

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-500123

(P2004-500123A)

(43) 公表日 平成16年1月8日(2004.1.8)

(51) Int.Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A
C 0 7 K 14/47	C O 7 K 14/47	2 G O 4 5
C 0 7 K 16/18	C O 7 K 16/18	4 B O 2 4
C 0 7 K 19/00	C O 7 K 19/00	4 B O 6 5
C 1 2 N 1/15	C 1 2 N 1/15	4 C O 8 4
	審査請求 未請求 予備審査請求 有	4 H O 4 5
		(全 70 頁) 最終頁に続く

(21) 出願番号	特願2001-575617 (P2001-575617)	(71) 出願人	591032596
(86) (22) 出願日	平成13年4月4日 (2001.4.4)		メルク パテント ゲゼルシャフト ミッ
(85) 翻訳文提出日	平成14年10月4日 (2002.10.4)		ト ベシュレンクテル ハフツング
(86) 国際出願番号	PCT/EP2001/003793		Merck Patent Gesell
(87) 国際公開番号	W02001/077147		schaft mit beschrae
(87) 国際公開日	平成13年10月18日 (2001.10.18)		nkter Haftung
(31) 優先権主張番号	00107352.7		ドイツ連邦共和国 デー-64293 ダ
(32) 優先日	平成12年4月5日 (2000.4.5)		ルムシュタット フランクフルター シュ
(33) 優先権主張国	欧州特許庁 (EP)		トラーセ 250
(81) 指定国	EP (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), CA, JP, US		Frankfurter Str. 25
			O, D-64293 Darmstadt
			, Federal Republic o
			f Germany
		(74) 代理人	100123788
			弁理士 宮崎 昭夫
			最終頁に続く

(54) 【発明の名称】 新しいプロモドメインタンパク質

(57) 【要約】

h u p o l y b r o m o 1 のポリペプチドおよびポリヌクレオチドならびにそのようなポリペプチドを組換え技術によって製造するための方法が開示される。h u p o l y b r o m o 1 のポリペプチドおよびポリヌクレオチドを診断アッセイにおいて用いるための方法もまた開示される。

【特許請求の範囲】

【請求項 1】

- (a) 配列番号 1 の配列を含むポリヌクレオチドによってコードされるポリペプチド、
(b) 配列番号 2 のポリペプチド配列に対して少なくとも 95 % の同一性を有するポリペプチド配列を含むポリペプチド、
(c) 配列番号 2 のポリペプチド配列に対して少なくとも 95 % の同一性を有するポリペプチド、
(d) 配列番号 2 のポリペプチド配列、および
(e) (a) ~ (d) に記載のかかるポリペプチドのフラグメントおよび変異体
からなる群から選択されるポリペプチド。

10

【請求項 2】

配列番号 2 のポリペプチド配列を含む、請求項 1 に記載のポリペプチド。

【請求項 3】

配列番号 2 のポリペプチド配列である、請求項 1 に記載のポリペプチド。

【請求項 4】

- (a) 配列番号 1 のポリヌクレオチド配列に対して少なくとも 95 % の同一性を有するポリヌクレオチド配列を含むポリヌクレオチド、
(b) 配列番号 1 のポリヌクレオチドに対して少なくとも 95 % の同一性を有するポリヌクレオチド、
(c) 配列番号 2 のポリペプチド配列に対して少なくとも 95 % の同一性を有するポリペプチド配列をコードするポリヌクレオチド配列を含むポリヌクレオチド、
(d) 配列番号 2 のポリペプチド配列に対して少なくとも 95 % の同一性を有するポリペプチド配列をコードするポリヌクレオチド配列を有するポリヌクレオチド、
(e) 配列番号 1 の配列または少なくとも 15 塩基長を有するそのフラグメントを有する標識されたプローブを用いた厳格なハイブリダイゼーション条件下でライブラリをスクリーニングすることにより得られる少なくとも 100 塩基長の塩基配列を有するポリヌクレオチド、
(f) (a) ~ (e) のポリヌクレオチドの RNA 等価体であるポリヌクレオチド、
(g) (a) ~ (f) のいずれかの前記ポリヌクレオチドに対して相補的なポリヌクレオチド配列、および
(h) (a) ~ (g) のいずれかのポリヌクレオチドの変異体またはフラグメントであるか、あるいは前記のポリヌクレオチドに対してその全長にわたって相補的であるポリヌクレオチド
からなる群から選択されるポリヌクレオチド。

20

30

【請求項 5】

- (a) 配列番号 1 のポリヌクレオチドを含むポリヌクレオチド、
(b) 配列番号 1 のポリヌクレオチド、
(c) 配列番号 2 のポリペプチドをコードするポリヌクレオチド配列を含むポリヌクレオチド、および
(d) 配列番号 2 のポリペプチドをコードするポリヌクレオチド
からなる群から選択される、請求項 4 に記載のポリヌクレオチド。

40

【請求項 6】

該発現ベクターが適合性宿主細胞に存在する際、請求項 1 ~ 3 のいずれかに記載のポリペプチドを産生し得るポリヌクレオチドを含む発現システム。

【請求項 7】

請求項 1 ~ 3 のいずれかに記載のポリペプチドを発現する、請求項 6 に記載される発現ベクターを含む組換え宿主細胞またはその膜。

【請求項 8】

請求項 1 ~ 3 のいずれかに記載のポリペプチドを製造する方法であって、請求項 7 に記載の宿主細胞を前記ポリペプチドの産生に十分な条件下で培養して、該ポリペプチドを該培

50

養培地から回収する工程を含む方法。

【請求項 9】

免疫グロブリンのFc領域と請求項1～3のいずれかに記載のポリペプチドとからなる融合タンパク質。

【請求項 10】

請求項1～3のいずれかに記載のポリペプチドに対して免疫特異的な抗体。

【請求項 11】

請求項1～3のいずれかに記載のポリペプチドの機能または濃度を刺激または阻害する化合物を同定するためのスクリーニング方法であって、

(a) 該ポリペプチド(または該ポリペプチドを発現する細胞もしくは膜)またはその融合タンパク質に対する候補化合物の結合を、該候補化合物に直接的または間接的に結合している標識によって定量的または定性的に測定または検出すること、 10

(b) 該ポリペプチド(または該ポリペプチドを発現する細胞もしくは膜)あるいはその融合タンパク質に対する候補化合物の結合の競合を、標識された競合剤の存在下で測定すること、

(c) 該ポリペプチドの活性化または阻害により発生するシグナルを該候補化合物がもたらすか否かを、該ポリペプチドを発現する細胞または細胞膜に適切な検出システムを使用して試験すること、

(d) 候補化合物を、請求項1～3のいずれかに記載のポリペプチドを含有する溶液と混合して混合物を作製し、該混合物中の該ポリペプチドの活性を測定し、そして、何らの候補化合物を含有しない対照混合物に対して、該混合物の活性を比較すること、または 20

(e) 細胞中における前記ポリペプチドをコードするmRNAまたは前記ポリペプチドの産生に対する候補化合物の作用を、例えばELISAアッセイを使用して検出すること、からなる群から選択される方法と、

(f) 生物工学的または化学的な標準的な技術に従って前記化合物を製造すること、を含む方法。

【発明の詳細な説明】

【0001】

(発明の分野)

本発明は、以降、時に「hupolybromo1」と称される新しく同定されたポリペプチド、およびかかるポリペプチドをコードするポリヌクレオチド、診断ならびに治療上潜在的に有用となるアゴニスト、アンタゴニストであり得る化合物を同定する際のそれらの使用、ならびにかかるポリペプチドおよびポリヌクレオチドの製造に関する。 30

【0002】

(発明の背景)

創薬プロセスは現在、「機能ゲノム科学」、すなわちハイ・スループットのゲノムまたは遺伝子に基づく生物学を取り入れ、根本的改革が進められている。この方法は、治療の標的として遺伝子および遺伝子産物を同定する手段として、迅速に、「ポジショナルクローニング」に基づいたより初期の方法の代わりになりつつある。生物機能または遺伝子の疾患である表現型を同定し、次いでこれに対する原因遺伝子とその遺伝子地図の位置に基づいて突き止められる。 40

【0003】

機能ゲノム科学は、ハイ・スループットDNA配列決定技術、および現在入手可能な多くの分子生物学データベースから、潜在的に注目される遺伝子配列を同定するためにバイオインフォマティクスの様々なツールに大きく依存している。創薬の標的として、さらなる遺伝子およびその関連するポリペプチド/タンパク質を同定し、その特定を行うことが引き続き求められている。

【0004】

(発明の概要)

本発明は、hupolybromo1に関し、特にhupolybromo1ポリペプ 50

チドおよび h u p o l y b r o m o 1 ポリヌクレオチド、組換え体、ならびにそれらの製造方法に関する。かかるポリペプチドおよびポリヌクレオチドは、一般にはガン、アレルギー、慢性関節リウマチ、自己免疫疾患、卒中、痴呆、精神遅滞、パーキンソン病、てんかん、精神分裂病、アルツハイマー病、鬱病、冠状動脈性心疾患、心不全、心筋症および嚢胞性線維症（これらに限定されない）を含むある種の疾患（これらは以降「本発明の疾患」と呼ばれる）を処置する方法に関連して注目されている。さらなる態様において、本発明は、本発明により提供される材料を用いてアゴニストおよびアンタゴニスト（例えば阻害剤）を同定する方法、ならびに同定された化合物を用いて、h u p o l y b r o m o 1 の活性に関連した状態を処置する方法に関する。さらにさらなる態様において、本発明は h u p o l y b r o m o 1 の不適切な活性および濃度に付随する疾患を検出する診断アッセイに関する。

10

【0005】

（発明の説明）

第一の態様において、本発明は h u p o l y b r o m o 1 ポリペプチドに関する。かかるポリペプチドには下記が含まれる：

- （a）配列番号 1 の配列を含むポリヌクレオチドによってコードされるポリペプチド、
- （b）配列番号 2 のポリペプチド配列に対して少なくとも 95 %、96 %、97 %、98 % または 99 % の同一性を有するポリペプチド配列を含むポリペプチド、
- （c）配列番号 2 のポリペプチド配列を含むポリペプチド、
- （d）配列番号 2 のポリペプチド配列に対して少なくとも 95 %、96 %、97 %、98 % または 99 % の同一性を有するポリペプチド、
- （e）配列番号 2 のポリペプチド配列、および
- （f）配列番号 2 のポリペプチド配列と比較して、0.95、0.96、0.97、0.98 または 0.99 の同一性指標を有するポリペプチド配列を有するか、またはそのようなポリペプチド配列を含むポリペプチド、
- （g）（a）～（f）に記載のかかるポリペプチドのフラグメントおよび変異体。

20

【0006】

本発明のポリペプチドは、ポリブロモ（P o l y b r o m o）ドメインタンパク質ファミリーのポリペプチドのメンバーであると考えられる。したがって、本発明のポリペプチドは、アダプタータンパク質として作用する多数の転写因子が、ブロモドメインと呼ばれるアミノ酸ドメインを含むことが見出されているので注目されている。ブロモドメインタンパク質は、調節性 DNA 配列への転写因子の接近を容易にするクロマチン再構成複合体である S W I / S N F 複合体のメンバーである。多数のブロモドメイン、短縮型 H M G ボックスおよび 2 つの反復を含有する p o l y b r o m o の分子クローニングが N i c o l a s , R . H . 他（G e n e , 1996、175：233～40）によって報告されている。このタンパク質複合体は、アダプタータンパク質によって決定される、異なったサブユニット組成を有する。この複合体は、最近では、染色体安定性の維持を調節し、そして DNA 修復の様々な局面を調節することによって細胞増殖の制御に関連している。L i u - Y 他（N u c l e i c - A c i d s - R e s . , 1998、26：1038～45）は、マウスの DNA メチルトランスフェラーゼ（D N A M T a s e）は p o l y b r o m o タンパク質によって DNA 複製の様々な部位に標的化されることを見出した。ブロモドメインタンパク質 b r m（S N F 2）の発現は細胞が形質転換されたときに高頻度にダウンレギュレーションされる。h u p o l y b r o m o 1 は、ヒト S N 2 - 4 タンパク質および B r m - d r o m e に対して相長的である。さらに、クロマチンの再構成に関与していると考えられるシゾサッカロミセス（S c h i z o s a c c h a r o m y c e）のタンパク質（G I 3 1 6 9 0 9 0）に対して相長的である。マウスにおいて、S N F 2 / S W I 2 ファミリーのメンバーである E T L 1 を有さない場合には、骨格の形成異常、生育遅延、低下した生後生存性および損なわれた生殖能力が認められている（S c h o o r , M . 他、M e c h . D e v . 1999、85：73～83）。最近、哺乳動物の S W I / S N F の増殖制御の様々な機能的側面が M u c h a r d t 他（S e m i n . C e l l D e v

30

40

50

・ Biol. 1999、10:189~95) によってまとめられている。これまでに記載された機能的側面は、クロマチンの再構成および遺伝子の示差的な転写調節に関して重要であり、したがってガン細胞および転移の発生に関与する機構の一部を構成する。さらに、これらの機構は、脳障害、自己免疫疾患およびいくつかの形態の脳疾患にもまた関与している。

【0007】

hupolybromo1の生物学的性質は、以降、「hupolybromo1の生物学的活性」または「hupolybromo1活性」と称する。好ましくは、本発明のポリペプチドは、少なくとも1つのhupolybromo1の生物学的活性を示す。

【0008】

本発明のポリペプチドには、すべての対立遺伝子形態およびスプライス変異体を含む上記ポリペプチドの変異体もまた含まれる。かかるポリペプチドは、挿入、欠失、ならびに保存的または非保存的であり得る置換、あるいはそれらの任意の組合せによって基準ポリペプチドから変異している。特に好ましい変異体は、いくつかの、例えば、50~30、30~20、20~10、10~5、5~3、3~2、2~1または1個のアミノ酸が任意の組合せで挿入、置換または欠失されている変異体である。

【0009】

本発明のポリペプチドの好ましいフラグメントには、配列番号2のアミノ酸配列に由来する少なくとも30、50または100個の連続したアミノ酸を有するアミノ酸配列を含むポリペプチド、あるいは配列番号2のアミノ酸配列から、少なくとも30、50または100個の連続したアミノ酸が短縮化または欠失しているアミノ酸配列を含むポリペプチドが含まれる。好ましいフラグメントは、hupolybromo1の生物学的活性をもたらす生物学的に活性なフラグメントである。これには、類似する活性または改善された活性を有するか、あるいは望ましくない活性が低下しているフラグメントが含まれる。また、動物、特にヒトにおいて抗原性または免疫原性であるそのようなフラグメントも好ましい。

【0010】

本発明のポリペプチドのフラグメントは、対応する完全長型ポリペプチドをペプチド合成によって製造するために用いることができる。従って、これらの変異体は本発明の完全長型ポリペプチドを製造するための中間体として用いることができる。本発明のポリペプチドは、「成熟型」タンパク質の形態であってもよく、あるいは前駆体または融合タンパク質などのより大きなタンパク質の一部であってもよい。分泌配列またはリーダー配列、プロ配列、精製を助ける配列、例えば反復するヒスチジン残基、あるいは組換え産生時の安定性に必要なさらなる配列を含有する付加的なアミノ酸配列を含むことは、しばしば有益である。

【0011】

本発明のポリペプチドは、任意の好適な方法で、例えば、天然に存在する供給源から、または発現システム(下記参照)を含む遺伝子操作された宿主細胞から単離することによって、あるいは例えば、自動化されたペプチド合成機を使用する化学合成によって、かかるそのような方法の組合せによって調製することができる。かかるポリペプチドを調製するための手段は当該分野では十分に理解されている。

【0012】

さらなる態様において、本発明はhupolybromo1ポリヌクレオチドに関する。そのようなポリヌクレオチドには下記が含まれる：

- (a) 配列番号1のポリヌクレオチド配列に対する同一性が少なくとも95%、96%、97%、98%または99%であるポリヌクレオチド配列を含むポリヌクレオチド、
- (b) 配列番号1のポリヌクレオチドを含むポリヌクレオチド、
- (c) 配列番号1のポリヌクレオチドに対する同一性が少なくとも95%、96%、97%、98%または99%であるポリヌクレオチド、
- (d) 配列番号1のポリヌクレオチド、

10

20

30

40

50

(e) 配列番号 2 のポリペプチド配列に対して少なくとも 95%、96%、97%、98% または 99% の同一性を有するポリペプチド配列をコードするポリヌクレオチド配列を含むポリヌクレオチド、

(f) 配列番号 2 のポリペプチドをコードするポリヌクレオチド配列を含むポリヌクレオチド、

(g) 配列番号 2 のポリペプチド配列に対して少なくとも 95%、96%、97%、98% または 99% の同一性を有するポリペプチド配列をコードするポリヌクレオチド配列を有するポリヌクレオチド、

(h) 配列番号 2 のポリペプチドをコードするポリヌクレオチド、

(i) 配列番号 1 のポリヌクレオチド配列と比較して、0.95、0.96、0.97、0.98 または 0.99 の同一性指標を有するポリヌクレオチド配列を有するか、またはそのようなポリヌクレオチド配列を含むポリヌクレオチド、 10

(j) 配列番号 2 のポリペプチド配列と比較して、0.95、0.96、0.97、0.98 または 0.99 の同一性指標を有するポリペプチド配列をコードするポリヌクレオチド配列を有するか、またはそのようなポリヌクレオチド配列を含むポリヌクレオチド、ならびに

上記ポリヌクレオチドのフラグメントおよび変異体であるポリヌクレオチド、あるいは上記ポリヌクレオチドに対してその全長にわたって相補的であるポリヌクレオチド。

【0013】

本発明のポリヌクレオチドの好ましいフラグメントには、配列番号 1 の配列に由来する少なくとも 15、30、50 または 100 個の連続した塩基を有する塩基配列を含むポリヌクレオチド、あるいは配列番号 1 の配列から、少なくとも 30、50 または 100 個の連続した塩基が短縮化または欠失している配列を含むポリヌクレオチドが含まれる。 20

【0014】

本発明のポリヌクレオチドの好ましい変異体には、スプライス変異体、対立遺伝子変異体、および 1 つまたは複数の単一塩基多型 (SNP) を有するポリヌクレオチドを含む多型体が含まれる。

【0015】

本発明のポリヌクレオチドには、配列番号 2 のアミノ酸配列を含み、かついくつもの、例えば、50 ~ 30、30 ~ 20、20 ~ 10、10 ~ 5、5 ~ 3、3 ~ 2、2 ~ 1 または 1 個のアミノ酸残基が任意の組合せで置換、欠失または付加されているポリペプチド変異体をコードするポリヌクレオチドもまた含まれる。 30

【0016】

さらなる態様において、本発明は、本発明の DNA 配列の RNA 転写物であるポリヌクレオチドを提供する。従って下記の RNA ポリヌクレオチドが提供される：

(a) 配列番号 2 のポリペプチドをコードする DNA 配列の RNA 転写物を含む RNA ポリヌクレオチド、

(b) 配列番号 2 のポリペプチドをコードする DNA 配列の RNA 転写物である RNA ポリヌクレオチド、

(c) 配列番号 1 の DNA 配列の RNA 転写物を含む RNA ポリヌクレオチド、または 40

(d) 配列番号 1 の DNA 配列の RNA 転写物である RNA ポリヌクレオチド、ならびにそれらに対して相補的である RNA ポリヌクレオチド。

【0017】

配列番号 1 のポリヌクレオチド配列は、ニワトリ (gallus gallus) の polybromotankin X90849 との非常に大きい相同性を示す。これはまた、C. elegans のプロモドメインタンパク質 (GI: 1301625) に対しても相同的である。配列番号 1 のポリヌクレオチド配列は、配列番号 2 のポリペプチドをコードする cDNA 配列である。配列番号 2 のポリペプチドをコードするポリヌクレオチド配列は、配列番号 1 の配列をコードするポリペプチドと同一であってもよく、あるいは遺伝子暗号の縮退 (縮重性) の結果として配列番号 2 のポリペプチドを同様にコードする、配列 50

番号1以外の配列であってもよい。配列番号2のポリペプチドは、ニワトリ (gallus gallus) の polybromotank 質 X90849 との相同性および/または構造的類似性を有する Polybromotank 質ファミリーの他のタンパク質との関連性を有する。これはまた、ヒトタンパク質の SN2-4 (P51532) および Brm-drome (P25435) ならびに命名されていない2つのタンパク質 GI7022804 および GI7023493 に対しても相同的である。

【0018】

本発明の好ましいポリペプチドおよびポリヌクレオチドは、特に、それらの相同的なポリペプチドおよびポリヌクレオチドと類似する生物学的な機能/性質を有することが期待される。さらに、本発明の好ましいポリペプチドおよびポリヌクレオチドは少なくとも1つの hupolybromo1 活性を有する。 10

【0019】

本発明のポリヌクレオチドは、ヒトの胚組織混合物から単離した mRNA に由来する cDNA ライブラリから標準的なクローニング技術およびスクリーニング技術を使用して得ることができる (例えば、Sambrook 他、Molecular Cloning: A Laboratory Manual、第2版、Cold Spring Harbor Laboratory Press、Cold Spring Harbor、N.Y. (1989) 参照)。本発明のポリヌクレオチドはまた、ゲノム DNA ライブラリなどの天然の供給源から得ることができ、あるいは周知の技術および市販の技術を用いて合成することができる。 20

【0020】

本発明のポリヌクレオチドを本発明のポリペプチドの組換え製造のために使用する際には、該ポリヌクレオチドは、成熟型ポリペプチドのコード配列、それ自体、あるいはリーダー配列もしくは分泌配列、プレタンパク質配列もしくはプロタンパク質配列もしくはブレプロタンパク質配列、または他の融合ペプチド部分をコードするコード配列などの他のコード配列と読み枠を合わせた成熟型ポリペプチドのコード配列を含むことができる。例えば、融合ポリペプチドの精製を容易にするマーカー配列をコードさせることができる。本発明のこの態様のいくつかの好ましい実施形態において、該マーカー配列は、pQEベクター (Qiagen, Inc.) において提供され、そして Gentz 他、Proc. Natl. Acad. Sci. USA、(1989) 86:821~824 に記載されているように、ヘキサ・ヒスチジン・ペプチド、あるいは HA タグである。該ポリヌクレオチドはまた、転写される非翻訳配列、スプライシングシグナルおよびポリアデニレーション・シグナル、リボソーム結合部位、ならびに mRNA を安定化させる配列などの 5' および 3' 非コード配列を含有することができる。 30

【0021】

配列番号1のポリヌクレオチド配列に対して同一であるか、または十分な同一性を有するポリヌクレオチドは、cDNA およびゲノム DNA に対するハイブリダイゼーション・プローブとして、あるいは核酸増幅反応 (例えば PCR) 用のプライマーとして使用することができる。かかるプローブおよびプライマーは、本発明のポリペプチドをコードする完全長型 cDNA およびゲノム・クローンを単離するために、そして配列番号1に対する高い配列類似性 (概して少なくとも95%の同一性) を有する他の遺伝子 (ヒト供給源に由来するパラログならびにヒト以外の種に由来するオルソログおよびパラログをコードする遺伝子を含む) の cDNA クローンおよびゲノム・クローンを単離するために使用することができる。好ましいプローブおよびプライマーは、一般には、少なくとも15塩基、好ましくは少なくとも30塩基を含み、そして少なくとも100塩基ではなくとも、少なくとも50塩基を有し得る。特に好ましいプローブは30~50塩基を有する。特に好ましいプライマーは20~25塩基を有する。 40

【0022】

ヒト以外の種に由来するホモログを含む、本発明のポリペプチドをコードするポリヌクレオチドは、配列番号1の配列またはそのフラグメントを有する、好ましくは少なくとも1 50

5塩基の標識されたプローブを用いて厳密なハイブリダイゼーション条件下でライブラリをスクリーニングする過程；および前記ポリヌクレオチド配列を含有する完全長型cDNAクローンおよびゲノム・クローンを単離する過程を含む方法によって得ることができる。かかるハイブリダイゼーション技術は当業者には十分に知られている。好ましい厳格なハイブリダイゼーション条件には、50%ホルムアミド、5×SSC(150mMのNaCl、15mMのクエン酸三ナトリウム)、50mMのリン酸ナトリウム(pH7.6)、5×デンハルト溶液、10%のデキストラン硫酸および20マイクログラム/mlの変性させた剪断サケ精子DNAを含む溶液において42℃で一晩インキュベーションし、その後、0.1×SSCにおいて約65℃でフィルターを洗浄することが含まれる。従って、本発明はまた、配列番号1の配列またはそのフラグメントを有する、好ましくは少なくとも15塩基の標識されたプローブを用いて厳格なハイブリダイゼーション条件下でライブラリをスクリーニングすることによって得られる単離されたポリヌクレオチド、好ましくは少なくとも100塩基の塩基配列を有する単離されたポリヌクレオチドを含む。

10

【0023】

当業者は、多くの場合において、該ポリペプチドをコードする領域が5'末端に至るまで必ずしも完全に伸長していない点で、単離されたcDNA配列が不完全であることを理解している。これは、本質的に「プロセシング能」(ポリメリゼーション反応の間、テンプレートに結合した状態を維持する酵素の能力の指標)が低い酵素である逆転写酵素が、第1鎖cDNA合成の際にmRNAテンプレートのDNAコピーを完成させることができない結果である。

20

【0024】

全長型cDNAを得るために、あるいは短いcDNAを伸長させるために利用でき、かつ当業者に周知の方法がいくつかある。例えば、cDNA末端の迅速な増幅(RACE)方法に基づく方法がある(例えば、Frohman他、Proc. Nat. Acad. Sci. USA、85、8998~9002、1988を参照)。例えば、Marathon(商標)法(Clontech Laboratories Inc.)によって例示されるこの技術の近年の改変により、より長いcDNAに対する探索が著しく単純化されている。Marathon(商標)法では、選択された組織から抽出されたmRNAとその両脇に連結された「アダプター」配列とからcDNAが調整される。その後、核酸増幅(PCR)を行い、遺伝子特異的オリゴヌクレオチドプライマーおよびアダプター特異的オリゴヌクレオチドプライマーの組合せを使用してcDNAの「失われている」5'末端を増幅する。その後、「ネスティッド(入れこ型)」プライマー、すなわち、増幅産物の内部にアニーリングするように設計されたプライマー(典型的には、アダプター配列においてさらに3'側にアニーリングするアダプター特異的プライマー、および既知の遺伝子配列においてさらに5'側にアニーリングする遺伝子特異的プライマー)を使用して、PCR反応を繰り返す。その後、この反応の生成物はDNA配列決定によって分析することができる。生成物を既存のcDNAに直接結合して完全な配列を得るか、または5'プライマーを設計するための新しい配列情報を使用して別の完全長PCRを行うことによって、完全長型のcDNAを構築することができる。

30

【0025】

本発明の組換えポリペプチドは、発現システムを含む遺伝子操作された宿主細胞から当該分野で周知の方法によって調製することができる。従って、さらなる態様において、本発明は、本発明のポリヌクレオチドの1つまたは複数を含む発現システムと、そのような発現システムで遺伝子操作されている宿主細胞と、および組換え技術による本発明のポリペプチドの製造とに関する。本発明のDNA構築物に由来するRNAを用いてかかるタンパク質を製造するために、無細胞翻訳システムを用いることもできる。

40

【0026】

組換え製造のために、宿主細胞を遺伝子操作し、本発明のポリヌクレオチドに対する発現システムまたはその一部を取り込ませることができる。ポリヌクレオチドは、Davis他、Basic Methods in Molecular Biology(198

50

6) および S a m b r o o k 他 (同 上) などの多くの標準的な実験室マニュアルに記載される方法によって宿主細胞に導入することができる。ポリヌクレオチドを宿主細胞に導入する好ましい方法には、例えば、リン酸カルシウム・トランスフェクション、D E A E - デキストラン媒介トランスフェクション、トランスベクション、マイクロインジェクション、カチオン性脂質媒介トランスフェクション、エレクトロポレーション、トランスダクション、スクレイプ負荷、バリスティック導入または感染が含まれる。

【 0 0 2 7 】

適切な宿主の代表的な例には、ストレプトコッカス属、スタフィロコッカス属、大腸菌、ストレプトミセス属ならびに枯草菌細胞などの細菌細胞；酵母細胞やアスペルギルス属細胞などの菌類細胞；ショウジョウバエ (D r o s o p h i l a) S 2 細胞や S p o d o p t e r a S f 9 細胞などの昆虫細胞；C H O、C O S、H e L a、C 1 2 7、3 T 3、B H K、H E K 2 9 3 および B o w e s メラノーマ細胞などの動物細胞；ならびに植物細胞が含まれる。

10

【 0 0 2 8 】

多様な発現システムを使用することができる。例えば、染色体、エピソームおよびウイルスに由来するシステム、例えば、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、バキュロウイルス、パポバウイルス (S V 4 0 など)、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などのウイルスに由来するベクター、ならびに、プラスミドおよびバクテリオファージの遺伝子エレメントから誘導されたものなどの、それらの組合せに由来するベクター (コスミドおよびファージミドなど) を使用することができる。該発現システムは、発現を生じさせるとともに、発現を調節するための制御領域を含有することができる。一般に、宿主においてポリペプチドを産生させるためにポリヌクレオチドを維持し、または伝搬させ、または発現させることができるシステムまたはベクターはどれも使用することができる。適切なポリヌクレオチド配列を、例えば、S a m b r o o k 他 (同 上) に示されている技術などの様々なよく知られている日常的な技術のいずれかによって発現システムに挿入することができる。適切な分泌シグナルを、所望するポリペプチドに組み込んで、翻訳されたタンパク質を小胞体の内腔、ペリプラズム腔または細胞外環境に分泌させることができる。これらのシグナルは、ポリペプチドに対して内因性であってもよく、あるいは異種のシグナルであってもよい。

20

30

【 0 0 2 9 】

スクリーニングアッセイで用いるために本発明のポリペプチドを発現させる場合、該ポリペプチドを細胞の表面に産生させることが一般には好ましい。この場合細胞を集菌し、その後、スクリーニングアッセイにおいて使用することができる。該ポリペプチドが培地中に分泌される場合には、培地を回収してポリペプチドの回収および精製をすることができる。細胞内に産生された場合には、細胞を最初に溶解して、その後、ポリペプチドを回収しなければならない。

【 0 0 3 0 】

本発明のポリペプチドは、硫酸アンモニウムまたはエタノール沈殿、酸抽出、アニオンまたはカチオン交換クロマトグラフィ、ホスホセルロース・クロマトグラフィ、疎水性相互作用クロマトグラフィ、アフィニティー・クロマトグラフィ、ヒドロキシルアパタイト・クロマトグラフィおよびレクチン・クロマトグラフィを含む周知の方法によって組換え細胞培養物から回収および精製することができる。最も好ましくは、ハイ・パフォーマンス液体クロマトグラフィが精製に用いられる。該ポリペプチドが細胞内合成、単離および/または精製のときに変性する際には、タンパク質をリフォールディングさせるために周知の技術を用いて、活性な立体配座を再生させることができる。

40

【 0 0 3 1 】

本発明のポリヌクレオチドは、関連遺伝子における変異を検出することによって診断試薬として使用することができる。c D N A 配列またはゲノム配列における配列番号 1 のポリヌクレオチドにより特定される遺伝子で、機能不全に関連している遺伝子の突然変異型を

50

検出することによって、その遺伝子の過少発現、過剰発現あるいは変化した空間的または時間的な発現から生じる疾患の診断またはそのような疾患に対する感受性の診断の一助になり得るか、またはそのような診断を規定し得る診断ツールが提供される。遺伝子に変異を有する個体を、当該分野で周知の様々な技術によってDNAレベルで検出することができる。

【0032】

診断に必要な核酸は、血液、尿、唾液、組織生検体または解剖体などの被験体の細胞から得ることができる。ゲノムDNAは、検出のために直接使用することができ、あるいは分析に先立ち、PCR、好ましくはRT-PCR、または他の増幅法を使用して酵素的に増幅してもよい。RNAまたはcDNAもまた同様の方法で使用することができる。欠失および挿入は、正常な遺伝子型と比較して増幅産物のサイズにおける変化によって検出することができる。点変異は、増幅されたDNAをhupolybromo1の標識されたヌクレオチド配列にハイブリダイゼーションさせることによって同定することができる。完全に一致する配列は、リボヌクレアーゼ(RNase)消化によって、あるいは融解温度の差によってミスマッチした二重鎖から区別することができる。DNA配列の違いはまた、変性剤の存在下または非存在下でのゲルにおけるDNAフラグメントの電気泳動移動度の変化によって、あるいは直接的なDNA配列決定によって検出することができる(例えば、Myers他、Science、(1985)230:1242)。特定の位置における配列の変化もまた、RNase保護またはS1保護などのヌクレアーゼ保護アッセイあるいは化学的な切断方法によって明らかにすることができる(Cotton他、Proc. Natl. Acad. Sci. USA、(1985)85:4397~4401)。

10

20

【0033】

hupolybromo1ポリヌクレオチド配列またはそのフラグメントを含むオリゴヌクレオチド・プローブのアレイを構築して、例えば、遺伝子変異の効率的なスクリーニングを行うことができる。かかるアレイは、好ましくは、高密度のアレイまたは格子状である。アレイ技術法は周知であり、一般的な適用性を有し、遺伝子発現、遺伝子連鎖および遺伝子変動性を含む分子遺伝学における様々な問題を検討するために使用することができる(例えば、M. Chee他、Science、274、610~613(1996))およびそれに引用されている他の参考文献を参照)。

【0034】

異常に、低下または増大しているポリペプチドまたはmRNAの発現レベルの検出もまた、本発明の疾患に対する被験体の感受性を診断または決定するために使用することができる。発現の減少または増加は、ポリヌクレオチドを定量することに関して当該分野で周知の方法、例えば、核酸増幅、例えば、PCR、RT-PCR、RNase保護、ノーザンブロッティングおよび他のハイブリダイゼーション法のいずれかを使用してRNAレベルで測定することができる。宿主に由来するサンプルにおける本発明のポリペプチドなどのタンパク質のレベルを決定するために使用することができるアッセイ手法は当業者には周知である。そのようなアッセイ方法には、放射免疫アッセイ、競合的結合アッセイ、ウエスタンブロット分析およびELISAアッセイが含まれる。

30

【0035】

従って、別の態様において本発明は下記を含む診断キットに関する：

- (a) 本発明のポリヌクレオチド、好ましくは、配列番号1のヌクレオチド配列またはそのフラグメントもしくはRNA転写物；
- (b) (a)のヌクレオチド配列に対して相補的なヌクレオチド配列、
- (c) 本発明のポリペプチド、好ましくは、配列番号2のポリペプチドまたはそのフラグメント、あるいは
- (d) 本発明のポリペプチドに対する抗体、好ましくは、配列番号2のポリペプチドに対する抗体。

40

【0036】

任意のかかるキットにおいて、(a)、(b)、(c)または(d)は実質的な成分を構

50

成することができることは理解される。かかるキットは、疾患または疾患に対する感受性、中でも特に本発明の疾患を診断する際に有用である。

【0037】

本発明のポリヌクレオチド配列は染色体局在化研究に有益である。該配列は、個々のヒト染色体における特定位置に対して特異的に標的化され、かつその特定の位置とハイブリダイゼーションし得る。本発明に従って関連する配列を染色体にマッピングすることは、そのような配列を遺伝子関連疾患と関連させる際の重要な最初の過程である。配列が正確な染色体位置にマッピングされると、染色体上における配列の物理的な位置を遺伝地図データと関連させることができる。かかるデータは、例えば、V. McKusickの「ヒトにおけるメンデル遺伝」において見出される（これはJohns Hopkins大学Welch Medical Libraryからオンラインで入手可能）。遺伝子と、同じ染色体領域にマッピングされている疾患との関係が、その後、連鎖解析（物理的に隣接する遺伝子の同時遺伝）によって同定される。ゲノム配列（遺伝子フラグメントなど）に関する正確なヒト染色体局在化を放射ハイブリッド（RH）マッピングを使用して決定することができる（Walter, M., Spillet, D., Thomas, P., Weissenbach, J. および Goodfellow, P. (1994)、ゲノム全体の放射ハイブリッドマップを構築する方法、Nature Genetics、7、22～28）。多数のRHパネルをResearch Genetics（Huntsville, AL、米国）から得ることができる。例えば、GeneBridge 4 RHパネル（Hum Mol Genet、1996、Mar; 5(3): 339～46、ヒトゲノムの放射ハイブリッドマップ。Gyapay G., Schmitt K., Fitzgerald C., Jones H., Vega-Czarny N., Spillet D., Musset D., Prud'Homme JF, Dib C., Auffray C., Morissette J., Weissenbach, J., Goodfellow, P.N.）。このパネルを使用して遺伝子の染色体上の位置を決定するために、93回のPCRが、RH DNAについて目的とする遺伝子から設計されたプライマーを用いて行われる。これらのDNAはそれぞれが、ハムスターのバックグラウンド（ヒト/ハムスターのハイブリッド細胞株）中に維持されたランダムなヒト・ゲノム・フラグメントを含有する。これらのPCRにより、目的とする遺伝子のPCR産物の存在の有無を示す93のスコアが得られる。これらのスコアを既知の位置のゲノム配列に由来するPCR産物を使用して作製されたスコアと比較される。この比較は、<http://www.genome.wi.mit.edu/>において行われる。

10

20

30

【0038】

本発明のポリヌクレオチド配列はまた、組織発現研究に対する有益なツールである。かかる研究は、それらをコードするmRNAを検出することによって、コードされたポリヌクレオチドの組織内の発現パターンに関する指標を与える、本発明のポリヌクレオチドの発現パターンの決定を可能とする。使用される技術は、当該分野では周知であり、cDNAマイクロアレイハイブリダイゼーション（Schenar他、Science、270、467～470、1995およびShalton他、Genome Res、6、639～645、1996）などの、格子上に配置された様々なクローンに対するインサイチュ（*in situ*）・ハイブリダイゼーション技術、およびPCRなどのヌクレオチド増幅技術を含む。好ましい方法では、Perkin Elmerから得られるTAQMAN（商標）法が使用される。これらの研究結果により、生物におけるポリペプチドの正常な機能の指標を得ることができる。さらに、mRNAの正常な発現パターンと、同じ遺伝子の別の形態（例えば、潜在的または調節的な差異をコードするポリペプチドにおける変化を有する形態）によってコードされるmRNAの発現パターンとの比較研究により、本発明のポリペプチドの役割に対する有益な洞察がもたらされ得るか、または疾患におけるその不適切な発現の役割に対する有益な洞察がもたらされ得る。そのような不適切な発現は、時間的、空間的または単に量的な性質であり得る。

40

【0039】

50

本発明のポリペプチドは、T細胞、B細胞、脾臓、リンパ節、胎盤、胚中心、子宮、頸部、乳房、精巣、肺、胃、脳、心臓、網膜、腎臓、メラノサイト、腸において発現している。

【0040】

本発明のさらなる態様は抗体に関する。本発明のポリペプチドまたはそのフラグメントあるいはそれらを発現する細胞は、本発明のポリペプチドに対して免疫特異的である抗体を産生させるための免疫原として使用することができる。用語「免疫特異的」は、抗体が、先行技術における他の関連するポリペプチドに対するその親和性よりも実質的に大きい親和性を本発明のポリペプチドに対して有することを意味する。

【0041】

本発明のポリペプチドに対して生じる抗体は、慣用的なプロトコルを使用して、該ポリペプチドまたはエピトープ保持したフラグメントまたは細胞を動物、好ましくは非ヒト動物に投与することによって得ることができる。モノクローナル抗体を調製する場合、継代的な細胞株培養物によって産生される抗体を提供するいかなる技術を使用することができる。例としては、ハイブリドーマ技術（Kohler, G. および Milstein, C., Nature (1975), 256: 495 ~ 497）、トリオーマ技術、ヒトB細胞ハイブリドーマ技術（Kozbor 他、Immunology Today (1983), 4: 72）、およびEBVハイブリドーマ技術（Cole 他、Monoclonal Antibodies and Cancer Therapy, 77 ~ 96, Alan R. Liss, Inc., 1985）が含まれる。

【0042】

米国特許第4,946,778号に記載される技術などの一本鎖抗体の製造技術もまた、本発明のポリペプチドに対する一本鎖抗体を製造するために適応させることができる。また、トランスジェニックマウスまたは他の哺乳動物を含む他の生物を使用して、ヒト化抗体を発現させることができる。

【0043】

上記の抗体を用いて、該ポリペプチドを発現するクローンを単離または同定する、もしくはアフィニティー・クロマトグラフィによって該ポリペプチドを精製することができる。本発明のポリペプチドに対する抗体はまた、特に本発明の疾患を処置するために用いることができる。

【0044】

本発明のポリペプチドおよびポリヌクレオチドはまたワクチンとして使用することができる。従って、さらなる態様において、本発明は、哺乳動物における免疫学的応答を誘導するための方法に関する。この方法は、前記動物を疾患から、その疾患が個体において既に慢性化しているか否かに関わらず保護するために抗体応答および/またはT細胞免疫応答（例えば、サイトカイン産生T細胞または細胞傷害性T細胞を含む）を生じさせるのに適切な本発明のポリペプチドを哺乳動物に接種することを含む。哺乳動物における免疫学的応答はまた、本発明にかかる疾患から前記動物を保護する抗体を産生させるようにかかる免疫学的応答を誘導するために、該ポリヌクレオチドの発現を支配し、かつ該ポリペプチドをコードするベクターによって本発明のポリペプチドをインビボで送達することを含む方法によって誘導され得る。そのようなベクターを投与する1つの方法は、粒子またはそれ以外のものの上へのコーティング物として所望する細胞中へのそれを促進することによる。かかる核酸ベクターは、DNA、RNA、修飾型核酸またはDNA/RNAハイブリッドを含むことができる。用途によって、ワクチン、ポリペプチドまたは核酸ベクターは、通常、ワクチン配合物（組成物）として提供される。該配合物は好適なキャリアをさらに含むことができる。ポリペプチドは胃で分解されることもあるため、それは、好ましくは非経口で投与される（例えば、皮下、筋肉内、静脈内または皮内注射）。非経口投与に好適な配合物には、抗酸化剤、緩衝剤、静菌剤、および配合物を接種者の血液と等張性にする溶質を含んでもよい、水性および非水性の無菌注射液；ならびに懸濁剤または増粘剤を含んでもよい、水性および非水性の無菌懸濁物が含まれる。該配合物は、単位用量容器

10

20

30

40

50

または多回用量容器で、例えば密封アンプルおよびバイアルに入れて提供することができ、そして使用直前に無菌の液体キャリアを添加することだけでよい、凍結乾燥状態で保存することができる。該ワクチン配合物はまた、水中油型システムおよび当該分野で既知のその他のシステムなどの、配合物の免疫原性を増強するアジュバント・システムを含むことができる。投与量は該ワクチンの比活性に依存するが、型どおりの実験によって容易に決定することができる。

【0045】

本発明のポリペプチドは、1つまたは複数の疾患状態、特に、前述の本発明の疾患に関連する1つまたは複数の生物学的機能を有する。従って、該ポリペプチドの機能または濃度を刺激または阻害する化合物を同定することは有用である。従って、さらなる態様において、本発明は、該ポリペプチドの機能または濃度を刺激または阻害する化合物を同定するために化合物をスクリーニングする方法を提供する。かかる方法により、本明細書前記のような本発明のそのような疾患に対する治療目的および予防目的のために用いることができるアゴニストまたはアンタゴニストが同定される。化合物は、様々な供給源、例えば、細胞、無細胞調製物、化学ライブラリ、化学化合物のコレクション、および天然産物の混合物から同定することができる。このように同定される、かかるアゴニストまたはアンタゴニストは、場合によっては、ポリペプチド自体の、天然または修飾された基質、リガンド、受容体、酵素など；その構造的または機能的な模倣体（Coligan他、Current Protocols in Immunology、1(2)：5章(1991)参照）あるいは小分子であり得る。

10

20

【0046】

該スクリーニング方法は、候補化合物に直接的または間接的に結合されている標識を利用して、該ポリペプチド、あるいは該ポリペプチドまたはその融合タンパク質を表出している細胞または膜に対する候補化合物の結合を単に測定することでもよい。あるいは、スクリーニング方法は、標識された競合剤（例えばアゴニストまたはアンタゴニスト）に対抗して候補化合物がポリペプチドに競合的に結合することを（定性的または定量的に）測定または検出することを含み得る。さらにこれらのスクリーニング方法では、該ポリペプチドを表出している細胞に適切な検出システムを使用して候補化合物がポリペプチドの活性化または阻害によって誘起されるシグナルをもたらすか否かを試験することもできる。一般には、活性化阻害剤が既知のアゴニストの存在下でアッセイされ、そして候補化合物の存在下でのアゴニストによる活性化に対する作用が観測される。さらに、該スクリーニング方法は、候補化合物を、本発明のポリペプチドを含有する溶液と混合して、混合物を形成させる工程と、混合物におけるhupolybromo1活性を測定する工程と、および混合物のhupolybromo1活性を、候補化合物を含有しないコントロール混合物と比較する工程とを単に含み得る。

30

【0047】

本発明のポリペプチドは、従来の低い容量のスクリーニング方法において、そしてまた高処理能（ハイ・スループット）スクリーニング（HTS）形式において用いることができる。かかるHTS形式には、十分に確立された96穴マイクロタイター（microtiter）プレートおよびより最近には384穴マイクロタイター（microtiter）プレートの使用だけでなく、Schullek他、Anal Biochem、246、20～29（1997）に記載のナノウエル法などの最近現れた方法もまた含まれる。

40

【0048】

Fc部分およびhupolybromo1ポリペプチドから生成される融合タンパク質などの融合タンパク質もまた、本明細書中前記のように、本発明のポリペプチドに対するアンタゴニストを同定するための高処理能スクリーニングアッセイに使用することができる（D. Bennett他、J Mol Recognition、8：52～58（1995）；K. Johanson他、J Biol Chem、270（16）：9459～9471（1995））。

【0049】

50

スクリーニング技術

本発明のポリヌクレオチド、ポリペプチドおよび該ポリペプチドに対する抗体はまた、細胞内の mRNA およびポリペプチドの産生に対する添加された化合物の作用を検出するためのスクリーニング方法を形成するために使用することができる。例えば、当該分野で公知の標準的な方法によってモノクローナル抗体およびポリクローナル抗体を使用してポリペプチドの分泌または細胞結合濃度を測定するために、ELISA アッセイを構築することができる。これは、好適に操作された細胞または組織からのポリペプチドの産生を阻害または増強し得る薬剤（これらはそれぞれアンタゴニストまたはアゴニストとも呼ばれる）を発見するために使用することができる。

【0050】

本発明のポリペプチドを用いて、当該分野で公知の標準的な受容体結合技術によって、存在する場合には膜結合型または可溶性の受容体を同定することができる。これらには、これらに限定されないものの、ポリペプチドを放射性同位体（例えば¹²⁵I）で標識、化学修飾（例えばビオチン化され）、あるいは検出または精製に好適なペプチド配列に融合して、そして推定される受容体の供給源（細胞、細胞膜、細胞上清、組織抽出物、体液）とインキュベーションするリガンド結合およびクロスリンク・アッセイが含まれる。他の方法には、表面プラズモン共鳴および分光測定法などの生物物理学的技術が含まれる。これらのスクリーニング方法はまた、それらが存在する場合にはポリペプチドのその受容体に対する結合と競合する該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使用することができる。かかるアッセイを行うための標準的な方法は当該分野では十分に理解されている。

【0051】

本発明のポリペプチドのアンタゴニストの例には、抗体、あるいは、ある場合には、該ポリペプチド自体のリガンド、基質、受容体、酵素などに密接に関連するオリゴヌクレオチドまたはタンパク質、場合により、例えば、該リガンド、基質、受容体、酵素などのフラグメント；あるいは本発明のポリペプチドに結合するものの、応答を誘発せず、その結果、ポリペプチドの活性が妨げられる、小分子が含まれる。

【0052】

スクリーニング方法はまた、トランスジェニック技術および *hupolybromo1* 遺伝子の使用を含んでもよい。トランスジェニック動物を構築する技術は十分に確立されている。例えば、受精した卵母細胞の雄性前核へのマイクロインジェクション、移植前または移植後の胚へのレトロウイルス移入、エレクトロポレーションなどにより遺伝子操作された胚性幹細胞の宿主胚盤胞への注入によって *hupolybromo1* 遺伝子を導入することができる。特に有用なトランスジェニック動物は、いわゆる「ノック・イン」動物であり、動物の遺伝子がその動物のゲノム内においてヒトの等価体によって置き換えられている。ノック・イン・トランスジェニック動物は、創薬プロセスにおいて、化合物がヒトの標的に対して特異的であるという、標的の妥当性検証用に有用である。他の有用なトランスジェニック動物は、いわゆる「ノック・アウト」動物であり、内因性の DNA 配列によってコードされている、本発明のポリペプチドに対する動物オルソログの発現が細胞内で部分的または完全に無効にされている。遺伝子のノック・アウトは、技術の限界の結果として、特異的な細胞または組織を対象とする、特定の細胞または組織においてのみ生じてよく、あるいは動物内のすべての細胞または実質的にすべての細胞において生じ得る。トランスジェニック動物の手法はまた、導入遺伝子が発現されて、大量の本発明のポリペプチドが得られる、動物全体の発現・クローニング・システムを提供する。

【0053】

上記の方法において使用されるスクリーニングキットは、本発明のさらなる態様をなす。かかるスクリーニングキットは下記のものを含む：

- (a) 本発明のポリペプチド、
- (b) 本発明のポリペプチドを発現する組換え細胞、
- (c) 本発明のポリペプチドを発現する細胞膜、または

10

20

30

40

50

(d) 本発明のポリペプチドに対する抗体、
そのようなポリペプチドは、好ましくは配列番号 2 のポリペプチドである。

【 0 0 5 4 】

かかるキット何れにおいても、(a)、(b)、(c) または (d) は実質的な構成要素を構成することは理解される。

【 0 0 5 5 】

(用語集)

下記の定義は、本明細書中で既に頻繁に使用されているいくつかの用語の理解を容易にするために提供される。

【 0 0 5 6 】

本明細書中に用いられる「抗体」は、ポリクローナルおよびモノクローナル抗体、キメラ、一本鎖、ならびにヒト化抗体、同様に F a b フラグメントをも包含し、F a b 発現ライブラリまたは他の免疫グロブリン発現ライブラリの生成物を包含する。

【 0 0 5 7 】

「単離(された)」は、その自然の状態から「ヒトの手によって」変化していること、すなわち、自然界に存在する場合、その本来の環境から変化、または取り出されている、もしくはその両方であることを意味する。例えば、生きた生物中に天然に存在するポリヌクレオチドまたはポリペプチドは「単離」されていないが、その天然状態の共存物質から分離された同じポリヌクレオチドまたはポリペプチドはこの用語が本明細書中で用法に従うと、「単離」されている。さらに、形質転換、遺伝子操作、または何らかの他の組換え方法によって生物に導入されているポリヌクレオチドまたはポリペプチドは、その生物が生存または非生存かのいずれでも、前記生物内に依然として存在している場合でさえ、「単離」されている。

【 0 0 5 8 】

「ポリヌクレオチド」は、一般には、任意のポリリボヌクレオチド(RNA)またはポリデオキシリボヌクレオチド(DNA)をいうが、非改変型または改変型のRNAまたはDNAであってもよい。「ポリヌクレオチド」には、一本鎖DNAおよび二本鎖DNA、一本鎖領域および二本鎖領域の混合であるDNA、一本鎖RNAおよび二本鎖RNA、ならびに一本鎖領域および二本鎖領域の混合であるRNA、一本鎖またはより典型的には二本鎖であり得るか、または一本鎖領域および二本鎖領域の混合であり得るDNAおよびRNAを含むハイブリッド分子が含まれるが、これらに限定されることはない。さらに、「ポリヌクレオチド」は、RNAまたはDNAあるいはRNAとDNAとの両方を含む三重鎖領域をもいう。用語「ポリヌクレオチド」はまた、1つまたは複数の修飾された塩基を含有するDNAまたはRNA、および安定性またはその他の理由のために修飾された骨格を有するDNAまたはRNAを含む。「修飾(された)」塩基には、例えば、トリチル化された塩基、およびイノシンなどの非通常型の塩基が含まれる。様々な修飾をDNAおよびRNAに対して行うことができる。従って、「ポリヌクレオチド」は、自然界に典型的に見出されるようなポリヌクレオチドの化学的、酵素的または代謝的に修飾された形態、ならびにウイルスおよび細胞に特徴的なDNAおよびRNAの化学的形態を包含する。「ポリヌクレオチド」はまた、しばしばオリゴヌクレオチドと称される比較的短いポリヌクレオチドを包含する。

【 0 0 5 9 】

「ポリペプチド」は、ペプチド結合または修飾されたペプチド結合(すなわちペプチド等配電子体)によって互いに連結された2つ以上のアミノ酸を含むポリペプチドの何れをもいう。「ポリペプチド」は、ペプチド、オリゴペプチドまたはオリゴマーと広く呼ばれる短い鎖、ならびに一般にはタンパク質と呼ばれるより長い鎖の両方をいう。ポリペプチドは遺伝子によってコードされる20のアミノ酸とは異なるアミノ酸を含有することができる。「ポリペプチド」は、翻訳後プロセッシングなどの天然のプロセス、または当該分野で周知の化学的な修飾方法のいずれかによって修飾されたアミノ酸配列を含む。かかる修飾は、基本的な教本、より詳細な専門書ならびに数多くの研究文献に広く記載されている。

10

20

30

40

50

修飾は、ペプチド骨格、アミノ酸側鎖およびアミノ末端またはカルボキシル末端を含むポリペプチド内の任意のところに存在させることができる。同じ型の修飾が所与ポリペプチド内のいくつかの部位に同じ程度または異なる程度、存在してもよいことが理解される。また、所与ポリペプチドは多くのタイプの修飾を含有することができる。ポリペプチドは、ユビキチン化の結果として分枝が成されてもよく、そして分枝型または非分枝型の環状であり得る。環状、分枝状および分枝した環状ポリペプチドは、翻訳に続く天然のプロセスに起因しても、あるいは合成的方法によって作製されてもよい。修飾には、アセチル化、アシル化、ADP-リボシル化、アミド化、ビオチン化、フラビンの共有結合、ヘム成分の共有結合、ヌクレオチドまたはヌクレオチド誘導体の共有結合、脂質または脂質誘導体の共有結合、ホスホチジルイノシトールの共有結合、クロス・リンク形成、環化、ジスルフィド結合の形成、脱メチル化、共有結合架橋の形成、システインの形成、ピログルタミン酸の形成、ホルミル化、 γ -カルボキシル化、グリコシル化、GPI・アンカーの形成、ヒドロキシル化、ヨウ素化、メチル化、ミリスチル化、酸化、タンパク質分解プロセッシング、リン酸化、プレニル化、ラセミ化、セレノイル化、硫酸化、アルギニル化などのタンパク質へのアミノ酸のトランスファー・RNA媒介付加、ならびユビキチン化が含まれる（例えば、Proteins - Structure and Molecular Properties、第2版、T. E. Creighton、W. H. Freeman and Company、New York、1993；Wold, F.、翻訳後のタンパク質修飾：全体像および展望、1~12、Post-translational Covalent Modification of Proteins、B. C. Johnson編、Academic Press、New York、1983；Seifter他、「タンパク質修飾および非タンパク質補助因子の分析」、Meth Enzymol、182、626~646、1990；Rattan他、「タンパク質合成：翻訳後修飾およびエージング」、Ann NY Acad Sci、663、48~62、1992）。

10

20

30

40

50

【0060】

ポリペプチド配列の「フラグメント」は、基準配列よりも短いが基準ポリペプチドとなるポリペプチドと同じ生物学的な機能または活性を本質的に保持しているポリペプチド配列をいう。ポリヌクレオチド配列の「フラグメント」は、配列番号1の基準配列よりも短いポリヌクレオチド配列をいう。

【0061】

「変異体」は、基準のポリヌクレオチドまたはポリペプチドとは異なるが、その本質的な性質を保持しているポリヌクレオチドまたはポリペプチドをいう。ポリヌクレオチドの典型的な変異体は、基準ポリヌクレオチドに対して塩基配列に相違がある。変異体の塩基配列における変化により、基準ポリヌクレオチドによってコードされるポリペプチドのアミノ酸配列が変化してもよく、あるいは変化しなくてもよい。ヌクレオチドの変化は、後述するとおり、基準配列によってコードされるポリペプチドにおけるアミノ酸の置換、付加、欠失、融合および短縮化をもたらし得る。ポリペプチドの典型的な変異体は、アミノ酸配列が基準ポリペプチドとは異なる。一般に、改変は基準ポリペプチドおよび変異体の配列が全体的に非常に類似し、そして多くの領域において同一であるように制限される。変異体ポリペプチドおよび基準ポリペプチドは、アミノ酸配列が、1つまたは複数の置換、挿入、欠失の任意の組合せによって異なりうる。置換または挿入されるアミノ酸残基は、遺伝子コードによってコードされるアミノ酸残基であっても、あるいはなくてもよい。典型的な保存的置換には、Gly、Ala；Val、Ile、Leu；Asp、Glu；Asn、Gln；Ser、Thr；Lys、Arg；ならびにPheおよびTyrが含まれる。ポリヌクレオチドまたはポリペプチドの変異体は、対立遺伝子などの天然に存在するものであってもよく、あるいは天然に存在することが知られていない変異体であってもよい。ポリヌクレオチドおよびポリペプチドの天然に存在しない変異体は、変異誘発方法によって、あるいは直接的な合成によって作製することができる。また、1つ以上の翻訳後修飾、例えば、グリコシル化、リン酸化、メチル化、ADPリボシル化などを有するポリ

ペプチドもまた変異体として含まれる。実施形態には、N末端アミノ酸のメチル化、セリンおよびトレオニンのリン酸化、ならびにC末端グリシンの修飾が含まれる。

【0062】

「対立遺伝子」は、ゲノム内の所与遺伝子座に存在する遺伝子の2つ以上の選択的な形態の1つをいう。

【0063】

「多型」は、集団内のゲノムにおける所与の位置での塩基配列（関連する場合にはコードされるポリペプチド配列）の変動をいう。

【0064】

「単一塩基多型」(SNP)は、集団内においてゲノム内の1つの塩基位置における塩基変動の発生をいう。SNPは、遺伝子内に、またはゲノムの遺伝子間領域内に存在し得る。SNPは、対立遺伝子特異的増幅(ASA)を使用してアッセイすることができる。該プロセスには少なくとも3つのプライマーが必要である。共通プライマーがアッセイされる多型に対して逆方向に相補となるように使用される。この共通プライマーは多型の塩基から50bpから1500bpの間で隔たったものとできる。それ以外の2つ（またはそれ以上）のプライマーは、最後の3'塩基が、多型を構成する2つ（またはそれ以上）の対立遺伝子の1つと一致するように変化している点を除いて互いに同一である。そして、それぞれ、共通プライマーおよび1つの対立遺伝子特異的プライマーを使用して、2つ（またはそれ以上）のPCR反応を、サンプルDNAについて行う。

10

【0065】

本発明において用いられる「スプライス変異体」は、同じゲノムDNA配列から一旦転写されたが、但し、択一的なRNAスプライシングを受けたRNA分子から産生したcDNA分子をいう。択一的なRNAスプライシングは、一般にはイントロンを除くために一次RNA転写物がスプライシングを受けているときに生じる。その結果、それぞれが異なるアミノ酸配列をコードし得る2つ以上のmRNA分子が生じる。スプライス変異体の用語はまた、上記のcDNA分子によってコードされるタンパク質をも示す。

20

【0066】

「同一性」は、その配列を比較することによって決定される、2つ以上のポリペプチド配列または2つ以上のポリヌクレオチド配列の間における相互関係を反映する。一般に、同一性とは、比較されている配列の長さに対して2つのポリヌクレオチド配列または2つのポリペプチド配列のそれぞれの塩基毎またはアミノ酸毎の厳密な一致をいう。

30

【0067】

「%同一性」- 正確な一致が存在しない配列については、「%同一性」を決定することができる。一般に、対比すべき2つの配列を、最大の相関が配列間に得られるようにアラインメントされる。これには、アラインメントの程度を高めるために「ギャップ」をいずれか一方または両方の配列に挿入することが含まれ得る。%同一性は、比較されている配列のそれぞれの長さ全体にわたって決定することができる（いわゆる全体的アラインメント）：これは同じ長さまたは非常に類似する長さの配列の場合には特に好適である；あるいはより短い規定された長さに対して決定することができる（いわゆる局所的アラインメント）：これは長さが等しくない配列の場合にはより好適である。

40

【0068】

「類似性」は、2つのポリペプチド配列の間における関係に対する、より精巧な、さらなる尺度である。一般に「類似性」は、残基毎に基づき、（同一性に関してと同様に）比較されている配列のそれぞれからの、相互に対を成す残基間における正確な一致だけでなく、正確な一致が存在しない場合にも、進化的な基準に基づいて、1つの残基は、他方に対する適当な置換であるかどうかをも考慮する、2つのポリペプチド鎖のアミノ酸間での比較を意味する。この蓋然性は付随した「スコア」を有し、2つの配列の「%類似性」はそれに基づき決定することができる。

【0069】

2つ以上の配列の同一性および類似性を比較する方法は当該分野では周知となっている。

50

従って、例えば、ウイスコンシン配列分析パッケージ（バージョン 9.1；D e v e r e u x J. 他、N u c l e i c A c i d s R e s . 12、387～395、1984；G e n e t i c s C o m p u t e r G r o u p (M a d i s o n、W i s c o n s i n、米国) から入手可能) において利用可能なプログラム、例えば、B E S T F I T および G A P プログラムを使用して、2つのポリヌクレオチド間の%同一性ならびに2つのポリペプチド配列間の%同一性および%類似性を決定することができる。B E S T F I Tでは、S m i t h および W a t e r m a n の「局所的相同性」アルゴリズム (J M o l B i o l、147、195～197、1981、A d v a n c e s i n A p p l i e d M a t h e m a t i c s、2、482～489、1981) を用いて、2つの配列間における類似性の最も良い1つの領域が見出される。B E S T F I Tは、長さが類似していない2つのポリヌクレオチド配列または2つのポリペプチド配列を比較することに対してより適している。該プログラムでは、より短い配列がより長い配列の一部を表すことが仮定されている。一方、G A Pは、N e d d l e m a n および W u n s c h のアルゴリズム (J M o l B i o l、48、443～453、1970) に従って2つの配列をアラインメントし、これにより「最大の類似性」を見出す。G A Pは、ほぼ同じ長さであり、また、アラインメントが長さ全体にわたって期待される配列を比較することに対してより適している。好ましくは、各プログラムにおいて使用される「ギャップ加重」および「長さ加重」のパラメーターは、それぞれ、ポリヌクレオチド配列の場合には50および3であり、ポリペプチド配列の場合には12および4である。好ましくは、比較されている2つの配列が最適にアラインメントした上で%同一性および類似性を決定する。

10

20

【0070】

配列間の同一性および/または類似性を決定するための他のプログラムもまた当該分野では知られている：例えば、B L A S Tファミリーのプログラム (A l t s c h u l S F 他、J M o l B i o l、215、403～410、1990；A l t s c h u l S F 他、N u c l e i c A c i d s R e s .、25：389～3402、1997；これはN a t i o n a l C e n t e r f o r B i o t e c h n o l o g y I n f o r m a t i o n (N C B I) (B e t h e s d a、M a r y l a n d、米国) から入手可能であり、www.ncbi.nlm.nih.govにおけるNCBIのホームページからアクセス可能である)、およびF A S T A (P e a r s o n W R、M e t h o d s i n E n z y m o l o g y、183、63～99、1990；P e a r s o n W R および L i p m a n D J、P r o c . N a t . A c a d . S c i . U S A、85、2444～2448、1988；これはウイスコンシン配列分析パッケージの一部として入手可能である)。

30

【0071】

好ましくは、B L O S U M 6 2 アミノ酸置換行列 (H e n i k o f f S および H e n i k o f f J G、P r o c . N a t . A c a d . S c i . U S A、89、10915～10919、1992) が、比較前にヌクレオチド配列がアミノ酸配列に最初に翻訳される場合をも含みポリペプチド配列比較において使用される。

【0072】

好ましくは、プログラムB E S T F I Tが、基準ポリヌクレオチド配列またはポリペプチド配列に関して、検討するポリヌクレオチドまたはポリペプチド配列の%同一性を決定するために使用される。この場合、前記したように、検討と基準とする配列とは、最適にアラインメントされ、そしてプログラムのパラメーターは暫定の値に設定されている。

40

【0073】

「同一性指標」は、候補配列 (ポリヌクレオチドまたはポリペプチド) と基準配列とを比較するために使用することができる、配列関連性の尺度である。したがって、例えば基準ポリヌクレオチド配列と比較したときに、例えば0.95の同一性指標を有する候補ポリヌクレオチド配列は、該候補ポリヌクレオチド配列が基準配列の各100塩基あたり平均して5個までの相違を含んでもよい点を除いて基準配列と同一である。かかる相違は、少なくとも1つの塩基の欠失、トランジションおよびトランスバージョンを含む置換、また

50

は挿入からなる群から選択される。これらの相違は、基準ポリヌクレオチド配列の5'または3'末端部位に、あるいはこれらの末端部位間における任意の位置で、基準配列内の塩基間に個々に、あるいは基準配列内に1つまたは複数の連続した群で点在して存在してもよい。すなわち、基準ポリヌクレオチド配列と比較したときに0.95の同一性指標を有するポリヌクレオチド配列を得るためには、既に記載のように、基準配列において塩基100個毎に平均して5個までが任意の組合せで欠失、置換または挿入されていてもよい。同じことが、同一性指標の他の値、例えば、0.96、0.97、0.98および0.99について必要に応じて変更して適用される。

【0074】

同様に、ポリペプチドの場合、基準ポリペプチド配列と比較したときに、例えば、0.95の同一性指標を有する候補ポリペプチド配列は、基準配列の各100アミノ酸あたり平均して5個までの相違をポリペプチド配列が含んでもよいことを除いて基準配列と同一である。かかる相違は、少なくとも1つのアミノ酸の欠失、保存的および非保存的な置換を含む置換、または挿入からなる群から選択される。これらの違いは、基準ポリペプチド配列のアミノ末端またはカルボキシ末端部位に、あるいはこれらの末端部位間の任意の位置に、基準配列内のアミノ酸間に個々に、あるいは基準配列内に1つまたは複数の連続した群中で点在して起きてもよい。すなわち、基準ポリペプチド配列と比較した際に0.95の同一性指標を有するポリペプチド配列を得るためには、既に記載のように、基準配列内のアミノ酸の100個毎に平均して5個まで任意の組合せで欠失、置換または挿入がなされてもよい。同じことが、同一性指標の他の値、例えば、0.96、0.97、0.98および0.99について必要に応じて変更して適用される。

【0075】

塩基またはアミノ酸の相違数と同一性指標との関係は下記の式で表すことができる：

$$n_a = x_a - (x_a \cdot I)$$

式中、

n_a は塩基またはアミノ酸の相違数であり、

x_a は配列番号1または配列番号2における塩基またはアミノ酸のそれぞれの総数であり、

I は同一性指標であり、

$\cdot$ は乗算演算子に対する記号であり、

この場合、 x_a と I との非整数の積は最も近い整数に切り捨てられ、その後 x_a からその値を引く。

【0076】

「ホモログ」は、基準配列に対する高度の配列関連性を有するポリヌクレオチド配列またはポリペプチド配列を示すために、当該分野で使用されている総称である。かかる関連性は、既に定義されているように2つの配列間の同一性および/または類似性の程度を決定することによって定量化することができる。この総称には、「オルソログ」および「パラログ」の用語が含まれる。「オルソログ」は、別の種中における該ポリヌクレオチドまたはポリペプチドの機能的等価体であるポリヌクレオチドまたはポリペプチドをいう。「パラログ」は、機能的に類似している同じ種内におけるポリヌクレオチドまたはポリペプチドをいう。

【0077】

「融合タンパク質」は、2つの関連しない融合された遺伝子またはそのフラグメントによってコードされるタンパク質をいう。その例が米国特許第5541087号、米国特許第5726044号に開示されている。Fc-hupolybromo1の場合、融合タンパク質の一部として免疫グロブリンのFc領域を用いることは、Fc-hupolybromo1またはhupolybromo1のフラグメントを機能的に発現させて、治療に使用する際にかかる融合タンパク質の薬物動態学的性質を改善するために、ならびに二量体のhupolybromo1を生成させるためには好都合である。Fc-hupolybromo1のDNA構築物は、5'から3'の方向で、分泌カセット、すなわち哺乳動物

10

20

30

40

50

細胞からの細胞外への輸送を引き起こすシグナル配列、融合パートナーとして免疫グロブリンのFc領域フラグメントをコードするDNA、およびhupolybromo1をコードするDNAまたはその断片を含む。ある用途では、融合タンパク質の残部には手を触れることなく、機能的なFc側を変異させ、その固有的な機能的性質（補体結合、Fc受容体結合）を変える、あるいは発現後にFc部分を完全に除くことを可能とすることが望ましい。

【0078】

特許および特許出願に限らず、これらを含む本明細書中で引用されている刊行物および参考文献の全ては、十分に述べているように、個々の刊行物または参考文献を、参照して組み込むために、明示的かつ個別的に示唆されているように、その全部を、参照することで、本明細書中に組み込まれる。本出願が優先権を主張する特許出願はいずれも、先に刊行物および参考文献に関して記載したと同様に、その全部を、参照することで、本明細書中に組み込まれる。

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
18 October 2001 (18.10.2001)

PCT

(10) International Publication Number
WO 01/77147 A2

- (51) International Patent Classification: **C07K 14/00** **WILBERT, Oliver** (DE/DE); Lichthausstrasse 54, 64289 Darmstadt (DE); **DÜCKER, Klaus** (DE/DE); Einselestrasse 5, 64291 Darmstadt (DE).
- (21) International Application Number: PCT/EP01/03793
- (22) International Filing Date: 4 April 2001 (04.04.2001) (74) Common Representative: **MERCK PATENT GMBH**, Frankfurter Strasse 250, 64293 Darmstadt (DE).
- (25) Filing Language: English
- (26) Publication Language: English (81) Designated States (*national*): CA, JP, US.
- (30) Priority Data: 00107352.7 5 April 2000 (05.04.2000) EP (84) Designated States (*regional*): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GR, GB, IE, IT, LU, MC, NL, PT, SE, TR).
- (71) Applicant (*for all designated States except US*): **MERCK PATENT GMBH** (DE/DE), Frankfurter Strasse 250, 64293 Darmstadt (DE). Published: *without international search report and to be republished upon receipt of that report*
- (72) Inventors; and (75) Inventors/Applicants (*for US only*): **BURGER, Christa** (DE/DE), Casanweg 23, 64289 Darmstadt (DE). *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 01/77147 A2

(54) Title: NEW BROMELONIN PROTEIN

(57) Abstract: Hypobromol polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing hypobromol polypeptides and polynucleotides in diagnostic assays.

WO 01/77147

PCT/EP01/03793

New Bromodomain Protein**Field of the Invention**

This invention relates to newly identified polypeptides and polynucleotides encoding such polypeptides sometimes hereinafter referred to as „hupolybromo1“, to their use in diagnosis and in identifying compounds that may be agonists, antagonists that are potentially useful in therapy, and to production of such polypeptides and polynucleotides.

Background of the Invention

The drug discovery process is currently undergoing a fundamental revolution as it embraces "functional genomics", that is, high throughput genome- or gene-based biology. This approach as a means to identify genes and gene products as therapeutic targets is rapidly superceding earlier approaches based on "positional cloning". A phenotype, that is a biological function or genetic disease, would be identified and this would then be tracked back to the responsible gene, based on its genetic map position.

Functional genomics relies heavily on high-throughput DNA sequencing technologies and the various tools of bioinformatics to identify gene sequences of potential interest from the many molecular biology databases now available. There is a continuing need to identify and characterise further genes and their related polypeptides/proteins, as targets for drug discovery.

Summary of the Invention

The present invention relates to hupolybromo1, in particular hupolybromo1 polypeptides and hupolybromo1 polynucleotides, recombinant materials and methods for their production. Such polypeptides and polynucleotides are of interest in relation to methods of treatment of certain diseases, including, but not limited to, Cancer in general, allergy,

CONFIRMATION COPY

WO 01/77147

PCT/EP01/03793

- 2 -

rheumatic arthritis, autoimmune diseases, stroke, dementia, mental retardation, parkinson, epilepsy, schizophrenia, alzheimer, depression, coronary heart disease, heart failure, myocarditis, and cystic fibrosis. , hereinafter referred to as "diseases of the invention". In a further aspect, the invention relates to methods for identifying agonists and antagonists (e.g., inhibitors) using the materials provided by the invention, and treating conditions associated with hupolybromo1 imbalance with the identified compounds. In a still further aspect, the invention relates to diagnostic assays for detecting diseases associated with inappropriate hupolybromo1 activity or levels.

Description of the Invention

In a first aspect, the present invention relates to hupolybromo1 polypeptides. Such polypeptides include:

(a) a polypeptide encoded by a polynucleotide comprising the sequence of SEQ ID NO:1;

(b) a polypeptide comprising a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO:2;

(c) a polypeptide comprising the polypeptide sequence of SEQ ID NO:2;

(d) a polypeptide having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO:2;

(e) the polypeptide sequence of SEQ ID NO:2; and

(f) a polypeptide having or comprising a polypeptide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to the polypeptide sequence of SEQ ID NO:2;

(g) fragments and variants of such polypeptides in (a) to (f).

Polypeptides of the present invention are believed to be members of the Polybromo domain protein family of polypeptides. They are therefore of interest because a number of transcription factors that act as adaptor proteins have been found to contain an amino acid domain called the

WO 01/77147

PCT/EP01/03793

- 3 -

bromodomain. Bromodomain proteins are members of the SWI/SNF complex which is a chromatin remodeling complex facilitating access of transcription factors to regulatory DNA sequences. The molecular cloning of polybromo containing multiple bromodomains, a truncated HMG-box, and two repeats has been reported by Nicolas, R.H. et al. (Gene. 1996, 175: 233-40). The protein complexes have different subunit compositions determined by the adaptor proteins. This complexes has recently been associated with the control of cell growth by regulating the maintenance of chromosome stability and various aspects of DNA repair (Liu-Y. et al., Nucleic-Acids-Res., 1998, 26: 1038-45) found that the mouse DNA methyltransferase (DNA MTase) is targeted to sites of DNA replication by a polybromo protein. Expression of the bromodomain protein brm (SNF2) is frequently down regulated upon cellular transformation. Hupolybromo 1 is homolog to the human SN2-4 protein and to Brm-drome. In addition it is homolog to a protein of Schizosaccharomyce which seems to be involved in remodelling of chromatin (GI 3169090). In mice, lacking the SNF2/SWI2 family member ETL1, skeletal dysplasias, growth retardation, reduced postnatal survival, and impaired fertility has been observed (Schoor, M. et al., Mech. Dev. 1999, 85: 73-83). Recently the functional aspects of growth control of the mammalian SWI/SNF have been summarized by Muchardt et al. (Semin. Cell Dev. Biol. 1999, 10: 189-95). The functional aspects described sofar are important for chromatin remodeling and differential regulation of transcription of genes thus they form part of a mechanism which is involved in the generation of cancer cells and metastasis. Furthermore these mechanism are as well involved in brain disorders, autoimmune diseases and some forms of brain diseases.

The biological properties of the hupolybromo1 are hereinafter referred to as "biological activity of hupolybromo1" or "hupolybromo1 activity". Preferably, a polypeptide of the present invention exhibits at least one biological activity of hupolybromo1.

Polypeptides of the present invention also includes variants of the aforementioned polypeptides, including all allelic forms and splice variants. Such polypeptides vary from the reference polypeptide by insertions, deletions, and substitutions that may be conservative or non-conservative, or any combination thereof. Particularly preferred variants are those in which several, for instance from 50 to 30, from 30 to 20, from 20 to 10, from

WO 01/77147

PCT/EP01/03793

- 4 -

10 to 5, from 5 to 3, from 3 to 2, from 2 to 1 or 1 amino acids are inserted, substituted, or deleted, in any combination.

Preferred fragments of polypeptides of the present invention include a polypeptide comprising an amino acid sequence having at least 30, 50 or 100 contiguous amino acids from the amino acid sequence of SEQ ID NO: 2, or a polypeptide comprising an amino acid sequence having at least 30, 50 or 100 contiguous amino acids truncated or deleted from the amino acid sequence of SEQ ID NO: 2. Preferred fragments are biologically active fragments that mediate the biological activity of hupolybromo1, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also preferred are those fragments that are antigenic or immunogenic in an animal, especially in a human.

Fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these variants may be employed as intermediates for producing the full-length polypeptides of the invention. The polypeptides of the present invention may be in the form of the "mature" protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence that contains secretory or leader sequences, pro-sequences, sequences that aid in purification, for instance multiple histidine residues, or an additional sequence for stability during recombinant production.

Polypeptides of the present invention can be prepared in any suitable manner, for instance by isolation from naturally occurring sources, from genetically engineered host cells comprising expression systems (*vide infra*) or by chemical synthesis, using for instance automated peptide synthesizers, or a combination of such methods. Means for preparing such polypeptides are well understood in the art.

In a further aspect, the present invention relates to hupolybromo1 polynucleotides. Such polynucleotides include:

WO 01/77147

PCT/EP01/03793

- 5 -

(a) a polynucleotide comprising a polynucleotide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polynucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the polynucleotide of SEQ ID NO:1;

5 (c) a polynucleotide having at least 95%, 96%, 97%, 98%, or 99% identity to the polynucleotide of SEQ ID NO:1;

(d) the polynucleotide of SEQ ID NO:1;

10 (e) a polynucleotide comprising a polynucleotide sequence encoding a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO:2;

(f) a polynucleotide comprising a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2;

15 (g) a polynucleotide having a polynucleotide sequence encoding a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO:2;

(h) a polynucleotide encoding the polypeptide of SEQ ID NO:2;

(i) a polynucleotide having or comprising a polynucleotide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to the polynucleotide sequence of SEQ ID NO:1;

20 (j) a polynucleotide having or comprising a polynucleotide sequence encoding a polypeptide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to the polypeptide sequence of SEQ ID NO:2; and

25 polynucleotides that are fragments and variants of the above mentioned polynucleotides or that are complementary to above mentioned polynucleotides, over the entire length thereof.

30 Preferred fragments of polynucleotides of the present invention include a polynucleotide comprising an nucleotide sequence having at least 15, 30, 50 or 100 contiguous nucleotides from the sequence of SEQ ID NO: 1, or a polynucleotide comprising an sequence having at least 30, 50 or 100

WO 01/77147

PCT/EP01/03793

- 6 -

contiguous nucleotides truncated or deleted from the sequence of SEQ ID NO: 1.

Preferred variants of polynucleotides of the present invention include splice variants, allelic variants, and polymorphisms, including polynucleotides having one or more single nucleotide polymorphisms (SNPs).

Polynucleotides of the present invention also include polynucleotides encoding polypeptide variants that comprise the amino acid sequence of SEQ ID NO:2 and in which several, for instance from 50 to 30, from 30 to 20, from 20 to 10, from 10 to 5, from 5 to 3, from 3 to 2, from 2 to 1 or 1 amino acid residues are substituted, deleted or added, in any combination.

In a further aspect, the present invention provides polynucleotides that are RNA transcripts of the DNA sequences of the present invention. Accordingly, there is provided an RNA polynucleotide that:

(a) comprises an RNA transcript of the DNA sequence encoding the polypeptide of SEQ ID NO:2;

(b) is the RNA transcript of the DNA sequence encoding the polypeptide of SEQ ID NO:2;

(c) comprises an RNA transcript of the DNA sequence of SEQ ID NO:1; or

(d) is the RNA transcript of the DNA sequence of SEQ ID NO:1;

and RNA polynucleotides that are complementary thereto.

The polynucleotide sequence of SEQ ID NO:1 shows greatest homology with gallus gallus polybromo protein X90849. It is as well homolog to C.elegans bromodomain protein (GI:1301625). The polynucleotide sequence of SEQ ID NO:1 is a cDNA sequence that encodes the polypeptide of SEQ ID NO:2. The polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may be identical to the polypeptide encoding sequence of SEQ ID NO:1 or it may be a sequence other than SEQ ID NO:1, which, as a result of the redundancy (degeneracy) of the genetic

WO 01/77147

PCT/EP01/03793

- 7 -

code, also encodes the polypeptide of SEQ ID NO:2. The polypeptide of the SEQ ID NO:2 is related to other proteins of the Polybromo domain protein family, having homology and/or structural similarity with * gallus gallus polybromo protein X90849, it is as well homolog to human protein SN2-4 (P51532) and Brm-drome (P25435) and the two unnamed proteins GI 7022804 and GI 7023493.

Preferred polypeptides and polynucleotides of the present invention are expected to have, *inter alia*, similar biological functions/properties to their homologous polypeptides and polynucleotides. Furthermore, preferred polypeptides and polynucleotides of the present invention have at least one hupolybromo1 activity.

Polynucleotides of the present invention may be obtained using standard cloning and screening techniques from a cDNA library derived from human mRNA isolated from mixed embryonic tissues, (see for instance, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)). Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

When polynucleotides of the present invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence for the mature polypeptide, by itself, or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, Proc Natl Acad Sci USA (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

WO 01/77147

PCT/EP01/03793

- 8 -

Polynucleotides that are identical, or have sufficient identity to a polynucleotide sequence of SEQ ID NO:1, may be used as hybridization probes for cDNA and genomic DNA or as primers for a nucleic acid amplification reaction (for instance, PCR). Such probes and primers may be used to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding paralogs from human sources and orthologs and paralogs from species other than human) that have a high sequence similarity to SEQ ID NO:1, typically at least 95% identity. Preferred probes and primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50, if not at least 100 nucleotides. Particularly preferred probes will have between 30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides.

A polynucleotide encoding a polypeptide of the present invention, including homologs from species other than human, may be obtained by a process comprising the steps of screening a library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO: 1 or a fragment thereof, preferably of at least 15 nucleotides; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to the skilled artisan. Preferred stringent hybridization conditions include overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.8), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA; followed by washing the filters in 0.1x SSC at about 65°C. Thus the present invention also includes isolated polynucleotides, preferably with a nucleotide sequence of at least 100, obtained by screening a library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:1 or a fragment thereof, preferably of at least 15 nucleotides.

The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide does not extend all the way through to the 5' terminus. This is a consequence of reverse transcriptase, an enzyme with inherently low "processivity" (a measure of the ability of the enzyme to remain attached

WO 01/77147

PCT/EP01/03793

- 9 -

to the template during the polymerisation reaction), failing to complete a DNA copy of the mRNA template during first strand cDNA synthesis.

There are several methods available and well known to those skilled in the art to obtain full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman et al., Proc Nat Acad Sci USA 85, 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon (trade mark) technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon (trade mark) technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using 'nested' primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the known gene sequence). The products of this reaction can then be analysed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

Recombinant polypeptides of the present invention may be prepared by processes well known in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems comprising a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

WO 01/77147

PCT/EP01/03793

- 10 -

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Polynucleotides may be introduced into host cells by methods described in many standard laboratory manuals, such as Davis et al., *Basic Methods in Molecular Biology* (1986) and Sambrook *et al. (ibid)*. Preferred methods of introducing polynucleotides into host cells include, for instance, calcium phosphate transfection, DEAE-dextran mediated transfection, transfection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as *Streptococci*, *Staphylococci*, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used, for instance, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector that is able to maintain, propagate or express a polynucleotide to produce a polypeptide in a host may be used. The appropriate polynucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al. (ibid)*. Appropriate secretion signals may be incorporated into the desired polypeptide to allow secretion of the translated protein into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular environment. These signals may be endogenous to the polypeptide or they may be heterologous signals.

WO 01/77147

PCT/EP01/03793

- 11 -

If a polypeptide of the present invention is to be expressed for use in screening assays, it is generally preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If the polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide. If produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and/or purification.

Polynucleotides of the present invention may be used as diagnostic reagents, through detecting mutations in the associated gene. Detection of a mutated form of the gene characterised by the polynucleotide of SEQ ID NO:1 in the cDNA or genomic sequence and 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 gene may be detected at the DNA level by a variety of techniques well known in the art.

Nucleic acids 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 it may be amplified enzymatically by using PCR, preferably RT-PCR, or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled hupolybromo1 nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures.

WO 01/77147

PCT/EP01/03793

- 12 -

DNA sequence difference 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 (see, for instance, 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 USA (1985) 85: 4397-4401).

An array of oligonucleotide probes comprising hupolybromo1 polynucleotide sequence or fragments thereof can be constructed to conduct efficient screening of *e.g.*, genetic mutations. Such arrays are preferably high density arrays or grids. 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, 274, 610-613 (1996) and other references cited therein.

Detection of abnormally decreased or increased levels of polypeptide or mRNA expression may also be used for diagnosing or determining susceptibility of a subject to a disease of the invention. 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. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagnostic kit comprising:

- (a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO: 1, or a fragment or an RNA transcript thereof;
- (b) a nucleotide sequence complementary to that of (a);
- (c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO:2 or a fragment thereof; or

WO 01/77147

PCT/EP01/03793

- 13 -

(d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO:2.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly diseases of the invention, amongst others.

The polynucleotide sequences of the present invention are valuable for chromosome localisation studies. The sequence 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 the present invention is an important first step in correlating those sequences with 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 University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (co-inheritance of physically adjacent genes). Precise human chromosomal localisations for a genomic sequence (gene fragment etc.) can be determined using Radiation Hybrid (RH) Mapping (Walter, M. Spillelt, D., Thomas, P., Weissenbach, J., and Goodfellow, P., (1994) A method for constructing radiation hybrid maps of whole genomes, Nature Genetics 7, 22-28). A number of RH panels are available from Research Genetics (Huntsville, AL, USA) e.g. the GeneBridge4 RH panel (Hum Mol Genet 1996 Mar;5(3):339-46 A radiation hybrid map of the human genome. Gyapay G, Schmitt K, Fizames C, Jones H, Vega-Czarny N, Spillelt D, Muselet D, Prud'Homme JF, Dib C, Auffray C, Morissette J, Weissenbach J, Goodfellow PN). To determine the chromosomal location of a gene using this panel, 93 PCRs are performed using primers designed from the gene of interest on RH DNAs. Each of these DNAs contains random human genomic fragments maintained in a hamster background (human / hamster hybrid cell lines). These PCRs result in 93 scores indicating the presence or absence of the PCR product of the gene of interest. These scores are compared

WO 01/77147

- 14 -

PCT/EP01/03793

with scores created using PCR products from genomic sequences of known location. This comparison is conducted at <http://www.genome.wi.mit.edu/>.

5 The polynucleotide sequences of the present invention are also valuable tools for tissue expression studies. Such studies allow the determination of expression patterns of polynucleotides of the present invention which may give an indication as to the expression patterns of the encoded polypeptides in tissues, by detecting the mRNAs that encode them. The techniques used are well known in the art and include in situ hybridisation techniques to clones arrayed on a grid, such as cDNA microarray hybridisation (Scherena *et al*, Science, 270, 467-470, 1995 and Shalon *et al*, Genome Res, 6, 639-645, 1996) and nucleotide amplification techniques such as PCR. A preferred method uses the TAQMAN (Trade mark) technology available from Perkin Elmer. Results from these studies can provide an indication of the normal function of the polypeptide in the organism. In addition, comparative studies of the normal expression pattern of mRNAs with that of mRNAs encoded by an alternative form of the same gene (for example, one having an alteration in polypeptide coding potential or a regulatory mutation) can provide valuable insights into the role of the polypeptides of the present invention, or that of inappropriate expression thereof in disease. Such inappropriate expression may be of a temporal, spatial or simply quantitative nature.

25 The polypeptides of the present invention are expressed in T cells, B cells, spleen, lymph node, placenta, germinal center, uterus, cervix, breast, testis, lung, stomach, brain, heart, retina, kidney, melanocytes, intestine.

30 A further aspect of the present invention relates to antibodies. The polypeptides of the invention or their fragments, or cells expressing them, can be used as immunogens to produce antibodies that are immunospecific for polypeptides of the present invention. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

WO 01/77147

PCT/EP01/03793

- 15 -

Antibodies generated against polypeptides of the present invention may be obtained by administering the polypeptides or epitope-bearing fragments, or cells to an animal, preferably a non-human animal, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, Immunology Today (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, Monoclonal Antibodies and Cancer Therapy, 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies, such as those described in U.S. Patent No. 4,946,778, can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms, including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography. Antibodies against polypeptides of the present invention may also be employed to treat diseases of the invention, amongst others.

Polypeptides and polynucleotides of the present invention may also be used as vaccines. Accordingly, in a further aspect, the present invention relates to a method for inducing an immunological response in a mammal that comprises inoculating the mammal with a polypeptide of the present invention, adequate to produce antibody and/or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said animal from disease, whether that disease is already established within the individual or not. An immunological response in a mammal may also be induced by a method comprises delivering a polypeptide of the present invention via a vector directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases of the invention. One way of administering the vector is by accelerating it into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may comprise DNA, RNA, a

modified nucleic acid, or a DNA/RNA hybrid. For use as a vaccine, a polypeptide or a nucleic acid vector will be normally provided as a vaccine formulation (composition). The formulation may further comprise a suitable carrier. Since a polypeptide may be broken down in the stomach, it is preferably administered parenterally (for instance, subcutaneous, intramuscular, intravenous, or intradermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions that may contain anti-oxidants, buffers, bacteriostats and solutes that render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions that may include suspending agents or thickening agents. The formulations 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 vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Polypeptides of the present invention have one or more biological functions that are of relevance in one or more disease states, in particular the diseases of the invention hereinbefore mentioned. It is therefore useful to identify compounds that stimulate or inhibit the function or level of the polypeptide. Accordingly, in a further aspect, the present invention provides for a method of screening compounds to identify those that stimulate or inhibit the function or level of the polypeptide. Such methods identify agonists or antagonists that may be employed for therapeutic and prophylactic purposes for such diseases of the invention as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, collections of chemical compounds, and natural product mixtures. Such agonists or antagonists so-identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of the polypeptide; a structural or functional mimetic thereof (see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991)) or a small molecule.

The screening method may simply measure the binding of a candidate compound to the polypeptide, or to cells or membranes bearing the polypeptide, or a fusion protein thereof, by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve measuring or detecting (qualitatively or quantitatively) the competitive binding of a candidate compound to the polypeptide against a labeled competitor (e.g. agonist or antagonist). Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells bearing the polypeptide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide of the present invention, to form a mixture, measuring a hupolybromo1 activity in the mixture, and comparing the hupolybromo1 activity of the mixture to a control mixture which contains no candidate compound.

Polypeptides of the present invention may be employed in conventional low capacity screening methods and also in high-throughput screening (HTS) formats. Such HTS formats include not only the well-established use of 96- and, more recently, 384-well microtiter plates but also emerging methods such as the nanowell method described by Schullek et al, Anal Biochem., 248, 20-29, (1997).

Fusion proteins, such as those made from Fc portion and hupolybromo1 polypeptide, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists for the polypeptide of the present invention (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).

Screening techniques

WO 01/77147

PCT/EP01/03793

- 18 -

The polynucleotides, polypeptides and antibodies to the polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents that may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

A polypeptide of the present invention may be used to identify membrane bound or soluble receptors, if any, through standard receptor binding techniques known in the art. These include, but are not limited to, ligand binding and crosslinking assays in which the polypeptide is labeled with a radioactive isotope (for instance, ^{125}I), chemically modified (for instance, biotinylated), or fused to a peptide sequence suitable for detection or purification, and incubated with a source of the putative receptor (cells, cell membranes, cell supernatants, tissue extracts, bodily fluids). Other methods include biophysical techniques such as surface plasmon resonance and spectroscopy. These screening methods may also be used to identify agonists and antagonists of the polypeptide that compete with the binding of the polypeptide to its receptors, if any. Standard methods for conducting such assays are well understood in the art.

Examples of antagonists of polypeptides of the present invention include antibodies or, in some cases, oligonucleotides or proteins that are closely related to the ligands, substrates, receptors, enzymes, etc., as the case may be, of the polypeptide, e.g., a fragment of the ligands, substrates, receptors, enzymes, etc.; or a small molecule that bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

Screening methods may also involve the use of transgenic technology and hupolybromo1 gene. The art of constructing transgenic animals is well established. For example, the hupolybromo1 gene may be introduced through microinjection into the male pronucleus of fertilized oocytes, retroviral transfer into pre- or post-implantation embryos, or injection of genetically modified, such as by electroporation, embryonic stem cells into host blastocysts. Particularly useful transgenic animals

WO 01/77147

- 19 -

PCT/EP01/03793

are so-called "knock-in" animals in which an animal gene is replaced by the human equivalent within the genome of that animal. Knock-in transgenic animals are useful in the drug discovery process, for target validation, where the compound is specific for the human target. Other useful transgenic animals are so-called "knock-out" animals in which the expression of the animal ortholog of a polypeptide of the present invention and encoded by an endogenous DNA sequence in a cell is partially or completely annulled. The gene knock-out may be targeted to specific cells or tissues, may occur only in certain cells or tissues as a consequence of the limitations of the technology, or may occur in all, or substantially all, cells in the animal. Transgenic animal technology also offers a whole animal expression-cloning system in which introduced genes are expressed to give large amounts of polypeptides of the present invention

Screening kits for use in the above described methods form a further aspect of the present invention. Such screening kits comprise:

- (a) a polypeptide of the present invention;
- (b) a recombinant cell expressing a polypeptide of the present invention;
- (c) a cell membrane expressing a polypeptide of the present invention; or
- (d) an antibody to a polypeptide of the present invention;

which polypeptide is preferably that of SEQ ID NO:2.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Glossary

The following definitions are provided to facilitate understanding of certain terms used frequently hereinbefore.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an

Fab or other immunoglobulin expression library.

"Isolated" means altered "by the hand of man" from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Moreover, a polynucleotide or polypeptide that is introduced into an organism by transformation, genetic manipulation or by any other recombinant method is "isolated" even if it is still present in said organism, which organism may be living or non-living.

"Polynucleotide" generally refers to any polyribonucleotide (RNA) or polydeoxynucleotide (DNA), which may be unmodified or modified RNA or DNA. "Polynucleotides" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term "polynucleotide" also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications may be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any polypeptide comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, *i.e.*, peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids.

"Polypeptides" include amino acid sequences modified either by natural processes, such as post-translational processing, or by chemical modification techniques that are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications may occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present to the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from post-translation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, biotinylation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination (see, for instance, *Proteins - Structure and Molecular Properties*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993; Wold, F., *Post-translational Protein Modifications: Perspectives and Prospects*, 1-12, in *Post-translational Covalent Modification of Proteins*, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol*, 182, 626-646, 1990, and Rattan *et al.*, "Protein Synthesis: Post-translational Modifications and Aging", *Ann NY Acad Sci*, 663, 48-62, 1992).

"Fragment" of a polypeptide sequence refers to a polypeptide sequence that is shorter than the reference sequence but that retains essentially the same biological function or activity as the reference polypeptide.

"Fragment" of a polynucleotide sequence refers to a polynucleotide sequence that is shorter than the reference sequence of SEQ ID NO:1..

"Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains the essential properties thereof. A typical variant of a polynucleotide differs in nucleotide sequence from the reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from the reference polypeptide. Generally, alterations are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, insertions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. Typical conservative substitutions include Gly, Ala, Val, Ile, Leu, Asp, Glu, Asn, Gln; Ser, Thr, Lys, Arg; and Phe and Tyr. A variant of a polynucleotide or polypeptide may be naturally occurring such as an allele, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis. Also included as variants are polypeptides having one or more post-translational modifications, for instance glycosylation, phosphorylation, methylation, ADP ribosylation and the like. Embodiments include methylation of the N-terminal amino acid, phosphorylations of serines and threonines and modification of C-terminal glycines.

"Allele" refers to one of two or more alternative forms of a gene occurring at a given locus in the genome.

"Polymorphism" refers to a variation in nucleotide sequence (and encoded polypeptide sequence, if relevant) at a given position in the genome within a population.

"Single Nucleotide Polymorphism" (SNP) refers to the occurrence of nucleotide variability at a single nucleotide position in the genome, within a population. An SNP may occur within a gene or within intergenic regions of the genome. SNPs can be assayed using Allele Specific Amplification (ASA). For the process at least 3 primers are required. A common primer is used in reverse complement to the polymorphism being assayed. This common primer can be between 50 and 1500 bps from the polymorphic base. The other two (or more) primers are identical to each other except that the final 3' base wobbles to match one of the two (or more) alleles that make up the polymorphism. Two (or more) PCR reactions are then conducted on sample DNA, each using the common primer and one of the Allele Specific Primers.

"Splice Variant" as used herein refers to cDNA molecules produced from RNA molecules initially transcribed from the same genomic DNA sequence but which have undergone alternative RNA splicing. Alternative RNA splicing occurs when a primary RNA transcript undergoes splicing, generally for the removal of introns, which results in the production of more than one mRNA molecule each of that may encode different amino acid sequences. The term splice variant also refers to the proteins encoded by the above cDNA molecules.

"Identity" reflects a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, determined by comparing the sequences. In general, identity refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of the two polynucleotide or two polypeptide sequences, respectively, over the length of the sequences being compared.

"% Identity" - For sequences where there is not an exact correspondence, a "% identity" may be determined. In general, the two sequences to be compared are aligned to give a maximum correlation between the sequences. This may include inserting "gaps" in either one or both sequences, to enhance the degree of alignment. A % identity may be determined over the whole length of each of the sequences being compared (so-called global alignment), that is particularly suitable for sequences of the same or very similar length, or over shorter, defined lengths (so-called local alignment), that is more suitable for sequences of unequal length.

WO 01/77147

PCT/EP01/03793

- 24 -

"Similarity" is a further, more sophisticated measure of the relationship between two polypeptide sequences. In general, "similarity" means a comparison between the amino acids of two polypeptide chains, on a residue by residue basis, taking into account not only exact correspondences between a between pairs of residues, one from each of the sequences being compared (as for identity) but also, where there is not an exact correspondence, whether, on an evolutionary basis, one residue is a likely substitute for the other. This likelihood has an associated "score" from which the "% similarity" of the two sequences can then be determined.

Methods for comparing the identity and similarity of two or more sequences are well known in the art. Thus for instance, programs available in the Wisconsin Sequence Analysis Package, version 9.1 (Devereux J et al, Nucleic Acids Res, 12, 387-395, 1984, available from Genetics Computer Group, Madison, Wisconsin, USA), for example the programs BESTFIT and GAP, may be used to determine the % identity between two polynucleotides and the % identity and the % similarity between two polypeptide sequences. BESTFIT uses the "local homology" algorithm of Smith and Waterman (J Mol Biol, 147,195-197, 1981, Advances in Applied Mathematics, 2, 482-489, 1981) and finds the best single region of similarity between two sequences. BESTFIT is more suited to comparing two polynucleotide or two polypeptide sequences that are dissimilar in length, the program assuming that the shorter sequence represents a portion of the longer. In comparison, GAP aligns two sequences, finding a "maximum similarity", according to the algorithm of Needleman and Wunsch (J Mol Biol, 48, 443-453, 1970). GAP is more suited to comparing sequences that are approximately the same length and an alignment is expected over the entire length. Preferably, the parameters "Gap Weight" and "Length Weight" used in each program are 50 and 3, for polynucleotide sequences and 12 and 4 for polypeptide sequences, respectively. Preferably, % identities and similarities are determined when the two sequences being compared are optimally aligned.

Other programs for determining identity and/or similarity between sequences are also known in the art, for instance the BLAST family of programs (Altschul S F et al, J Mol Biol, 215, 403-410, 1990, Altschul S F et al, Nucleic Acids Res., 25:389-3402, 1997, available from the National

WO 01/77147

PCT/EP01/03793

- 25 -

Center for Biotechnology Information (NCBI), Bethesda, Maryland, USA and accessible through the home page of the NCBI at www.ncbi.nlm.nih.gov) and FASTA (Pearson W R, Methods in Enzymology, 183, 63-99, 1990; Pearson W R and Lipman D J, Proc Nat Acad Sci USA, 85, 2444-2448, 1988, available as part of the Wisconsin Sequence Analysis Package).

Preferably, the BLOSUM62 amino acid substitution matrix (Henikoff S and Henikoff J G, Proc. Nat. Acad Sci. USA, 89, 10915-10919, 1992) is used in polypeptide sequence comparisons including where nucleotide sequences are first translated into amino acid sequences before comparison.

Preferably, the program BESTFIT is used to determine the % identity of a query polynucleotide or a polypeptide sequence with respect to a reference polynucleotide or a polypeptide sequence, the query and the reference sequence being optimally aligned and the parameters of the program set at the default value, as hereinbefore described.

"Identity Index" is a measure of sequence relatedness which may be used to compare a candidate sequence (polynucleotide or polypeptide) and a reference sequence. Thus, for instance, a candidate polynucleotide sequence having, for example, an Identity Index of 0.95 compared to a reference polynucleotide sequence is identical to the reference sequence except that the candidate polynucleotide sequence may include on average up to five differences per each 100 nucleotides of the reference sequence. Such differences are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion. These differences may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between these terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polynucleotide sequence having an Identity Index of 0.95 compared to a reference polynucleotide sequence, an average of up to 5 in every 100 of the nucleotides of the in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The same applies *mutatis mutandis* for other values of the Identity Index, for instance 0.96, 0.97, 0.98 and 0.99.

WO 01/77147

PCT/EP01/03793

- 26 -

Similarly, for a polypeptide, a candidate polypeptide sequence having, for example, an Identity Index of 0.95 compared to a reference polypeptide sequence is identical to the reference sequence except that the polypeptide sequence may include an average of up to five differences per each 100 amino acids of the reference sequence. Such differences are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion. These differences may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between these terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polypeptide sequence having an Identity Index of 0.95 compared to a reference polypeptide sequence, an average of up to 5 in every 100 of the amino acids in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The same applies *mutatis mutandis* for other values of the Identity Index, for instance 0.96, 0.97, 0.98 and 0.99.

The relationship between the number of nucleotide or amino acid differences and the Identity Index may be expressed in the following equation:

$$n_a \leq x_a - (x_a \bullet I),$$

in which:

n_a is the number of nucleotide or amino acid differences,

x_a is the total number of nucleotides or amino acids in SEQ ID NO:1 or SEQ ID NO:2, respectively,

I is the Identity Index ,

$\bullet$ is the symbol for the multiplication operator, and

in which any non-integer product of x_a and I is rounded down to the nearest integer prior to subtracting it from x_a .

"Homolog" is a generic term used in the art to indicate a polynucleotide or polypeptide sequence possessing a high degree of sequence relatedness

WO 01/77347

PCT/EP01/03793

- 27 -

to a reference sequence. Such relatedness may be quantified by determining the degree of identity and/or similarity between the two sequences as hereinbefore defined. Falling within this generic term are the terms "ortholog", and "paralog". "Ortholog" refers to a polynucleotide or polypeptide that is the functional equivalent of the polynucleotide or polypeptide in another species. "Paralog" refers to a polynucleotide or polypeptide that within the same species which is functionally similar.

"Fusion protein" refers to a protein encoded by two, unrelated, fused genes or fragments thereof. Examples have been disclosed in US 5541087, 5726044. In the case of Fc-hupolybromo1, employing an immunoglobulin Fc region as a part of a fusion protein is advantageous for performing the functional expression of Fc-hupolybromo1 or fragments of hupolybromo1, to improve pharmacokinetic properties of such a fusion protein when used for therapy and to generate a dimeric hupolybromo1. The Fc-hupolybromo1 DNA construct comprises in 5' to 3' direction, a secretion cassette, i.e. a signal sequence that triggers export from a mammalian cell, DNA encoding an immunoglobulin Fc region fragment, as a fusion partner, and a DNA encoding hupolybromo1 or fragments thereof. In some uses it would be desirable to be able to alter the intrinsic functional properties (complement binding, Fc-Receptor binding) by mutating the functional Fc sides while leaving the rest of the fusion protein untouched or delete the Fc part completely after expression.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

WO 01/77147

PCT/EP01/03793

- 28 -

Claims

1. A polypeptide selected from the group consisting of:
 - (a) a polypeptide encoded by a polynucleotide comprising the sequence of SEQ ID NO:1;
 - 5 (b) a polypeptide comprising a polypeptide sequence having at least 95% identity to the polypeptide sequence of SEQ ID NO:2;
 - c) a polypeptide having at least 95% identity to the polypeptide sequence of SEQ ID NO:2;
 - d) the polypeptide sequence of SEQ ID NO:2 and
 - 10 (e) fragments and variants of such polypeptides in (a) to (d).
2. The polypeptide of claim 1 comprising the polypeptide sequence of SEQ ID NO:2.
- 15 3. The polypeptide of claim 1 which is the polypeptide sequence of SEQ ID NO:2.
4. A polynucleotide selected from the group consisting of:
 - (a) a polynucleotide comprising a polynucleotide sequence having at least 95% identity to the polynucleotide sequence of SEQ ID NO:1;
 - 20 (b) a polynucleotide having at least 95% identity to the polynucleotide of SEQ ID NO:1;
 - (c) a polynucleotide comprising a polynucleotide sequence encoding a polypeptide sequence having at least 95% identity to the polypeptide sequence of SEQ ID NO:2;
 - 25 (d) a polynucleotide having a polynucleotide sequence encoding a polypeptide sequence having at least 95% identity to the polypeptide sequence of SEQ ID NO:2;

WO 01/77147

- 29 -

PCT/EP01/03793

- (e) a polynucleotide with a nucleotide sequence of at least 100 nucleotides obtained by screening a library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO: 1 or a fragment thereof having at least 15 nucleotides;
- 5 (f) a polynucleotide which is the RNA equivalent of a polynucleotide of (a) to (e);
- (g) a polynucleotide sequence complementary to said polynucleotide of any one of (a) to (f), and
- (h) polynucleotides that are variants or fragments of the polynucleotides of any one of (a) to (g) or that are complementary to above mentioned polynucleotides,
- 10 over the entire length thereof.
5. A polynucleotide of claim 4 selected from the group consisting of:
- (a) a polynucleotide comprising the polynucleotide of SEQ ID NO:1;
- (b) the polynucleotide of SEQ ID NO:1;
- 15 (c) a polynucleotide comprising a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2; and
- (d) a polynucleotide encoding the polypeptide of SEQ ID NO:2.
6. An expression system comprising a polynucleotide capable of producing a
- 20 polypeptide of any one of claim 1-3 when said expression vector is present in a compatible host cell.
7. A recombinant host cell comprising the expression vector of claim 6 or a membrane thereof expressing the polypeptide of any one of claim 1-3.
- 25
8. A process for producing a polypeptide of any one of claim 1-3 comprising the step of culturing a host cell as defined in claim 7 under conditions sufficient for

WO 01/77147

PCT/EP01/03793

- 30 -

the production of said polypeptide and recovering the polypeptide from the culture medium.

9. A fusion protein consisting of the Immunoglobulin Fc-region and a polypeptide
5 any one one of claims 1-3.
10. An antibody immunospecific for the polypeptide of any one of claims 1 to 3.
11. A method for screening to identify compounds that stimulate or inhibit the
function or level of the polypeptide of any one of claim 1-3 comprising a method
selected from the group consisting of:
- 10 (a) measuring or, detecting, quantitatively or qualitatively, the binding of a
candidate compound to the polypeptide (or to the cells or membranes expressing
the polypeptide) or a fusion protein thereof by means of a label directly or
indirectly associated with the candidate compound;
- (b) measuring the competition of binding of a candidate compound to the
15 polypeptide (or to the cells or membranes expressing the polypeptide) or a fusion
protein thereof in the presence of a labeled competitor;
- (c) testing whether the candidate compound results in a signal generated by
activation or inhibition of the polypeptide, using detection systems appropriate to
the cells or cell membranes expressing the polypeptide;
- 20 (d) mixing a candidate compound with a solution containing a polypeptide of any
one of claims 1-3, to form a mixture, measuring activity of the polypeptide in the
mixture, and comparing the activity of the mixture to a control mixture which
contains no candidate compound; or
- (e) detecting the effect of a candidate compound on the production of mRNA
25 encoding said polypeptide or said polypeptide in cells, using for instance, an
ELISA assay, and
- (f) producing said compound according to biotechnological or chemical standard
techniques.

WO 01/77147

PCT/EP01/03793

- 1 -

SEQUENCE LISTING

<110> Merck Patent GmbH

5 <120> New protein with polybromo domains

<130> hupolybromolcbws

<140>

10 <141>

<160> 2

<170> PatentIn Ver. 2.1

15 <210> 1

<211> 4902

<212> DNA

<213> Homo sapiens

20 <220>

<221> CDS

<222> {1}..(4899)

25 <400> 1

atg ggt tcc aag aga aga aga gct acc tcc cct tcc agc agt gtc ayc	48
Met Gly Ser Lys Arg Arg Arg Ala Thr Ser Pro Ser Ser Ser Val Ser	
1 5 10 15	

ggc ggc gac ttt gat gat ggc cac cat tct gtg tca aca cca ggc coa	96
Gly Gly Asp Phe Asp Asp Gly His His Ser Val Ser Thr Pro Gly Pro	
20 25 30	

agc agg aaa agg agy aga ctt tcc aat ctt cca act gla gat cct att	144
Ser Arg Lys Arg Arg Arg Leu Ser Asn Leu Pro Thr Val Asp Pro Ile	
35 40 45	

gcc gtg tgc cat gac ctc tat aat acc atc cga gac tat aag gat gaa	192
Ala Val Cys His Glu Leu Tyr Asn Thr Ile Arg Asp Tyr Lys Asp Glu	
50 55 60	

cag ggc aga ctt ctc tgt gag ctc ttc att agg gca cca aag cga aga	240
Gln Gly Arg Leu Leu Cys Glu Leu Phe Ile Asn Ala Pro Lys Arg Arg	
65 70 75 80	

aat caa cca gac tac tat gaa gtg gtt tct cag ccc att gac ttg atg	288
Asn Gln Pro Asp Tyr Tyr Glu Val Val Ser Gln Pro Ile Asp Leu Met	
85 90 95	

aaa atc caa cag aaa cta caa atg gaa gag tat gat gat gtt aat ttg	336
Lys Ile Gln Gln Lys Leu Lys Met Glu Glu Tyr Asp Asp Val Asn Leu	
100 105 110	

ctg act gct gac ttc cag ctt ctt ttt aac aat gca aag tcc tat tat	384
Leu Thr Ala Asp Phe Gln Leu Leu Phe Asn Asn Ala Lys Ser Tyr Tyr	
115 120 125	

aag cca gat tct cct gaa tat aaa gcc got tgc aaa ctc tgg gat ttg	432
Lys Pro Asp Ser Pro Glu Tyr Lys Ala Ala Cys Lys Leu Trp Asp Leu	

WO 01/77147

- 2 -

PCT/EP01/03793

	130	135	140	
	tac ctt cga aca aga aat gag ttt gtt cag aca gga gaa gca gat gac			480
5	Tyr Leu Arg Thr Arg Asn Glu Phe Val Gln Lys Gly Glu Ala Asp Asp	150	155	160
	gaa gat gat gat gaa gat ggg caa gac aat cag ggc aca tct tct cca			528
10	Glu Asp Asp Asp Glu Asp Gly Gln Asp Asn Gln Gly Thr Ser Ser Pro	165	170	175
	gct tac ttg aag gag atc ctg gag cag ctt ctt gaa gcc ata gtt gta			576
	Ala Tyr Leu Lys Glu Ile Leu Glu Gln Leu Leu Glu Ala Ile Val Val	180	185	190
15	gct aca aat cca tca gga cgt ctc att agc gaa ctt ttt cag aaa ctg			624
	Ala Thr Asn Pro Ser Gly Arg Leu Ile Ser Glu Leu Phe Gln Lys Leu	195	200	205
20	cct tct aaa gtg caa tat cca gat tat tat gca ata att aag gag cct			672
	Pro Ser Lys Val Gln Tyr Pro Asp Tyr Tyr Ala Ile Ile Lys Glu Pro	210	215	220
	ata gat ctc aag acc att gcc cag agg ata cag aat gga agc tac aaa			720
25	Ile Asp Leu Lys Thr Ile Ala Gln Arg Ile Gln Asn Gly Ser Tyr Lys	225	230	235
	agt att cat gca atg gcc aaa gat ata gat ctc ctc gca aaa aat gcc			768
30	Ser Ile His Ala Met Ala Lys Asp Ile Asp Leu Leu Ala Lys Asn Ala	245	250	255
	aaa act tat aat gag cct ggc tct caa gta ttc aag gat gca aat tca			816
	Lys Thr Tyr Asn Glu Pro Gly Ser Gln Val Phe Lys Asp Ala Asn Ser	260	265	270
35	att aac aaa ata ttt tat atg aac aag gct gaa att gaa cat cat gaa			864
	Ile Lys Lys Ile Phe Tyr Met Lys Lys Ala Glu Ile Glu His His Glu	275	280	285
40	atg gct aag tca agt ctt cga atg agg act cca tcc aac ctg gct gca			912
	Met Ala Lys Ser Ser Leu Arg Met Arg Thr Pro Ser Asn Leu Ala Ala	290	295	300
	gcc aga ctg aca ggt cct tca tca cac agt aac ggc agc ctt ggt gaa			960
45	Gla Arg Leu Thr Gly Pro Ser Ser His Ser Lys Gly Ser Leu Gly Glu	305	310	315
	gag aga aat ccc act agc aag tat tac cgt aat aaa aga gca gta caa			1008
50	Glu Arg Asn Pro Thr Ser Lys Tyr Tyr Arg Asn Lys Arg Ala Val Gln	325	330	335
	gga ggt cgt tta tca gca att aca atg gca ctt caa tat ggc tca gaa			1056
	Gly Gly Arg Leu Ser Ala Ile Thr Met Ala Leu Gln Tyr Gly Ser Glu	340	345	350
55	agt gaa gaa gat gct gct tta gct gct gca cgc tat gaa gag gga gag			1104
	Ser Glu Glu Asp Ala Ala Leu Ala Ala Ala Arg Tyr Glu Glu Gly Glu	355	360	365
60	tca gaa gca gaa agc atc act tcc ttt atg gat gtt tca aat cct ttt			1152
	Ser Glu Ala Glu Ser Ile Thr Ser Phe Met Asp Val Ser Asn Pro Phe	370	375	380

WO 01/77147

PCT/EP01/03793

- 3 -

	tat cag ctt tat gac aca gtt agg agt tgt cgg aat aac caa ggg cag	1200
	Tyr Gln Leu Tyr Asp Thr Val Arg Ser Cys Arg Asn Asn Gln Gly Gln	
	385 390 395 400	
5	cta ata got gaa cct ttt tac cat ttg cct tca aag aca aaa tac cct	1248
	Leu Ile Ala Glu Pro Phe Tyr His Leu Pro Ser Lys Lys Lys Tyr Pro	
	405 410 415	
10	gat tat tac cag caa att aaa atg ccc ata tca cta caa cag atc cga	1296
	Asp Tyr Tyr Gln Gln Ile Lys Met Pro Ile Ser Leu Gln Gln Ile Arg	
	420 425 430	
15	aca aaa ctg aag aat caa gaa tat gaa act tta gat cat ttg gag tgt	1344
	Thr Lys Leu Lys Asn Gln Glu Tyr Glu Thr Leu Asp His Leu Glu Cys	
	435 440 445	
20	gat ctg aat tta atg ttt gaa aat gcc aaa cgc tac aat gtg ccc aac	1392
	Asp Leu Asn Leu Met Phe Glu Asn Ala Lys Arg Tyr Asn Val Pro Asn	
	450 455 460	
25	tca gcc atc tac aag cga gtt cta aaa ttg cag caa gtt atg cag gaa	1440
	Ser Ala Ile Tyr Lys Arg Val Leu Lys Leu Gln Gln Val Met Gln Ala	
	465 470 475 480	
	aag aag aag gag ctt gcc agg aga gac gat atc gag gac gga gac agc	1488
	Lys Lys Lys Glu Leu Ala Arg Arg Asp Asp Ile Glu Asp Gly Asp Ser	
	485 490 495	
30	atg atc tct tca gcc acc tct gat act ggt agt gcc aaa aga aaa agt	1536
	Met Ile Ser Ser Ala Thr Ser Asp Thr Gly Ser Ala Lys Arg Lys Ser	
	500 505 510	
35	aaa aag aac ata aga aag cag cga atg aaa atc tta ttc aat gtt gtt	1584
	Lys Lys Asn Ile Arg Lys Gln Arg Met Lys Ile Leu Phe Asn Val Val	
	515 520 525	
40	ctt caa got cga gag cca ggt tca gcc aga aga ctt tgt gac cta ttt	1632
	Leu Glu Ala Arg Glu Pro Gly Ser Gly Arg Arg Leu Cys Asp Leu Phe	
	530 535 540	
45	atg gtt aaa caa tcc aaa aag gac tat cct gat tat tat aaa atc atc	1680
	Met Val Lys Pro Ser Lys Lys Asp Tyr Pro Asp Tyr Tyr Lys Ile Ile	
	545 550 555 560	
	ttg gag cca atg gac ttg aaa ata att gag cat aac atc cgc aat gac	1728
	Leu Glu Pro Met Asp Leu Lys Ile Ile Glu His Asn Ile Arg Asn Asp	
	565 570 575	
50	aaa tat got ggt gaa gag gga atg ata gaa gac atg aag ctg atg ttc	1776
	Lys Tyr Ala Gly Glu Glu Gly Met Ile Glu Asp Met Lys Leu Met Phe	
	580 585 590	
55	cgg aat gcc ags cac tat aat gag gag ggc tcc cag gtt tat aat gat	1824
	Arg Ser Ala Arg His Tyr Asn Glu Glu Gly Ser Gln Val Tyr Asn Asp	
	595 600 605	
60	gca cat atc ctg gag aag tta ctg aag gag aaa aag aaa gag ctg gcc	1872
	Ala His Ile Leu Glu Lys Leu Leu Lys Glu Lys Arg Lys Glu Leu Gly	
	610 615 620	

WO 01/77147

PCT/EP01/03793

- 4 -

	cca ctg cct gat gat gat gac atg gct tct ccc aaa ctc aag ctg agt	1920
	Pro Leu Pro Asp Asp Asp Met Ala Ser Pro Lys Leu Lys Leu Ser	
	625 630 635 640	
5	agg aag agt ggc att tct cct aaa aaa tca aac tac atg act cca atg	1968
	Arg Lys Ser Gly Ile Ser Pro Lys Lys Ser Lys Tyr Met Thr Pro Met	
	645 650 655	
10	cag cag aaa cta aat gag gtc tat gaa gct gta aag aac tat act gat	2016
	Gln Gln Lys Leu Asn Glu Val Tyr Glu Ala Val Lys Asn Tyr Thr Asp	
	660 665 670	
15	aag agg ggt cgc cgc ctc agt gcc ata ttt ctg agg ctt ccc tct aga	2064
	Lys Arg Gly Arg Arg Leu Ser Ala Ile Phe Leu Arg Pro Ser Arg	
	675 680 685	
	tct gag ttg cct gac tac tat ctg act att aaa aag ccc atg gac atg	2112
	Ser Glu Leu Pro Asp Tyr Tyr Leu Thr Ile Lys Pro Met Asp Met	
	690 695 700	
20	gaa aaa att cga agt cac atg atg gcc aac aag tac caa gat att gac	2160
	Glu Lys Ile Arg Ser His Met Met Ala Asn Lys Tyr Gln Asp Ile Asp	
	705 710 715 720	
25	tct atg gtt gag gac ttt gtc atg atg ttt aat aat gcc tgt aca tac	2208
	Ser Met Val Glu Asp Phe Val Met Met Phe Asn Asn Ala Cys Thr Tyr	
	725 730 735	
30	aat gag ccg gag tct ttg atc tac aaa gat gct ctt gtt cta cag aaa	2256
	Asn Glu Pro Glu Ser Leu Ile Tyr Lys Asp Ala Leu Val Leu His Lys	
	740 745 750	
35	gtc ctg ctt gaa aca cgc aga gac ctg gag gga gat gag gac tct cat	2304
	Val Leu Leu Glu Thr Arg Arg Asp Leu Glu Gly Asp Glu Asp Ser His	
	755 760 765	
	gtc cca aat gtg act ttg ctg att caa gag ctt atc cac aat ctt ttt	2352
	Val Pro Asn Val Thr Leu Leu Ile Gln Glu Leu Ile His Asn Leu Phe	
	770 775 780	
40	gtg tca gtc atg agt cat cag gat gat gag gga aga tgc tac agc gat	2400
	Val Ser Val Met Ser His Gln Asp Asp Glu Gly Arg Cys Tyr Ser Asp	
	785 790 795 800	
45	tct tta gca gaa att cct gct gtg gat ccc aac ttt cct aac aaa cca	2448
	Ser Leu Ala Glu Ile Pro Ala Val Asp Pro Asn Phe Pro Asn Lys Pro	
	805 810 815	
50	ccc ctt aca ttt gac cta att agg aag aat gtt gaa aat aat cgc tac	2496
	Pro Leu Thr Phe Asp Ile Ile Arg Lys Asn Val Glu Asn Asn Arg Tyr	
	820 825 830	
55	ogt ogg ctt gat tta ttt caa gag cat atg ttt gaa gta ttg gaa cga	2544
	Arg Arg Leu Asp Leu Phe Gln Glu His Met Phe Glu Val Leu Glu Arg	
	835 840 845	
	gca aga agg atg aat cgg aca gat tca gaa ata tat gaa gat gca gta	2592
	Ala Arg Arg Met Asn Arg Thr Asp Ser Glu Ile Tyr Glu Asp Ala Val	
	850 855 860	
60	gaa ctt cag cag ttt ttt att aac att cgc gat gaa ctc tgc aac aat	2640

WO 01/77147

PCT/EP01/03793

- 5 -

	Glu Leu Gln Gln Phe Phe Ile Lys Ile Arg Asp Glu Leu Cys Lys Asn	
	865 870 875 880	
5	gga gag att ctt ctt tca cag qca ctc agc tat acc aca aaa cat ttg	2688
	Gly Gln Ile Leu Leu Ser Pro Ala Leu Ser Tyr Thr Lys His Leu	
	885 890 895	
10	cat aat gat gtg gag aaa gag aga aag gaa aaa ttg cca aaa gaa ata	2736
	His Asn Asp Val Glu Lys Glu Arg Lys Glu Lys Leu Pro Lys Glu Ile	
	900 905 910	
15	gag gaa gat aaa cta aaa cga gaa gaa gaa aaa aga gaa gct gaa aag	2784
	Ser Gln Asp Lys Leu Lys Arg Glu Glu Glu Lys Arg Glu Ala Glu Lys	
	915 920 925	
20	agt gaa gat tcc tct ggt gct gca ggc ctc tca ggc tta cat cgc aca	2832
	Tyr Ser Gln Asp Ser Ser Gly Ala Ala Gly Leu Ser Gly Leu His Arg Thr	
	930 935 940	
25	tac agc cag gac tgt agc ttt aaa aac agc atg tac cat gtt gga gat	2880
	Tyr Ser Gln Asp Cys Ser Phe Lys Asn Ser Met Tyr His Val Gly Asp	
	945 950 955 960	
30	tac gtc tat gtg gaa cct gca gag gcc aac cta caa cca cat atc gtc	2928
	Tyr Val Tyr Val Glu Pro Ala Glu Ala Asn Leu Gln Pro His Ile Val	
	965 970 975	
35	tgt att gaa aga ctg tgg gag gat tca got ggt gaa aaa tgg ttg tat	2976
	Cys Ile Glu Arg Leu Trp Glu Asp Ser Ala Gly Glu Lys Trp Leu Tyr	
	980 985 990	
40	ggc tgt tgg ttt tac cga cca aat gaa aca ttc cac ctg gct aca cga	3024
	Gly Cys Trp Phe Tyr Arg Pro Asn Glu Thr Phe His Leu Ala Thr Arg	
	995 1000 1005	
45	aaa ttt cta gaa aaa gaa gtt ttt aag agt gac tat tac aac aaa gtt	3072
	Lys Phe Leu Glu Lys Glu Val Phe Lys Ser Asp Tyr Tyr Asn Lys Val	
	1010 1015 1020	
50	caa gtt agt aaa att cta ggc aag tgt gtg gtc atg ttt gtc aag gaa	3120
	Pro Val Ser Lys Ile Leu Gly Lys Cys Val Val Met Phe Val Lys Glu	
	1025 1030 1035 1040	
55	tac ttt aag tta tgc cca gaa aac ttc cga gat gag gat gtt ttt gtc	3168
	Tyr Phe Lys Leu Cys Pro Glu Asn Phe Arg Asp Glu Asp Val Phe Val	
	1045 1050 1055	
60	tgt gaa tca cgg tat tct gcc aaa acc aaa tct ttt aag aaa att aaa	3216
	Cys Glu Ser Arg Tyr Ser Ala Lys Thr Lys Ser Phe Lys Lys Ile Lys	
	1060 1065 1070	
65	ctg tgg acc atg ccc atc agc tca gtc agg ttt gtc cct cgg gat gtg	3264
	Leu Trp Thr Met Pro Ile Ser Ser Val Arg Phe Val Pro Arg Asp Val	
	1075 1080 1085	
70	cct ctg cct gtg gtt cgc gtg gcc tct gta ttt gca aat gca gat aaa	3312
	Pro Leu Pro Val Val Arg Val Ala Ser Val Phe Ala Asn Ala Asp Lys	
	1090 1095 1100	
75	ggt gat gat gag aag aat aca gac aac tca gag gac agt cga gct gaa	3360
	Gly Asp Asp Glu Lys Asn Thr Asp Asn Ser Glu Asp Ser Arg Ala Glu	

WO 01/77147

PCT/EP01/03793

- 6 -

	1105	1110	1115	1120	
	gac aat ttt tca aac ttg gaa aag gaa aaa gaa gat gtc cct gtg gaa				3408
5	Asp Asn Phe Ser Asn Leu Glu Lys Glu Lys Glu Asp Val Pro Val Glu	1125	1130	1135	
	atg tcc aat ggt gaa cca ggt tgc cac tac ttt gag cag ctc cat tcc				3456
	Met Ser Asn Met Trp Leu Lys Val Gly Asp Cys His Tyr Phe Glu Glu Leu His Tyr	1140	1145	1150	
10	aat gac atg tgg ctg aag gtt ggc gac tgt gtc ttc atc aag tcc cat				3504
	Asn Asp Met Trp Leu Lys Val Gly Asp Cys Val Phe Ile Lys Ser His	1155	1160	1165	
15	ggc ctg gtg cgt cct cgt gtc ggc aga att gaa aaa gta tgg gtt cga				3552
	Gly Leu Val Arg Pro Arg Val Gly Arg Ile Glu Lys Val Trp Val Arg	1170	1175	1180	
20	gat gga gct gca tat ttt tat ggc ccc atc ttc att cac cca gaa gaa				3600
	Asp Gly Ala Ala Tyr Phe Tyr Gly Pro Ile Phe Ile His Pro Glu Glu	1185	1190	1195	1200
	aca gag cat gag ccc acc aaa atg ttc tcc aaa aaa gaa gta ttt ctg				3648
25	Thr Glu His Glu Pro Thr Lys Met Phe Tyr Lys Lys Glu Val Phe Leu	1205	1210	1215	
	agt aat ctg gaa gaa acc tgc ccc atg acc tgt att ctc gga aag tgt				3696
	Ser Asn Leu Glu Glu Thr Cys Pro Met Thr Cys Ile Leu Gly Lys Cys	1220	1225	1230	
30	gct gtg ttg tca ttc aag gac ttc ctc tcc tgc agg cca act gaa ata				3744
	Ala Val Leu Ser Phe Lys Asp Phe Leu Ser Cys Arg Pro Thr Glu Ile	1235	1240	1245	
35	cca gaa aat gac att ctg att tgt gag agc cgc tac aat gag agc gac				3792
	Pro Glu Asn Asp Ile Leu Leu Cys Glu Ser Arg Tyr Asn Glu Ser Asp	1250	1255	1260	
40	aag cag atg aag aaa ttc aaa gga ttg aag agg ttt tca ctc tct gct				3840
	Lys Glu Met Lys Lys Phe Lys Gly Leu Lys Arg Phe Ser Leu Ser Ala	1265	1270	1275	1280
	aaa glg gta gat gat gaa att tac tac ttc aga aaa cca att gtt cct				3888
45	Lys Val Val Asp Asp Glu Ile Tyr Tyr Phe Arg Lys Pro Ile Val Pro	1285	1290	1295	
	cag aag gag cca tca cct ttg ctg gaa aag aag atc cag ttg cta gaa				3936
	Gln Lys Glu Pro Ser Pro Leu Leu Glu Lys Lys Ile Gln Leu Leu Glu	1300	1305	1310	
50	gct aaa ttt gcc gag tta gaa ggt gga gat gat gat att gaa gag atg				3984
	Ala Lys Phe Ala Glu Leu Glu Gly Gly Asp Asp Asp Ile Glu Glu Met	1315	1320	1325	
55	gga gaa gaa gat agt gag gtc att gaa cct cct tot cta cct cag att				4032
	Gly Glu Glu Asp Ser Glu Val Ile Glu Pro Pro Ser Leu Pro Gln Leu	1330	1335	1340	
60	gag acc acc ctg gcc agt gag ctg gac ctc atg ccc tac aca ccc ccc				4080
	Gln Thr Pro Leu Ala Ser Glu Leu Asp Leu Met Pro Tyr Thr Pro Pro	1345	1350	1355	1360

WO 01/77147

PCT/EP01/03793

- 7 -

	cag tct acc cca aag tct gcc aaa ggc agt gca aag aag gaa ggc tcc	4128
	Gln Ser Thr Pro Lys Ser Ala Lys Gly Ser Ala Lys Lys Glu Gly Ser	
	1365 1370 1375	
5	aaa cgg aaa atc aac atg agt ggc tac atc ctg ttc agc agt gag atg	4176
	Lys Arg Lys Ile Asn Met Ser Gly Tyr Ile Leu Phe Ser Ser Glu Met	
	1380 1385 1390	
10	agg gct gtg att aag gcc aca cac cca gac tac tct ttc ggg gag ctg	4224
	Arg Ala Val Ile Lys Ala Gln His Pro Asp Tyr Ser Phe Gly Glu Leu	
	1395 1400 1405	
15	agc cgc ctg gtg ggg aca gaa tgg aga aat ctg gag aca gcc aag aaa	4272
	Ser Arg Leu Val Gly Thr Glu Trp Arg Asn Leu Glu Thr Ala Lys Lys	
	1410 1415 1420	
20	gca gaa tat gaa ggc atg atg ggt ggc tat ccg cca ggc ctt cca cct	4320
	Ala Glu Tyr Glu Gly Met Met Gly Gly Tyr Pro Pro Gly Leu Pro Pro	
	1425 1430 1435 1440	
	ttg cag ggc cca gtt gat ggc ctt gtt ago atg ggc agc atg cag cca	4368
	Leu Gln Gly Pro Val Asp Gly Leu Val Ser Met Gly Ser Met Gln Pro	
	1445 1450 1455	
25	ctt ccc cct ggg ggg cct cca ccc ccc cat ctt ccg cca ggt gtg cct	4416
	Leu His Pro Gly Gly Pro Pro Pro His His Leu Pro Pro Gly Val Pro	
	1460 1465 1470	
30	ggc ctg ccg ggc atc cca cca ccg ggt gtg atg aac cca gga gtg gcc	4464
	Gly Leu Pro Gly Ile Pro Pro Pro Gly Val Met Asn Gln Gly Val Ala	
	1475 1480 1485	
35	cct atg gta ggg act cca gca cca ggt gga agt cca tat gga caa cag	4512
	Pro Met Val Gly Thr Pro Ala Pro Gly Gly Ser Pro Tyr Gly Gln Gln	
	1490 1495 1500	
40	gtg gga gtt ttg ggg cct cca ggg cag cag gca cca cct cca tat ccc	4560
	Val Gly Val Leu Gly Pro Pro Gly Gln Gln Ala Pro Pro Pro Tyr Pro	
	1505 1510 1515 1520	
	ggc cca cat cca gct gga ccc cct gtc ata cag cag cca aca aca ccc	4608
	Gly Pro His Pro Ala Gly Pro Pro Val Ile Gln Gln Pro Thr Thr Pro	
	1525 1530 1535	
45	atg ttt gta gct ccc cca cca aag acc cag cgg ctt ctt cac tca gag	4656
	Met Phe Val Ala Pro Pro Pro Lys Thr Gln Arg Leu Leu His Ser Glu	
	1540 1545 1550	
50	gcc tac ctg aaa tac att gaa gga ctg agt gcg gag tcc aac agc att	4704
	Ala Tyr Leu Lys Tyr Ile Glu Gly Leu Ser Ala Glu Ser Asn Ser Ile	
	1555 1560 1565	
55	agc aag tgg gat cag aca ctg gca gct cga aga cgc gac gtc cat ttg	4752
	Ser Lys Trp Asp Gln Thr Leu Ala Ala Arg Arg Arg Asp Val His Leu	
	1570 1575 1580	
60	tcg aaa gaa cag gag agc cgc cta ccc tct cac tgg ctg aaa agc aaa	4800
	Ser Lys Glu Gln Glu Ser Arg Leu Pro Ser His Trp Leu Lys Ser Lys	
	1585 1590 1595 1600	

WO 01/77147

PCT/EP01/03793

- 8 -

```

ggg gcc cac acc acc atg gca gat gcc ctc tgg cgc ctt cga gat ttg 4848
Gly Ala His Thr Thr Met Ala Asp Ala Leu Trp Arg Leu Arg Asp Leu
1605 1610 1615

5 atg ctc cgg gac acc ctc aac att cgc caa gca tac aac cta gaa aat 4896
Met Leu Arg Asp Thr Leu Asn Cile Arg Glu Ala Tyr Asn Leu Glu Asn
1620 1625 1630

gtt taa 4902
10 Val

<210> 2
<211> 1633
15 <212> PRT
    <213> Homo sapiens

<400> 2
Met Gly Ser Lys Arg Arg Arg Ala Thr Ser Pro Ser Ser Ser Val Ser
20 1 5 10 15

Gly Gly Asp Phe Asp Asp Gly His His Ser Val Ser Thr Pro Gly Pro
25 20 25 30

Ser Arg Lys Arg Arg Arg Leu Ser Asn Leu Pro Thr Val Asp Pro Ile
25 35 40 45

Ala Val Cys His Glu Leu Tyr Asn Thr Ile Arg Asp Tyr Lys Asp Glu
50 55 60

30 Gln Gly Arg Leu Leu Cys Glu Leu Phe Ile Arg Ala Pro Lys Arg Arg
65 70 75 80

Asn Gln Pro Asp Tyr Tyr Glu Val Val Ser Gln Pro Ile Asp Leu Met
35 85 90 95

Lys Ile Gln Gln Lys Leu Lys Met Glu Glu Tyr Asp Asp Val Asn Leu
100 105 110

40 Leu Thr Ala Asp Phe Gln Leu Leu Phe Asn Asn Ala Lys Ser Tyr Tyr
115 120 125

Lys Pro Asp Ser Pro Glu Tyr Lys Ala Ala Cys Lys Leu Trp Asp Leu
130 135 140

45 Tyr Leu Arg Thr Arg Asn Glu Phe Val Gln Lys Gly Glu Ala Asp Asp
145 150 155 160

Glu Asp Asp Asp Glu Asp Gly Gln Asp Asn Gln Gly Thr Ser Ser Pro
50 165 170 175

Ala Tyr Leu Lys Glu Ile Leu Gln Gln Leu Leu Glu Ala Ile Val Val
180 185 190

55 Ala Thr Asn Pro Ser Gly Arg Leu Ile Ser Glu Leu Phe Gln Lys Leu
195 200 205

Pro Ser Lys Val Gln Tyr Pro Asp Tyr Tyr Ala Ile Ile Lys Glu Pro
210 215 220

60 Ile Asp Leu Lys Thr Ile Ala Gln Arg Ile Gln Asn Gly Ser Tyr Lys

```

PCT/EP01/03793

	225				230				235				240				
	Ser	Ile	His	Ala	Met	Ala	Lys	Asp	Ile	Asp	Leu	Leu	Ala	Lys	Asn	Ala	245
5					245					250						255	
	Lys	Thr	Tyr	Asn	Glu	Pro	Gly	Ser	Gln	Val	Phe	Lys	Asp	Ala	Asn	Ser	
				260					265				270				
10	Ile	Lys	Lys	Ile	Phe	Tyr	Met	Lys	Lys	Ala	Glu	Ile	Glu	His	His	Glu	
				275					280				285				
	Met	Ala	Lys	Ser	Ser	Leu	Arg	Met	Arg	Thr	Pro	Ser	Asn	Leu	Ala	Ala	
		290					295					300					
15	Ala	Arg	Leu	Thr	Gly	Pro	Ser	Ser	His	Ser	Lys	Gly	Ser	Leu	Gly	Glu	
		305				310					315					320	
	Glu	Arg	Asn	Pro	Thr	Ser	Lys	Tyr	Tyr	Arg	Asn	Lys	Arg	Ala	Val	Gln	
				325					330						335		
20	Gly	Gly	Arg	Leu	Ser	Ala	Ile	Thr	Met	Ala	Leu	Gln	Tyr	Gly	Ser	Glu	
				340					345				350				
	Ser	Glu	Glu	Asp	Ala	Ala	Leu	Ala	Ala	Ala	Arg	Tyr	Glu	Glu	Gly	Glu	
		355					360						365				
25	Ser	Glu	Ala	Glu	Ser	Ile	Thr	Ser	Phe	Met	Asp	Val	Ser	Asn	Pro	Phe	
		370					375					380					
30	Tyr	Gln	Leu	Tyr	Asp	Thr	Val	Arg	Ser	Cys	Arg	Asn	Asn	Gln	Gly	Gln	
		385				390					395					400	
	Leu	Ile	Ala	Glu	Pro	Phe	Tyr	His	Leu	Pro	Ser	Lys	Lys	Lys	Tyr	Pro	
				405						410						415	
35	Asp	Tyr	Tyr	Gln	Gln	Ile	Lys	Met	Pro	Ile	Ser	Leu	Gln	Gln	Ile	Arg	
				420					425				430				
	Thr	Lys	Leu	Lys	Asn	Gln	Glu	Tyr	Glu	Thr	Leu	Asp	His	Leu	Glu	Cys	
		435					440					445					
40	Asp	Leu	Asn	Leu	Met	Phe	Glu	Asn	Ala	Lys	Arg	Tyr	Asn	Val	Pro	Asn	
		450					455					460					
45	Ser	Ala	Ile	Tyr	Lys	Arg	Val	Leu	Lys	Leu	Gln	Gln	Val	Met	Gln	Ala	
		465				470					475					480	
	Lys	Lys	Lys	Glu	Leu	Ala	Arg	Arg	Asp	Ile	Glu	Asp	Gly	Asp	Ser		
				485					490					495			
50	Met	Ile	Ser	Ser	Ala	Thr	Ser	Asp	Thr	Gly	Ser	Ala	Lys	Arg	Lys	Ser	
				500					505					510			
	Lys	Lys	Asn	Ile	Arg	Lys	Gln	Arg	Met	Lys	Ile	Leu	Phe	Asn	Val	Val	
				515					520				525				
55	Leu	Glu	Ala	Arg	Glu	Pro	Gly	Ser	Gly	Arg	Arg	Leu	Cys	Asp	Leu	Phe	
				530			535					540					
60	Met	Val	Lys	Pro	Ser	Lys	Lys	Asp	Tyr	Pro	Asp	Ser					

WO 01/77147

- 10 -

PCT/EP01/03793

Leu Glu Pro Met Asp Leu Lys Ile Ile Glu His Asn Ile Arg Asn Asp
 565 570 575
 5 Lys Tyr Ala Gly Glu Glu Gly Met Ile Glu Asp Met Lys Leu Met Phe
 580 585 590
 Arg Asn Ala Arg His Tyr Asn Glu Glu Gly Ser Gln Val Tyr Asn Asp
 595 600 605
 10 Ala His Ile Leu Glu Lys Leu Leu Lys Glu Lys Arg Lys Glu Leu Gly
 610 615 620
 Pro Leu Pro Asp Asp Asp Met Ala Ser Pro Lys Leu Lys Leu Ser
 625 630 635 640
 Arg Lys Ser Gly Ile Ser Pro Lys Lys Ser Lys Tyr Met Thr Pro Met
 645 650 655
 20 Gln Gln Lys Leu Asn Glu Val Tyr Glu Ala Val Lys Asn Tyr Thr Asp
 660 665 670
 Lys Arg Gly Arg Arg Leu Ser Ala Ile Phe Leu Arg Leu Pro Ser Arg
 675 680 685
 25 Ser Glu Leu Pro Asp Tyr Tyr Leu Thr Ile Lys Lys Pro Met Asp Met
 690 695 700
 Glu Lys Ile Arg Ser His Met Met Ala Asn Lys Tyr Gln Asp Ile Asp
 705 710 715 720
 Ser Met Val Glu Asp Phe Val Met Met Phe Asn Asn Ala Cys Thr Tyr
 725 730 735
 35 Asn Glu Pro Glu Ser Leu Ile Tyr Lys Asp Ala Leu Val Leu His Lys
 740 745 750
 Val Leu Leu Glu Thr Arg Arg Asp Leu Glu Gly Asp Glu Asp Ser His
 755 760 765
 40 Val Pro Asn Val Thr Leu Leu Ile Gln Glu Leu Ile His Asn Leu Phe
 770 775 780
 Val Ser Val Met Ser His Gln Asp Asp Glu Gly Arg Cys Tyr Ser Asp
 785 790 795 800
 45 Ser Leu Ala Glu Ile Pro Ala Val Asp Pro Asn Phe Pro Asn Lys Pro
 805 810 815
 Pro Leu Thr Phe Asp Ile Ile Arg Lys Asn Val Glu Asn Asn Arg Tyr
 820 825 830
 Arg Arg Leu Asp Leu Phe Gln Glu His Met Phe Glu Val Leu Glu Arg
 835 840 845
 55 Ala Arg Arg Met Asn Arg Thr Asp Ser Glu Ile Tyr Glu Asp Ala Val
 850 855 860
 Glu Leu Gln Gln Phe Phe Ile Lys Ile Arg Asp Glu Leu Cys Lys Asn
 865 870 875 880

WO 01/77147

- 11 -

PCT/EP01/03793

Gly Glu Ile Leu Leu Ser Pro Ala Leu Ser Tyr Thr Thr Lys His Leu
 895 890 895
 His Asn Asp Val Glu Lys Glu Arg Lys Glu Lys Leu Pro Lys Glu Ile
 5 900 905 910
 Glu Glu Asp Lys Leu Lys Arg Glu Glu Glu Lys Arg Glu Ala Glu Lys
 915 920 925
 10 Ser Glu Asp Ser Ser Gly Ala Ala Gly Leu Ser Gly Leu His Arg Thr
 930 935 940
 Tyr Ser Gln Asp Cys Ser Phe Lys Asn Ser Met Tyr His Val Gly Asp
 945 950 955 960
 15 Tyr Val Tyr Val Glu Pro Ala Glu Ala Asn Leu Gln Pro His Ile Val
 965 970 975
 Cys Ile Glu Arg Leu Trp Glu Asp Ser Ala Gly Glu Lys Trp Leu Tyr
 20 980 985 990
 Gly Cys Trp Phe Tyr Arg Pro Asn Glu Thr Phe His Leu Ala Thr Arg
 995 1000 1005
 25 Lys Phe Leu Glu Lys Glu Val Phe Lys Ser Asp Tyr Tyr Asn Cys Val
 1010 1015 1020
 Pro Val Ser Lys Ile Leu Gly Lys Cys Val Val Met Phe Val Lys Glu
 1025 1030 1035 1040
 30 Tyr Phe Lys Leu Cys Pro Glu Asn Phe Arg Asp Glu Asp Val Phe Val
 1045 1050 1055
 Cys Glu Ser Arg Tyr Ser Ala Lys Thr Lys Ser Phe Lys Lys Ile Lys
 35 1060 1065 1070
 Leu Trp Thr Met Pro Ile Ser Ser Val Arg Phe Val Pro Arg Asp Val
 1075 1080 1085
 40 Pro Leu Pro Val Val Arg Val Ala Ser Val Phe Ala Asn Ala Asp Lys
 1090 1095 1100
 Gly Asp Asp Glu Lys Asn Thr Asp Asn Ser Glu Asp Ser Arg Ala Glu
 1105 1110 1115 1120
 45 Asp Asn Phe Ser Asn Leu Glu Lys Glu Lys Glu Asp Val Pro Val Glu
 1125 1130 1135
 Met Ser Asn Gly Glu Pro Gly Cys His Tyr Phe Glu Gln Leu His Tyr
 50 1140 1145 1150
 Asn Asp Met Trp Leu Lys Val Gly Asp Cys Val Phe Ile Lys Ser His
 1155 1160 1165
 55 Gly Leu Val Arg Pro Arg Val Gly Arg Ile Glu Lys Val Trp Val Arg
 1170 1175 1180
 Asp Gly Ala Ala Tyr Phe Tyr Gly Pro Ile Phe Ile His Pro Glu Glu
 1185 1190 1195 1200
 60 Thr Glu His Glu Pro Thr Lys Met Phe Tyr Lys Lys Glu Val Phe Leu

WO 01/77147

PCT/EP01/03793

- 12 -

	1205	1210	1215
	Ser Asn Leu Glu Glu Thr Cys Pro Met Thr Cys Ile Leu Gly Lys Cys		
	1220	1225	1230
5	Ala Val Leu Ser Phe Lys Asp Phe Leu Ser Cys Arg Pro Thr Glu Ile		
	1235	1240	1245
10	Pro Glu Asn Asp Ile Leu Leu Cys Glu Ser Arg Tyr Asn Glu Ser Asp		
	1250	1255	1260
	Lys Glu Met Lys Lys Phe Lys Gly Leu Lys Arg Phe Ser Leu Ser Ala		
	1265	1270	1275
15	Lys Val Val Asp Asp Glu Ile Tyr Tyr Phe Arg Lys Pro Ile Val Pro		
	1285	1290	1295
	Gln Lys Glu Pro Ser Pro Leu Leu Glu Lys Lys Ile Gln Leu Leu Glu		
	1300	1305	1310
20	Ala Lys Phe Ala Glu Leu Glu Gly Gly Asp Asp Asp Ile Glu Glu Met		
	1315	1320	1325
	Gly Glu Glu Asp Ser Glu Val Ile Glu Pro Pro Ser Leu Pro Gln Leu		
	1330	1335	1340
	Gln Thr Pro Leu Ala Ser Glu Leu Asp Leu Met Pro Tyr Thr Pro Pro		
	1345	1350	1355
30	Gln Ser Thr Pro Lys Ser Ala Lys Gly Ser Ala Lys Lys Glu Gly Ser		
	1365	1370	1375
	Lys Arg Lys Ile Asn Met Ser Gly Tyr Ile Leu Phe Ser Ser Glu Met		
	1380	1385	1390
35	Arg Ala Val Ile Lys Ala Gln His Pro Asp Tyr Ser Phe Gly Glu Leu		
	1395	1400	1405
	Ser Arg Leu Val Gly Thr Glu Trp Arg Asn Leu Glu Thr Ala Lys Lys		
	1410	1415	1420
	Ala Glu Tyr Glu Gly Met Met Gly Gly Tyr Pro Pro Gly Leu Pro Pro		
	1425	1430	1435
45	Leu Glu Gly Pro Val Asp Gly Leu Val Ser Met Gly Ser Met Gln Pro		
	1445	1450	1455
	Leu His Pro Gly Gly Pro Pro Pro His His Leu Pro Pro Gly Val Pro		
	1460	1465	1470
50	Gly Leu Pro Gly Ile Pro Pro Pro Gly Val Met Asn Gln Gly Val Ala		
	1475	1480	1485
	Pro Met Val Gly Thr Pro Ala Pro Gly Gly Ser Pro Tyr Gly Gln Gln		
	1490	1495	1500
	Val Gly Val Leu Gly Pro Pro Gly Gln Gln Ala Pro Pro Pro Tyr Pro		
	1505	1510	1515
60	Gly Pro His Pro Ala Gly Pro Pro Val Ile Gln Gln Pro Thr Thr Pro		
	1525	1530	1535

WO 01/77147

- 13 -

PCT/EP01/03793

Met Phe Val Ala Pro Pro Pro Lys Thr Gln Arg Leu Leu His Ser Glu
 1540 1545 1550

5 Ala Tyr Leu Lys Tyr Ile Glu Gly Leu Ser Ala Glu Ser Asn Ser Ile
 1555 1560 1565

Ser Lys Trp Asp Gln Thr Leu Ala Ala Arg Arg Arg Asp Val His Leu
 1570 1575 1580

10 Ser Lys Glu Gln Glu Ser Arg Leu Pro Ser His Trp Leu Lys Ser Lys
 1585 1590 1595 1600

Gly Ala His Thr Thr Met Ala Asp Ala Leu Trp Arg Leu Arg Asp Leu
 1605 1610 1615

Met Leu Arg Asp Thr Leu Asp Ile Arg Gln Ala Tyr Asn Leu Glu Asn
 1620 1625 1630

20 Val

【国際公開パンフレット（コレクトバージョン）】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
18 October 2001 (18.10.2001)

PCT

(11) International Publication Number
WO 01/77147 A3

(51) International Patent Classification: **C07K 14/00** (64289 Darmstadt (DE); DÜCKER, Klaus (DE/DF);
Eltterstrasse 5, 64291 Darmstadt (DE)).

(21) International Application Number: PCT/EP01/03793

(22) International Filing Date: 4 April 2001 (04.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 09107552.7 5 April 2000 (05.04.2000) EP

(71) Applicant (for all designated States except US): **MERCK PATENT GMBH** (DE/DF); Frankfurter Strasse 250, 64293 Darmstadt (DE).

(72) Inventors; and
(73) Inventors/Applicants (for US only): **BÜRGER, Christa** (DE/DE); Cusanweg 25, 64289 Darmstadt (DE);
WILBERT, Oliver (DE/DE); Lohmattenstrasse 54,

(74) Common Representative: **MERCK PATENT GMBH**; Frankfurter Strasse 250, 64293 Darmstadt (DE).

(81) Designated States (national): CA, JP, US.

(84) Designated States (regional): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published:
— with international search report

(88) Date of publication of the international search report: 14 March 2002



WO 01/77147 A3

(54) Title: NEW BROMODOMAIN PROTEIN

(57) Abstract: Lipopolybromal polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing lipopolybromal polypeptides and polynucleotides in diagnostic assays.

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/EP 01/03793
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C12N15/62 C07K14/47 G01N33/68		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched: classification system followed by classification codes IPC 7 C12N C07K G01N		
Documentation searched other than minimum documentation to the extent that such documents are indicated in the fields searched		
Findings data base consulted during the international search (name of data base and/or other practical search terms used) EMBL, EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE, EMBASE		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim no.
X	DATABASE EM_EST 'Online! EMBL, ID HS1246021, AC AA456653, 7 June 1997 (1997-06-07) HILLIER L ET AL.: "aa01d07.s1 Soares_NHMPu.S1 Homo sapiens cDNA clone IMAGE:811981 3' similar to IR:6951231 G951231 POLYBROMO 1 PROTEIN. ; mRNA sequence" XP002179222 Note: 100.0% nt seq identity with SEQ ID NO:1 in 472 nt overlap (1-472:1525-1054) the whole document --- -/--	1-11
☒ Further documents are listed in the continuation of box C ☒ Patent family members are listed in annex.		
* Special categories of cited documents: "A" documents defining the general state of the art which is not considered to be of particular relevance; "B" earlier document but published on or after the international filing date; "C" documents which may throw doubts on priority claims or which are cited to establish the publication date of another citation or other special reason (as specified); "D" documents relating to an oral disclosure, use, exhibition or other means; "E" documents published prior to the international filing date but later than the priority date claimed; "F" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention; "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when this document is taken alone; "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art; "Z" document member of the same patent family.		
Date of the actual completion of the international search		Date of issuing of the international search report
4 October 2001		16/10/2001
Name and mailing address of the ISA European Patent Office, P. B. 5616 Dattelnstr. 2 W - 52001 EU Datteln Tel. (+49-701-340-2040), Tx. 21 851 epo nl, Fax: (+49-701-340-3016)		Authorized officer van de Kamp, M

3
 1. See PCT-ISA/2001 (July 1997).

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/EP 01/03793

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication where appropriate, of the relevant passages	Relevance to claim No.
X	DATABASE EM_EST 'Online! EMBL; ID A1830417, AC A1830417, 13 July 1999 (1999-07-13) NATIONAL CANCER INSTITUTE, CANCER GENOME ANATOMY PROJECT (CGAP): "wh94d12.x1 NCI_CGAP_CELL1 Homo sapiens cDNA clone IMAGE:2388407 3' similar to TR:Q90941 Q90941 POLYBROMO I PROTEIN. ; mRNA sequence" XP002179223 Note: 99.8% nt seq identity (100.0% ungapped) with SEQ ID NO:1 in 450 nt overlap (25-474:1971-2420) the whole document	1-11
X	DATABASE EM_QV 'Online! EMBL; ID GGPB1GENE, AC X90849, 22 August 1995 (1995-08-22) NICOLAS R H ET AL.: "G. gallus P81 gene" XP002179224 cited in the application Note: 81.7% nt seq identity of P81 gene with SEQ ID NO:1 in 4903 nt overlap (94-4994:1-4901), 91.0% aa seq identity of P81 protein with SEQ ID NO:2 in 1633 aa overlap (1-1633:1-1633) the whole document	1,4,6-11
X	-& NICOLAS R H ET AL.: "Molecular cloning of polybromo, a nuclear protein containing multiple domains including five bromodomains, a truncated HMG-box, and two repeats of a novel domain" GENE, vol. 175, no. 1, 10 October 1996 (1996-10-10), pages 233-240, XP004043320 cited in the application the whole document	1,4,6-11
X	DATABASE EM_HTG 'Online! EMBL; ID AC012522, AC AC012522, 2 November 1999 (1999-11-02) MUZNY D M ET AL.: "Homo sapiens chromosome 3 clone RP1-86505, WORKING DRAFT SEQUENCE, 21 unordered pieces" XP002179225 Note: 100.0% nt seq identity with SEQ ID NO:1 in 644 nt overlap (114609-115252:2562-1919) page 30	4

3

Form PCT/EP 01/03793 (Continuation of second sheet) (July 1995)

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/EP 01/03793

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Outline of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 978 519 A (CHUGAI RES INST MOLECULAR MED) 9 February 2000 (2000-02-09) page 2, line 56 -page 3, line 6 claims 1-11 figures 1,6,14-18 ---	6-11
A	JEANMOUGIN F ET AL.: "The bromodomain revisited" TIBS TRENDS IN BIOCHEMICAL SCIENCES, vol. 22, no. 6, 1 May 1997 (1997-05-01), pages 151-153. XP004061751 ISSN: 0968-0004 the whole document page 151, left-hand column, line 1-12 ---	
P,X	DATABASE EM_HUM "Online!" EMBL: ID AF197569, AC AF197569, 30 November 2000 (2000-11-30) XUE Y ET AL.: "Homo sapiens BAF180 (BAF180) mRNA, complete cds" XP002179226 Note: 90.7% nt seq identity of BAF180 cDNA with SEQ ID NO:1 in 3013 nt overlap (97-3082:1-2981), 96.1% aa seq identity of BAF180 protein with SEQ ID NO:2 in 1637 aa overlap (1-1582:1-1633) the whole document page 13015, right-hand column, line 21-34 -& XUE Y ET AL.: "The human SWI/SNF-B chromatin-remodeling complex is related to yeast Rsc and localizes at kinetochores of mitotic chromosomes." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 97, no. 24, 21 November 2000 (2000-11-21), pages 13015-13020, XP002179233 the whole document figures 2,3	1,4
P,X	---	
P,X	DATABASE EM_HUM "Online!" EMBL: ID AF177387, AC AF177387, 12 January 2001 (2001-01-12) HORIKAWA I ET AL.: "Homo sapiens polybromo-1 (PBI) mRNA, complete cds, alternatively spliced" XP002179227 Note: 94.6% nt seq identity of PBI cDNA with SEQ ID NO:1 in 4934 nt overlap (4-4908:1-4902), 99.6% aa seq identity of PBI protein with SEQ ID NO:2 in 1637 aa overlap (1-1634:1-1633) the whole document -----	1,4,6-8, 10
P,X	---	
P,X	DATABASE EM_HUM "Online!" EMBL: ID AF177387, AC AF177387, 12 January 2001 (2001-01-12) HORIKAWA I ET AL.: "Homo sapiens polybromo-1 (PBI) mRNA, complete cds, alternatively spliced" XP002179227 Note: 94.6% nt seq identity of PBI cDNA with SEQ ID NO:1 in 4934 nt overlap (4-4908:1-4902), 99.6% aa seq identity of PBI protein with SEQ ID NO:2 in 1637 aa overlap (1-1634:1-1633) the whole document -----	1,4

INTERNATIONAL SEARCH REPORT				International Application No.	
Information on patent family members				PCT/EP 01/03793	
Patent document cited in search report	Publication date	Patent family member(s)	Publication date	Publication date	Publication date
EP 0978519	A	09-02-2000	AU	6854098 A	13-11-1998
			EP	0978519 A1	09-02-2000
			WO	9847920 A1	29-10-1998

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード(参考)
C 1 2 N 1/19	C 1 2 N 1/19	
C 1 2 N 1/21	C 1 2 N 1/21	
C 1 2 N 5/10	G 0 1 N 33/15	Z
G 0 1 N 33/15	G 0 1 N 33/50	Z
G 0 1 N 33/50	G 0 1 N 33/53	D
G 0 1 N 33/53	G 0 1 N 33/53	M
G 0 1 N 33/566	G 0 1 N 33/566	
// A 6 1 K 45/00	C 1 2 N 5/00	A
A 6 1 P 9/00	A 6 1 K 45/00	
A 6 1 P 9/04	A 6 1 P 9/00	
A 6 1 P 9/10	A 6 1 P 9/04	
A 6 1 P 25/00	A 6 1 P 9/10	1 0 3
A 6 1 P 25/08	A 6 1 P 25/00	
A 6 1 P 25/16	A 6 1 P 25/08	
A 6 1 P 25/18	A 6 1 P 25/16	
A 6 1 P 25/24	A 6 1 P 25/18	
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 35/00	A 6 1 P 29/00	1 0 1
A 6 1 P 37/06	A 6 1 P 35/00	
A 6 1 P 37/08	A 6 1 P 37/06	
	A 6 1 P 37/08	

(74)代理人 100088328

弁理士 金田 暢之

(74)代理人 100106297

弁理士 伊藤 克博

(74)代理人 100106138

弁理士 石橋 政幸

(72)発明者 ブルゲル、 クリスタ

ドイツ連邦共和国 6 4 2 8 9 ダルムシュタット カルサンヴェーク 2 3

(72)発明者 ヴィルベルト、 オリベル

ドイツ連邦共和国 6 4 2 8 9 ダルムシュタット リーフフラウエンシュトラッセ 5 4

(72)発明者 ドゥーカー、 クラウス

ドイツ連邦共和国 6 4 2 9 1 ダルムシュタット エテステルシュトラッセ 5

F ターム(参考) 2G045 AA40 DA36 FB01 FB03 FB07

4B024 AA01 AA11 BA80 CA04 CA12 HA11 HA17

4B065 AA93Y AB01 CA24 CA44 CA46

4C084 AA17 NA14 ZA022 ZA062 ZA122 ZA152 ZA162 ZA182 ZA362 ZA402

ZB082 ZB132 ZB152 ZB262

4H045 AA10 AA11 AA20 AA30 BA10 BA41 CA40 DA75 EA20 EA50

FA74

专利名称(译)	新的溴结构域蛋白		
公开(公告)号	JP2004500123A	公开(公告)日	2004-01-08
申请号	JP2001575617	申请日	2001-04-04
申请(专利权)人(译)	默克专利GESELLSCHAFT手套Beshurenkuteru Hafutongu		
[标]发明人	ブルゲルクリスタ ヴィルベルトオリベル ドゥーカークラウス		
发明人	ブルゲル、 クリスタ ヴィルベルト、 オリベル ドゥーカー、 クラウス		
IPC分类号	G01N33/50 A61K45/00 A61P9/00 A61P9/04 A61P9/10 A61P25/00 A61P25/08 A61P25/16 A61P25/18 A61P25/24 A61P25/28 A61P29/00 A61P35/00 A61P37/06 A61P37/08 C07K14/00 C07K14/47 C07K16/18 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 G01N33/15 G01N33/53 G01N33/566		
CPC分类号	A61P9/00 A61P9/04 A61P9/10 A61P25/00 A61P25/08 A61P25/16 A61P25/18 A61P25/24 A61P25/28 A61P29/00 A61P35/00 A61P37/06 A61P37/08 C07K14/47 C07K14/4702 C07K2319/00		
FI分类号	C12N15/00.ZNA.A C07K14/47 C07K16/18 C07K19/00 C12N1/15 C12N1/19 C12N1/21 G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A A61K45/00 A61P9/00 A61P9/04 A61P9/10.103 A61P25/00 A61P25/08 A61P25/16 A61P25/18 A61P25/24 A61P25/28 A61P29/00.101 A61P35/00 A61P37/06 A61P37/08		
F-TERM分类号	2G045/AA40 2G045/DA36 2G045/FB01 2G045/FB03 2G045/FB07 4B024/AA01 4B024/AA11 4B024/BA80 4B024/CA04 4B024/CA12 4B024/HA11 4B024/HA17 4B065/AA93Y 4B065/AB01 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA17 4C084/NA14 4C084/ZA022 4C084/ZA062 4C084/ZA122 4C084/ZA152 4C084/ZA162 4C084/ZA182 4C084/ZA362 4C084/ZA402 4C084/ZB082 4C084/ZB132 4C084/ZB152 4C084/ZB262 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/BA41 4H045/CA40 4H045/DA75 4H045/EA20 4H045/EA50 4H045/FA74		
代理人(译)	宫崎昭雄 伊藤 克博		
优先权	2000107352 2000-04-05 EP		
外部链接	Espacenet		

摘要(译)

公开了Hupolybromo1多肽和多核苷酸以及通过重组技术产生这种多肽的方法。还公开了在诊断测定中利用hupolybromo1多肽和多核苷酸的方法。

