02 ZA05

专利名称(译)	<无法获取翻译>		
公开(公告)号	<a href="#">JP2005500012A5</a>	公开(公告)日	2005-12-22
申请号	JP2002569878	申请日	2002-03-05
[标]申请(专利权)人(译)	Inpharmatica公司有限公司		
申请(专利权)人(译)	Inpharmatica公司有限公司		
[标]发明人	ファガンリチャードジョセフ フェルプスクリストファーベンジャミン		
发明人	ファガン リチャード ジョセフ フェルプス クリストファー ベンジャミン		
IPC分类号	A01K67/027 A61K31/7088 A61K35/76 A61K38/00 A61K39/00 A61K39/395 A61K45/00 A61K48/00 A61P1/00 A61P1/18 A61P3/00 A61P3/04 A61P3/06 A61P3/10 A61P3/14 A61P5/00 A61P5/14 A61P7 /00 A61P7/02 A61P9/00 A61P9/06 A61P9/08 A61P9/10 A61P9/12 A61P11/00 A61P11/06 A61P13/12 A61P15/00 A61P17/00 A61P17/02 A61P17/04 A61P17/06 A61P17/10 A61P19/00 A61P19/02 A61P19 /10 A61P21/00 A61P25/00 A61P25/02 A61P25/04 A61P25/22 A61P25/24 A61P25/28 A61P29/00 A61P31/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37 /00 A61P37/02 A61P37/06 A61P37/08 A61P43/00 C07K14/705 C07K14/72 C07K16/28 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12Q1/68 G01N33/15 G01N33/50 G01N33/53 G01N37/00		
CPC分类号	A61K38/00 A61P1/00 A61P1/18 A61P3/00 A61P3/04 A61P3/06 A61P3/10 A61P3/14 A61P5/00 A61P5 /14 A61P7/00 A61P7/02 A61P9/00 A61P9/06 A61P9/08 A61P9/10 A61P9/12 A61P11/00 A61P11/06 A61P13/12 A61P15/00 A61P17/00 A61P17/02 A61P17/04 A61P17/06 A61P17/10 A61P19/00 A61P19 /02 A61P19/10 A61P21/00 A61P25/00 A61P25/02 A61P25/04 A61P25/22 A61P25/24 A61P25/28 A61P29/00 A61P31/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35 /02 A61P37/00 A61P37/02 A61P37/06 A61P37/08 A61P43/00 C07K14/72		
FI分类号	C12N15/00.ZNA.A A01K67/027 A61K31/7088 A61K35/76 A61K39/00.H A61K45/00 A61K48/00 A61P1 /00 A61P3/04 A61P3/06 A61P3/10 A61P3/14 A61P5/14 A61P7/00 A61P9/00 A61P9/06 A61P9/10 A61P9/10.101 A61P9/10.103 A61P9/12 A61P11/00 A61P11/06 A61P13/12 A61P17/02 A61P17/04 A61P17/06 A61P17/10 A61P19/02 A61P19/10 A61P25/00 A61P25/02 A61P25/04 A61P25/22 A61P25 /24 A61P25/28 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/02 A61P37/06 A61P37/08 A61P43/00.105 C07K14/72 C12N1/15 C12N1/19 C12N1 /21 C12Q1/68.A G01N33/15.Z G01N33/50.Z C12N5/00.A A61K37/02		
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/DA13 2G045/DA36 2G045/FB02 4B024/AA01 4B024/AA11 4B024 /AA12 4B024/BA63 4B024/CA04 4B024/CA09 4B024/EA04 4B024/GA11 4B024/HA12 4B063/QA18 4B063/QA19 4B063/QQ53 4B063/QQ79 4B063/QR32 4B063/QR38 4B063/QR55 4B063/QR60 4B063 /QR62 4B063/QR80 4B063/QS11 4B063/QS24 4B063/QS25 4B063/QS33 4B063/QS34 4B065/AA93Y 4B065/AB01 4B065/AC14 4B065/BA02 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA02 4C084 /AA03 4C084/AA07 4C084/AA13 4C084/AA17 4C084/BA01 4C084/BA22 4C084/BA23 4C084/BA35 4C084/BA44 4C084/CA01 4C084/CA18 4C084/CA53 4C084/DB63 4C084/MA13 4C084/MA17 4C084 /MA22 4C084/MA23 4C084/MA31 4C084/MA35 4C084/MA37 4C084/MA41 4C084/MA43 4C084/MA52 4C084/MA59 4C084/MA60 4C084/MA65 4C084/MA66 4C084/NA14 4C084/ZA022 4C084/ZA032 4C084/ZA082 4C084/ZA122 4C084/ZA152 4C084/ZA162 4C084/ZA202 4C084/ZA222 4C084/ZA362 4C084/ZA402 4C084/ZA422 4C084/ZA452 4C084/ZA512 4C084/ZA592 4C084/ZA622 4C084/ZA662 4C084/ZA702 4C084/ZA812 4C084/ZA892 4C084/ZA962 4C084/ZB072 4C084/ZB092 4C084/ZB112 4C084/ZB132 4C084/ZB262 4C084/ZB322 4C084/ZB332 4C084/ZB352 4C084/ZB372 4C084/ZC062 4C084/ZC212 4C084/ZC332 4C084/ZC352 4C084/ZC412 4C084/ZC422 4C085/AA03 4C085/BB11 4C085/BB24 4C085/CC21 4C085/CC24 4C085/DD62 4C085/EE01 4C085/EE06 4C085/GG02 4C085		

/GG03 4C085/GG04 4C085/GG05 4C086/AA01 4C086/AA02 4C086/AA03 4C086/EA16 4C086/MA01  
4C086/MA04 4C086/MA13 4C086/MA17 4C086/MA22 4C086/MA23 4C086/MA31 4C086/MA35 4C086  
/MA37 4C086/MA41 4C086/MA43 4C086/MA52 4C086/MA59 4C086/MA60 4C086/MA65 4C086/MA66  
4C086/NA14 4C086/ZA02 4C086/ZA05 4C086/ZA08 4C086/ZA12 4C086/ZA15 4C086/ZA16 4C086  
/ZA20 4C086/ZA22 4C086/ZA36 4C086/ZA40 4C086/ZA42 4C086/ZA45 4C086/ZA51 4C086/ZA59  
4C086/ZA66 4C086/ZA70 4C086/ZA81 4C086/ZA89 4C086/ZA96 4C086/ZB05 4C086/ZB07 4C086  
/ZB09 4C086/ZB11 4C086/ZB13 4C086/ZB26 4C086/ZB32 4C086/ZB33 4C086/ZB35 4C086/ZB37  
4C086/ZC06 4C086/ZC21 4C086/ZC33 4C086/ZC35 4C086/ZC41 4C086/ZC42 4C087/AA01 4C087  
/AA02 4C087/BC83 4C087/CA12 4C087/CA47 4C087/MA01 4C087/MA13 4C087/MA17 4C087/MA22  
4C087/MA23 4C087/MA31 4C087/MA35 4C087/MA37 4C087/MA41 4C087/MA43 4C087/MA52 4C087  
/MA59 4C087/MA60 4C087/MA65 4C087/MA66 4C087/NA14 4C087/ZA02 4C087/ZA05

---

代理人(译)

中村稔  
小川伸男  
西岛隆义

---

优先权

2001005402 2001-03-05 GB

---

其他公开文献

JP2005500012A

---

摘要(译)

本发明涉及一种新型蛋白质，称为LBDG3，在本文中被鉴定为核激素受体配体结合域，并且涉及该蛋白质和来自编码基因的核酸序列在疾病的诊断，预防和治疗中的用途。