

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3761784号

(P3761784)

(45) 発行日 平成18年3月29日(2006.3.29)

(24) 登録日 平成18年1月20日(2006.1.20)

(51) Int. Cl.

F I

C 1 2 N 15/09 (2006.01)
 C O 7 K 14/47 (2006.01)
 C O 7 K 16/18 (2006.01)
 C O 7 K 16/40 (2006.01)
 C 1 2 N 1/15 (2006.01)

C 1 2 N 15/00 Z N A A
 C O 7 K 14/47
 C O 7 K 16/18
 C O 7 K 16/40
 C 1 2 N 1/15

請求項の数 18 (全 120 頁) 最終頁に続く

(21) 出願番号 特願2000-601171 (P2000-601171)
 (86) (22) 出願日 平成12年2月24日(2000.2.24)
 (65) 公表番号 特表2003-525023 (P2003-525023A)
 (43) 公表日 平成15年8月26日(2003.8.26)
 (86) 国際出願番号 PCT/IB2000/000324
 (87) 国際公開番号 W02000/050607
 (87) 国際公開日 平成12年8月31日(2000.8.31)
 審査請求日 平成13年8月24日(2001.8.24)
 (31) 優先権主張番号 60/121,483
 (32) 優先日 平成11年2月24日(1999.2.24)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 501337203
 サウス、フアン
 スペイン国 エー46005 ヴァレンシ
 ア カッレ コンデ デ アルテア 8-
 7アー
 (74) 代理人 100083806
 弁理士 三好 秀和
 (72) 発明者 サウス、フアン
 スペイン国 エー46005 ヴァレンシ
 ア カッレ コンデ デ アルテア 8-
 7アー

審査官 飯室 里美

最終頁に続く

(54) 【発明の名称】 グッドパスチャー抗原結合タンパク質

(57) 【特許請求の範囲】

【請求項1】

配列番号：1、配列番号：3、配列番号：5、配列番号：7、配列番号：9、配列番号：11および配列番号：25からなる群から選択される核酸に高ストリンジエンシー条件下でハイブリダイズする、ヒトグッドパスチャー(GP)抗原のN末端領域に特異的に結合しセリン/スレオニンキナーゼ活性を持つタンパク質をコードする核酸を含む、単離された核酸。

【請求項2】

配列番号：1、配列番号：3、配列番号：5、配列番号：7、配列番号：9、配列番号：11および配列番号：25からなる群から選択される核酸を含む、単離された核酸。

【請求項3】

配列番号：2、配列番号：4、配列番号：6、配列番号：8、配列番号：10および配列番号：12からなる群から選択されるタンパク質と、少なくとも90%以上の同一性をもち、ヒトGP抗原のN末端領域に特異的に結合しセリン/スレオニンキナーゼ活性を持つタンパク質を、コードする核酸を含む、単離された核酸。

【請求項4】

配列番号：2、配列番号：4、配列番号：6、配列番号：8、配列番号：10および配列番号：12からなる群から選択されるタンパク質をコードする核酸を含む、単離された核酸。

【請求項5】

10

20

請求項 1 から 4 のいずれかに記載の単離された核酸を含む、組換え発現ベクター。

【請求項 6】

配列番号：1、配列番号：3、配列番号：5、配列番号：7、配列番号：9、配列番号：11 および配列番号：25 からなる群から選択される単離された核酸を含む、組換え発現ベクター。

【請求項 7】

請求項 5 または 6 に記載の組換え発現ベクターでトランスフェクトした宿主細胞。

【請求項 8】

配列番号：2、配列番号：4、配列番号：6、配列番号：8、配列番号：10 および配列番号：12 からなる群から選択されるポリペプチドと、少なくとも 90% 以上の同一性をもち、ヒト GP 抗原の N 末端領域に特異的に結合しセリン/スレオニンキナーゼ活性を持つポリペプチドを含む、実質的に精製されたポリペプチド。

10

【請求項 9】

配列番号：2、配列番号：4、配列番号：6、配列番号：8、配列番号：10 および配列番号：12 からなる群から選択されるポリペプチドを含む、実質的に精製されたポリペプチド。

【請求項 10】

請求項 8 または 9 のいずれかに記載の実質的に精製されたポリペプチドに選択的に結合する抗体。

【請求項 11】

抗体がポリクローナル抗体である請求項 10 に記載の抗体。

20

【請求項 12】

抗体がモノクローナル抗体である請求項 10 に記載の抗体。

【請求項 13】

配列番号：2、配列番号：4、配列番号：6、配列番号：8、配列番号：10 および配列番号：12 からなる群から選択されるポリペプチドと、少なくとも 90% 以上の同一性をもち、ヒト GP 抗原の N 末端領域に特異的に結合しセリン/スレオニンキナーゼ活性を持つタンパク質の存在を検出する方法であって；

a) スクリーニングするタンパク質サンプルを調製し；

b) スクリーニングするタンパク質サンプルを抗体 - 抗原複合体形成を促進する条件下で請求項 10 から 12 のいずれかの抗体と接触させること；および

30

c) 抗体 - 抗原複合体の存在が、配列番号：2、配列番号：4、配列番号：6、配列番号：8、配列番号：10 および配列番号：12 からなる群から選択されるポリペプチドと、少なくとも 90% 以上の同一性を持つタンパク質の存在を示す、抗体 - 抗原複合体形成の検出；

からなる方法。

【請求項 14】

検出が、免疫局在化、免疫蛍光分析、ウェスタンブロット分析、ELISA および核酸発現ライブラリースクリーニングからなる群から選択される方法を含む、請求項 13 に記載の方法。

40

【請求項 15】

配列番号：1、配列番号：3、配列番号：5、配列番号：7、配列番号：9、配列番号：11 および配列番号：25 からなる群から選択されるポリヌクレオチドに、高ストリンジエンシー条件下でハイブリダイズする、ヒト GP 抗原の N 末端領域に特異的に結合しセリン/スレオニンキナーゼ活性を持つタンパク質をコードする核酸をサンプル中で検出する方法であって、サンプルを高ストリンジエンシー条件下で請求項 1 から 4 のいずれかに記載の単離された核酸と接触させること、および複合体形成を検出することからなり、ここで複合体形成は、配列番号：1、配列番号：3、配列番号：5、配列番号：7、配列番号：9、配列番号：11 および配列番号：25 からなる群から選択されるポリヌクレオチドに、高ストリンジエンシー条件下でハイブリダイズする核酸の、サンプル中の存在を示

50

す方法。

【請求項 16】

ハイブリダイゼーション、逆転写、PCR、逆転写を組み合わせたPCR、ノーザンブロットリング、サザンブロットリング、およびDNAライブラリースクリーニングからなる群から選択される方法により検出を実施する請求項 15 に記載の方法。

【請求項 17】

患者の自己免疫状態を検出する方法であって、
患者からの組織または体液サンプルを調製し；
自己免疫状態が存在しない対照組織または体液サンプルを調製し；および
対照サンプルと比較しての組織または体液サンプル中のGPBP RNAまたはタンパク質発現の変化の検出であって、ここで対照と相対するGPBP RNAまたはタンパク質発現の変化は自己免疫状態の存在を示し、またこのときGPBP RNAは、配列番号：1、配列番号：3、配列番号：5、配列番号：7、配列番号：9および配列番号：11からなる群から選択される核酸と少なくとも90%以上一致する、ヒトGP抗原のN末端領域に特異的に結合しセリン/スレオニンキナーゼ活性を持つタンパク質をコードする核酸から発現し、GPBP タンパク質は、配列番号：2、配列番号：4、配列番号：6、配列番号：8、配列番号：10および配列番号：12からなる群から選択されるタンパク質と少なくとも90%以上一致しセリン/スレオニンキナーゼ活性を持つ；
ことを含む方法。

【請求項 18】

組織または体液サンプル中のアポトーシスまたは癌形質転換が行われている細胞を検出する方法であって、
患者からの組織または体液サンプルを調製し；
正常な対照組織または体液サンプルを調製し；および
対照サンプルと比較して組織または体液サンプル中のGPBP RNAまたはタンパク質発現の変化の検出であって、ここで対照と相対するGPBP RNAまたはタンパク質発現の変化がアポトーシスまたは癌形質転換が行われている細胞の存在を示し、このときGPBP RNAは、配列番号：1、配列番号：3、配列番号：5、配列番号：7、配列番号：9および配列番号：11からなる群から選択される核酸と少なくとも90%以上一致する、ヒトGP抗原のN末端領域に特異的に結合しセリン/スレオニンキナーゼ活性を持つタンパク質をコードする核酸から発現し、GPBP タンパク質は、配列番号：2、配列番号：4、配列番号：6、配列番号：8、配列番号：10および配列番号：12からなる群から選択されるタンパク質と少なくとも90%以上一致しセリン/スレオニンキナーゼ活性を持つ；
ことを含む方法。

【発明の詳細な説明】

【0001】

(相互参照)

本出願は1999年2月24日に提出された米国仮出願番号60/121483の優先権を請求する。

【0002】

(政府権限の記載)

本研究はComision Interministerial de Ciencia y Tecnologia (CICYT、スペイン)のPlan Nacional I+Dからの許可番号SAL91/0513、SAF94/1051およびSAF97/0065、Fondo de Investigaciones Sanitarias (FISs、スペイン)からの許可番号93/0343、並びにla Direcció General d'Ensenyaments Universitaris i Investigació (Comunitat Valenciana、スペイン)からの許可番号GV-3366/95、GV-C-VS-21-118-96により援助を

10

20

30

40

50

部分的に受けている；従ってスペイン政府は本発明の権利を保持する。

【0003】

(発明の分野)

本発明はプロテインキナーゼ、自己免疫疾患、アポトーシスおよび癌の分野に関する。

【0004】

(発明の背景)

グッドパスチャー(GP)病はヒトのみで報告されている自己免疫疾患である。GP患者では、IV型コラーゲン鎖の非コラーゲン性C末端ドメイン(NC1) (「グッドパスチャー抗原」)に対する自己抗体が急速な進行性糸球体腎炎およびしばしば肺出血を引き起こし、これはGP症候群の二つの重要な臨床発現である(1を参照のこと。このセクションでの参照番号は実施例1の参考文献リストに対応する)。

10

【0005】

一つ以上の自己免疫疾患を表す有意数の患者により、またこれらの疾患のいくつかが特異的MHCハプロタイプに対して示す強力な共通した連鎖により、少なくともいくつかの自己免疫障害では共通の発病原因が存在するという考えが示唆される(31、32)。自己免疫においては自己抗原が先導部分であり、自己抗原性である自己成分の数は限られている(31)という実験観察は、これらの自己成分が、免疫系による自己/非自己認識の重要な結果と生物学的特徴を共有することを示唆している。異なるが特異的である自己成分を変化させることにより、引き金となる事象が結果的に異常な抗原処理をする可能性がある。特定のMHC特異性を発現するある個体では、異常ペプチドが非耐性Tセルにより認識され、免疫応答を引き起こす(1)。

20

【0006】

我々は以前にGP抗原を探索し、自己免疫発病原因において関連性のある生物学的特徴を同定した。NC1ドメインは種間で、および異なるIV型コラーゲン鎖(1から6)間で高度に保存されているので(2)、天然の自己免疫応答におけるヒト3(IV)NC1の排他的な関与は、このドメインが発病原因に関連する構造的および/または生物学的特徴を有することを示唆している。これに一致して、ヒト抗原のN末端は高度に分岐しており、A型プロテインキナーゼの機能的リン酸化部位を確認する独特の5個の残基モチーフ(KRGDS⁹)を含有する(3、4)。さらに、ヒト3遺伝子は、別法としてスプライシングされ、リン酸化可能なN末端領域をも含有する複数の転写物を生じるが、その他のヒト関連または別の種の相同性遺伝子にはそれはない(5から7)。最近の研究ではcAMP依存性プロテインキナーゼによるGP抗原のN末端のリン酸化が別の生成物の存在により上方制御されることが示されている(以下の実施例3を参照のこと)。特異的セリンリン酸化およびプレmRNA代替スプライシングはまたアセチルコリンレセプターおよびミエリン塩基性タンパク質(MBP)を含有するその他の自己抗原の生物学とも関連する(4)。後者は免疫系が中枢神経系の白質を標的とする別の排他的ヒト自己免疫疾患である多発性硬化症(MS)における主要な抗原であると考えられる。GP病およびMSは同一のHLA-IIクラスIIタイプ(HLA-DRB1*1501)と強力な関連性を示すヒト疾患である(32、33)。このHLA特異性を担持するMS患者のGP病による死亡が最近報告されたのに加えて(34)、これはこれらのヒト疾患における共通の発病原因の存在を支持している。

30

40

【0007】

このように、特異的セリン/スレオニンリン酸化がヒトGP抗原、その他の種のGP抗原、およびその他のヒト(IV)鎖からの相同性ドメイン間の主要な生物学的差異であり、発病原因において重要であろう(1、4)。

【0008】

従って、ヒトGP抗原のN末端領域をリン酸化する特異的セリン/スレオニンキナーゼの同定および単離はGP症候群および恐らくその他の自己免疫障害の診断および処置に非常に有利であろう。

【0009】

50

(発明の開示)

本発明はヒトGP抗原の独特なN末端領域に特異的に結合し、リン酸化するセリン/スレオニンキナーゼの同定および単離に関して当該分野の必要性を満足させるものである。一つの態様では、本発明は種々の形態のグッドパスチャー抗原結合タンパク質(GPBP)をコードする核酸配列、およびGPBPコード化配列に機能的に結合した組換え発現ベクターを提供する。

【0010】

別の態様では、本発明は組換え発現ベクターでトランスフェクトされた宿主細胞を提供する。さらなる態様では、本発明は実質的に精製されたGPBPおよびGPBPに選択的に結合する抗体を提供する。さらに別の態様では、本発明はGPBPまたはGPBPをコードするの核酸の存在を検出する方法を提供する。

10

【0011】

さらなる態様では、本発明は自己免疫状態またはアポトーシスの存在を検出する方法であって、対照組織に比較して、組織におけるGPBPの発現の増加の検出からなる方法を提供する。

【0012】

別の態様では、本発明は自己免疫障害、アポトーシス、または腫瘍を処置するための方法および医薬用組成物であって、GPBPの発現または活性を改変する必要性を有する患者において、GPBPの発現または活性を改変することからなる方法および医薬用組成物を提供する。

20

【0013】

(発明の詳細な説明)

本発明に引用する全ての参考文献は出展明示により本明細書の一部とする。

【0014】

本明細書で用いる略語は：bp、塩基対；DTT、ジチオスレイトール；DMEM、ダルベッコ改変イーグル培地；EDTA、エチレンジアミン四酢酸；EGTA、エチレングリコールビス(アミノエチルエーテル)N,N,N',N'-四酢酸；GP、グッドパスチャー；rGP III、rGP III/IV/VおよびrGP V、各々スプライシングしたエキソンIII、エキソンIII, IVおよびVまたはエキソンVから得られたグッドパスチャー抗原の代替の形態を表す組換え物質；GPBPおよびrGPBP、元来のおよび組換えグッドパスチャー抗原結合タンパク質；GPBP 26およびrGPBP 26、GPBPの元来のおよび組換え代替形態；GST、グルタチオンS-トランスフェラーゼ；HLA、ヒトリンパ球抗原；HPLC、高速液体クロマトグラフィ；Kb、千塩基対；kDa、千ダルトン；MBP、rMBP、元来のおよび組換え21kDaミエリン塩基性タンパク質；MBP IIおよびrMBP II、エキソンIIをスプライシングして得られた元来のおよび組換え18.5kDaミエリン塩基性タンパク質；MBP VおよびMBP II/V、各々エキソンV並びにエキソンIIおよびVをスプライシングして得られたミエリン塩基性タンパク質の代替形態；MHC、主要組織適合性複合体；NC1、非コラーゲン性ドメイン；PH、プレクストリン相同性；PKA、cAMP依存性プロテインキナーゼ；PMSF、フッ化フェニルメチルスルフォニル；SDS-PAGE、ドデシル硫酸ナトリウムポリアクリルアミドゲル電気泳動；TBS、トリス緩衝生理食塩水：である。

30

40

【0015】

本出願では特記しない場合、利用される技術はいくつかの既知の参考文献例えばMolecular Cloning: A Laboratory Manual (Sambrookら、コールドスプリングハーバーラボラトリープレス(1989))、Gene Expression Technology (Methods in Enzymology、185巻、D. Goeddel編、アカデミックプレス、サンディエゴ、カリフォルニア州(1991))、Methods in Enzymologyの「Guide to Protein Purification」(M. P. Deutscher編

50

、アカデミックプレスインコーポレーティッド(1990)) ; PCR Protocols: A Guide to Methods and Applications (Innisら、アカデミックプレス、サンディエゴ、カリフォルニア州(1990))、Culture of Animal Cells: A Manual of Basic Technique、第2版(R. I. Freshney、リスインコーポレーティッド、ニューヨーク、ニューヨーク州(1987))、Gene Transfer and Expression Protocols、109-128頁、(E. J. Murray編、ザヒューマナプレスインコーポレーティッド、クリフトン、ニュージャージー州)、およびアンピオン1998年カタログ(アンピオン、オースティン、テキサス州)のいずれかに見出すことができる。

10

【0016】

本明細書で用いる「GPBP」なる用語はグッドパスチャー結合タンパク質を意味し、その単量体およびオリゴマーの両方を含む。ヒト(配列番号: 2)、マウス(配列番号: 4)、およびウシGPBP配列(配列番号: 6)を本明細書にて提供する。

【0017】

本明細書で用いる「GPBP 26」なる用語は配列番号: 14に示す26個のアミノ酸配列を削除したグッドパスチャー結合タンパク質を意味し、その単量体およびオリゴマーの両方を含む。ヒト(配列番号: 8)、マウス(配列番号: 10)、およびウシGPBP配列(配列番号: 12)を本明細書にて提供する。

【0018】

本明細書で用いる「GPBP pep 1」なる用語は配列番号: 14に示す26個のアミノ酸ペプチドを意味し、その単量体およびオリゴマーの両方を含む。

20

【0019】

本明細書で用いる用語「GP抗原」はIV型コラーゲンの3NC1ドメインを意味する。

【0020】

本明細書で用いる「MBP」はミエリン塩基性タンパク質を意味する。

【0021】

一つの態様では、本発明はGPBP、GPBP 26、およびGPBP pep 1並びにその変異体またはフラグメントをコードする単離された核酸を提供する。一つの態様では、単離された核酸は配列番号: 1、配列番号: 3、配列番号: 5、配列番号: 7、配列番号: 9、配列番号: 11、配列番号: 13、配列番号: 15、配列番号: 17、配列番号: 19、配列番号: 21、配列番号: 23もしくは配列番号: 25またはそのフラグメントと実質的に類似する配列を含む。

30

【0022】

別の態様では、本発明はGP抗原またはMBPの代替生成物をコードする単離された核酸を提供する。一つの態様では、単離された核酸は配列番号: 43および配列番号: 44に実質的に類似するペプチドをコードする配列を含む。

【0023】

本明細書で用いる「実質的に類似した」なる用語は、欠失、付加、または置換など(これらに限定するものではない)の、開示する配列からの保存または非保存変種の一つまたはそれ以上を有するDNAもしくはRNA、またはタンパク質のアミノ酸配列のヌクレオチドであって、得られた核酸および/またはアミノ酸配列は本明細書に開示する配列と機能的に同等であるヌクレオチド配列を意味する。機能的に同等な配列は実質的に同一の方法で機能して実質的に同一の本明細書に開示するタンパク質を製造する。例えば機能的に同等のDNAは本明細書に開示するタンパク質と同一であるか、または一つまたはそれ以上の保存アミノ酸変種、例えば別の非極性残基に関して非極性残基の置換、もしくは荷電した残基に関しては同様に荷電した残基の置換体を有するタンパク質をコードする。これらの変化には当業者によりタンパク質の3次構造を実質的に変化しない置換体として認識されている変化などがある。

40

50

【0024】

実際には、実質的に類似したなる用語は高ストリンジェンシーに対して中程度の条件下で互いにハイブリダイズする2個のタンパク質をコードするDNAを意味し、同一のアミノ酸配列を有するか、またはその構造もしくは機能を变化しない配列における変化を有するタンパク質をコードする。本明細書で用いる実質的に類似したヌクレオチドまたはアミノ酸配列は少なくとも約70%の同一性、より好ましくは少なくとも約80%の同一性、および最も好ましくは少なくとも約90%の同一性を共有する。しかしながら、スプライシング変種として生じる前記のレベルの相同性以下であるタンパク質（およびかかるタンパク質をコードするDNAまたはRNA）、または保存アミノ酸置換（または縮重コドンの置換）により改変されたものは本発明の範囲内であることを意図することは認識される。

10

【0025】

本明細書で用いるハイブリダイゼーションのストリンジェンシーは核酸ハイブリッドが安定である条件を意味する。当業者に既知であるように、ハイブリッドの安定性はハイブリッドの融点(T_M)に反映する。配列相同性が1%減少する毎に T_M はおよそ1から1.5低下する。概して、ハイブリッドの安定性はナトリウムイオン濃度および温度の相関関係である。典型的には、低ストリンジェンシー条件下でハイブリダイゼーション反応を行い、続いてより高度なストリンジェンシー条件に変化して洗浄する。ハイブリダイゼーションストリンジェンシーに対する言及はかかる洗浄条件に関係する。このように本明細書で用いる中程度のストリンジェンシーとは0.1%SSPE中37または55で安定したハイブリッドを形成するこれらの核酸配列がハイブリダイズできる条件を意味し、一方高度なストリンジェンシーとは0.1%SSPE中65で安定したハイブリッドを形成するこれらの核酸配列がハイブリダイズできる条件を意味する。これらの条件が種々のバッファーおよび温度を用いて繰り返すことができ、正確である必要はないことは理解されよう。デンハーツ溶液およびSSPE（例えばSambrook、FritschyおよびManiatis、Molecular Cloning; A Laboratory Manual、コールドスプリングハーバーラトリブレス（1989））はその他の適当なハイブリダイゼーションバッファーとして当業者に既知である。

20

【0026】

単離された核酸配列は一つまたはそれ以上のイントロンを有するRNA、cDNAまたはゲノムクローンを含んでもよい。単離された配列はさらにポリA配列、改変コザック配列、およびエピトープタグをコードする配列、搬出シグナル、および分泌シグナル、核局在化シグナル、および原形質膜局在化シグナルなど（これらに限定するものではない）のコードされたタンパク質の発現および/または精製を促進するのに有用なさらなる配列を含んでもよい。

30

【0027】

別の態様では、本発明はGPBP、GPBP 26またはGPBP pep 1、およびその変異体またはフラグメントを発現する核酸配列を含む組換え発現ベクターを提供する。一つの態様では、ベクターは配列番号：1、配列番号：3、配列番号：5、配列番号：7、配列番号：9、配列番号：11、配列番号：13、配列番号：15、配列番号：17、配列番号：19、配列番号：21、配列番号：23もしくは配列番号：25に示す配列またはそのフラグメントと実質的に類似する核酸配列を含む。

40

【0028】

別の態様では、本発明は配列番号：43、配列番号：44に示すアミノ酸配列またはそのペプチドフラグメントに実質的に類似するペプチドを発現する核酸配列を含む組換え発現ベクターを提供する。

【0029】

「組換え発現ベクター」には核酸コーディング領域または遺伝子を遺伝子生成物の発現に影響できるいずれかのプロモーターに機能的に連結するベクターをなどがある。哺乳動物系において開示した核酸配列の発現を誘導するのに用いられるプロモーター配列は構成性(CMV、SV40、RSV、アクチン、EFなどの種々のプロモーターのいずれかに

50

より誘導されるが、これらに限定するものではない)かまたは誘導性(テトラサイクリン、エクジソン、ステロイド応答性などの多くの誘導性プロモーターのいずれかにより誘導されるが、これらに限定するものではない)でよい。原核細胞をトランスフェクトするのに用いられる発現ベクターの構築はまた当業者に既知であり、従って標準的な技術により達成できる。(例えば Sambrook、Fritsch および Maniatis、Molecular Cloning; A Laboratory Manual、コールドスプリングハーバーラボラトリープレス(1989); E. J. Murray 編、Gene Transfer and Expression Protocols、109-128頁、ザヒューマナプレスインコーポレーティッド、クリフトン、ニュージャージー州)、アンピオン1998年カタログ(アンピオン、オースチン、テキサス州を参照のこと)。

10

【0030】

発現ベクターはエピソームとしてかまたは宿主染色体DNAに組み込むことにより宿主生物内で複製可能でなければならない。好ましい態様では、発現ベクターはプラスミドを含む。しかしながら、本発明は同等の機能を提供するその他の発現ベクター、例えばウイルスベクターを含むことを意図する。

【0031】

さらなる態様では、本発明は本明細書に開示する組換え発現ベクターでトランスフェクトされた宿主細胞を提供し、ここで宿主細胞は原核細胞かまたは真核細胞でよい。細胞を一過性でまたは安定してトランスフェクトできる。当業者に既知のいずれかの技術、例えば標準細菌性形質転換、リン酸カルシウム共沈、エレクトロポレーション、またはリポソーム媒介、DEAEデキストラン媒介、ポリカチオン媒介またはウイルス媒介トランスフェクションにより(これらに限定するものではない)発現ベクターの原核および真核細胞へのトランスフェクションを達成できる。(例えばMolecular Cloning; A Laboratory Manual (Sambrookら(1989)、コールドスプリングハーバーラボラトリープレス); Culture of Animal Cells: A Manual of Basic Technique、第2版(R. I. Freshney(1987)、リズインコーポレーティッド、ニューヨーク、ニューヨーク州)を参照のこと。)

20

【0032】

さらに別の態様では、本発明は実質的に精製されたGPBP、GPBP 26およびGPBP pep1並びにその変異体またはフラグメントを提供する。一つの態様では、実質的に精製されたタンパク質のアミノ酸配列は配列番号: 2、配列番号: 4、配列番号: 6、配列番号: 8、配列番号: 10、配列番号: 12、配列番号: 14、配列番号: 16、配列番号: 18、配列番号: 20、配列番号: 22、配列番号: 24またはそのペプチドフラグメントに実質的に類似している。

30

【0033】

別の態様では、本発明はGP抗原およびMBPの実質的に精製された代替生成物を提供する。一つの態様では、実質的に精製されたポリペプチドのアミノ酸配列は配列番号: 43、配列番号: 44またはそのペプチドフラグメントに実質的に類似している。

40

【0034】

本明細書で用いる「実質的に精製された」なる用語はそのin vivo細胞環境から分離されているタンパク質を意味する。従って、タンパク質を天然供給源から精製できるか、または組換えタンパク質を前記で開示したトランスフェクトされた宿主細胞から精製できる。好ましい態様では、前記で開示したトランスフェクトされた細胞によりタンパク質を製造し、標準的な技術を用いて精製する。(例えばMolecular Cloning; A Laboratory Manual (Sambrookら(1989)、コールドスプリングハーバーラボラトリープレス)を参照のこと。)このように原核または真核生物供給源からタンパク質を精製できる。さらに好ましい種々の態様において、タンパク質を細菌、酵母または哺乳動物細胞から精製する。

50

【 0 0 3 5 】

タンパク質は、タンパク質の精製を促進するのに有用なさらなる配列、例えばエピトープタグおよび輸送シグナルを含んでもよい。かかるエピトープタグの例としては、FLAG (シグマケミカルズ、セントルイス、ミズーリー州)、myc (9 E 1 0) (インビトロゲン; カールスバッド、カリフォルニア州)、6 - His (インビトロゲン、ノバゲン、マジソン、ウィスコンシン州) およびHA (ベーリンガーマンハイムバイオケミカルズ) などがあるが、これらに限定するものではない。かかる輸送シグナルの例としては搬出シグナル、分泌シグナル、核局在化シグナルおよび原形質膜局在化シグナルなどがあるが、これらに限定するものではない。

【 0 0 3 6 】

別の態様では、本発明はGPBP、GPBP 26またはGPBP pep 1に選択的に結合する抗体を提供する。一つの態様では、抗体は配列番号：2、配列番号：4、配列番号：6、配列番号：8、配列番号：10、配列番号：12、配列番号：14、配列番号：16、配列番号：18、配列番号：20、配列番号：22、配列番号：24またはそのペプチドフラグメントからなる群から選択される配列を含むタンパク質に選択的に結合する。宿主動物を完全GPBPでか、またはその抗原性ペプチドで免疫することによりかかる抗体を製造できる。抗体はポリクローナル、またはモノクローナル抗体でよい。

【 0 0 3 7 】

別の態様では、本発明は配列番号：43、配列番号：44、配列番号：46、配列番号：48、配列番号：50、配列番号：54またはその抗原性フラグメントからなる群から選択される配列に実質的に類似したアミノ酸配列を含むポリペプチドに選択的に結合する抗体を提供する。

【 0 0 3 8 】

既知の方法、例えばHarlowおよびLane、Antibodies ; Laboratory Manual、コールドスプリングハーバーラボラトリー、コールドスプリングハーバー、ニューヨーク(1988)に記載の方法により抗体を製造できる。一例を挙げると、最初の免疫の前に前免疫血清を収集する。実質的に精製した本発明のタンパク質またはその抗原性フラグメントを適当なアジュバントと共に、免疫応答を引き出すのに十分な量および間隔で動物に注射する。一定の間隔で、好ましくは週毎に動物を出血させ、抗体力価を決定する。1次免疫後動物にブースター注射をしてもしなくてもよい。各ブースター免疫後約7日に、または1回免疫後約毎週、動物を出血させ、血清を収集し、一定分量を約-20 で保存する。次いで動物から収集した血清を、GPBP関連タンパク質を含まない同一の発現系から製造した非抗原性関連タンパク質が結合するカラムに通すことにより、本発明のタンパク質およびペプチドに対するポリクローナル抗体を直接的に精製できる。

【 0 0 3 9 】

動物から脾臓細胞を得ることによりモノクローナル抗体を製造できる(KohlerおよびMilstein、Nature, 256:495-497(1975)を参照のこと)。一例では、近交系マウスを本発明のタンパク質もしくはペプチドまたはその抗原性フラグメントで免疫することにより目的のモノクローナル抗体(mAb)を製造する。腹腔内(IP)または皮下(SC)経路により、免疫応答を引き出すのに十分な量および間隔でマウスを免疫する。0日にマウスに1次免疫し、約3から約30週休ませる。免疫したマウスに静脈内(IV)経路により1回またはそれ以上ブースター免疫する。当業者に既知の標準的な方法により免疫マウスから脾臓を除去することにより、抗体陽性マウスからリンパ球を入手する。安定したハイブリドーマを形成できる条件下で適当な融合パートナーと脾臓リンパ球を混合することによりハイブリドーマ細胞を製造する。抗体製造細胞および融合パートナー細胞をポリエチレングリコール中約30%から約50%の濃度で融合させる。当業者に既知の方法により、ヒポキサンチン、チミジンおよびアミノプテリンを補充したダルベッコ改変イーグル培地(DMEM)中成長させることにより融合したハイブリドーマを選別する。成長陽性ウェルから上澄液を収集し、固相イムノラジオアッセイ

10

20

30

40

50

のごときイムノアッセイにより抗体生成に関してスクリーニングする。MacPhersonの軟寒天技術(Tissue Culture Methods and Applications、KruseおよびPaterson編、アカデミックプレス(1973))のごとき技術により抗体陽性ウェルからのハイブリドーマ細胞をクローニングする。

【0040】

かかる抗体応答を生じるために、典型的には本発明のタンパク質を医薬上許容される非経口投与用担体と共に処方する。かかる許容されるアジュバントにはフロイント完全、フロイント不完全、アルミニウム沈殿、コリネバクテリウムパルヴムおよびtRNAを含有する油中水乳剤などがあるが、これらに限定するものではない。ポリペプチドの濃度およびペヒクルおよびその他の構成成分の選択などのかかる組成物の処方 は当業者の範囲内である。

10

【0041】

本明細書で用いる抗体なる用語は本発明のタンパク質およびペプチドまたはそのフラグメントと選択的に反応するその抗体フラグメントを含むことを意図する。慣用的な技術を用いて抗体をフラグメント化することができ、そのフラグメントを抗体全体に関して前記したのと同じの方法で有用性に関してスクリーニングできる。例えば抗体をペプシン消化することによりF(ab')₂フラグメントを生じることができる。得られたF(ab')₂フラグメントを反応させてジスルフィド結合を還元して、Fab'フラグメントを作ることができる。

20

【0042】

さらなる態様では本発明は、タンパク質サンプル中の本発明のタンパク質またはペプチドの存在を検出するための方法であって、スクリーニングするタンパク質サンプルを調製し、スクリーニングするタンパク質サンプルを本発明のタンパク質またはペプチドに対する抗体と接触させること、および抗体-抗原複合体の形成を検出することからなる方法を提供する。抗体は前記するポリクローナルまたはモノクローナルのいずれかであり、モノクローナル抗体が好ましい。本明細書で用いる「タンパク質サンプル」なる用語は本発明のタンパク質またはペプチドおよびそのフラグメントを含有し得るいずれかのサンプルを意味し、組織およびその部分、組織切片、完全な細胞、細胞抽出物、精製または部分的に精製したタンパク質サンプル、体液、核酸発現ライブラリーなどがあるが、これらに限定するものではない。従って、本発明のこの態様を用いて、標準的な技術、例えば免疫局在化、免疫蛍光分析、ウェスタンブロット分析、ELISAおよび核酸発現ライブラリースクリーニングなど(これらに限定するものではない)によりこれらの種々タンパク質サンプル中のGPBP、GPBP₂₆、GPBP_{pep1}、またはGP抗原の代替生成物の存在に関して試験できる(例えばSambrookら(1989)を参照のこと)。一つの具体例では、その技術により目的のタンパク質またはペプチドの存在または不在のみを決定できる。別法として、その技術により定量し、サンプル中の目的のタンパク質またはペプチドの相対量についての情報を提供できる。定量目的ではELISAが好ましい。

30

【0043】

標準的な検出技術により本発明のタンパク質もしくはペプチド、またはそのフラグメントおよびその抗体またはそのフラグメント間の免疫複合体形成を検出できる。例えば標識抗体または2次抗体を用いることにより免疫複合体を検出できる。標識の選択などのかかる方法は当業者に既知である。(HarlowおよびLane、前記で引用)。別法として、ポリクローナルまたはモノクローナル抗体を検出可能な物質に結合できる。用語「結合した」は検出可能な物質が物理的に抗体に連結しているという意味で用いる。適当な検出可能な物質には種々酵素、補欠分子団、蛍光性物質、発光性物質、および放射性物質などがある。適当な酵素の例としては、西洋ワサビペルオキシダーゼ、アルカリ性ホスファターゼ、ガラクトシダーゼ、またはアセチルコリンエステラーゼなどがある。適当な補欠分子団複合体の例としてはストレプトアビジン/ビオチンおよびアビジン/ビオチンなどがある。適当な蛍光性物質の例としてはアンペリフェロン、フルオレセイン、フルオレセ

40

50

インイソチオシアネート、ローダミン、ジクロロトリアジニルアミンフルオレセイン、塩化ダンシルまたはフィコエリトリンなどがある。発光性物質の例としてはルミノールなどがある。適当な放射性物質の例としては¹²⁵I、¹³¹I、³⁵Sまたは³Hなどがある。

【0044】

かかる検出方法は種々の目的で有用であり、自己免疫状態の検出、標的とされるまたはアポトーシスが起きている細胞の同定、組織サンプル中の目的のタンパク質の免疫局在化、ウェスタンブロット分析、関連タンパク質を見出すための発現ライブラリーのスクリーニングなどがあるが、これらに限定するものではない。

【0045】

さらに別の態様では、本発明はサンプル中のGPBP、GPBP 26、GPBP pep 1、またはGP抗原の代替生成物をコードする核酸配列の存在を検出する方法であって、スクリーニングする核酸サンプルを調製し、サンプルを本発明の単離された核酸配列またはそのフラグメントから誘導した核酸プローブと接触させること、および複合体形成を検出することからなる方法を提供する。

【0046】

本明細書で用いる用語「サンプル」はGPBP関連核酸を含有し得るいずれかのサンプルを意味し、組織およびその一部、組織切片、完全な細胞、細胞抽出物、精製または部分的に精製した核酸サンプル、DNAライブラリー、並びに体液などがあるが、これらに限定するものではない。従って、本発明のこの態様を用いて、切片上ハイブリッド形成法、ノーザンブロットティング、サザンブロットティング、DNAライブラリースクリーニング、ポリメラーゼ連鎖反応(PCR)または逆転写PCR(RT-PCR)など(これらに限定するものではない)の標準的な技術により、これらの種々サンプル中のGPBP mRNAまたはDNAの存在に関して試験できる。(例えばSambrookら(1989)を参照のこと。)一つの具体例では、この技術により目的の核酸の存在または不在のみを決定できる。別法として、この技術により定量でき、サンプル中の目的の核酸の相対量に関する情報を提供できる。定量目的では、定量用PCRおよびRT-PCRが好ましい。従って、一例では、RNAをサンプルから単離し、目的の核酸配列から誘導したオリゴヌクレオチドと接触させ、逆転写酵素と共に適当なバッファーおよび温度条件下でGPBP関連RNAからcDNAを製造する。次いでcDNAを目的の核酸配列から誘導したプライマー対を用いるPCRに供する。好ましい態様では、配列番号：5のRNA発現生成物の存在を検出するようにプライマーを設計し、サンプル中のGPBP遺伝子発現の量を対照サンプルのレベルと比較する。

【0047】

目的の核酸配列を検出するために、標準的な標識技術を用いてプローブ、目的の核酸、またはプローブおよび目的の核酸間の複合体を標識でき、その標識技術には放射性、酵素、化学ルミネサンスまたはアビジンもしくはビオチン標識技術などがあるが、これらに限定するものではなく、これらは全て当業者に既知である。(例えばMolecular Cloning: A Laboratory Manual (Sambrookら、(1989)、コールドスプリングハーバートラクトプレス)、Gene Expression Technology (Methods in Enzymology、185巻、D. Goeddel編(1991)、アカデミックプレス、サンディエゴ、カリフォルニア州); PCR Protocols: A Guide to Methods and Applications (Innisら、(1990)、アカデミックプレス、サンディエゴ、カリフォルニア州)を参照のこと。)

【0048】

かかる核酸検出法は自己免疫状態の診断、標的になるまたはアポトーシスが行われている細胞の同定、切片上ハイブリッド形成法、ノーザンおよびサザンブロット分析およびDNAライブラリースクリーニングなど(これらに限定するものではない)の種々の目的に有用である。

10

20

30

40

50

【0049】

以下の実施例で示すように、GPBPは天然発生自己免疫応答において一般に標的とされる組織構造で優先的に発現することが示され、グッドパスチャー症候群(GP)、全身性エリテマトーデス(SLE)および扁平苔癬など(これらに限定するものではない)のいくつかの自己免疫状態において高度に発現する。さらに、後記する類似の実験研究法に従ってGP病(3IV型コラーゲン)、SLE(P1リボソームリントアンパク質およびSm-D1小型核リボヌクレオタンパク質)および皮膚筋炎(ヒスチジル-tRNAシターゼ)の自己抗原を表す組換えタンパク質が*in vitro*でGPBPの基質であることが示された。

【0050】

このように、好ましい態様では、GPBP発現の検出を用いて自己免疫状態を検出する。GPBP発現に関して、試験しているサンプルを対照サンプルと比較し、ここでGPBP発現レベルの上昇は自己免疫状態の存在を示す。この具体例では、これはGPBPには存在するが、GPBP 26には存在しないGPBP p e p 1に選択的に結合する抗体を用いるのが好ましい。

【0051】

さらに、添付の実施例に示すように、GPBPは腫瘍の培養細胞株において下方制御され、データはGPBP / GPBP 26がアポトーシスを誘起する細胞シグナル発生経路に参与している可能性があり、これは自己免疫の病理発生過程において上方制御され、腫瘍成長中に形質転換された細胞を攻撃する自己免疫攻撃を防御するために細胞を形質転換中には下方制御される。このように、本明細書に開示する検出方法を用いて標的化されたまたはアポトーシスが行われている細胞を検出することができる。

【0052】

別の態様では、本発明は自己免疫障害、腫瘍、を処置する、またはGPBP、GPBP 26もしくはGPBP p e p 1に実質的に類似するポリペプチドを含むタンパク質の発現または活性を改変することからなる細胞アポトーシスを防御する必要がある患者においてかかる処置または防御をするための方法を提供する。GPBP、GPBP 26もしくはGPBP p e p 1に実質的に類似するポリペプチドを含むタンパク質の発現または活性を改変することは、GPBP発現もしくは活性の特異的誘発因子もしくは阻害因子、GPBP抗体、GPもしくはミエリン塩基性タンパク質代替生成物を用いる遺伝子もしくはタンパク質治療、GPもしくはミエリン塩基性タンパク質代替生成物を発現する宿主細胞を用いる細胞治療、または当業者に既知のその他の技術を用いて達成できる。好ましい具体例では、方法はさらにGP抗原またはMBPの、実質的に精製された代替生成物を投与し、GPBP、GPBP 26もしくはGPBP p e p 1に実質的に類似するポリペプチドを含むタンパク質の発現または活性を改変することを含む。本明細書で用いる「発現または活性を改変すること」とはRNAまたはタンパク質生成物のいずれかの発現または活性を改変することを意味する。

【0053】

別の態様では、本発明はGPBP RNAまたはタンパク質の発現または活性を改変するための、実質的に精製された有効量のGP抗原またはMBPの代替生成物および医薬上許容される担体を含む医薬用組成物を提供する。

【0054】

投与用に、活性物質を通常指示された投与経路に適当な一つまたはそれ以上のアジュバントと組み合わせる。化合物をラクトース、スクロース、デンプン粉、アルカノン酸のセルロースエステル、ステアリン酸、タルク、ステアリン酸マグネシウム、酸化マグネシウム、リン酸および硫酸のナトリウムおよびカルシウム塩、アカシア、ゼラチン、アルギン酸ナトリウム、ポリビニルピロリドンおよび/またはポリビニルアルコールと混合でき、慣用される投与用に錠剤化またはカプセル化できる。別法として、本発明の化合物を生理食塩水、水、ポリエチレングリコール、プロピレングリコール、カルボキシメチルセルロースコロイド溶液、エタノール、コーン油、落花生油、綿実油、ゴマ油、トラガントゴム、

10

20

30

40

50

および/またはその他のバッファーに溶解できる。その他のアジュバントおよび投与様式は医薬の分野で既知である。担体または希釈剤には遅延物質、例えばモノステアリン酸グリセリル、ジステアリン酸グリセリル単独もしくはワックスと共に、または当業者に既知のその他の物質などがある。

【0055】

本発明は説明のみを目的とする添付の実施例を参考としてよりよく理解されるであろうが、本明細書に添付の特許請求の範囲で定義するように本発明の範囲を制限すると解釈すべきではない。

【0056】

(実施例1: GPBPの特徴づけ)

ここでヒトGP抗原の独特のN末端領域に特異的に結合し、リン酸化する新規の型のセリン/スレオニンキナーゼのクローニングおよび特徴づけについて報告する。

【0057】

材料および方法

合成重合体 - ペプチド。ヒトGP抗原の3から23残基を表すGPpep1、KGKRGDSGSPATWTTRGFVFT(配列番号: 26)、およびSer⁹からAla⁹へのその変異体であるGPpep1Ala⁹KGKRGDA⁹SGSPATWTTRGFVFT(配列番号: 27)をメッドプローブおよびチロンFLAGペプチド(シグマより入手)により合成された。

【0058】

オリゴヌクレオチド。以下は、いくつかのその他のGPBP特異的オリゴヌクレオチドと同様、ゲノシスおよびGIBCO-BRLにより合成された:

ON-GPBP-54m:

TCGAATTTCACCATGGCCCCACTAGCCGACTACAAGGACGACGATGACAAAG(配列番号: 28)

ON-GPBP-55c:

CCGAGCCCCGACGAGTTCCAGCTCTGATTATCCGACATCTTGTCATCGTTCG(配列番号: 29)

ON-HNC-B-N-14m: CGGGATCCGCTAGCTAAGCCAGGCAAGGATGG(配列番号: 30)

ON-HNC-B-N-16c: CGGGATCCATGCATAAATAGCAGTCTGCTGT(配列番号: 31)

【0059】

ヒトGPBPをコードするcDNAクローンの単離および特徴づけ - いくつかのヒト - g t 1 1 cDNA発現ライブラリー(眼、胎児および成人肺、腎臓およびHeLa S3、クロンテックより)をGPpep1と相互作用するタンパク質をコードするcDNAに関してプローブする。標準的な免疫スクリーニング法に従って調製したニトロセルロースフィルター(ミリポア)を遮断し、GPpep1 mlあたり1から10ナノモルと共に37でインキュベートした。特異的に結合したGPpep1をM3/1Aモノクローナル抗体を用いて検出した(7)。HeLa誘導ライブラリーで1個のクローンを同定した(HeLa1)。融合タンパク質結合の特異性を組換え真核ヒトGP抗原への類似の結合により確認した。HeLa1(0.5kb)のEcoRI cDNAインサートを用いて同一のライブラリーをさらにスクリーニングし、重複cDNAを単離した。HeLa1(n4')の全cDNAを含有する最大のcDNA(2.4kb)全長を配列決定した。

【0060】

ノーザンおよびサザンプロット - 製造者の指示書に従って、予め作成したノーザンおよびサザンプロット(クロンテック)をHeLa1 cDNAでプローブした。

【0061】

組換えタンパク質のプラスミド構築、発現および精製 - GPBP誘導した物質。元来の - g t 1 1 HeLa1クローンを大腸菌(E. coli) Y1089中溶原として発現

10

20

30

40

50

した(8)。A P T Gアガロースカラム(ベーリンガー)を用いて、細胞可溶化液から G P B PのN末端150残基を含有する対応するガラクトシダーゼ誘導融合タンパク質を精製した。n4'のE c o R I 2.4kbフラグメントをブルースクリプトM13+ベクター(ストラッタジーン)(B S - n4')中サブクローニングし、増幅し、続いてクローニングに使用した。E c o R I制限部位、翻訳開始のための標準コザックコンセンサス、タグペプチド配列をコードする領域(F L A G、D Y K D D D D K(配列番号:32))、および予測M e t_iおよびB a n I I制限部位を含むG P B Pの最初の11個の残基をコードする配列、(5'から3'で)含有するDNAフラグメントをO N - G P B P - 54mおよびO N - G P B P - 55cをハイブリダイズし、修飾T₇ DNAポリメラーゼ(アマシャム)で伸長することにより得た。得られたDNA生成物をE c o R IおよびB a n I Iで消化し、p H I L - D 2(インビトロゲン)のE c o R I部位におけるB S - n4'のB a n I I / E c o R I c DNAフラグメントとライゲートし、p H I L - F L A G - n4'を製造した。このプラスミドを用いてP i c h i a p a s t o r i sのG S 115株のM u t 5形質転換体を得、慣用的な液体培養かまたは発酵法(P i c h i a発現キット、インビトロゲン)のいずれかによりF L A Gタグ化組換えG P B P(r G P B P)を発現した。細胞可溶化液を抗F L A G M 2カラム(シグマ)に負荷し、非結合性物質をトリス緩衝生理食塩水(T B S、50mM トリスH C l、p H 7.4、150mM N a C l)または塩を補充したT B S(2M N a C lまで)で洗い出し、組換え物質をF L A Gペプチドと共に溶出した。培養ヒト腎臓誘導293細胞(A T C C 1573 - C R L)中に発現するために、B S - n4'またはp H I L - F L A G - n4'のいずれかの2.4または2.0kb E c o R I c DNAインサートをp c D N A 3(インビトロゲン)にサブクローニングし、p c - n4'およびp c - F L A G - n4'を各々製造した。一過性の発現のために用いる場合、トランスフェクションの8時間後、可溶化液を各々S D S - P A G Eに用いるのかまたはF L A G精製に用いるのかに依存して0.1% S D S含有または不含で、3.5から4μl/cm²の冷却溶解バッファー(T B S中1% ノニデットP - 40またはトリトンX 100、5mM E D T Aおよび1mM P M S F)を用いて細胞を溶解した。F L A G精製用には、6個に対して4個の175cm²培養皿の可溶化液を溶解バッファーで50mlまで希釈し、前記のように精製した。安定発現用には、p c - n4'で細胞を同様にトランスフェクトし、800μg/mlのG 4 18で3週間選別した。細菌組換え発現用には、p H I L - F L A G - n4'の2.0kb E c o R I c DNAフラグメントをp G E X - 5 x - 1(ファルマシア)のグルタチオンSトランスフェラーゼ(G S T)コード化c D N Aの下流でフレーム内クローニングした。得られた構築物を用いてD H 5細胞中にG S T - G P B P融合タンパク質を発現した(9)。

【0062】

G P抗原誘導物質。O N - H N C - B - N - 14mとO N - H N C - B - N - 16cの間で3特異的c D N Aを含有するp R c / C M V - B M 40発現ベクターを用いて293細胞中にヒト組換えG P抗原(r G P)を生成した。発現ベクターは、B i l l y G . H u d s o n(カンサス大学医療センター)により提供された、開始M e t、タグペプチド配列(F L A G)に続くB M 40シグナルペプチドおよびポリリンカークローニング部位をコードするc D N Aを含有するp R c / C M V(インビトロゲン)誘導ベクターである。3特異的c D N Aを得るために、前記のオリゴヌクレオチドおよび鋳型(クローンC 2)として前記で報告した3(I V)c D N A配列(3)を含有するプラスミドを用いてポリメラーゼ連鎖反応を実施した。r G Pの安定発現のために293細胞を得られた構築物(f 3 V L C)でトランスフェクトし、400μg/mlのG 4 18で選別した。回収したr G Pを抗F L A G M 2カラムを用いて精製した。

【0063】

全ての構築物を制限マッピングおよびヌクレオチド配列決定により確認した。

【0064】

細胞培養およびDNAトランスフェクション - ヒト293細胞を10% ウシ胎児血清を

補充したダルベッコ改変イーグル培地 (DMEM) 中で成長させた。プロフェクション哺乳類トランスフェクションシステム (プロメガ) のリン酸カルシウム沈殿法を用いてトランスフェクションを実施した。G418に対する抵抗性により安定してトランスフェクトされた細胞を選別した。生存細胞群の中心部分を単離し、クローン化しおよび増幅した。

【0065】

抗体生成 - GPBPのN末端領域に対するポリクローナル抗体。溶原としてとしてHeLa 1-gt11を発現する細胞をラエムリサンプルバッファの存在下で超音波処理により溶解し、7.5%のアクリルアミド調製用ゲル中電気泳動に供した。ゲルをクーマシブルーで染色し、目的の融合タンパク質を含有するバンドを切除し、ウサギ免疫に用いた(10)。APTGAフィニティ精製抗原を用いて反応性に関して抗血清を試験した。アフィニティ精製抗体を得るために、抗血清をTBSで1:5に希釈し、共有結合したアフィニティ精製抗原を含有するセファロース4Bカラムに負荷した。結合物質を溶出し、特記しない場合、免疫化学研究に用いた。

10

【0066】

GPBPに対するモノクローナル抗体。基本的には前記のように(7)GST-GPBPを用いてモノクローナル抗体を製造した。rGPBPに対する抗体に関して個々のクローンの上澄を分析した。

【0067】

in vitroリン酸化アッセイ - タンパク質基質0.5から1 μ gまたは合成ペプチド10ナノモルの存在下または不在下、25mM グリセロホスフェート(pH7.0)、0.5mM EDTA、0.5mM EGTA、8mM MgCl₂、5mM MnCl₂、1mM DTTおよび0.132 μ M ³²P-ATP中、全量50 μ lでrGPBP約200ngを30で一晚インキュベートした。

20

【0068】

in vivoリン酸化アッセイ - 24ウェル皿の個々のウェルに、正常のまたは安定にpc-n4'トランスフェクトした293細胞を播種した。細胞が望ましい密度に成長したとき、正常293細胞の多くのウェルをpc-FLAG-n4'でトランスフェクトした。12時間後、培養培地を除去し、リン酸不含DMEM100 μ l中20 μ Ci/ウェルのH₃³²PO₄を加え、4時間インキュベーションを続けた。1% トリトンX-100、2mM EDTA、1mM PMSF、50mM NaFおよび0.2mM バナジン酸塩を含有するTBS 300 μ l/ウェルで細胞を溶解し、特異的抗体およびプロテインA-セファロースで抽出した。抗GPBP血清を用いた場合、前免疫血清およびプロテインA-セファロースを用いて可溶化液を予め清澄化した。

30

【0069】

rGPBPのin vitro脱リン酸化 - 10mM 酢酸トリス(pH7.5)、10mM 酢酸マグネシウムおよび50mM 酢酸カリウム100 μ l中rGPBP1 μ gを0.85単位のウシ腸管アルカリ性ホスファターゼ(ファルマシア)で30で30分間脱リン酸化した。

【0070】

再生アッセイ - 前記するようにrGPBP1から5 μ gを用いてインプロット再生アッセイを実施した(11)。

40

【0071】

ヌクレオチド配列分析 - []³⁵S-dATP、修飾T₇ DNAポリメラーゼ(アマシャム)およびユニバーサルまたはGPBP特異的プライマーを用いてジデオキシ鎖終止法によりcDNA配列分析を実施した(8から10)。

【0072】

³²Pホスホアミノ酸分析 - 免疫精製したrGPBPまたは目的の物質を含有するそのHPLCゲル濾過分画をリン酸化し、加水分解し、1次元(4)または2次元薄層クロマトグラフィ(12)で分析した。2次元分析を実施する場合、第1の次元のためのバッファは葉酸:酢酸:水(1:3.1:35.9)(pH1.9)であり、第2の次元のため

50

のバッファーは酢酸：ピリジン：水（2：0.2：37.8）（pH 3.5）であった。アミノ酸をニンヒドリンで現し、³²Pホスホアミノ酸をオートラジオグラフィにより現した。

【0073】

物理学的方法および免疫化学技術 - (4) のように SDS - PAGE およびウェスタンブロットングを実施した。ABCペルオキシダーゼ法を用いて、ヒト多重組織対照スライド（バイオメディア、バイオジェネックス）で免疫化学研究を行った（13）。

【0074】

コンピューター分析 - 相同性はNCBIサーバーでBLAST 2.0（14）を用いてジーンバンクおよびスイスプロットデータベースに対して、およびゲノミックリサーチサーバーのためのインスティテュートを用いて発現配列タグに関してTIGRヒューマンジーンインデックスデータベースに対して検索した。スイスインスティテュートオブバイオインフォマティクスでプロフィールスキャンプログラムを用いて、機能的なパターンおよびプロフィールに関する検索をPROSITEデータベースに対して行った（15）。21および28の両方の残基のウィンドウを有するプログラムコイル（16）を用いて実験的癌研究に関してスイスインスティテュートで高次コイル構造の予測を行った。

【0075】

結果

GPBPの分子クローニング - ヒトGP抗原の分岐したN末端領域と特異的に相互作用するタンパク質を検索するために、この領域および隣接配列を包含する21残基ペプチド（GPpep1；配列番号：26）、並びにそれに対する特異的モノクローナル抗体を組み合わせいくつかのヒトcDNA発現ライブラリーをスクリーニングした。5 × 10⁶以上のファージをスクリーニングし、抗体結合を妨害せずにGPpep1と特異的に相互作用する融合タンパク質をコードする単一のHeLa誘導組換え体を同定した。

【0076】

元来のクローンのcDNAインサート（HeLa1）を用いて5'非翻訳配列の408bp、624残基をコードする1872bpのオープンリーディングフレーム（ORF）、および3'非翻訳配列の109bpを含有する2.4kb cDNA（n4'）を単離した（図1）（配列番号：1から2）。その他の構造的特徴が興味深い。第1に予測されるポリペプチド（本明細書で以後GPBPと称する）は多くのリン酸化可能な（17.9%）および酸性（16%）残基を有し、配列に沿って偏って分布している。最も豊富にある残基である（9.3%）セリンはタンパク質の2個の短い領域に偏好性を示し、これは残りのポリペプチド鎖全体で平均7%以下であるのに比較して、ここではアミノ酸のほぼ40%からなる。N末端の、よりセリンに富む領域は主にSer - Xaa - Yaa反復を含むことも注目値する。酸性残基はポリペプチドのN末端の4分の3に優先的に位置し、ほとんど18%の残基が酸性である。これらの残基は最大のポリペプチドのC末端の4分の1で9%しか現れず、2個の電気的に対立するドメインを有するポリペプチド鎖になる。N末端ではポリペプチドはプレクストリン相同性（PH）ドメインを含有し、これは生物学的活性を呈する細胞膜への多くのシグナル発生タンパク質の補給を意味している（17）。最終的に、2分した核標的化配列（18）は高次コイルを形成する全ての構造的要件に全て合致する7個の反復領域の必須部分として存在する（16）。

【0077】

タンパク質データベース検索によりPHドメインを有するN末端でおよそ100個の残基内に相同性がほとんど独占されていることが示された。オキシステロール結合タンパク質のPHドメインが最も類似しており、GPBPとの全体の同一性が33.5%および類似性が65.2%である。加えて、Caenorhabditis elegans（線虫）コスミドF25H2（信託番号Q93569）はタンパク質全配列を通して全体の同一性が26.5%および類似性が61%を示す仮定ORFを含有し、これは類似のタンパク質が下等の無脊椎動物に存在することを示している。いくつかのヒト発現配列タグ（信託番号AA287878、AA287561、AA307431、AA331618、AA

10

20

30

40

50

040134、AA158618、AA040087、AA122226、AA158617、AA121104、AA412432、AA412433、AA282679、およびN27578)は対応するGPBP cDNAの枝分かれを有する高度な配列同一性(98%以上)を有し、これらがヒトGPBPに相当することを示唆している。興味深いことに、AA287878ESTはGPBP 5'非翻訳領域に対応する配列内で67個のヌクレオチドのギャップを示し、これはGPBP pre-mRNAが別法としてヒト組織においてスプライシングされることを示唆している(図示せず)。

【0078】

ヒト組織におけるGPBP遺伝子の分布および発現を最初にノーザンブロット分析により評価した(図2、パネルA)。遺伝子は4.4kbおよび7.5kbの長さの間で2個の主要なmRNA種として、およびより短い少数のその他の種として発現される。これらの複数のmRNA種間の構造上の関係は解っておらず、その相関的な発現は組織間で変化する。最も高い発現レベルは横紋筋(骨格および心臓の筋肉)に見られ、肺および肝臓では発現レベルは最も低いことが示される。

10

【0079】

異なる種からのゲノムDNAのサザンブロット研究分析により、相同性遺伝子は系統発生を通して存在することが示された(図2、パネルB)。ヒト由来のプロープに合致するように、ゲノムDNAの起源として累進的に減少するハイブリダイゼーション強度は、進化においてヒトから排除された。

【0080】

翻訳開始部位の実験的決定 - 予測ORFを実験的に確認するために、n4'のcDNAの2.4kbか、またはFLAG配列でタグ化された予測ORF(図3A)のみを含有する真核発現ベクターを293細胞における一過性発現に使用した。GPBPまたはFLAG特異的抗体を用いて免疫ブロットにより対応する抽出物を分析した。GPBP特異的抗体は両方のトランスフェクトされた細胞において類似の主要ポリペプチドに結合するが、操作された構築物により製造されるポリペプチドのみがFLAG配列を発現した(図3B)。これは予測Metでn4' cDNAの翻訳開始部位に位置し、提示された1次構造を確認した。さらに、組換えポリペプチドは予測された分子量よりも高い分子量を示し(80対71kDa)、これはGPBPが翻訳後修飾されていることを示唆している。

20

【0081】

酵母rGPBPの発現および特徴づけ - 酵母発現およびFLAG基盤のアフィニティ精製を組み合わせることでrGPBPを製造した(図4A)。低Mrを示す複数の関連生成物と共に、~89kDaの主要なポリペプチドが得られた。抗FLAGおよびGPBP特異的抗体の両方により組換え物質が認識され、発現系の適合性が保証された。しかしながら、再度主要生成物により示されたMrが予測されたMrよりも著明に高く、293細胞由来の組換え物質のMrよりも高く、GPBPが重要な区別される翻訳後修飾を受けているという考えが支持された。ポリペプチド鎖にはリン酸化可能な残基が豊富であるので、組換え物質におけるホスホアミノ酸の存在を調べた。ホスホセリン(Pser)、ホスホスレオニン(PThr)およびホスホチロシン(PTyr)に対するモノクローナルまたはポリクローナル抗体を用いることにより(図示せず)、酵母rGPBP(図4B)または293細胞由来物質(図示せず)のいずれかの三つの全てのリン酸残基の存在を同定した。さらに5から10mMの対応するホスホアミノ酸の添加によりその結合を部分的に阻止することにより抗体の特異性を評価した(図示せず)。これはリン酸残基含量が細胞発現系に依存して変化し、Mrの差は主にリン酸化によるものであることを示唆している。脱リン酸化酵母誘導物質は一貫して293細胞に由来する物質に類似のMrを示し、ホスホアミノ酸含量はSDS-PAGE移動性に相関する(図4C)。in vivo測定のように、293細胞におけるrGPBPのリン酸化を評価した(図4D)。対照細胞(レーン1)および安定した(レーン2)または一過性(レーン3)様式のrGPBP発現細胞を $H_3^{32}PO_4$ の存在下培養した。する免疫沈澱組換え物質は ^{32}P を含有し、これはin vivoでGPBPのリン酸化を生じ、従ってこれは物理学的過程である可能性があるこ

30

40

50

とを示した。

【0082】

rGPBPはヒトGP抗原のN末端領域をリン酸化するセリン/スレオニンキナーゼ - GPBPはプロテインキナーゼの古典的な触媒ドメインを定義するのに必要な保存構造を含有しないが、新規非保存プロテインキナーゼ(19から27)の最近の同定および特徴づけによりそのリン酸化活性の研究が促進された。Mn²⁺およびMg²⁺の存在下[³²P]ATPをrGPBP(酵母または293細胞(図示せず)のいずれかに由来)に添加した結果、抗FLAGおよび特異的抗体の両方により認識される主要なおよび関連する生成物においてP_{Ser}およびP_{Thr}として³²Pが組み込まれ(図5AおよびB)、これはアフィニティ精製物質がSer/Thrプロテインキナーゼを含有することを示した。この活性をさらに特徴づけるために、GPpep1, GPpep1Ala⁹(Ser⁹がAlaで置換されたGPpep1変異体)、元来のおよび組換えヒトGP抗原、並びに元来のウシGP抗原を検定した(図5C)。アフィニティ精製rGPBPは異なった程度で全てのヒト由来物質をリン酸化した。しかしながら、類似の条件でウシ由来基質において測定可能な³²P組み込みは観察されなかった。GPpep1に比較した場合、GPpep1Ala⁹により示された³²P組み込みは低度であり、ウシ抗原のリン酸化は欠如し、これはrGPBPに存在するキナーゼがヒトおよびウシ抗原を認識し、Ser⁹がキナーゼの標的であることを示している。

10

【0083】

精製系により高品質の物質が提供されるが、プロテインキナーゼ活性と共に夾雑物が存在することは排除できない。夾雑物の存在はまたFLAG含有40kDaポリペプチドの存在によっても示唆され、これはリン酸化アッセイにおいて特異的抗体または³²Pの組み込みのいずれとも反応性を示さなかった(図4Aおよび5A)。プロテインキナーゼ活性を留めるポリペプチドを正確に同定するために、SDS-PAGEおよびウェスタンブロットの後in vitroキナーゼ再生アッセイを実施した(図6)。我々は特異的抗体の使用(レーン1)および切片上³²P組み込みのオートラジオグラフィによる検出(レーン2)を組み合わせるのに成功し、Ser/Thrキナーゼ活性を留める1次ポリペプチドとして89kDaのrGPBP物質を同定した。rGPBP由来生成物および40kDa夾雑物において³²P組み込みが欠如しているということはさらに再生アッセイの特異性を支持し、キナーゼ活性が89kDaポリペプチドに存在することを確認する。最近、ポリペプチドに親密に関連するプロテインキナーゼの痕跡がブロット膜から遊離され、標識ステップにおいてポリペプチドに結合し、リン酸化できることが示された(28)。我々の系においてこの可能性を評価するために、89kDaのポリペプチドを含有する膜小片を単独でかまたはブロットレーンの別の領域を表す膜片と一緒に用いて復元実験を行った。同時にインキュベートした小片に関わらず(図示せず)、89kDaポリペプチドで類似の³²P組み込みが観察され、これは我々のサンプルにおいて同時に精製されたプロテインキナーゼがある場合、これらが同時移動しない限り、これらは再生アッセイにおいて89kDaポリペプチドをリン酸化していないことを示している。同時移動が関係しないと考えられるが、しかしながら、異なる移動性を示すrGPBP欠損変異体(GPBP 26およびR3; 以下を参照のこと)もまたキナーゼ活性を有し、インブロットで同様に復元され得る(図示せず)。

20

30

40

【0084】

新規キナーゼの免疫組織化学的局在 - ヒト組織におけるGPBP発現を研究するために、特異的ポリクローナル(図7)またはモノクローナル抗体(図示せず)を用いて免疫組織化学実験を行った。GPBPはヒト組織において広範に発現するが、それは組織および細胞特異性を示す。腎臓では主に細管上皮細胞および糸球体メサンギウム細胞および有足細胞に見出される。肺胞では抗体は肺細胞の染色に加えて基底膜局在化を示唆する直線的パターンを表す。肝臓は柔組織においては低発現を示すが、胆管では高発現を示す。中枢神経系における発現は白質で観察されるが、脳のニューロンでは観察されない。精巣では精原細胞において高発現し、対照的にセルトリ細胞においては発現しない。副腎では髄質に

50

対して皮質細胞において高レベル発現を示す。膵臓では、G P B P は外分泌部分に対してランゲルハンス島で優先的に発現される。前立腺では、G P B P は上皮細胞で発現されるが、間質細胞においては発現されない(図7)。その他の場所では、G P B P は横紋筋、腸管上皮細胞、および小脳のプルキンエ細胞で高発現される(図示せず)。概してG P B P を高発現する組織では染色パターンは主に細胞基質に拡散している。しかしながら、特定の場所では、加えて核(精原細胞)で、細胞質膜(肺細胞、肝細胞、前立腺上皮細胞、白質)で、または細胞外マトリックス(肺胞)で重要な染色増強がある(図7)。

【0085】

考察

我々のデータはG P B P が新規で、非保存セリン/スレオニンキナーゼであることを示している。我々はまたG P B P がヒトおよびウシG P 抗原を認識し、*in vivo*でヒトG P 抗原のリン酸化可能な領域を標的化することを明らかにした。いくつかの証拠のラインが、89 k D a ポリペプチドがアフィニティ精製r G P B P において唯一のキナーゼであることを示している。第1に、我々は150 mM、0.5 N、1 Mまたは2 Mの塩の存在下で精製されたr G P B P サンプル間で自己およびトランスリン酸化における差異を見出さず(データを示さず)、これはr G P B P が親密な結合性キナーゼを担持しないことを示唆している。第2に、G P B P 特異的抗体を用いて精製した物質がリン酸化において差異を示さない(図示せず)、我々のサンプルではF L A G 含有、酵母誘導キナーゼがない。第3に欠失変異体(G P B P 26; 以下を参照のこと)は自己およびトランスリン酸化活性を減じ(図示せず)、これは89 k D a ポリペプチドがリン酸移動を実施する能力を有するr G P B P のほんの一部であることを示している。

【0086】

G P B P はその他の非保存性キナーゼと相同ではないが、N末端ヘリックス高次コイル(26、27)、セリンに富むモチーフ(24)、ホスホアミノ酸の高含量(27)、2分裂した核局在化シグナル(27)、および典型的なヌクレオチドまたはA T P 結合モチーフの不在(24、27)などのいくつかの構造上の特徴を共有している。

【0087】

免疫組織化学研究はG P B P が細胞基質性ポリペプチドであり、原形質膜に結合して核においても見出され、基底膜に結合して細胞外マトリックスに見出される可能性もあり、これはこれらの全方向に到達するための構造上の要件を有していることを示している。核局在シグナルおよびP H ドメインはそれに各々核および細胞膜に到達する可能性を付与する(17、29、30)。G P B P は搬出される構造上の要件を有していないが、そのm R N A の5'末端非翻訳領域が初めにフレーム内に停止コドンを含む(130残基の上流O R F を含む(図1))。単一の塩基対(U)を挿入するm R N A 編集過程が機能的なフレーム内出発部位および編集M e t のすぐ下流に搬出シグナルを含有する754残基のO R F を生じるのであろう(図示せず)。この仮説上の余剰配列の一部(P R S A R C Q A R R R R G G R T S S (配列番号: 33))を示す合成ペプチドに対するポリクローナル抗体はヒト組織において直線的な血管反応性を示し、これは細胞外基底膜局在を示唆する(データは示さず)。

【0088】

別法として、スプライシング現象により搬出のための構造上の要件を提供するさらなる未同定エキソンを有する転写物を生じることができた。複数の細胞局在、P T y r の高含量、および*in vitro*チロシンキナーゼ活性の欠損はG P B P 自体が特異的チロシンキナーゼの標的であり、従って、特異的シグナル発生カスケードに関与している可能性がある。

【0089】

前記するように、特異的セリンリン酸化およびプレm R N A 代替スプライシングは、G P 抗原、アセチルコリンレセプターおよびミエリン塩基性タンパク質(M B P)などのいくつかの自己抗原の生物学に関連している(4)。後者は多発性硬化症(M S)、免疫系が中枢神経系の白質を標的化するその他の排他的ヒト自己免疫疾患における主要な抗原であ

10

20

30

40

50

ることが疑われる。G P病およびMSは同一のHLA IIクラスIIタイプ(HLA DRB1*1501)と強力な関連性を示すヒト障害である(32、33)。これは、このHLA特異性を担持するMS患者のG P病による死亡の最近の報告に加えて(34)、これらのヒト障害における共通の発病原因が存在することを支持している。

【0090】

特異的なセリンのリン酸化が細胞内タンパク質溶解を変化させることが示されている(35から40)。考えられることは、タンパク質リン酸化における変化はプロセッシングおよびペプチド展示に影響し、従って自己免疫を媒介する。HLA-DR15によるG P抗原誘導ペプチド展示は、相対的に乱雑なDR15分子の優先性よりもプロセッシングにより依存し(41)、これはプロセッシングが異常なリン酸化により影響される場合、得られるペプチドはこのHLAにより提示されることを示唆している。我々の最近のデータでは、G PおよびMBP系の両方で、代替スプライシング生成物の生成が特異的および構造上相同であるPKA部位のリン酸化を制御するために提供され、これはこのまたは非常に関連性のあるキナーゼが*in vivo*リン酸化酵素であることを示唆している。抗原リン酸化の程度の変化は代替生成物における不均衡によるか、または同一のモチーフのリン酸化を脱制御する侵入性のキナーゼの作用により引き起こされ、素因のある個体において自己免疫応答を導くことができる。G P抗原の特異的代替生成物の存在がGPBPによる1次抗原のリン酸化に影響しなかったため、rGPBPは見かけ上制御されない様式で主要PKAリン酸化部位でヒトG P抗原をリン酸化する(図示せず)。

【0091】

GPBPは偏在して発現されるが、特定の器官および組織では共通の自己免疫応答の標的となる細胞および組織構造の偏好性が示される：ランゲルハンス細胞(I型糖尿病)；中枢神経系の白質(多発性硬化症)；胆管(胆汁性肝硬変)；副腎の皮質細胞(アジソン病)；横紋筋細胞(重症筋無力症)；精原細胞(男性不妊症)；小脳のプルキンエ細胞(腫瘍随伴性小脳変性症候群)；および腸管上皮細胞(悪性貧血、自己免疫性胃炎および腸炎)。ヒト自己免疫疾患のモデルを想定する場合、前記の観察はすべてこの新規キナーゼが魅力的な候補物質であることを示している。

【0092】

背景および実施例1の参考文献

- 1 Saus, J. (1998) グッドパスチャー症候群. *Encyclopedia of Immunology* 第二版, Delves, P.J., Roitt, I.M. 編, Academic Press Limited, ロンドン, イギリス 30
- 2 Leinonen, A., Mariyama, M., Mochizuki, T., Tryggvason, K. および Reeders, S.T. (1994) *J. Biol. Chem.* 269, 26172 - 26177
- 3 Quinones, S., Bernal, D., Garcia-Sogo, M., Elena, S.F. および Saus, J. (1992) *J. Biol. Chem.* 267, 19780 - 19784
- 4 Revert, F., Penades J.R., Plana, M., Bernal, D., Johansson, C., Itarte, E., Cervera, J., Wieslander, J., Quinones, S. および Saus, J. (1995) *J. Biol. Chem.* 270, 13254 - 13261 40
- 5 Bernal, D., Quinones, S. および Saus, J. (1993) *J. Biol. Chem.* 268, 12090 - 12094
- 6 Feng, L., Xia, Y. および Wilson, C.B. (1994) *J. Biol. Chem.* 269, 2342 - 2348
- 7 Penades, J.R., Bernal, D., Revert, F., Johansson, C., Fresquet, V.J., Cervera, J., Wieslander, J., Fritsch, E.F. および Man 50

- iatis, T. (1989) *Molecular Cloning; A Laboratory Manual*, 第二版, コールドスプリングハーバーラボラトリー, コールドスプリングハーバー, ニューヨーク州
- 9 Coligan, J.E., Dunn, B.N., Ploegh, H.L., Speicher, D.W. および Winfield, P.T. (1995-97) *Current protocols in Protein Science*, John Wiley & Sons 編, ニューヨーク, ニューヨーク州
- 10 Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Deidman, J.G., Smith, J.A. および Struhl, K. (1994-98) *Current Protocols in Molecular Biology*, John Wiley & Sons 編, ニューヨーク, ニューヨーク州
- 11 Ferrel, J.E. および Martin, G.S. (1991) *Methods in Enzymology* 201, 110-149
- 12 Boyle, W.J., van der Geer, P. および Hunter, T. (1991) *Methods in Enzymology* 201, 110-149
- 13 Hsu, S.M., Raine, L. および Fanger, H. (1981) *J. Histochem. Cytochem.* 29, 577-580
- 14 Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. および Lipman, D.J. (1997) *Nucleic Acids Res.* 25, 3389-3402
- 15 Bairoch, A., Bucher, P. および Hofmann, K. (1997) *Nucleic Acids Res.* 25, 217-221
- 16 Lupas, A. (1996) *Trends Biochem. Sci.* 21, 375-382
- 17 Lemmon, M.A., Falasca, M., Ferguson, K.M. および Schlessinger, J. (1997) *Trends Cell Biol.* 7, 237-242
- 18 Bouliskas, T. (1993) *Crit. Rev. Eukaryot. Gene Expr.* 3, 193-227
- 19 Csermely, P. および Kahn, C.R. (1991) *J. Biol. Chem.* 266, 4943-4950
- 20 Maru, Y. および Witte, O.N. (1991) *Cell* 67, 459-468
- 21 Beeler, J.F., LaRoche, W.J., Chedd, M., Tronick, S.R. および Aaronson, S.A. (1994) *Mol. Cell. Biol.* 14, 982-988
- 22 Csermely, P., Miyata, Y., Schneider, T. および Yahara, I. (1995) *J. Biol. Chem.* 270, 6381-6388
- 23 Dikstein, R., Ruppert, S. および Tjian, R. (1996) *Cell* 84, 781-790
- 24 Eichinger, L., Bomblies, L., Vandekerckhove, J., Schleicher, M. および Gettermans, J. (1996) *EMBO J.* 15, 5547-5556
- 25 Cote, G.P., Luo, X., Murphy, M.B. および Egelfhoff, T.T. (1997) *J. Biol. Chem.* 272, 6846-6849

- 26 Ryazanov, A.G., Ward, M.D., Mendola, C.E., Paver, K.S., Dorovkov, M.V., Wiedmann, M., Erdjument-Bromage, H., Tempst, P., Parmer, T.G., Prostoko, C.R., Germino, F.J. および Hait, W.N. (1997) Proc. Natl. Acad. Sci. USA 94, 4884 - 4889
- 27 Fraser, R.A., Heard, D.J., Adam, S., Lavigne, A.C., Le Douarin, B., Tora, L., Losson, R., Rochette-Egly, C. および Chambon, P. (1998) J. Biol. Chem. 273, 16199 - 16204 10
- 28 Langelier, Y., Champoux, L., Hamel, M., Guilbault, C., Lamarche, N., Gaudreau, P. および Massie, B. (1998) J. Biol. Chem. 273, 1435 - 1443
- 29 Lemmon, M.A. および Ferguson, K.M. (1998) Curr. Top. Microbiol. Immunol. 228, 39 - 74
- 30 Rebecchi, M.J. および Scarlata, S. (1998) Annu. Rev. Biophys. Biomol. Struct. 27, 503 - 528
- 31 Roitt, I. (1994) Autoimmune diseases in Essential Immunology, 383 - 439, 第8版, Blackwell Scientific, Oxford, UK 20
- 32 Erlich, H. および Apple, R. (1998) MHC疾患関連。Encyclopedia of Immunology 第二版, Delves, P. J., Roitt, I.M. 編, Academic Press Limited, ロンドン、イギリス
- 33 Phelps, R.G., Turner, A.N. および Rees, A. J. (1996) J. Biol. Chem. 271, 18549 - 18553
- 34 Henderson, R.D., Saltissi, D. および Pender, M.P. (1998) Acta Neurol. Scand. 98, 134 - 135 30
- 35 Litersky, J.M. および Johnson, G.V.W. (1992) J. Biol. Chem. 267, 1563 - 1568.
- 36 Brown, K., Gerstberger, S., Carlson, L., Franzoso, G. および Siebenlist, U. (1995) Science 267, 1485 - 1488
- 37 Chen, Z.J., Parent, L. および Maniatis, T. (1996) Cell 84, 853 - 862
- 38 Aberle, H., Bauer, a., Stappert, J., Kispert, A. および Kemler, R. (1997) EMBO J. 16 40, 3797 - 3804
- 39 Regnier, C.H., Song, H.Y., Gao, X., Goeddel, D.V., Cao, Z. および Rorthe, M. (1997) Cell 90, 373 - 383
- 40 Vlach, J., Hennecke, S. および Amati, B. (1998) EMBO J. 16, 5334 - 5344
- 41 Phelps, R.G., Jones, V.L., Coughlan, M., Turner, A.N. および Rees, A.J. (1998) J. Biol. Chem. 273, 11440 - 11447

(実施例2: GPBP代替スプライシング)

ここで26残基のセリンに富むモチーフをコードする78塩基対(bp)の長さのエキシソンの代替スプライシングにより生じるGPBPの二つのアイソフォームの存在を報告する。両方のアイソフォーム、GPBPおよびGPBP₂₆は、ポリペプチド自己凝集の結果である高分子凝集体として存在する。ポリペプチド鎖における26残基ペプチドの存在はより有効に自己相互作用し、より高い特異活性を有する凝集体を形成する分子種に至る。最終的には、GPBPがヒト自己免疫の発病原因に関与するという観察を支持する証拠を提示する。

【0094】

材料および方法

合成ポリマー:

ペプチド: GP₁pep1、KGKRGDSGSPATWTTTRGFVFT(配列番号: 26)を実施例1で記載する。GPBP₁pep1、PYSRSSSMSSIDLVSASDDVHRFSSQ(配列番号: 14)はGPBPの371から396残基を表し、ゲノシスにより合成された。

【0095】

オリゴヌクレオチド。以下のオリゴヌクレオチドをライフテクノロジーズインコーポレーティッドにより合成した(5'から3'へ); ON-GPBP-11m, GCGG GAC TCA GCG GCC GGA TTT TCT(配列番号: 34); ON-GPBP-15m, AC AGC TGG CAG AAG AGA C(配列番号: 35); ON-GPBP-20c, C ATG GGT AGC TTT TAA AG(配列番号: 36) ON-GPBP-22m, TA GAA GAA CAG TCA CAG AGT GAA AAG G(配列番号: 37); ON-GPBP-53c, GAA TTC GAA CAA AAT AGG CTT TC(配列番号: 38); ON-GPBP-56m, CCC TAT AGT CGC TCT TC(配列番号: 39); ON-GPBP-57c, CTG GGA GCT GAA TCT GT(配列番号: 40); ON-GPBP-62c, GTG GTT CTG CAC CAT CTC TTC AAC(配列番号: 41); ON-GPBP-26, CA CAT AGA TTT GTC CAA AAG GTT GAA GAG ATG GTG CAG AAC(配列番号: 42)。

【0096】

逆転写酵素およびポリメラーゼ連鎖反応(RT-PCR)。全RNAを(15)に記載するように、異なる対照およびGP組織から調製した。全RNA 5 μgをレディートウーゴユープライム ファーストストランドビーズ(アマシャム ファルマシア バイオテック)、および40ピコモルのON-GPBP-53cを用いて逆転写した。対応するcDNAをプライマーON-GPBP-11m/ON-GPBP-53cまたはON-GPBP-15m/ON-GPBP-62cを用いるPCRに供した。15m-62cで得られた生成物の同一性をさらにAlu I制限により確認した。GPBP転写物を特異的に増幅するために、プライマーON-GPBP-15m/ON-GPBP-57cを用いてPCRを実施した。

【0097】

ノーザンハイブリダイゼーション研究。予め製造したヒト多重組織および腫瘍培養細胞株ノーザンプロット(クロンテック)をGPBPにのみ存在する78bpエキソンを含有するcDNAで、または両方のアイソフォームを表すcDNAでプローブした。プライマー対ON-GPBP-56mおよびON-GPBP-57cを用いて、GPBPを鋳型として用いて、またはプライマー対ON-GPBP-22mおよびON-GPBP-20cを用いて、GPBP₂₆を鋳型として用いてPCRにより対応するcDNAを得た。得られた生成物をランダム標識し、および製造者指示書に従ってハイブリダイズした。

【0098】

プラスミド構築、組換えタンパク質の発現および精製。Pichia pastoris

10

20

30

40

50

におけるFLAGタグ化GPBPの組替え発現に用いられたプラスミドpHIL-FLAG-n4'は別に報告されている(4)。ON-GPBP 26を用いて78bpエキソンをコードする配列を位置指定突然変異誘発により欠失させ、プラスミドpHIL-FLAG-n4' 26を生じた。組換えGPBPおよびGPBP 26の発現およびアフィニティ精製を(4)の通りに行った。

【0099】

ゲル濾過HPLC。250 μ lのサンプルを50mMのトリスHCl、pH7.5、150mMのNaClで平衡にしたゲル濾過PE-TSK-G4000SW HPLCカラムに注入した。カラムから0.5ml/分で物質を溶出し、220nmでモニター観察し、微小な分画を収集した。

10

【0100】

in vitroリン酸化アッセイ。自己、トランスリン酸化およびインプロット還元実験を実施例1に記載のとおり実施した。

【0101】

抗体および免疫化学技術。ポリクローナル抗体をニワトリにおいて78bpエキソン(ゲノシス)によりコードされる配列を表す合成ペプチド(GPBP pep1)に対して産生させた。卵黄を水で1:10に希釈し、pHを5.0に調整した。4 $^{\circ}$ Cで6時間後、溶液を遠心(4 \times 10000 \times gで25分間)により清澄化し、20 \times 20000 \times gで20重量/容量% 硫酸ナトリウムを添加することにより抗体を沈殿させた。ペレットをPBS(卵黄あたり1ml)に溶解し、免疫化学実験に用いた。GPBP/GPBP 26に対するまたは3(IV)NC1ドメインに対する抗体の生成については前記している(また4,13を参照のこと)。

20

【0102】

沈降速度。VIS-UVスキャナーを装備したオプティマXL-A分析用超遠心(ベックマンインストラメンツインコーポレーティッド)で、Ti60ローターおよび12mm光学距離のイーボンチャーコールの2重セクターセルを用いて沈降速度の決定を行った。およそ400 μ lのサンプルを20 \times 30000rpmで遠心し、220nmでラジアル走査を5分毎に取った。XLAVEL(ベックマン提供による)プログラムを用いて溶質結合の動きの速度から沈降係数を得た。

【0103】

沈降平衡。沈降平衡実験を70 μ lのサンプルを用いる速度実験に関して前記で記載するように実施し、8000rpmで遠心した。平衡時の実験濃度勾配をEQASSOCプログラム(ベックマン)を用いて分析し、対応する平均分子量を決定した。GPBPに関して0.711 cm^3/g およびGPBP 26に関して0.729 cm^3/g の部分的特異的容量を対応するアミノ酸組成物から算出した。

30

【0104】

物理的方法および免疫化学技術。(3)で前記するように還元条件下SDS-PAGEおよびウェスタンブロッティングを実施した。ホルマリン固定パラフィン包埋組織に関して、ABC過酸化法(4)を用いて、または冷アセトンで固定した凍結ヒト生検に関して間接的免疫蛍光の標準的な方法を用いて免疫組織化学実験を実施した。

40

【0105】

二つのハイブリッド実験。pGBT9およびpGAD424(クロンテック)を用いてSaccharomyces cerevisiae(酵母)(HF7c)において自己相互作用実験を実施し、各々GAL4結合およびドメイン融合タンパク質の活性化を生じた。製造者の推奨に従って相互作用を評価した。ガラクトシダーゼ活性を切片上ではX-GAL(0.75mg/ml)を用いて、および溶液中決定ではオルトニトロフェニル-Dガラクトピラノシド(0.64mg/ml)を用いて評価した。

【0106】

結果

二つのスプライシングGPBP変種の同定。正常ヒト組織におけるGPBP種を特徴づけ

50

るために、G P B Pの全オープンリーディングフレームを隣接する特異的オリゴヌクレオチドを用いて、異なる組織からの全RNAに関してポリメラーゼ連鎖反応に逆転写を連動させた。骨格筋由来RNAから(図8A)、および腎臓、肺、皮膚、または副腎由来RNAから(図示せず)、予測されたよりも小型であることを示す単一のcDNAフラグメントを得た。入れ子式PCR増幅およびエンドヌクレアーゼ制限マッピングを組み合わせ、全RT-PCR生成物が同一の分子種に対応することを決定した(図示せず)。ヒト筋肉からの2.2kbのcDNAを全て配列決定し、26残基モチーフ(アミノ酸371-396)をコードする78ヌクレオチド(1519-1596位置)が存在しない以外は、HeLa誘導物質と同一であることを見出した(図8B)。従って、このより一般的であるG P B PのアイソフォームをG P B P 26と称した。

10

【0107】

78bpがブレムRNAプロセッシング中に転写をスキップするエキソンを表すのかどうかを調べるために、このcDNAを用いてヒト由来のゲノムライブラリーをプローブし、~14Kbのクローンを単離した。サザンブロットハイブリダイゼーションおよびPCRを組み合わせ、ゲノムクローンを特徴づけし、12482bpの隣接DNAフラグメントを全て配列決定した(配列番号:25)。配列は(5'から3'で)767bpのイントロン配列、93bpのエキソン、818bpのイントロン、78bpの目的のエキソン配列、9650bpのイントロン、96bpのエキソンおよび980bpのイントロン配列(図8C)を含有した。対応するDNAおよびcDNA配列を比較することにより決定したエキソンイントロン境界は5'および3'スプライシング部位の標準的なコンセンサスに合致し(図8C)(5)、78bp配列のエキソン特性を確認する。プローブとしてゲノムクローンをを用いる蛍光切片上ハイブリッド形成法(FISH)によりG P B P遺伝子は染色体5q13に局在した(図示せず)。

20

【0108】

G P B PおよびG P B P 26の両方に存在する78bpエキソンかまたは78bpの隣接配列を表す260bpのcDNA(103bp5'および157bp3')を用いるノーザンブロット分析によりヒト由来の標本におけるG P B Pの相対的発現を評価した(図9)。目的の分子種を含有する78bpは横紋筋(骨格筋および心筋の両方)および脳に好んで発現し、胎盤、肺および肝臓にはあまり発現しなかった。G P B P 26とは対照的に、G P B Pは腎臓、膵臓および癌の培養細胞株では非常に低レベルでしか発現しなかつた。

30

【0109】

前記の全ては、G P B Pが正常ヒト組織では低レベルで発現され、G P B PのRT-PCRにより最初に検出されなかったのはより豊富なG P B P 26が優先的に増幅されたためであろうということを示している。実際、78bpエキソン特異的オリゴヌクレオチドを用いて特異的RT-PCR増幅を行った場合、G P B PのcDNAをヒト組織(骨格筋、肺、腎臓、皮膚および副腎)から増幅できた(図示せず)。これはまた、G P B PおよびG P B P 26の両方を表すcDNAプローブを用いる場合、G P B P 26 mRNAがノーザンブロット実験で検出された主要な転写物であることを示唆している。

【0110】

G P B P 26の組換え発現および機能上の特徴づけ。26残基のセリンに富むモチーフの欠如がG P B Pの生化学的特性に影響するかどうかを調べるために、両方のアイソフォーム(r G P B Pおよびr G P B P 26)を発現し、精製し、自己およびトランスリン酸化活性を評価した(図10)。r G P B Pに関して前記で報告するように(4をも参照のこと)、単一の主要なポリペプチドおよびいくつかの関連する微少な生成物としてr G P B P 26を精製した(図10A)。しかしながら、r G P B Pと比較して、誘導生成物の数および相対量は変化し、SDS-PAGEで単純に26残基の削除のせいにはできないMrが表示された。これは26残基モチーフが重要な構造を有し、および溶液中自己およびトランスリン酸化活性の低減の原因となる機能上の成果がr G P B Pにより表されたことを示唆する(図10B)。興味深いことに、リン酸化がSDS-PAGE後インプ

40

50

ロットで評価され、復元された場合、溶液中アッセイにおいて示された特異的活性における差異は明白ではなく、これは26残基モチーフが4次構造レベルで重要な機能上の成果を有する可能性があることを示唆した。さらに復元実験によりリン酸塩移動活性が提示されたオープンリーディングフレームを表す主要なポリペプチドに存在し、誘導された微量生成物では検出されないことを示した。

【0111】

rGPBPおよびrGPBP-26は非常に活性な高分子量凝集体として存在する。アフィニティ精製rGPBPまたはrGPBP 26のゲル濾過分析により、2個のクロマトグラフィピーク(IおよびII)を生じ、SAS-PAGE実験で決定されるように、共に個々の分子種で予測されたものよりも高MWを示した(各々89kDa、84kDa) (図11)。大部分の組換え物質を158kDaおよび669kDa分子量マーカーの間で単一のピークとして溶出し、一方限定量のrGPBPおよび極微量のrGPBP 26をピークIで溶出した(>1000kDa)。各クロマトグラフィプロフィールを表す分画の一定分量をSDS-PAGEに供し、染色するか、または³²P[]ATPの存在下インキュベートし、イムノプロットおよびオートラジオグラフィにより分析した。主要な1次ポリペプチドに加えて、各クロマトグラフィのピークはより大きいまたはより小さい大きさの複数の誘導生成物を含有し、これは1次ポリペプチドが結合して、共有および非共有結合により安定化された高分子量凝集体を形成することを示している(図示せず)。キナーゼ活性もまたクロマトグラフィプロフィールと一致する2個のピークを呈した。しかしながら、ピークIはピークIIよりもより高度な特異的活性を示し、これはこれらの高分子量凝集体がキナーゼのより活性な形態を含有していることを示した。等容量のrGPBP分画番号13および20は、タンパク質含量では分画13はウェスタンプロットおよびクーマシーブルー染色により推測されるより20倍低いのが、匹敵するリン酸化活性を呈した(図11A)。rGPBPおよびrGPBP 26のピークIIでの特異的活性もまた異なり、物質全体に関して示された実験と合致し、従って26残基のセリンに富むモチーフの存在によりキナーゼがより活性になるにするとする仮説を支持している。これらの結果はまたrGPBPおよびrGPBP 26の両方が元来の条件下でオリゴマーとして存在し、高分子量凝集体の形成および特異的活性の両方が26残基のセリンに富むモチーフの存在に大いに依存することを示唆している。rGPBPの分析用遠心分析によりピークIが大きな凝集体(10⁷Da以上)を含有することが表された。rGPBPのピークIIは220±10kDa凝集体の相同性集合体を含有し、11Sの沈降係数を有する三量体を表す可能性がある。しかしながら、rGPBP 26のピークIIは、いくつかのオリゴマー種を含有する可能性のあるより異種性の集合体からなる。主要な集合体(約80%)は平均分子量310±10kDaおよび14Sの沈降係数を示した。

【0112】

酵母2ハイブリッド系におけるGPBPおよびGPBP 26自己相互作用。自己凝集の生理学的関係の評価するために、および26残基モチーフの役割を決定するために、酵母における2ハイブリッド相互作用系を用いて比較実験を実施した。この型の実験では、GAL4転写因子の活性化または結合ドメインのいずれかを含有する融合タンパク質の一部として、相互作用を実験するポリペプチドを発現する。実験するポリペプチドを通して2個の融合タンパク質間の有効な相互作用が転写活性化因子の再構築および引き続いて2個のレポーター遺伝子、LacZおよびHis3の発現に至り、各々コロニーの色の検出およびHis欠損培地における成長が可能になる。異なる濃度のHis3遺伝子生成物(3-A T)の競合阻害因子の存在下、ヒスチジン欠損培地における成長速度による相互作用の強度、および定量比色液体ガラクトシダーゼアッセイを評価した。代表的な実験を図12に表す。自己相互作用に関してGPBP 26を検定する場合、レポーター遺伝子の有意な誘導が観察されたが、各融合タンパク質が単独でまたは対照融合タンパク質と共に発現された場合、発現は検出されなかった。GPBPを得るためのポリペプチドにおける26残基モチーフの挿入により、ポリペプチド相互作用において著明に増加した。前記の全データはGPBP 26が*in vivo*で自己結合し、26残基のポリペプチド鎖へ

10

20

30

40

50

の挿入により、より相互に影響する分子種を生じることが示す。

【0113】

G P B P はヒト系球体および肺胞において発現されるが、ウシおよびネズミ系球体および肺胞では発現されない。G P B P / G P B P 26 は天然に発生する自己免疫応答において一般的に標的化されるヒト細胞および組織において優先的に発現される。G P B P の発現を具体的に調べるために、このキナーゼアイソフォームの26残基モチーフの特徴を表す合成ペプチドに対してポリクローナル抗体を産生させ、凍結またはホルマリン固定し、パラフィン包埋したヒト組織において免疫組織化学実験に用いた(図13)。概して、これらの抗体は、自己免疫応答の標的である組織構造、例えば胆管、ランゲルハンス島、または中枢神経系の白質に関して両方のアイソフォームを認識する抗体よりも高い特異性を示す(図示せず)。それにも関わらず、最も著明な知見は、試験したいずれの組織でも小さい血管の周囲のG P B P 選択的抗体の直線的な沈着の存在であり(A)、これはG P B P が内皮細胞基底膜と結合していることを示唆している。結果的に、系球体では、抗G P B P 抗体は3(IV)NC1を特異的に認識するモノクローナル抗体(13Bを13Cおよび13Dと比較する)によるか、または循環G P 自己抗体により(13Eおよび13Fを比較する)生じる系球体基底膜染色に非常に類似した血管パターンを示した。これらの観察によりG P B P が天然の自己免疫応答において標的化される組織構造において発現されるといふ最初の観察がさらに支持され、これはG P B P の発現は危険因子であり、自己免疫攻撃に対して宿主組織を脆弱にすることを示唆している。

10

【0114】

この仮説をさらに評価するために、ヒトと比較して天然にはG P 病にならない2種の哺乳動物の系球体においてG P B P およびG P B P 26 の存在を調べた(図14)。G P B P 特異的抗体はウシおよびネズミの両方の標本の系球体を染色せず(14Aを14Bおよび14Cと比較する)、G P B P およびG P B P 26 の両方に共通するN末端配列を認識する抗体は三つ全ての種のこれらの構造を染色したが、分布および強度は異なっていた(14Dから14F)。ウシ腎皮質では、G P B P 26 はヒトよりも低率で発現されたが、類似の組織分布が示された。しかしながらネズミサンプルでは、G P B P 26 はヒト系球体におけるG P B P の組織分布と非常に類似した組織分布が示された。異なる3種において肺胞を試験した場合、類似の結果が得られた(図示せず)。抗体検出における差異が差次発現よりもむしろ1次構造の差異によるものであることを排除するために、c D N A 配列決定により、これらの2種において対応する1次構造を決定した。ウシおよびマウスG P B P (配列番号:3から6、および9から12)は全体的にヒト物質と各々97.9%および96.6%の同一性を示した。さらに、マウス26残基モチーフはヒトと同一であったが、ウシは1個の残基においてのみ分岐した。最終的には、ヒトに類似して、対応する78bpエキソンに特異的なオリゴヌクレオチドを用いてマウスまたはウシ腎臓全RNAからG P B P c D N A をうまく増幅し、これによりG P B P が非常に低レベルで発現され、免疫化学技術により検出できないことが示された。

20

30

【0115】

G P B P はいくつかの自己免疫状態において高度に発現される。特異的R T - P C R により異なるG P 患者からのいくつかの組織を分析し、G P B P / G P B P 26 m R N A レベルを評価した。対照の腎臓と同様に、G P 腎臓で主に発現されたアイソフォームはG P B P 26 であった。しかしながら、1人の患者の筋肉ではG P B P が優先的に発現され、一方G P B P 26 は対照筋肉サンプルにおいて検出された唯一のアイソフォームであった(図15A)。我々はこの特定の患者からの腎臓サンプルを有していなかったため、対応する標的器官におけるG P B P / G P B P 26 発現を評価できなかった。類似の理由から、腎臓を試験した患者の筋肉におけるG P B P / G P B P 26 レベルを評価できなかった。筋肉細胞は高レベルのG P B P / G P B P 26 を発現し(図9におけるノーザンプロットを参照のこと)、これらは多量の組織を含む。対照的に腎臓におけるG P B P / G P B P 26 の発現はもっと少なく、系球体は実際に、G P B P アイソフォームを発現する唯一の腎臓構造である(図13を参照のこと)。系球体は、筋細胞が筋肉にあ

40

50

るよりも相対的に腎臓に余り豊富でない構造であり、糸球体はG P 発病過程において免疫攻撃により標的化される構造である。R T - P C R 実験を行う場合により豊富でより短いメッセージが優先的に増幅されるのに伴い、これらの因子により正常およびG P 腎臓の両方でG P B P が検出されないことが説明され、従って発病過程における糸球体でのG P B P 発現の評価を除外できる。それにも関わらず、G P 患者におけるG P B P レベルの増加はG P B P / G P B P 26 発現がG P 発病過程において変化し、G P B P 発現の増強がG P 病において重要な病因であることを示唆している。

【 0 1 1 6 】

自己免疫発病過程におけるG P B P およびG P B P 26 の発現を調べるために、皮下自己免疫過程を試験し、これらを正常皮膚または非自己免疫性皮膚炎を表す対照サンプルと比較した(図15)。対照サンプルは大抵の末梢ケラチノサイトにおいて限定的なG P B P 発現を示すが(15B、15E)、基底層から角質層に広がるケラチノサイトは全身性エリテマトーデス(SLE)(15C、15F)または扁平苔癬(15D、15G)の影響をうける皮膚において豊富なG P B P を発現した。UV照射およびアポトーシス誘導時の培養ケラチノサイトで前記した小疱に非常に類似した細胞表面構造においてG P B P が優先的に発現された(6)。対照的に、G P B P およびG P B P 26 の両方を認識する抗体は、自己免疫の影響を受けたまたは対照のサンプルの両方において表皮全体を通じて細胞基質性拡散パターンを生じた(図示せず)。これらのデータは対照および自己免疫の影響を受けたケラチノサイト両方において、G P B P 26 が細胞基質で発現され、細胞分化の間に発現は有意には変化しなかったことを示している。対照的に、成熟ケラチノサイトは実際に唯一のG P B P 発現細胞であった。しかしながら、小疱形成およびG P B P 発現は自己免疫応答により影響を受けた表皮において初期分化段階で観察された(15C、15D、15F、15G)。これはさらに基底ケラチノサイトで異常なアポトーシスが皮膚に影響する自己免疫の発病過程に関係することを示す以前の観察を支持し(7)、これはアポトーシスおよびG P B P 発現がこのヒト細胞系に関連していることを示唆している。

【 0 1 1 7 】

考察

代替プレmRNAスプライシングは組織分布、細胞内局在、および異なるプロテインキナーゼの機能を制御することが報告されている差次遺伝子発現の基礎的なメカニズムである(8から11)。この点で、およびG P B P に非常に類似して、B - R a f は複数のスプライシング変種として存在し、ここで特異的エキソンの存在により、より相互に影響し、有効な腫瘍形成性キナーゼになる(12)。

【 0 1 1 8 】

r G P B P 26 がなおこの新規キナーゼの特徴づけされていない触媒性ドメインを担持することは明らかであるが、r G P B P と比較した場合、自己およびトランスリン酸化活性の両方が大きく低下する。ゲル濾過および2ハイブリッド実験により、このように低下したリン酸移動活性の基礎となるメカニズムにいくらかの見識が提供される。r G P B P の約1から2%がr G P B P のリン酸化活性の約3分の1をあらわす非常に高分子量の凝集体に構成され、これは高分子量凝集体がより有効な4次構造を付与することを示している。組換えG P B P 26 は実施にピークIの物質を伴わず、一貫して低下したキナーゼ活性を示す。しかしながら、ピークIIに存在するr G P B P 26 物質もまたr G P B P の相同性分画と比較した場合、リン酸化活性の低下を示すので、凝集体が、26残基が特異活性を増加させる唯一のメカニズムではないようである。一つの可能性は、r G P B P 由来の凝集体が26残基モチーフの挿入により引き起こされる4次構造強化のために、より高度な特異活性を表すということである。大量の物質が非還元SDS-PAGEにおいて単一のポリペプチドとして現れ、8Mの尿素または6Mのグアニジンのいずれかの存在はクロマトグラフィゲル濾過プロフィールにおいてほとんど影響がないので、オリゴマーは非常に強力な非共有結合により主に繋がれている(図示せず)。どのように26残基モチーフがより強力な、活性な構造を付与するのかを明らかにすべきであろう。キナーゼ

の活性化状態を変化させるモチーフをコードするエキソンの存在により誘導される立体配座の変化が B - R a f の S r c タンパク質 (2 4) およびエキソン 8 b および 1 0 (1 2) のリンカードメインとして提示されている。別法として、26 残基モチーフが構造上の要件、例えばそのリン酸化が G P B P の全活性化に必要である残基を提供できる。

【 0 1 1 9 】

G P 抗原 (3 (I V) N C 1) の 1 次構造が複数の立体配座を生じる複合体折り畳み過程の標的であることを報告した (1 3) 。単離された立体配列は超分子凝集体のリン酸化および恐らく 4 次構造形成により特異的に活性化される非最低エネルギー構造である。G P 患者では 3 (I V) N C 1 が立体配座の変化およびコラーゲン I V ネットワークのジスルフィド安定化を媒介する能力の低下を示す。そして G P 抗体が患者 3 (I V) N C 1 配座異性体に対してより強力な親和性を示し、これは立体配座が変化した物質が自己免疫応答を引き起こすことを示している。従って、G P 病では 3 (I V) N C 1 の立体配座形成過程における初期の変化が、免疫系が耐性でない配座異性体の変化を生じ、このように自己免疫応答を媒介するようである。

10

【 0 1 2 0 】

その他の証拠 (R a y a ら、結果は未発表) により、リン酸化が非最低エネルギー末端に 3 (I V) N C 1 の折り畳みを誘導するシグナルであることが示されている。このシナリオでは、ヒト 3 (I V) N C 1 系の三つの特徴はウシまたはネズミのような自然には G P 病にならない種からの対応する抗原系と比較した場合、発病原因と特別な関係にある。第 1 にヒト 3 (I V) N C 1 の N 末端は P K A によりおよび G P B P によりリン酸化されるモチーフを含有する (前記、および 2 から 4 をも参照のこと) 。第 2 にヒト遺伝子は代替エキソンスプライシングにより複数の代替生成物を生じる (1 4 、 1 5) 。エキソンスキッピングにより、1 次 3 (I V) N C 1 生成物の *i n v i t r o* での P K A リン酸化を上方制御する分岐した C 末端を有する代替生成物を生じる (以下の実施例 3 を参照のこと) 。第 3 にヒト G P B P は、G P 病の二つの主要な標的である系球体および肺胞基底膜に結合して発現される。リン酸化依存性立体配座形成過程はまた非病原性 N C 1 ドメインの特徴でもあり (1 3) 、これはリン酸化可能な N 末端、代替スプライシングによる多様化および系球体および肺胞基底膜での G P B P の発現は全て、3 (I V) N C 1 の立体配座形成過程を脆弱な状態に置く全ての排他的なヒトの特徴であることを示唆している。四つの別個の G P 腎臓実験では高レベルお G P 抗原代替生成物を発現し (1 5 ; B e r n a l および S a u s 、結果は未発表) 、G P 患者では G P B P 発現の増加が見出された (前記を参照のこと) 。代替 G P 抗原生成物および G P B P の両方のレベル増加は 3 (I V) N C 1 のリン酸化依存性立体配座形成過程における結果である、従って発病の可能性を有していると予測される。

20

30

【 0 1 2 1 】

G P B P は天然の自己免疫応答により標的化される皮膚において高度に発現される。表皮では G P B P はケラチノサイトが成熟中に基底から角質層になるアポトーシス媒介分化過程の細胞表面小胞と結合している (2 2 、 2 3) 。S L E 患者からのケラチノサイトは U V 誘起のアポトーシスに対して感受性が著明に高められており (6 、 1 8 、 2 0) 、S L E および皮膚筋炎では増強されたおよび時期尚早なケラチノサイトアポトーシスの存在が報告されている (7) 。一貫して、円盤状狼瘡 (図示せず) 、S L E および扁平苔癬などのいくつかの皮膚自己免疫状態において基底から表皮の周辺層までアポトーシス体が広がっているのが見出された。自己抗原およびその修飾体はアポトーシスケラチノサイトの細胞表面小胞に密集する (6 、 1 8 、 2 0) 。アポトーシス表面小胞は自己抗原 (2 1) を提供し、循環系に修飾体を放出する可能性がある。 (1 6 から 2 0) 。これはアポトーシス体からの修飾自己抗原の放出が全身性自己免疫応答媒介 S L E および強皮症を媒介する免疫化であり得ることを示唆している (1 8 、 1 9) 。

40

【 0 1 2 2 】

G P B P および G P B P 2 6 の両方が *i n v i t r o* でプロテインキナーゼとして作用でき、G P B P は G P B P 2 6 よりもより活性なアイソフォームであるということが

50

示される。さらに酵母からまたはヒト293細胞から精製されたGPBPまたはGPBP 26を表す組換え物質が3(IV)NC1ドメインを特異的に減成する、関連するタンパク質溶解活性を有した(結果は未発表)。タンパク質溶解活性は真核細胞発現系において生成された3(IV)NC1に機能するが、細菌において生成された組換え物質には機能せず(結果は未発表)、これは3(IV)NC1プロセッシングが原核細胞組換え物質には存在しないいくつかの立体配座形成または翻訳後要件を有することを示している。最終的に、いくつかの自己抗原がアポトーシス性ケラチノサイトにおいてリン酸化および減成を行うことが報告されている(20)。正確なメカニズムにまで限定はしていないが、前記の全データに鑑み、我々はアポトーシス性小胞で機械的に集めたGPBPは、リン酸化、立体配座変化および減成などの自己抗原の複雑な改変を行う可能性があることを提示する。従って、SLE(P1リボソームリントタンパク質およびSm-D1小型核リボヌクレオタンパク質)および皮膚筋炎(ヒスチジルtRNAシンセターゼ)において自己抗原を表す組換えタンパク質は*in vitro*でGPBPの基質であった(結果は未発表)。

10

【0123】

GPBPの培養癌細胞株における下方制御はGPBP/GPBP 26を機械的に繋げた細胞がプログラムされた細胞死を誘起するシグナル発現経路に参与している可能性があることを示唆している。対応するアポトーシス経路は自己免疫発病過程で上方制御して特異的MHCハロタイプを担持する個体の抗原展示を変化させ;細胞の形質転換中に下方制御して腫瘍成長中の形質転換された細胞への自己免疫攻撃を防御できる。

20

【0124】

実施例2の参考文献

- 1 Saus, J. (1998) グッドパスチャー症候群について: *Encyclopedia of Immunology* 第二版, Vol. 2, Delves, P. J., Roitt, I.M. 編, (Academic Press Limited, ロンドン) pp. 1005-1011.
- 2 Quinones, S., Bernal, D., Garcia-Sogo, M., Elena, S.F. および Saus, J. (1992) *J. Biol. Chem.* 267, 19780-19784
- 3 Revert, F., Penades J.R., Plana, M., Bernal, D., Johansson, C., Itarte, E., Cervera, J., Wieslander, J., Quinones, S. および Saus, J. (1995) *J. Biol. Chem.* 270, 13254-13261
- 4 Raya, A., Revert, F., Navarro, S. および Saus, J. (1999) *J. Biol. Chem.* 274, 12642-12649.
- 5 Green, M.R. (1986) *Ann. Rev. Genet.* 20, 671-708.
- 6 Casciola-Rosen, L.A., Anhalt, G. および Rosen, A. (1994) *J. Exp. Med.* 179: 1317-1330.
- 7 Pablos, J.L., Santiago, B., Galindo, M., Carreira, P.E., Ballesin, C. および Gomez-Reino, J.J. (1999) *J. Pathol.* 188: 63-68.
- 8 Srinivasan, M., Edman, C.F. および Schulman, H. (1994) *J. Cell. Biol.* 126, 839-852.
- 9 Naito, Y., Watanabe, Y., Yokokura, H., Sugita, R., Nishio, M. および Hidaka, H. (1997) *J. Biol. Chem.* 272, 32704-32708.
- 10 Bayer, K.-U., Loehler, J. および Harbers, K. (1996) *Mol. Cell. Biol.* 16, 29-36.

30

40

50

- 11 Mandaule, P., Eda, M., Watanabe, N, Fujisawa, K., Matsuoka, T., Bitto, H., Ishizaki, T.およびNarumiya, S. (1998) Nature 394, 491 - 494.
- 12 Papin, C., Denouel - Galy, A., Laugier, D., Calothy, G.およびEychene, A. (1998) J. Biol. Chem. 273, 24939 - 24947.
- 13 米国仮特許出願, 出願番号割り当て未定, 2000年2月11日登録 (事件番号98, 723 - C)
- 14 Penades J.R., Bernal, D., Revert, F., Johansson, C., Fresquet, V.J., Cervera, J., Wieslander, J., Quinones, S.およびSaus, J. (1995) Eur. J. Biochem. 229, 754 - 760. 10
- 15 Bernal, D., Quinones, S.およびSaus, J. (1993) J. Biol. Chem., 268, 12090 - 12094.
- 16 Casciola - Rosen, L.A., Anhalt, G.J.およびRosen, A. (1995) J. Exp. Med. 182: 1625 - 1634.
- 17 Casiano, C.A., Martin, S.J., Green, D.R.およびTan, E.M. (1996) J. Exp. Med. 184: 765 - 770. 20
- 18 Casciola - Rosen, L.およびRosen, A. (1997) Lupus 6: 175 - 180.
- 19 Bolivar, J., Guelman, S., Iglesias, C., Ortiz, M.およびVardivia, M. (1998) J. Biol. Chem. 273: 17122 - 17127.
- 20 Utz, P.J.およびAnderson, P. (1998) Arthritis Rheum. 41: 1152 - 1160.
- 21 Golan, T.D., Elkon, K.B., Ghavari, A.E.およびKrueger, J.G. (1992) J. Clin. Invest. 90: 1067 - 1076. 30
- 22 Polalowska, R.R., Piacentini, M., Bartlett, R., Goldsmith, L.A.およびHaake, A.R. (1994) Dev. Dinam. 199: 176 - 188.
- 23 Maruoka, Y., Harada, H., Mitsuyasu他 (1997) Biochem. Biophys. Res. Commun. 238: 886 - 890.
- 24 Xu, W., Harrison, S.C.およびEck, M.J. (1997) Nature 385, 595 - 602.

【0125】

(実施例3. エキソンスプライシングによるヒト自己抗原リン酸化の制御) 40
導入

GP病ではIV型コラーゲンの3鎖(GP抗原)の非コラーゲン性C末端ドメイン(NC1)に対する自己抗体により免疫系攻撃が媒介される(1)。ヒト3(IV)NC1のN末端は独特な構造モチーフ、KRGD⁹を有する高度に分岐した親水性領域を含有し、A型プロテインキナーゼの機能的リン酸化部位に必要な部分として細胞付着シグナルを保持している(2, 3)。さらに特徴的には、ヒトGP抗原をコードする遺伝子領域は代替エキソンスプライシングにより複数のmRNAを生じる(4, 5)。代替生成物はC末端で分岐し、一つを除いて全てがN末端KRGD⁹を共有する(4, 5)。

【0126】

多発性硬化症(MS)は中枢神経系で炎症性脱ミエリン化プラークの存在により特徴づけ 50

られる排他的ヒト神経学的疾患である(6)。この疾患が、ミエリン塩基性タンパク質(MBP)などの白質の特異的構成成分に対する、細胞毒性Tセルに媒介される自己免疫攻撃により誘起されるということがいくつかの証拠により示されている(7、8)。ヒトでは、プレmRNAプロセシング中の代替エキソンスプライシングの結果、MBP遺伝子が四つの生成物(MBP、MBP^{II}、MBP^VおよびMBP^{II/V})を生じる(9)。これらの中でMBP^{II}が成熟中枢神経系では最も豊富であり、一方全てのエキソンを含有するMBP体は実際には存在しない(9)。

【0127】

自己免疫応答媒介GP病およびMSにはいくつかの生物学的類似性が存在し、主に；1) 両者ともに排他的な疾患であり典型的にはウイルス性インフルエンザ様の疾患後に始まる；2) IIクラス MHCのHLA-DR領域の同一のハロタイプに強力な連関が存在する；3) 代替スプライシングによりいくつかの生成物を生じる；および4) GP病によるMS患者の死亡が最近報告されている(10)。

【0128】

材料および方法

合成ポリマー：GP^{IIII}由来のペプチド、QRAHGQDLDA LFVKVLRSP (配列番号：43) およびGP^{IIII/IV/V}由来のペプチド、QRAHGQDLES LFHQL (配列番号：44) をBoc-(メドブローブ)またはFmoc-(カイトン、リボテック)化学のいずれかを用いて合成した。

【0129】

プラスミド構築および組換え発現

GP由来物質：他の場合に修飾pet15bベクターのBamHI部位に、対応する組換えタンパク質を発現するcDNAをサブクローニングすることにより(5)、異なるGPスプライシング形態を表す構築物を得、ここで開始Metを除く外来性ベクター由来アミノ末端配列を排除した。ベクターをNcoIおよびBamHIで切断することにより余分な配列を除去し、遊離末端をクレノウで埋め、再度ライゲートした。これにより両方の制限部位を再形成し、アミノ終止Metのコドンのすぐ下流にBamHI部位を置いた。

【0130】

GPまたはGP^V(配列番号：46)を表す組換えタンパク質を沈殿により精製した(5)。GP^{IIII}(配列番号：48)またはGP^{IIII/IV/V}(配列番号：50)を表す組換えタンパク質を含有する細菌性ペレットを40mMのトリスHCl(pH6.8)中で8Mの尿素、および超音波処理により溶解した。40000xgで遠心した後、上澄を0.22μmフィルターに通し、FPLC用にリソースQカラムに注入した。溶出物をHClでpH6まで酸性にし、2回目のFPLC精製用に40mMのMES(pH6)で予め平衡にしたリソースSカラムに注入した。得られた溶出物中の物質をin vivoリン酸化に用いた。

【0131】

MBP由来物質：中枢神経系からの全RNAを用いてRT-PCRによりヒトMBP^{II}I(配列番号：51)を表すcDNAを得た。ヒトMBPを表すcDNAはC.Campagnoni(UCLA)から進呈された。両方のフラグメントをC末端で6xHisコーディング配列を含有するpHIL-D2(インビトロゲン)の修飾体にクローニングし、各々pHIL-MBP^{II}-HisおよびpHIL-MBP-Hisを作った。(12)に記載するようにこれらのプラスミドを用いて組換え体をPichia pastorisに発現させた。固定金属アフィニティクロマトグラフィ(TALON樹脂、クロンテック)を用いて、変性条件下(8M尿素)組換えタンパク質を精製し、製造者の指示書に従って300mMイミダゾールで溶出した。次いでアフィニティ精製物質を50mMトリスHCl(pH8.0)、10mM CHAPS、400mM NaCl、2mM DTTの80容量に希釈することにより復元し、YM10型膜(アミコン)を通して限外濾過することにより50倍に濃縮した。クロンテックの形質転換体変異誘発キットを用いて構築物を含有する元来の配列で位置指定変異誘発によりSerからAlaへの変異

10

20

30

40

50

体を作り、その結果得られるタンパク質は同様に製造された。

【0132】

リン酸化実験。基本的に前記のように(3および12も参照のこと)リン酸化実験を行った。いくつかの実験では基質をインプロット還元し、次いでrGPBP0.5μgを含有するリン酸化バッファー100μlを積層することにより、室温で30分間リン酸化した。V8エンドペプチターゼを用いる消化および免疫沈澱を(3)に記載するように実施した。

【0133】

抗体生成。各々配列番号：43または配列番号：44に含まれるGP IIIIまたはGP IIII/IV/VのC末端分岐末端を表す合成ペプチドをグルタルアルデヒドカップリング標準法を用いてチトロームC、BSAまたはオブアルブミンに抱合させた。得られたタンパク質抱合体をマウス免疫に用いてGP IIIIに特異的なポリクローナル抗体およびGP IIII/IV/Vに特異的なモノクローナル抗体(Mab153)を得た。GP Vに特異的なモノクローナル抗体(Mab5A)を得るために、配列番号50を含んだ対応する代替型を表現する組換え細菌性タンパク質を用い、マウスを免疫した。GP代替体のN末端領域を表す配列番号26に対するモノクローナル(M3/1、P1/2)またはポリクローナル(抗GP pep1)抗体の生成については前記した(3、5)。

【0134】

Boc基に基づくペプチド合成

構築。Boc-ベンジル法を用いて段階的固相合成によりペプチドを構築した。用いた出発樹脂はBoc-Pro-PAM樹脂(0.56meq/g、R4108バッチ)であった。用いた脱保護/カップリング方法は：TFA(1×1分)、TFA(1×3分)、DCM(フローフラッシュ)、イソプロピルアルコール(1×30秒)、DMF(3×1分)、カップリング/DMF(1×10分)、DMF(1×1分)、カップリング/DMF(1×10分)、DMF(2×1分)、DCM(1×1分)：であった。各ステップでペプチド樹脂グラムあたり10mlを用いた。DMF中BOP、HobtおよびDIEAの存在下全アミノ酸(5倍過剰)のカップリングを実施した。以下の側鎖保護基を合成するために：セリンにベンジル；リジンに2クロロベンジルオキシカルボニル；アスパラギン酸およびグルタミン酸にはシクロヘキシル；ヒスチジンおよびアルギニンにはトシル：を用いた。

【0135】

切断。樹脂からペプチドを切断し、液体フッ化水素(HF)と反応させることにより十分に脱保護した：ペプチド樹脂グラムあたりHF10mlを加え、捕捉剤としてp-クレゾールの存在下で、混合物を0℃下45分間保持した。HFを蒸発させた後、粗製反応混合物をエーテルで洗浄し、TFAに溶解し、エーテルで沈殿させ、乾燥した。

【0136】

精製。固定相：シリカC18、15μm、120Å；移動相：溶媒A：水0.1% TFAおよび溶媒B：アセトニトリル/A、60/40(容量/容量)；勾配：20から60%線形Bで30分間；流速：40ml/分；および検出はU.V.(210nm)。純度80%以上の分画をプールし、凍結乾燥した。純度および同一性の対照を分析用HPLCおよびES/MSにより実施した。最終生成物の純度は88%であり、実験上の分子量は2192.9であった。

【0137】

Fmoc-基盤のペプチド合成

構築。標準的なFmoc/tBu化学を用いてPro-クロロトリチル樹脂(0.685meq/g)で段階的直線的固相によりペプチドを合成した。用いた脱保護/カップリング方法は：Fmoc-aa(0.66g)、HOBt(0.26g)、DIPCDI(0.28ml)で40分間であり、カイザー試験による対照に従った。試験が陽性であった場合、陰性に変化するまで時間を延長した。次いでDMF(31分)、ピペリジン/DMF 20%(11分)、ピペリジン/DMF 20%(15分)およびDMF(41分)

10

20

30

40

50

。側鎖保護体は：アルギニンにPmc（ペンタメチルクロメーンスルホニル）、リジンにBcc（tert-ブトキシカルボニル）、アスパルギン酸およびセリンにtBu（tert-ブチル）、並びにヒスチジンにTrl（トリチル）：であった。

【0138】

切断。ペプチドを切断し、処理切断物をTFA/水 90/10と反応させることにより十分に脱保護した：樹脂グラムあたりTFA 10mlを加えた。水は捕捉剤として作用する。2時間後、樹脂を濾過し、得られた溶液を冷ジエチルエーテルで5回沈殿させた。最終沈殿物を乾燥した。

【0139】

精製。固定相：クロマシルC18、10 μ m；移動相：溶媒A：水0.1%TFAおよび溶媒B：アセトニトリル0.1%TFA；定組成：28%B；流速：55ml/分；検出：220nm。高純度の分画をプールし、凍結乾燥し、第2巡のHPLC精製を行った。純度および同一性の対照を分析用HPLCおよびES/MSにより実施した。最終生成物の純度は97%であり、実験上の分子量は2190.9であった。

【0140】

結果

代替スプライシングによるヒトGP抗原のリン酸化の制御。1次抗原（GP）または個々の代替生成物GPV（配列番号：46）、GPIII（配列番号：48）およびGPIII/IV/V（配列番号：50）を表す細菌性組換えタンパク質を製造し、PKAによるリン酸化される能力を試験した（図16、左のパネル）。標準的なATP濃度（150 μ M）を使用して四つの組換え抗原全てがリン酸化されたが、程度は異なった。代替型は1次GP抗原よりも効率よく³²Pを取り込んだが、このことはこれらがよりよい基質であることを示唆している。これらの抗原が細胞外区画にあると考えられるので、より生理学的なATP濃度（0.1から0.5 μ M）でまたリン酸化能力を検定した。これらの条件下では、1次および代替生成物間の³²P組み込みにおける差異がより明確であり、これは低ATP濃度では1次GP抗原がキナーゼにはあまりよくない基質であることを示した。GP抗原に存在する三つのPKAリン酸化部位の中で、N末端Ser⁹およびSer²⁶が主要なものであり、検定した全ての生成物に共通である（3、5）。従って、全ポリペプチドのリン酸化で観察された差異はまた特異的V8消化および免疫沈澱後に決定されるように個々のN末端領域間にも存在した（図示せず）。これはリン酸化における差異が代替生成物における異なるC末端配列の存在によるものであることを強く示唆している。GPIIIおよびGPIII/IV/VがGPVよりも著明に高い³²P組み込み率を示し、これらはより短い分岐C末端領域を有するので（5）、別個にこれらのC末端配列を表す合成ペプチド（配列番号：43、配列番号：44）を用いてさらに元来の抗原のin vivoリン酸化における制御的役割を試験した。IV型コラーゲンは三つの絡み合った鎖からなる三量体分子である。基底膜では、二つのIV型コラーゲン分子がNC1ドメインを介して集まり、細菌性コラゲナーゼ消化により溶解できる六量体NC1構造を生じる（1）。六量体構造の解離により単量体のGP抗原およびジスルフィド関連の二量体を放出する（1）。以下の一連の実験のために、細胞外様の低ATP濃度の存在下、単量体または六量体の両方の元来のGP抗原を用いてリン酸化を実施した（図16、右パネル）。各特異的ペプチドは存在するが対照ペプチドは存在しない場合（図示せず）、22kDaの見かけのMWを表示する単一のポリペプチドのリン酸化が誘起された。特異的V8消化および免疫沈澱により対応するポリペプチドが、以前に特徴づけられ、PKAに最適の基質であると同定された3（IV）NC1の22kDaの配座異性体として同定された（11）。

【0141】

代替スプライシングによるMBPのリン酸化の制御。MBPはN末端領域で、GP抗原生成物に存在するN末端部位（Ser⁹）に構造的に類似する2個のPKAリン酸化部位（Ser⁸、Ser⁵⁷）を含有する（図17）。全てのMBPタンパク質に存在するSer⁸部位はGP由来のポリペプチドのSer⁹よりも類似の位置にある。加えて、MBP

10

20

30

40

50

および GP I I I Ser⁸ および Ser⁹ は各々対応するエキソン I I に存在する高度に相関したモチーフの1次構造で類似の距離にある(図17の曲がった矢印)。GP I I I由来のモチーフはGP抗原系においてSer⁹のPKAリン酸化を上方制御するC末端分岐領域と一致する(図16)。MBPにおける制御様配列はエキソン I I に位置し、最終生成物中のその存在は代替エキソンスプライシングメカニズムに依存する。従って、GP I I Iに構造比較することにより同定されるMBPモチーフはまたSer⁸のPKAリン酸化をも制御できる。我々はMBPおよびMBP I I (配列番号: 54)を表す組換えタンパク質並びに対応するSerからAlaへの変異体を製造し、エキソン I に存在する2個のPKAリン酸化部位の各々(Ser⁸およびSer⁹)をロックアウトした。続いてPKAによる*in vitro*リン酸化を評価した(図18)。MBP I IはMBPよりも良好な基質であり、Ser⁸は主要なリン酸化部位であり、これはGP抗原系に類似して、代替エキソンスプライシングが全てのMBP由来代替形態に共通するN末端領域に位置する特異的部位のPKAリン酸化を制御することを示した。

10

【0142】

組換えMBPタンパク質のGPBPリン酸化を評価する類似の実験では、GPBPは優先的にMBPをリン酸化し、一方MBP I Iのリン酸化はほとんど観察されなかった(図19)。さらに、組換えSerからAlaへの変異体は³²P組み込みにおいて有意な減少を示さず、これはGPBPが、PKAとは反対の方法でMBP/MBP I Iをリン酸化し、これらの2個のキナーゼがMBPタンパク質における主要なリン酸化部位を共有しないことを示した。

20

【0143】

これらの全データから、我々はMBP系においてPKAまたはGPBPのいずれかにより、代替スプライシングが特異的セリンのリン酸化を制御すると結論づけた。

【0144】

GP I I IのC末端領域を表す合成ペプチドがGPBPリン酸化に影響する。GPBP活性に及ぼすGP I I IのC末端領域の影響を評価するために、この領域を表すペプチドを二種の異なる化学反応(BocまたはFmoc)を用いて合成し、GPBPを含有するリン酸化混合物に別個に添加した(図20)。Boc基盤の合成ペプチドはGPBP自己リン酸化に陽性に影響するが、Fmoc基盤のものはGPBP自己リン酸化を阻止し、これはGPおよびMBP抗原系のいずれかで、代替生成物に由来する制御配列がGPBPのキナーゼ活性に影響しうることを示唆している。

30

【0145】

考察

我々は $\alpha 3(I V)NC1$ ドメインが二つの異なるメカニズム: 1) 代替スプライシング(4、5)および2) 1次生成物の立体配座異性体形成(11)により複雑な構造多様性を作ること示した。両方のメカニズムによりPKAにより区別される生成物が生じるが、これはPKAリン酸化が $\alpha 3(I V)NC1$ ドメインの生物学において重大な事象であることを示している。リン酸化は少なくとも部分的な折り畳みのみならず、IV型コラーゲンネットワークの $\alpha 3(I V)NC1$ ドメインの超分子集合体をも導く(11およびRayら、結果は未発表)。 $\alpha 3(I V)NC1$ の配座異性体の変化によりGP病を媒介する自己免疫応答に至り(11)、これは抗原リン酸化の変化が疾患の発病における1次的事象であり得ることを示唆している。従って、いくつかのGP腎臓においてGP I I Iの発現の増加が認められ(4およびBernallおよびSaus、結果は未発表)、および別のグッドパスチャー患者においてGPBPの発現の増加が検出された(図15)。代替GP抗原生成物の、およびGPBPの両方の発現増加は $\alpha 3(I V)NC1$ のリン酸化の安定した状態、従って対応する立体配座形成過程にある結果であると考えられる。PKAによる異なる構造生成物間の区別はこのキナーゼ、または別の構造的に類似するキナーゼが生理学的抗原立体配座形成過程に関与し、GPBPによる抗原リン酸化が発病原因の重要性を有することが強く示唆される。発病過程においてGPBPは介入性のキナーゼであり、リン酸化依存的立体配座形成過程に干渉する。従って、天然の自己免疫応答によ

40

50

り標的化される組織構造においてG P B Pが発現され、G P B Pの発現増加はいくつかの自己免疫状態に係る(前記の実施例1および2を参照のこと)。

【0146】

代替スプライシングメカニズムはまたM B P抗原系において特異的セリンのP K Aリン酸化を制御する。M B PはまたG P B Pの基質であり、これはG P B Pが多発性硬化症、およびその他の自己免疫応答において病因的役割を果たし得ることを示唆している。

【0147】

前記の全データによりG P B Pが自己免疫疾患の治療に可能性のある標的として同定される。図20ではG P I I I(配列番号:43)のC末端領域を表す合成ペプチドが*in vitro*でG P B Pの作用を変調することを示し、従ってこの、および関連の配列が、*in vivo*でG P B Pの活性を変調するためのペプチド基盤の化合物として同定された。P K AによるG P抗原リン酸化の誘導はB o c基盤のペプチドを用いたときに達成されたが、類似のF m o c基盤のペプチドを用いたときには達成されなかった。さらに、F m o c基盤ではなくB o c基盤のペプチドはP K Aの*in vitro*基質であり(図示せず)、これは両方の生成物間に重要な構造上の差異が存在することを示している。両生成物が質量分析法において有意な差異を示さなかったので、一つの可能性は、用いた異なる脱保護方法が生物学的活性に重大な2次構造における立体配座の差異の原因となり得るということである。従って、B o c基盤のペプチドは低温で長期保存したとき、P K Aを誘起する能力を喪失する。

【0148】

実施例3の参考文献

1 Saus, J. (1998) グッドパスチャー症候群について: *Encyclopedia of Immunology* 第二版, Vol. 2, Delves, P. J. および Roitt, I. M. 編, (Academic Press Ltd., ロンドン), pp. 1005 - 1011.

2 Quinones, S., Bernal, D., Garcia-Sogo, M., Elena, S. F. および Saus, J. (1992) *J. Biol. Chem.* 267, 19780 - 19784

3 Revert, F., Penades J. R., Plana, M., Bernal, D., Johansson, C., Itarte, E., Cervera, J., Wieslander, J., Quinones, S. および Saus, J. (1995) *J. Biol. Chem.* 270, 13254 - 13261

4 Bernal, D., Quinones, S. および Saus, J. (1993) *J. Biol. Chem.* 268, 12090 - 12094

5 Penades J. R., Bernal, D., Revert, F., Johansson, C., Fresquet, V. J., Cervera, J., Wieslander, J., Quinones, S. および Saus, J. (1995) *Eur. J. Biochem.* 229, 754 - 760.

6 Raus, J. C. M., 多発性硬化症について: *Encyclopedia of Immunology* 第二版, Vol. 3, (Delves, P. J. および Roitt, I. M. 編) 1786 - 1789 (Academic Press Ltd., ロンドン, 1998).

7 Pette, M., Fujita, K., Wilkinson, D., Altmann, D. M., Trowsdale, J., Giegerich, G., Hinkkanen, A., Epplen, J. T., Kappos, L. および Wekerle, H. (1994) *Proc. Natl. Acad. Sci. USA* 87, 7968 - 7972

8 Tschida, T., Parker, K. C., Turner, R. V., McFarland, H. F., Coligan, J. E. および Bidd

10

20

30

40

50

ison, W. E. (1994) Proc. Natl. Acad. Sci. U S A 91, 10859 - 10863.

9 Campagnoni, A. T. (1988) J. Neurochem. 51, 1 - 14.

10 Henderson, R. D., Saltissi, D. および Pender, M. P. (1998) Acta Neurol. Scand. 98, 134 - 135

11 米国仮特許出願, 出願番号割り当て未定, 2000年2月11日登録(事件番号98, 723 - C)

12 Raya, A., Revert, F., Navarro, S. および Saus, J. (1999) J. Biol. Chem. 274, 12642 - 12649. 【0149】

本発明は前記の特定の好ましい態様により制限されない。通常の熟練した当業者には、本発明の概念から逸脱することなく、開示された好ましい態様に種々の改変を加えることができるだろう。かかる改変は全て本発明の範囲内であることを意図する。

【図面の簡単な説明】

【図1】 n4' のヌクレオチドおよび誘導アミノ酸配列。示した構造上の特徴は5' から3' 末端で: 由来のクローン (HeLa1) に存在する cDNA は (点線の囲み)、これは PH 相同性ドメイン (黒色) および Ser - Xaa - Yaa 反復 (灰色) を含有する; 2分された核局在化シグナルを含有する予想できる高次らせん構造 (開いた囲み) の7個の反復 (灰色); 並びにセリンに富むドメイン (灰色ボックスで埋められている) である。アスタリスクはフレーム停止コドンを示す。

【図2】 ヒト組織 (ノーザンプロット) および真核生物種 (サザンプロット) における GPBP の分布。ランダムプライミング^{3 2} P 標識 HeLa1 cDNA プロブを用いて、提示したヒト組織 (パネル A) からのポリ (A⁺) RNA のノーザンプロットにおいて、または提示した真核生物種 (パネル B) からのゲノム DNA のサザンプロットにおいて相同性メッセージを同定した。高度に厳密な条件下でノーザンハイブリダイゼーションを実施して完全な対合メッセージを検出し、サザンの低厳密性で誤対合のメッセージを検出できた。変性ゲル電気泳動により、または同一のロットからの代表的なプロットをヒトアクチン cDNA でプロブする場合、個々のポリ A + RNA の品質および量において評価できる差異は観察されなかった。数字は用いた RNA または DNA マーカーの位置および大きさを kb で示す。

【図3】 翻訳出発部位の実験的決定。A) では一過性の発現に用いられる pc - n4' および pc - FLAG - n4' プラスミドに存在する2個の cDNA が黒線で示されている。対応する予想される (n4') または操作された (FLAG - n4') 翻訳出発部位の相対位置を示す (Met)。 (B) では対照 (-)、pc - n4'、pc - FLAG - n4' または (FLAG - n4') トランスフェクトした 293 細胞からの抽出物を還元条件下 10% ゲル中 SDS - PAGE に供した。分離されたタンパク質を PVDF 膜 (ミリポア) に移し、指示した抗体でプロットングした。数字および棒線は分子量を kDa で、および分子量マーカーの相対位置を各々示す。

【図4】 酵母および 293 細胞からの rGPBP の特徴づけ。 (A) では酵母 rGPBP の 1 μg (レーン 1) または 100 μg (レーン 2 および 3) を 10% ゲル中還元 SDS - PAGE により分析した。分離したタンパク質をクーマシーブルーで染色する (レーン 1) かまたは移動させて抗 FLAG 抗体 (レーン 2) かまたは Mab 14 すなわち GPBP に対するモノクローナル抗体 (レーン 3) でプロットングした。 (B) では、GPBP 発現酵母からの細胞抽出物を A のように分析し、各々抗 FLAG (レーン 1)、抗 P Ser (レーン 2)、抗 P Thr (レーン 3) または抗 P Tyr (レーン 4) モノクローナル抗体でプロットングした。 (C) では酵母 rGPBP (レーン 1)、脱リン酸化酵母 rGPBP (レーン 2) または 293 細胞由来の rGPBP (レーン 3) のいずれか 200 ng を B のように指示した抗体を用いて分析した。 (D) では類似の量の H₃^{3 2}

10

20

30

40

50

PO₄ 標識トランスフェクトされていない(レーン1)、安定したpc-n4'トランスフェクトされた(レーン2)、一過性のpc-FLAG-n4'発現(レーン3)293細胞を溶解し、指示した抗体で沈殿させ、SDS-PAGEおよびオートラジオグラフィにより分析した。分子量マーカーを図3のように数字および棒線で表す。矢印はrGPBPの位置を示す。

【図5】 組換えGPBPはヒトGP抗原のN末端領域を特異的にリン酸化するセリン/スレオニンキナーゼを含有する。リン酸化を評価するために、酵母rGPBPおよそ200ngを[]³²P-ATPと共にGP抗原誘導物質の不在下(AまたはB)または存在下(C)インキュベートした。(A)では混合物を還元SDS-PAGE(10%ゲル)およびオートラジオグラフィに供した。(B)では混合物を2次元薄層クロマトグラフィによる³²Pホスホアミノ酸分析に供した。ニンヒドリン染色ホスホアミノ酸の位置を点線の円で示す。(C)では指示したGP誘導物質のリン酸化混合物をSDS-PAGE(15%ゲル)およびオートラジオグラフィにより分析する(GP pep1およびGP pep1 Ala⁹)か、またはMab17すなわちヒトおよびウシ由来のGP抗原を特異的に認識するモノクローナル抗体で免疫沈澱させ、SDS-PAGE(12.5%)およびオートラジオグラフィにより分析した(rGP、GP)。rGPBPの相対位置(A)、rGP抗原および元来のヒトおよびウシGP抗原(C)を矢印で示す。数字および棒線は前記の図面のように分子量マーカーを意味する。

10

【図6】 rGPBPに存在するセリン/スレオニンキナーゼのインプロット復元。酵母のrGPBP5μgをインプロット復元した。抗FLAG抗体(レーン1)およびオートラジオグラフィにより検出される切片上³²P組み込み(レーン2)により組換え物質を特異的に同定した。数字および棒線は前記の図面のように分子量マーカーを意味する。矢印は89kDa rGPBPポリペプチドの位置を示す。

20

【図7】 ヒト組織におけるGPBPの免疫学的局在。GPBPのN末端領域に対するウサギ血清(1:50)を用いてヒト組織においてGPBPを局在化した。示した組織は腎臓(A)、糸球体(B)、肺(C)、肺胞(D)、肝臓(E)、脳(F)、精巣(G)、副腎(H)、膵臓(I)および前立腺(J)である。抗GPBPアフィニティ精製抗体または7個の異なるGPBP特異的モノクローナル抗体からの培養培地のプール(抗GPBP Mabs 3、4、5、6、8、10および14)を用いて類似の結果が得られた。ウサギ前免疫血清はパラレル対照実験においていずれの組織構造をも染色しなかった。拡大図はBおよびDは100倍で、それ以外は40倍であった。

30

【図8】 GPBP 26はGPBPのスプライシング変種。(A)正常の骨格筋からの全RNAをプライマー53cを用いて逆転写し、続いてプライマー11m-53c(レーン2)または15m-62c(レーン4)を用いるPCRに供した。同一のプライマー対を用いるGPBP cDNA含有プラスミドの対照増幅をレーン1およびレーン3に示す。左および右の数字は塩基対の分子量を意味する。正常な筋肉転写物で喪失した領域を同定し、そのヌクレオチド配列(小文字)および推定アミノ酸(大文字)を(B)に示す。目的のcDNA領域を含むゲノムDNAのクローンを配列決定し、その構造を(C)に描き、GPBP 26でスプライシングした78bpエキソン(黒色ボックス)、隣接エキソン(灰色ボックス)およびイントロン(線)の位置および相対的な大きさを示す。イントロンおよびエキソンの両方の大きさが得られ、イントロン-エキソン境界のヌクレオチド配列を示し、5'および3'スプライシング部位のコンセンサスを太字で示す。

40

【図9】 GPBPおよびGPBP 26の差次的発現。78bpエキソン(GPBP)または両方のアイソフォーム(GPBP/GPBP 26)に共通の隣接配列を示すフラグメントを³²P標識し、これを用いてヒト組織および腫瘍の培養細胞株ノーザンブロット(クローンテック)をハイブリダイズした。膜を最初にGPBP特異的プローブでハイブリダイズし、剥離し、次いでGPBP/GPBP 26プローブで再分析した。洗浄条件はGPBP特異的プローブ用(0.1% SSPE、37 または55)にはGPBPおよびGPBP 26各々のシグナルを増強するGPBP/GPBP 26(0.1% SSPE、68)用ほど厳密ではない。洗浄プログラムが68 の場合、GPBPプロ

50

ープではシグナルが検出されなかった(図示せず)。

【図10】 GPBP 26はGPBPよりも低いリン酸化活性を示す。(A)組換え発現、アフィニティ精製GPBP(rGPBP)(レーン1)またはrGPBP 26(レーン2)を還元条件下SDS-PAGEに供し、クーマシーブルー染色する(レーンあたり2 μ g)かまたはFLAG配列(-FLAG)またはGPBP/GPBP 26(Mab14)を認識するモノクローナル抗体でプロットイング(レーンあたり200ng)した。(B)rGPBP(レーン1)またはrGPBP 26(レーン2)200ngを、基質を用いずにin vitroリン酸化し、自己リン酸化を評価する(左)かまたは、5ナノモルGPpep1を用いてトランスリン酸化活性(右)を測定した。矢印はペプチドの位置を示す。(C)rGPBP(レーン1)またはrGPBP 26(レーン2)3 μ gを材料および方法で記載するようにインプロット復元した。数字および棒線は各々分子量をkDaで、および分子量マーカーの相対位置を示す。

10

【図11】 rGPBPおよびrGPBP 26は非常に活性な高分子量凝集体を形成する。rGPBP(A)またはrGPBP 26(B)約300 μ gを材料および方法で記載したゲル濾過HPLCに供した。垂直方向の矢印および数字は各々用いた分子量標準の溶出プロフィールおよび分子量(kDa)を示す。大きな凝集体は空隙容量で溶出され(I)、サンプルに存在する物質の大部分は669および158kDaマーカーの間の第2のピークとしてカラムの分画範囲に溶出された(II)。指示した微小分画の15 μ lをSDS-PAGEおよびクーマシーブルー染色に供した。同一分画の5 μ lを材料および方法に記載するようにin vitroリン酸化し、SDSサンプルバッファー中沸騰させることにより反応を停止させる。分画をSDS-PAGEに負荷し、PVDFに移し、コダックX-オーマットフィルムを用いて1または2時間オートラジオグラフィに供し、抗FLAGモノクローナル抗体(シグマ)を用いてプロットイングした。

20

【図12】 酵母2ハイブリッド系によるGPBPおよびGPBP 26の自己相互作用。(A)指示した組み合わせのプラスミドでトランスフェクトした細胞をロイシントリプトファン欠損培地(-Trp、-Leu)で選別し、別個の形質転換体を1mMの3-アミノトリアゾール(3-AT)の存在下または不在下ヒスチジン欠損プレート(-Trp、-Leu、-His)上で再度画線培養し、相互作用を評価した。画線培養の3日後に写真撮影した。(B)棒線は4個の別個のガラクトシダーゼ溶液内アッセイのガラクトシダーゼの任意のユニットの平均値を表す。

30

【図13】 GPBPは内皮および糸球体基底膜に結合して発現される。ヒト筋肉(A)または腎臓皮質(B、C)のパラフィン包埋切片をGPBP特異的抗体(A、B)またはMab189すなわちヒト3(IV)NC1に特異的なモノクローナル抗体(C)でプローブした。ヒト腎臓の凍結切片(DからF)をMab17すなわちヒト3(IV)NC1ドメインに特異的なモノクローナル抗体(D)、GPBP特異的抗体(E)、またはGP患者からの血清(F)でプローブした。対照血清(ニワトリ前免疫およびヒト対照)はパラレル実験で組織結合性を示さなかった(図示せず)。

【図14】 GPBPはヒト腎臓皮質において発現するがウシおよびネズミ腎臓皮質では発現しない。ヒト(A、D)、ウシ(B、E)またはネズミ(C、F)腎臓の皮質をパラフィン包埋し、GPBP特異的抗体(AからC)またはGPBP/GPBP 26特異的抗体(DからF)のいずれかでプローブした。

40

【図15】 GPBPはいくつかの自己免疫状態で高度に発現される。増幅プログラムにおいてオリゴヌクレオチド15mおよび62cを用いて、対照の個体から(レーン1)またはGP患者から(レーン2)の骨格筋全RNAを図8のようにRT-PCRに供した。凍結(BからD)またはパラフィン包埋(EからG)ヒト対照皮膚(B、E)またはSLE(C、F)または扁平苔癬(D、G)の影響を受けた皮膚をGPBP特異的抗体でプローブした。

【図16】 PKAによるGP代替スプライシング生成物のリン酸化。左のパネルでは等モル量のrGPの(レーン1)、rGP V(レーン2)、rGP III(レーン3)またはrGP III/IV/V(レーン4)(GP500ngと等価)を指示したAT

50

P濃度でリン酸化した。全リン酸化反応混合物の5分の1をゲル電気泳動により分離し、P V D Fに移し、オートラジオグラフィに供し(図示せず)、四つ全種を認識する特異的モノクローナル抗体であるM3/1で、または個々のC末端領域に特異的な抗体を用いてタンパク質をブロッティングした(図示せず)。矢印は上から下にG P、G P VおよびG P I I I - G P I I I / I V / Vの各組換えタンパク質の位置を示し、これらは同一の移動性を表した。右パネル:精製3(I V) N C 1ドメインまたはヘキサマーをG P I I I (レーン2)またはG P I I I / I V / V (レーン3)のいずれかのC末端領域を表すペプチド10ナノモルの存在下または不在下(レーン1)、P K Aおよび0.1 μ M A T Pでリン酸化した。指示する場合、精製3(I V) N C 1ドメインのリン酸化混合物をV8消化し、ヒト3(I V) N C 1ドメインのN末端に特異的な抗体で免疫沈澱させた(3)。棒線および数字は分子量マーカの位置および大きさ(k D a)を示す。

10

【図17】 G P I I IおよびM B Pの配列アラインメント。P K Aのリン酸化部位(囲み)およびM B PおよびG P I I IのS e r 8および9の部位に関する構造類似性を各々示す(下線)。両方の分子種の対応するエキソンI I (曲がった矢印)の同一性(垂直棒線)および化学的相同性(点線)を示す。コラゲナーゼ切断部位(72残基)からのG P I I I完全配列を、エキソンIおよびエキソンI Iの10個の残基を含むM B Pの69N末端残基で整列させた。

【図18】 P K Aによる組換えM B Pタンパク質のリン酸化。約200 n gのr M B P (レーン1)、もしくは8位置(レーン2)もしくは57位置(レーン3)でのそのS e rからA l aへの変異体、またはr M P B I I (レーン4)もしくは8位置(レーン5)もしくは57位置(レーン6)でのそのS e rからA l aへの変異体をP K Aおよび0.1 μ M A T Pでリン酸化した。混合物をS D S - P A G Eに供し、P V D Fに移し、オートラジオグラフィに供し(リン酸化)、個々の分子種をロッシュモレキュラーバイオケミカルズ(ウェスタン)から入手のヒトM B Pに対するモノクローナル抗体でブロッティングした。

20

【図19】 G P B Pによる組換えM B Pタンパク質のリン酸化。約200 n gのr M B P (レーン1)、もしくは8位置(レーン2)もしくは57位置(レーン3)でのそのS e rからA l aへの変異体、またはr M P B I I (レーン4)もしくは8位置(レーン5)もしくは57位置(レーン6)でのそのS e rからA l aへの変異体をS D S - P A G Eに供し、P V D Fに移し、タンパク質を含有する部分をボンソーで可視化し、剥離した。固定したタンパク質を材料および方法で記載したG P B Pで切片上でリン酸化し、オートラジオグラフィに供し(リン酸化)、続いて図18のようにブロッティングした(ウェスタン)。

30

【図20】 G P I I IのC末端領域によるG P B Pの制御。約200 n gのr G P B Pを、150 μ MのA T Pを用い、B o c - (レーン2)またはF m o c - (レーン3)化学反応のいずれかを用いて合成したG P I I I誘導ペプチド5ナノモルの存在下または不在下で(レーン1)、i n v i t r oリン酸化した。反応混合物をS D S - P A G Eに供し、P V D Fに移し、オートラジオグラフィに供して自己リン酸化を評価し、続いて抗F L A Gモノクローナル抗体(シグマ)を用いてブロッティングし、存在する組換え物質の量を決定した(ウェスタン)。

40

【配列表】

SEQUENCE LISTING

<110> Saus, Juan

<120> Goodpasture Binding Protein

<130> 98-723-B

<140> To Be Assigned

<141> Filed Herewith

<160> 54

<170> PatentIn Ver. 2.0

<210> 1

<211> 2389

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (409)..(2280)

<400> 1

```

gcaggaagat ggcggcggtta gcggaggtgt gaggggacgc gggactcagc ggccggattt 60
tctcttccct tcttttccct tttcttccc tatttgaat tggcatcgag ggggctaagt 120
tcgggtggca gcgccgggca caacgcaggg gtcacggcga cggcggcggc ggctgacggc 180
tggaagggta ggcttcattc accgctcgtc ctcttctctc gctcgcctcg gtgtcagggc 240
cggcggcggc gcggcgggca gacttcgtcc ctctctctgc tccccccac accggagcgg 300
gcactctctg ctctgcctac ccccgacct tcaccccgag gactgggcgc ctctctcggc 360
gcagctgagg gagcgggggc cggctctctg ctcggttgtc gaggctcc atg tcg gat 417
                                         Met Ser Asp
                                         1
aat cag agc tgg aac tcg tcg ggc tcg gag gag gat cca gag acg gag 465
Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu Thr Glu
      5                10                15
tct ggg ccg cct gtg gag cgc tgc ggg gtc etc agt aag tgg aca aac 513
Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp Thr Asn
      20                25                30                35
tac att cat ggg tgg cag gat cgt tgg gta gtt ttg aaa aat aat gct 561
Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys Asn Asn Ala
      40                45                50
ctg agt tac tac aaa tct gaa gat gaa aca gag tat ggc tgc aga gga 609
Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys Arg Gly
      55                60                65
tcc atc tgt ctt agc aag gct gtc atc aca cct cac gat ttt gat gaa 657
Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe Asp Glu
      70                75                80

```

10

20

30

tgt cga ttt gat att agt gta aat gat agt gtt tgg tat ctt cgt gct	705
Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu Arg Ala	
85 90 95	
cag gat cca gat cat aga cag caa tgg ata gat gcc att gaa cag cac	753
Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu Gln His	
100 105 110 115	
aag act gaa tct gga tat gga tct gaa tcc agc ttg cgt cga cat ggc	801
Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg Arg His Gly	
120 125 130	
tca atg gtg tcc ctg gtg tct gga gca agt ggc tac tct gca aca tcc	849
Ser Met Val Ser Leu Val Ser Gly Ala Ser Gly Tyr Ser Ala Thr Ser	
135 140 145	
acc tct tca ttc aag aaa ggc cac agt tta cgt gag aag ttg gct gaa	897
Thr Ser Ser Phe Lys Lys Gly His Ser Leu Arg Glu Lys Leu Ala Glu	
150 155 160	
atg gaa aca ttt aga gac atc tta tgt aga caa gtt gac acg cta cag	945
Met Glu Thr Phe Arg Asp Ile Leu Cys Arg Gln Val Asp Thr Leu Gln	
165 170 175	
aag tac ttt gat gcc tgt gct gat gct gtc tct aag gat gaa ctt caa	993
Lys Tyr Phe Asp Ala Cys Ala Asp Ala Val Ser Lys Asp Glu Leu Gln	
180 185 190 195	
agg gat aaa gtg gta gaa gat gat gaa gat gac ttt cct aca acg cgt	1041
Arg Asp Lys Val Val Glu Asp Asp Glu Asp Asp Phe Pro Thr Thr Arg	
200 205 210	
tct gat ggt gac ttc ttg cat agt acc aac ggc aat aaa gaa aag tta	1089
Ser Asp Gly Asp Phe Leu His Ser Thr Asn Gly Asn Lys Glu Lys Leu	
215 220 225	
ttt cca cat gtg aca cca aaa gga att aat ggt ata gac ttt aaa ggg	1137
Phe Pro His Val Thr Pro Lys Gly Ile Asn Gly Ile Asp Phe Lys Gly	
230 235 240	
gaa gcg ata act ttt aaa gca act act gct gga atc ctt gca aca ctt	1185
Glu Ala Ile Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu Ala Thr Leu	
245 250 255	
tct cat tgt att gaa cta atg gtt aaa cgt gag gac agc tgg cag aag	1233
Ser His Cys Ile Glu Leu Met Val Lys Arg Glu Asp Ser Trp Gln Lys	
260 265 270 275	
aga ctg gat aag gaa act gag aag aaa aga aga aca gag gaa gca tat	1281
Arg Leu Asp Lys Glu Thr Glu Lys Lys Arg Arg Thr Glu Glu Ala Tyr	
280 285 290	
aaa aat gca atg aca gaa ctt aag aaa aaa tcc cac ttt gga gga cca	1329
Lys Asn Ala Met Thr Glu Leu Lys Lys Lys Ser His Phe Gly Gly Pro	
295 300 305	
gat tat gaa gaa ggc cct aac agt ctg att aat gaa gaa gag ttc ttt	1377
Asp Tyr Glu Glu Gly Pro Asn Ser Leu Ile Asn Glu Glu Glu Phe Phe	
310 315 320	
gat gct gtt gaa gct gct ctt gac aga caa gat aaa ata gaa gaa cag	1425

10

20

30

Glu Leu Gln Arg Asp Lys Val Val Glu Asp Asp Glu Asp Asp Phe Pro
 195 200 205
 Thr Thr Arg Ser Asp Gly Asp Phe Leu His Ser Thr Asn Gly Asn Lys
 210 215 220
 Glu Lys Leu Phe Pro His Val Thr Pro Lys Gly Ile Asn Gly Ile Asp
 225 230 235 240
 Phe Lys Gly Glu Ala Ile Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu
 245 250 255
 Ala Thr Leu Ser His Cys Ile Glu Leu Met Val Lys Arg Glu Asp Ser
 260 265 270
 Trp Gln Lys Arg Leu Asp Lys Glu Thr Glu Lys Lys Arg Arg Thr Glu
 275 280 285
 Glu Ala Tyr Lys Asn Ala Met Thr Glu Leu Lys Lys Lys Ser His Phe
 290 295 300
 Gly Gly Pro Asp Tyr Glu Glu Gly Pro Asn Ser Leu Ile Asn Glu Glu
 305 310 315 320
 Glu Phe Phe Asp Ala Val Glu Ala Ala Leu Asp Arg Gln Asp Lys Ile
 325 330 335
 Glu Glu Gln Ser Gln Ser Glu Lys Val Arg Leu His Trp Pro Thr Ser
 340 345 350
 Leu Pro Ser Gly Asp Ala Phe Ser Ser Val Gly Thr His Arg Phe Val
 355 360 365
 Gln Lys Pro Tyr Ser Arg Ser Ser Ser Met Ser Ser Ile Asp Leu Val
 370 375 380
 Ser Ala Ser Asp Asp Val His Arg Phe Ser Ser Gln Val Glu Glu Met
 385 390 395 400
 Val Gln Asn His Met Thr Tyr Ser Leu Gln Asp Val Gly Gly Asp Ala
 405 410 415
 Asn Trp Gln Leu Val Val Glu Glu Gly Glu Met Lys Val Tyr Arg Arg
 420 425 430
 Glu Val Glu Glu Asn Gly Ile Val Leu Asp Pro Leu Lys Ala Thr His
 435 440 445
 Ala Val Lys Gly Val Thr Gly His Glu Val Cys Asn Tyr Phe Trp Asn
 450 455 460
 Val Asp Val Arg Asn Asp Trp Glu Thr Thr Ile Glu Asn Phe His Val
 465 470 475 480
 Val Glu Thr Leu Ala Asp Asn Ala Ile Ile Ile Tyr Gln Thr His Lys
 485 490 495
 Arg Val Trp Pro Ala Ser Gln Arg Asp Val Leu Tyr Leu Ser Val Ile
 500 505 510
 Arg Lys Ile Pro Ala Leu Thr Glu Asn Asp Pro Glu Thr Trp Ile Val

10

20

30

gat gaa aca gaa tat ggc tgt agg gga tcc atc tgt ctt agc aag gct	665
Asp Glu Thr Glu Tyr Gly Cys Arg Gly Ser Ile Cys Leu Ser Lys Ala	
60 65 70	
gtg atc acg cct cac gat ttt gat gaa tgc cgg ttt gat atc agt gta	713
Val Ile Thr Pro His Asp Phe Asp Glu Cys Arg Phe Asp Ile Ser Val	
75 80 85 90	
aat gat agt gtt tgg tac ctt cga gct cag gac ccg gag cac aga cag	761
Asn Asp Ser Val Trp Tyr Leu Arg Ala Gln Asp Pro Glu His Arg Gln	
95 100 105	
caa tgg gta gac gcc att gaa cag cac aag act gaa tcg gga tat gga	809
Gln Trp Val Asp Ala Ile Glu Gln His Lys Thr Glu Ser Gly Tyr Gly	
110 115 120	
tct gag tcc agc ttg cgt aga cat ggc tca atg gtg tca ctg gtg tct	857
Ser Glu Ser Ser Leu Arg Arg His Gly Ser Met Val Ser Leu Val Ser	
125 130 135	
gga gcg agt ggc tat tct gct acg tcc acc tct tct ttc aag aaa ggc	905
Gly Ala Ser Gly Tyr Ser Ala Thr Ser Thr Ser Ser Phe Lys Lys Gly	
140 145 150	
cac agt tta cgt gag aaa ctg gct gaa atg gag aca ttt cgg gac atc	953
His Ser Leu Arg Glu Lys Leu Ala Glu Met Glu Thr Phe Arg Asp Ile	
155 160 165 170	
ctg tgc cgg cag gtt gat act ctc cag aag tac ttt gat gtc tgt gct	1001
Leu Cys Arg Gln Val Asp Thr Leu Gln Lys Tyr Phe Asp Val Cys Ala	
175 180 185	
gac gct gtc tcc aag gat gag ctt cag agg gat aaa gtc gta gaa gat	1049
Asp Ala Val Ser Lys Asp Glu Leu Gln Arg Asp Lys Val Val Glu Asp	
190 195 200	
gat gaa gat gac ttc cct aca act cgt tct gat gga gac ttt ttg cac	1097
Asp Glu Asp Asp Phe Pro Thr Thr Arg Ser Asp Gly Asp Phe Leu His	
205 210 215	
aat acc aat ggt aat aaa gaa aaa tta ttt cca cat gta aca cca aaa	1145
Asn Thr Asn Gly Asn Lys Glu Lys Leu Phe Pro His Val Thr Pro Lys	
220 225 230	
gga att aat ggc ata gac ttt aaa ggg gaa gca ata act ttt aaa gca	1193
Gly Ile Asn Gly Ile Asp Phe Lys Gly Glu Ala Ile Thr Phe Lys Ala	
235 240 245 250	
act act gct gga atc ctt gct aca ctt tct cat tgt att gaa tta atg	1241
Thr Thr Ala Gly Ile Leu Ala Thr Leu Ser His Cys Ile Glu Leu Met	
255 260 265	
gta aaa cgg gaa gag agc tgg caa aaa aga cac gat agg gaa gtg gaa	1289
Val Lys Arg Glu Glu Ser Trp Gln Lys Arg His Asp Arg Glu Val Glu	
270 275 280	
aag agg aga cga gtg gag gaa gcg tac aag aat gtg atg gaa gaa ctt	1337
Lys Arg Arg Arg Val Glu Glu Ala Tyr Lys Asn Val Met Glu Glu Leu	
285 290 295	
aag aag aaa ccc cgt ttc gga ggg ccg gat tat gaa gaa ggt cca aac	1385

10

20

30

Lys	Lys	Lys	Pro	Arg	Phe	Gly	Gly	Pro	Asp	Tyr	Glu	Glu	Gly	Pro	Asn		
300						305					310						
agt	ctg	att	aat	gag	gaa	gag	ttc	ttt	gat	gct	gtt	gaa	gct	gct	ctt	1433	
Ser	Leu	Ile	Asn	Glu	Glu	Glu	Phe	Phe	Asp	Ala	Val	Glu	Ala	Ala	Leu		
315				320					325						330		
gac	aga	caa	gat	aaa	ata	gag	gaa	cag	tca	cag	agt	gaa	aag	gtc	agg	1481	
Asp	Arg	Gln	Asp	Lys	Ile	Glu	Glu	Gln	Ser	Gln	Ser	Glu	Lys	Val	Arg		
				335				340						345			
tta	cac	tgg	ccc	aca	tca	ttg	cca	tct	gga	gac	acc	ttt	tct	tct	gtc	1529	
Leu	His	Trp	Pro	Thr	Ser	Leu	Pro	Ser	Gly	Asp	Thr	Phe	Ser	Ser	Val		
			350					355						360			
ggg	acg	cat	aga	ttt	gta	caa	aag	ccc	tat	agt	cgc	tct	tcc	tcc	atg	1577	
Gly	Thr	His	Arg	Phe	Val	Gln	Lys	Pro	Tyr	Ser	Arg	Ser	Ser	Ser	Met		
		365				370						375					
tct	tcc	att	gat	cta	gtc	agt	gcc	tct	gac	gat	gtt	cac	aga	ttc	agc	1625	
Ser	Ser	Ile	Asp	Leu	Val	Ser	Ala	Ser	Asp	Asp	Val	His	Arg	Phe	Ser		
		380				385					390						
tcc	cag	gtt	gaa	gaa	atg	gta	cag	aac	cac	atg	aac	tat	tca	tta	cag	1673	
Ser	Gln	Val	Glu	Glu	Met	Val	Gln	Asn	His	Met	Asn	Tyr	Ser	Leu	Gln		
395				400					405						410		
gat	gta	ggt	ggt	gat	gca	aat	tgg	caa	ctg	gtt	gtt	gaa	gaa	gga	gaa	1721	
Asp	Val	Gly	Gly	Asp	Ala	Asn	Trp	Gln	Leu	Val	Val	Glu	Glu	Gly	Glu		
				415				420						425			
atg	aag	gta	tac	aga	aga	gaa	gtg	gaa	gaa	aat	gga	att	gtt	ctg	gat	1769	
Met	Lys	Val	Tyr	Arg	Arg	Glu	Val	Glu	Glu	Asn	Gly	Ile	Val	Leu	Asp		
		430						435					440				
cct	ttg	aaa	gct	act	cat	gca	gtt	aaa	ggt	gtt	aca	gga	cat	gag	gtc	1817	
Pro	Leu	Lys	Ala	Thr	His	Ala	Val	Lys	Gly	Val	Thr	Gly	His	Glu	Val		
		445					450					455					
tgc	aat	tac	ttt	tgg	aat	gtt	gat	gtt	cgc	aat	gac	tgg	gaa	act	act	1865	
Cys	Asn	Tyr	Phe	Trp	Asn	Val	Asp	Val	Arg	Asn	Asp	Trp	Glu	Thr	Thr		
	460				465						470						
ata	gaa	aac	ttt	cat	gtg	gtg	gaa	aca	tta	gct	gat	aat	gca	atc	atc	1913	
Ile	Glu	Asn	Phe	His	Val	Val	Glu	Thr	Leu	Ala	Asp	Asn	Ala	Ile	Ile		
475				480					485					490			
gtt	tat	caa	acg	cac	aag	aga	gta	tgg	ccc	gct	tct	cag	aga	gac	gta	1961	
Val	Tyr	Gln	Thr	His	Lys	Arg	Val	Trp	Pro	Ala	Ser	Gln	Arg	Asp	Val		
				495				500						505			
ctg	tat	ctt	tct	gct	att	cga	aag	atc	cca	gcc	ttg	act	gaa	aat	gat	2009	
Leu	Tyr	Leu	Ser	Ala	Ile	Arg	Lys	Ile	Pro	Ala	Leu	Thr	Glu	Asn	Asp		
			510				515						520				
cct	gaa	act	tgg	ata	gtt	tgt	aat	ttt	tct	gtg	gat	cat	gat	agt	gct	2057	
Pro	Glu	Thr	Trp	Ile	Val	Cys	Asn	Phe	Ser	Val	Asp	His	Asp	Ser	Ala		
		525					530					535					
cct	ctg	aac	aat	cga	tgt	gtc	cgt	gcc	aaa	atc	aat	att	gct	atg	att	2105	
Pro	Leu	Asn	Asn	Arg	Cys	Val	Arg	Ala	Lys	Ile	Asn	Ile	Ala	Met	Ile		

10

20

30

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

10

20

30

gcctccaccg gcgcagctca gggagcgggg gccgggtctcc tgctcggctg tcgcgctcc 420
 atg tcg gat aac cag agc tgg aac tcg tcg ggc tcg gag gag gat ccg 468
 Met Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro
 1 5 10 15
 gag acg gag tcc ggg ccg ccg gtg gag cgc tgc gga gtc ctc aac aag 516
 Glu Thr Glu Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Asn Lys
 20 25 30
 tgg aca aac tat att cat ggg tgg cag gat cgc tgg gta gtt ttg aaa 564
 Trp Thr Asn Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys
 35 40 45
 aat aac act ctg agt tac tac aaa tct gaa gat gag aca gag tat ggc 612
 Asn Asn Thr Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly 10
 50 55 60
 tgc aga gga tcc atc tgt ctt agc aag gct gtc atc acg cct cat gat 660
 Cys Arg Gly Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp
 65 70 75 80
 ttt gat gaa tgc cga ttt gat att agt gta aat gat agt gtt tgg tat 708
 Phe Asp Glu Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr
 85 90 95
 ctt cgt gct caa gat cca gat cac aga cag cag tgg ata gat gcc att 756
 Leu Arg Ala Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile
 100 105 110
 gaa cag cac aag act gaa tct gga tat gga tct gaa tcc agc ttg cgt 804
 Glu Gln His Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg 20
 115 120 125
 cga cat ggc tcc atg gta tca ttg gta tcc gga gca agt ggc tat tct 852
 Arg His Gly Ser Met Val Ser Leu Val Ser Gly Ala Ser Gly Tyr Ser
 130 135 140
 gca aca tcc acc tcc tca ttc aag aag ggc cac agt tta cgt gag aaa 900
 Ala Thr Ser Thr Ser Ser Phe Lys Lys Gly His Ser Leu Arg Glu Lys
 145 150 155 160
 ctg gct gaa atg gaa acc ttt aga gat ata ctg tgt aga caa gtt gat 948
 Leu Ala Glu Met Glu Thr Phe Arg Asp Ile Leu Cys Arg Gln Val Asp
 165 170 175
 acc cta cag aag ttc ttt gat gcc tgt gct gat gct gtc tcc aag gat 996
 Thr Leu Gln Lys Phe Phe Asp Ala Cys Ala Asp Ala Val Ser Lys Asp 30
 180 185 190
 gaa ttt caa agg gat aaa gtg gta gaa gat gat gaa gat gac ttt cct 1044
 Glu Phe Gln Arg Asp Lys Val Val Glu Asp Asp Glu Asp Asp Phe Pro
 195 200 205
 acg aca cgt tct gat gga gac ttc ttg cat aat acc aat ggc aat aag 1092
 Thr Thr Arg Ser Asp Gly Asp Phe Leu His Asn Thr Asn Gly Asn Lys
 210 215 220
 gaa aag gta ttt cca cat gta aca cca aaa gga att aat ggt ata gac 1140
 Glu Lys Val Phe Pro His Val Thr Pro Lys Gly Ile Asn Gly Ile Asp
 225 230 235 240

ttt aaa ggt gag gcg ata act ttt aaa gca act act gcc gga atc ctt	1188
Phe Lys Gly Glu Ala Ile Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu	
245	250 255
gct aca ctt tct cat tgt att gag ctg atg gta aaa cgt gag gac agc	1236
Ala Thr Leu Ser His Cys Ile Glu Leu Met Val Lys Arg Glu Asp Ser	
260	265 270
tgg caa aag aga atg gac aag gaa act gag aag aga aga aga gtg gag	1284
Trp Gln Lys Arg Met Asp Lys Glu Thr Glu Lys Arg Arg Arg Val Glu	
275	280 285
gaa gca tac aaa aat gcc atg aca gaa ctt aag aaa aaa tcc cac ttt	1332
Glu Ala Tyr Lys Asn Ala Met Thr Glu Leu Lys Lys Lys Ser His Phe	
290	295 300
gga gga cca gat tat gag gaa ggc cca aac agt ttg att aat gaa gag	1380
Gly Gly Pro Asp Tyr Glu Glu Gly Pro Asn Ser Leu Ile Asn Glu Glu	
305	310 315 320
gag ttc ttt gat gct gtt gaa gct gct ctt gac aga caa gat aaa ata	1428
Glu Phe Phe Asp Ala Val Glu Ala Ala Leu Asp Arg Gln Asp Lys Ile	
325	330 335
gaa gaa cag tcg cag agt gaa aag gtc agg tta cat tgg tct act tca	1476
Glu Glu Gln Ser Gln Ser Glu Lys Val Arg Leu His Trp Ser Thr Ser	
340	345 350
atg cca tct gga gat gcc ttt tct tct gtg ggg act cat aga ttt gtc	1524
Met Pro Ser Gly Asp Ala Phe Ser Ser Val Gly Thr His Arg Phe Val	
355	360 365
caa aag ccc tat agt cgc tct tcc tcc atg tct tcc att gat cta gtc	1572
Gln Lys Pro Tyr Ser Arg Ser Ser Ser Met Ser Ser Ile Asp Leu Val	
370	375 380
agt gcc tct gac ggt gtt cac aga ttc agc tcc cag gtt gaa gag atg	1620
Ser Ala Ser Asp Gly Val His Arg Phe Ser Ser Gln Val Glu Glu Met	
385	390 395 400
gtg cag aac cac atg acc tat tca ttg cag gat gta ggt ggg gac gcc	1668
Val Gln Asn His Met Thr Tyr Ser Leu Gln Asp Val Gly Gly Asp Ala	
405	410 415
aac tgg cag ttg gtt gta gaa gaa ggg gag atg aag gta tat aga aga	1716
Asn Trp Gln Leu Val Val Glu Glu Gly Glu Met Lys Val Tyr Arg Arg	
420	425 430
gaa gta gaa gaa aat ggg att gtt ctg gat cct ttg aaa gct acc cat	1764
Glu Val Glu Glu Asn Gly Ile Val Leu Asp Pro Leu Lys Ala Thr His	
435	440 445
gca gtt aaa ggc gtt aca gga cac gag gtc tgc aat tac ttc tgg aat	1812
Ala Val Lys Gly Val Thr Gly His Glu Val Cys Asn Tyr Phe Trp Asn	
450	455 460
gct gat gtt cgc aat gat tgg gaa aca act ata gaa aac ttt cat gtg	1860
Val Asp Val Arg Asn Asp Trp Glu Thr Thr Ile Glu Asn Phe His Val	
465	470 475 480

10

20

30

gtg gaa aca tta gct gat aat gca atc atc att tat caa acg cac aag 1908
 Val Glu Thr Leu Ala Asp Asn Ala Ile Ile Ile Tyr Gln Thr His Lys
 485 490 495

aga gtg tgg cca gcc tct cag cgg gat gtc tta tat ctg tct gcc att 1956
 Arg Val Trp Pro Ala Ser Gln Arg Asp Val Leu Tyr Leu Ser Ala Ile
 500 505 510

cga aag ata cca gct ttg aat gaa aat gac ccg gag act tgg ata gtt 2004
 Arg Lys Ile Pro Ala Leu Asn Glu Asn Asp Pro Glu Thr Trp Ile Val
 515 520 525

tgt aat ttt tct gta gat cac agc agt gct cct cta aac aat cga tgt 2052
 Cys Asn Phe Ser Val Asp His Ser Ser Ala Pro Leu Asn Asn Arg Cys
 530 535 540

gtc cgt gcc aaa ata aac gtt gct atg att tgt cag acc ttg gtg agc 2100
 Val Arg Ala Lys Ile Asn Val Ala Met Ile Cys Gln Thr Leu Val Ser
 545 550 555 560

ccc cca gag gga aac cag gag att agc agg gac aac att cta tgc aag 2148
 Pro Pro Glu Gly Asn Gln Glu Ile Ser Arg Asp Asn Ile Leu Cys Lys
 565 570 575

att aca tac gtg gcc aat gta aac cct gga gga tgg gcc cca gcc tca 2196
 Ile Thr Tyr Val Ala Asn Val Asn Pro Gly Gly Trp Ala Pro Ala Ser
 580 585 590

gtg tta cgg gca gtg gca aag cga gaa tat cca aag ttt cta aag cgt 2244
 Val Leu Arg Ala Val Ala Lys Arg Glu Tyr Pro Lys Phe Leu Lys Arg
 595 600 605

ttt act tct tac gta caa gaa aaa act gca gga aaa cct att ttg ttc 2292
 Phe Thr Ser Tyr Val Gln Glu Lys Thr Ala Gly Lys Pro Ile Leu Phe
 610 615 620

tagtattaac agtgactgaa gcaaggctgt gtgacattcc atgttggagg aaaaaaaaaa 2352

aaaaaaaaa 2361

10

20

<210> 6
 <211> 624
 <212> PRT
 <213> Bos taurus

<400> 6
 Met Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro
 1 5 10 15
 Glu Thr Glu Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Asn Lys
 20 25 30
 Trp Thr Asn Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys
 35 40 45
 Asn Asn Thr Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly
 50 55 60
 Cys Arg Gly Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp
 65 70 75 80

30

Val Gln Asn His Met Thr Tyr Ser Leu Gln Asp Val Gly Gly Asp Ala
 405 410 415

Asn Trp Gln Leu Val Val Glu Glu Gly Glu Met Lys Val Tyr Arg Arg
 420 425 430

Glu Val Glu Glu Asn Gly Ile Val Leu Asp Pro Leu Lys Ala Thr His
 435 440 445

Ala Val Lys Gly Val Thr Gly His Glu Val Cys Asn Tyr Phe Trp Asn
 450 455 460

Val Asp Val Arg Asn Asp Trp Glu Thr Thr Ile Glu Asn Phe His Val
 465 470 475 480

Val Glu Thr Leu Ala Asp Asn Ala Ile Ile Ile Tyr Gln Thr His Lys
 485 490 495

Arg Val Trp Pro Ala Ser Gln Arg Asp Val Leu Tyr Leu Ser Ala Ile
 500 505 510

Arg Lys Ile Pro Ala Leu Asn Glu Asn Asp Pro Glu Thr Trp Ile Val
 515 520 525

Cys Asn Phe Ser Val Asp His Ser Ser Ala Pro Leu Asn Asn Arg Cys
 530 535 540

Val Arg Ala Lys Ile Asn Val Ala Met Ile Cys Gln Thr Leu Val Ser
 545 550 555 560

Pro Pro Glu Gly Asn Gln Glu Ile Ser Arg Asp Asn Ile Leu Cys Lys
 565 570 575

Ile Thr Tyr Val Ala Asn Val Asn Pro Gly Gly Trp Ala Pro Ala Ser
 580 585 590

Val Leu Arg Ala Val Ala Lys Arg Glu Tyr Pro Lys Phe Leu Lys Arg
 595 600 605

Phe Thr Ser Tyr Val Gln Glu Lys Thr Ala Gly Lys Pro Ile Leu Phe
 610 615 620

10

20

<210> 7
 <211> 2187
 <212> DNA
 <213> Homo sapiens

30

<220>
 <221> CDS
 <222> (391)..(2187)

<400> 7
 tagcgggaggt gtgagtgac gcggaactca gcgccggat tttctctcc cttctttcc 60
 ctttctctc cctatttgaa attggcatcg agggggctaa gttcgggtgg cagcggcggg 120
 cgcaacgcag gggtcacggc gacggcggcg gcggtgacg gctggaaggg taggcttcat 180
 tcaccgctcg tctctctcc tcgctccgct cggtgtcagg cgcggcggcg gcgcggggg 240

Leu	His	Ser	Thr	Asn	Gly	Asn	Lys	Glu	Lys	Leu	Phe	Pro	His	Val	Thr		
			220					225					230				
cca	aaa	gga	att	aat	ggt	ata	gac	ttt	aaa	ggg	gaa	gcg	ata	act	ttt	1134	
Pro	Lys	Gly	Ile	Asn	Gly	Ile	Asp	Phe	Lys	Gly	Glu	Ala	Ile	Thr	Phe		
		235					240					245					
aaa	gca	act	act	gct	gga	atc	ctt	goa	aca	ctt	tct	cat	tgt	att	gaa	1182	
Lys	Ala	Thr	Thr	Ala	Gly	Ile	Leu	Ala	Thr	Leu	Ser	His	Cys	Ile	Glu		
	250					255					260						
cta	atg	gct	aaa	cgt	gag	gac	agc	tgg	cag	aag	aga	ctg	gat	aag	gaa	1230	
Leu	Met	Val	Lys	Arg	Glu	Asp	Ser	Trp	Gln	Lys	Arg	Leu	Asp	Lys	Glu		
265					270				275						280		
act	gag	aag	aaa	aga	aga	aca	gag	gaa	gca	tat	aaa	aat	gca	atg	aca	1278	
Thr	Glu	Lys	Lys	Arg	Arg	Thr	Glu	Glu	Ala	Tyr	Lys	Asn	Ala	Met	Thr		
				285					290					295			
gaa	ctt	aag	aaa	aaa	tcc	cac	ttt	gga	gga	cca	gat	tat	gaa	gaa	ggc	1326	
Glu	Leu	Lys	Lys	Lys	Ser	His	Phe	Gly	Gly	Pro	Asp	Tyr	Glu	Glu	Gly		
			300					305					310				
cct	aac	agt	ctg	att	aat	gaa	gaa	gag	ttc	ttt	gat	gct	ggt	gaa	gct	1374	
Pro	Asn	Ser	Leu	Ile	Asn	Glu	Glu	Glu	Phe	Phe	Asp	Ala	Val	Glu	Ala		
		315				320						325					
gct	ctt	gac	aga	caa	gat	aaa	ata	gaa	gaa	cag	tca	cag	agt	gaa	aag	1422	
Ala	Leu	Asp	Arg	Gln	Asp	Lys	Ile	Glu	Glu	Gln	Ser	Gln	Ser	Glu	Lys		
		330				335					340						
gtg	aga	tta	cat	tgg	cct	aca	tcc	ttg	ccc	tct	gga	gat	gcc	ttt	tct	1470	
Val	Arg	Leu	His	Trp	Pro	Thr	Ser	Leu	Pro	Ser	Gly	Asp	Ala	Phe	Ser		
345					350					355					360		
tct	gtg	ggg	aca	cat	aga	ttt	gtc	caa	aag	ggt	gaa	gag	atg	gtg	cag	1518	
Ser	Val	Gly	Thr	His	Arg	Phe	Val	Gln	Lys	Val	Glu	Glu	Met	Val	Gln		
			365						370					375			
aac	cac	atg	act	tac	tca	tta	cag	gat	gta	ggc	gga	gat	gcc	aat	tgg	1566	
Asn	His	Met	Thr	Tyr	Ser	Leu	Gln	Asp	Val	Gly	Gly	Asp	Ala	Asn	Trp		
			380					385					390				
cag	ttg	ggt	gta	gaa	gaa	gga	gaa	atg	aag	gta	tac	aga	aga	gaa	gta	1614	
Gln	Leu	Val	Val	Glu	Glu	Gly	Glu	Met	Lys	Val	Tyr	Arg	Arg	Glu	Val		
		395					400					405					
gaa	gaa	aat	ggg	att	ggt	ctg	gat	cct	tta	aaa	gct	acc	cat	gca	ggt	1662	
Glu	Glu	Asn	Gly	Ile	Val	Leu	Asp	Pro	Leu	Lys	Ala	Thr	His	Ala	Val		
		410				415					420						
aaa	ggc	gtc	aca	gga	cat	gaa	gtc	tgc	aat	tat	ttc	tgg	aat	ggt	gac	1710	
Lys	Gly	Val	Thr	Gly	His	Glu	Val	Cys	Asn	Tyr	Phe	Trp	Asn	Val	Asp		
					430					435				440			
ggt	cgc	aat	gac	tgg	gaa	aca	act	ata	gaa	aac	ttt	cat	gtg	gtg	gaa	1758	
Val	Arg	Asn	Asp	Trp	Glu	Thr	Thr	Ile	Glu	Asn	Phe	His	Val	Val	Glu		
				445					450					455			
aca	tta	gct	gat	aat	gca	atc	atc	att	tat	caa	aca	cac	aag	agg	gtg	1806	
Thr	Leu	Ala	Asp	Asn	Ala	Ile	Ile	Ile	Tyr	Gln	Thr	His	Lys	Arg	Val		

10

20

30

460	465	470	
tgg cct gct tct cag cga gac gta tta tat ctt tct gtc att cga aag			1854
Trp Pro Ala Ser Gln Arg Asp Val Leu Tyr Leu Ser Val Ile Arg Lys			
475	480	485	
ata cca gcc ttg act gaa aat gac cct gaa act tgg ata gtt tgt aat			1902
Ile Pro Ala Leu Thr Glu Asn Asp Pro Glu Thr Trp Ile Val Cys Asn			
490	495	500	
ttt tct gtg gat cat gac agt gct cct cta aac aac cga tgt gtc cgt			1950
Phe Ser Val Asp His Asp Ser Ala Pro Leu Asn Asn Arg Cys Val Arg			
505	510	515	520
gcc aaa ata aat gtt gct atg att tgt caa acc ttg gta agc cca cca			1998
Ala Lys Ile Asn Val Ala Met Ile Cys Gln Thr Leu Val Ser Pro Pro			
525	530	535	10
gag gga aac cag gaa att agc agg gac aac att cta tgc aag att aca			2046
Glu Gly Asn Gln Glu Ile Ser Arg Asp Asn Ile Leu Cys Lys Ile Thr			
540	545	550	
tat gta gct aat gtg aac cct gga gga tgg gca cca gcc tca gtg tta			2094
Tyr Val Ala Asn Val Asn Pro Gly Gly Trp Ala Pro Ala Ser Val Leu			
555	560	565	
agg gca gtg gca aag cga gag tat cct aaa ttt cta aaa cgt ttt act			2142
Arg Ala Val Ala Lys Arg Glu Tyr Pro Lys Phe Leu Lys Arg Phe Thr			
570	575	580	
tct tac gtc caa gaa aaa act gca gga aag cct att ttg ttc tag			2187
Ser Tyr Val Gln Glu Lys Thr Ala Gly Lys Pro Ile Leu Phe			
585	590	595	20
<210> 8			
<211> 598			
<212> PRT			
<213> Homo sapiens			
<400> 8			
Met Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro			
1	5	10	15
Glu Thr Glu Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys			
20	25	30	
Trp Thr Asn Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys			
35	40	45	30
Asn Asn Ala Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly			
50	55	60	
Cys Arg Gly Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp			
65	70	75	80
Phe Asp Glu Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr			
85	90	95	
Leu Arg Ala Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile			
100	105	110	

Glu Gln His Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg
 115 120 125
 Arg His Gly Ser Met Val Ser Leu Val Ser Gly Ala Ser Gly Tyr Ser
 130 135 140
 Ala Thr Ser Thr Ser Ser Phe Lys Lys Gly His Ser Leu Arg Glu Lys
 145 150 155 160
 Leu Ala Glu Met Glu Thr Phe Arg Asp Ile Leu Cys Arg Gln Val Asp
 165 170 175
 Thr Leu Gln Lys Tyr Phe Asp Ala Cys Ala Asp Ala Val Ser Lys Asp
 180 185 190
 Glu Leu Gln Arg Asp Lys Val Val Glu Asp Asp Glu Asp Asp Phe Pro
 195 200 205
 Thr Thr Arg Ser Asp Gly Asp Phe Leu His Ser Thr Asn Gly Asn Lys
 210 215 220
 Glu Lys Leu Phe Pro His Val Thr Pro Lys Gly Ile Asn Gly Ile Asp
 225 230 235 240
 Phe Lys Gly Glu Ala Ile Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu
 245 250 255
 Ala Thr Leu Ser His Cys Ile Glu Leu Met Val Lys Arg Glu Asp Ser
 260 265 270
 Trp Gln Lys Arg Leu Asp Lys Glu Thr Glu Lys Lys Arg Arg Thr Glu
 275 280 285
 Glu Ala Tyr Lys Asn Ala Met Thr Glu Leu Lys Lys Lys Ser His Phe
 290 295 300
 Gly Gly Pro Asp Tyr Glu Glu Gly Pro Asn Ser Leu Ile Asn Glu Glu
 305 310 315 320
 Glu Phe Phe Asp Ala Val Glu Ala Ala Leu Asp Arg Gln Asp Lys Ile
 325 330 335
 Glu Glu Gln Ser Gln Ser Glu Lys Val Arg Leu His Trp Pro Thr Ser
 340 345 350
 Leu Pro Ser Gly Asp Ala Phe Ser Ser Val Gly Thr His Arg Phe Val
 355 360 365
 Gln Lys Val Glu Glu Met Val Gln Asn His Met Thr Tyr Ser Leu Gln
 370 375 380
 Asp Val Gly Gly Asp Ala Asn Trp Gln Leu Val Val Glu Glu Gly Glu
 385 390 395 400
 Met Lys Val Tyr Arg Arg Glu Val Glu Glu Asn Gly Ile Val Leu Asp
 405 410 415
 Pro Leu Lys Ala Thr His Ala Val Lys Gly Val Thr Gly His Glu Val
 420 425 430

10

20

30

Cys Asn Tyr Phe Trp Asn Val Asp Val Arg Asn Asp Trp Glu Thr Thr
 435 440 445

Ile Glu Asn Phe His Val Val Glu Thr Leu Ala Asp Asn Ala Ile Ile
 450 455 460

Ile Tyr Gln Thr His Lys Arg Val Trp Pro Ala Ser Gln Arg Asp Val
 465 470 475 480

Leu Tyr Leu Ser Val Ile Arg Lys Ile Pro Ala Leu Thr Glu Asn Asp
 485 490 495

Pro Glu Thr Trp Ile Val Cys Asn Phe Ser Val Asp His Asp Ser Ala
 500 505 510

Pro Leu Asn Asn Arg Cys Val Arg Ala Lys Ile Asn Val Ala Met Ile
 515 520 525

Cys Gln Thr Leu Val Ser Pro Pro Glu Gly Asn Gln Glu Ile Ser Arg
 530 535 540

Asp Asn Ile Leu Cys Lys Ile Thr Tyr Val Ala Asn Val Asn Pro Gly
 545 550 555 560

Gly Trp Ala Pro Ala Ser Val Leu Arg Ala Val Ala Lys Arg Glu Tyr
 565 570 575

Pro Lys Phe Leu Lys Arg Phe Thr Ser Tyr Val Gln Glu Lys Thr Ala
 580 585 590

Gly Lys Pro Ile Leu Phe
 595

10

20

<210> 9
 <211> 2684
 <212> DNA
 <213> Mus musculus

<220>
 <221> CDS
 <222> (444)..(2237)

<400> 9
 cgggccacca cgtgtaaata gtatcggacc cggcaggaag atggcggctg tagcggaggt 60
 gtgagtgagt ggatctgggt ctctgccggt ggcttggtc ttcccgtt cctcccctcc 120
 tccctccctg actgaggttg gcatctaggg ggccgagttc aggtggcggc gccgggagca 180
 ggcaggggt caccggccacg gcggctgacg gctggaaggg caggctttct tcgccctcg 240
 tctccttcc ccggtccgct cgggtgcagg cgcggcggcg gcggcggc gggcgcgctt 300
 cgtccccttt cctgttccct cactcccagg agcgggctct cttggcggtg ccatcccccg 360
 acccttacc ccagggacta ggcgctgca ctggcgcagc tcgaggagcg ggggcccgtc 420
 tctgctcgg ctgtgcgctc tcc atg tcg gat aac cag agc tgg aac tcg tcg 473
 Met Ser Asp Asn Gln Ser Trp Asn Ser Ser
 1 5 10

30

act acc gct gga atc ctt gct aca ctt tct cat tgt att gaa tta atg 1241
 Thr Thr Ala Gly Ile Leu Ala Thr Leu Ser His Cys Ile Glu Leu Met
 255 260 265

gta aaa cgg gaa gag agc tgg caa aaa aga cac gat agg gaa gtg gaa 1289
 Val Lys Arg Arg Glu Glu Ser Trp Gln Lys Arg His Asp Arg Glu Val Glu
 270 275 280

aag agg aga cga gtg gag gaa gcg tac aag aat gtg atg gaa gaa ctt 1337
 Lys Arg Arg Arg Val Glu Glu Ala Tyr Lys Asn Val Met Glu Glu Leu
 285 290 295

aag aag aaa ccc cgt ttc gga ggg cgg gat tat gaa gaa ggt cca aac 1385
 Lys Lys Lys Pro Arg Phe Glu Gly Pro Asp Tyr Glu Glu Gly Pro Asn
 300 305 310

agt ctg att aat gag gaa gag ttc ttt gat gct gtt gaa gct gct ctt 1433
 Ser Leu Ile Asn Glu Glu Phe Phe Asp Ala Val Glu Ala Ala Leu
 315 320 325 330

gac aga caa gat aaa ata gag gaa cag tca cag agt gaa aag gtc agg 1481
 Asp Arg Gln Asp Lys Ile Glu Glu Gln Ser Gln Ser Glu Lys Val Arg
 335 340 345

tta cac tgg ccc aca tca ttg cca tct gga gac acc ttt tct tct gtc 1529
 Leu His Trp Pro Thr Ser Leu Pro Ser Gly Asp Thr Phe Ser Ser Val
 350 355 360

ggg acg cat aga ttt gta caa aag gtt gaa gaa atg gta cag aac cac 1577
 Gly Thr His Arg Phe Val Gln Lys Val Glu Glu Met Val Gln Asn His
 365 370 375

atg aac tat tca tta cag gat gta ggt ggt gat gca aat tgg caa ctg 1625
 Met Asn Tyr Ser Leu Gln Asp Val Gly Gly Asp Ala Asn Trp Gln Leu
 380 385 390

gtt gtt gaa gaa gga gaa atg aag gta tac aga aga gaa gtg gaa gaa 1673
 Val Val Glu Glu Gly Glu Met Lys Val Tyr Arg Arg Glu Val Glu Glu
 395 400 405 410

aat gga att gtt ctg gat cct ttg aaa gct act cat gca gtt aaa ggt 1721
 Asn Gly Ile Val Leu Asp Pro Leu Lys Ala Thr His Ala Val Lys Gly
 415 420 425

gtt aca gga cat gag gtc tgc aat tac ttt tgg aat gtt gat gtt cgc 1769
 Val Thr Gly His Glu Val Cys Asn Tyr Phe Trp Asn Val Asp Val Arg
 430 435 440

aat gac tgg gaa act act ata gaa aac ttt cat gtg gtg gaa aca tta 1817
 Asn Asp Trp Glu Thr Thr Ile Glu Asn Phe His Val Val Glu Thr Leu
 445 450 455

gct gat aat gca atc atc gtt tat caa acg cac aag aga gta tgg ccc 1865
 Ala Asp Asn Ala Ile Ile Val Tyr Gln Thr His Lys Arg Val Trp Pro
 460 465 470

gct tct cag aga gac gta ctg tat ctt tct gct att cga aag atc cca 1913
 Ala Ser Gln Arg Asp Val Leu Tyr Leu Ser Ala Ile Arg Lys Ile Pro
 475 480 485 490

gcc ttg act gaa aat gat cct gaa act tgg ata gtt tgt aat ttt tct 1961

10

20

30

Ala Leu Thr Glu Asn Asp Pro Glu Thr Trp Ile Val Cys Asn Phe Ser
 495 500 505

gtg gat cat gat agt gct cct ctg aac aat cga tgt gtc cgt gcc aaa 2009
 Val Asp His Asp Ser Ala Pro Leu Asn Asn Arg Cys Val Arg Ala Lys
 510 515 520

atc aat att gct atg att tgt caa act tta gta agc cca cca gag gga 2057
 Ile Asn Ile Ala Met Ile Cys Gln Thr Leu Val Ser Pro Pro Glu Gly
 525 530 535

gac cag gag ata agc aga gac aac att ctg tgc aag atc acg tat gta 2105
 Asp Gln Glu Ile Ser Arg Asp Asn Ile Leu Cys Lys Ile Thr Tyr Val
 540 545 550

gct aat gtg aac cca gga gga tgg gcg cca gct tcg gtc tta aga gca 2153
 Ala Asn Val Asn Pro Gly Gly Trp Ala Pro Ala Ser Val Leu Arg Ala
 555 560 565 570

gtg gca aag cga gaa tac cct aag ttt cta aaa cgt ttt act tct tat 2201
 Val Ala Lys Arg Glu Tyr Pro Lys Phe Leu Lys Arg Phe Thr Ser Tyr
 575 580 585

gtc caa gaa aaa act gca gga aaa cca att ttg ttt tagtattaac 2247
 Val Gln Glu Lys Thr Ala Gly Lys Pro Ile Leu Phe
 590 595

agtgactgaa gcaaggctgc gtgacgttcc atggttgaga aaggagggaa aaaataaaaa 2307

gaatcctcta agctggaacg taggatctac agccttgtct gtggoccaag aagaaacatt 2367

gcaatcgtaa agctgggtat ccagcactag ccattctctg ctaggcctcc tcgctcagcg 2427

tgtaactata aatacatgta gaatcacatg gatatggcta tatttttatt tgcttgctcc 2487

ttggagtgaa aacaaataac tttgaattac aactaggaat taaccgatgc ttttaattttg 2547

aggaactttt tcagaatttt ttatttacca tggccaacc taagatcttc agttgtatca 2607

agttttttgtg cacaaaagaa aagcacaaaa gttgaacgca cctgaaggca tgtgctctct 2667

gtgcaacaaa tactcag 2684

10

20

<210> 10
 <211> 598
 <212> PRT
 <213> Mus musculus

30

<400> 10
 Met Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro
 1 5 10 15

Glu Thr Glu Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys
 20 25 30

Trp Thr Asn Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys
 35 40 45

Asn Asn Thr Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly
 50 55 60

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

10

20

30

Asp Val Gly Gly Asp Ala Asn Trp Gln Leu Val Val Glu Glu Gly Glu
 385 390 395 400
 Met Lys Val Tyr Arg Arg Glu Val Glu Glu Asn Gly Ile Val Leu Asp
 405 410 415
 Pro Leu Lys Ala Thr His Ala Val Lys Gly Val Thr Gly His Glu Val
 420 425 430
 Cys Asn Tyr Phe Trp Asn Val Asp Val Arg Asn Asp Trp Glu Thr Thr
 435 440 445
 Ile Glu Asn Phe His Val Val Glu Thr Leu Ala Asp Asn Ala Ile Ile
 450 455 460
 Val Tyr Gln Thr His Lys Arg Val Trp Pro Ala Ser Gln Arg Asp Val
 465 470 475 480
 Leu Tyr Leu Ser Ala Ile Arg Lys Ile Pro Ala Leu Thr Glu Asn Asp
 485 490 495
 Pro Glu Thr Trp Ile Val Cys Asn Phe Ser Val Asp His Asp Ser Ala
 500 505 510
 Pro Leu Asn Asn Arg Cys Val Arg Ala Lys Ile Asn Ile Ala Met Ile
 515 520 525
 Cys Gln Thr Leu Val Ser Pro Pro Glu Gly Asp Gln Glu Ile Ser Arg
 530 535 540
 Asp Asn Ile Leu Cys Lys Ile Thr Tyr Val Ala Asn Val Asn Pro Gly
 545 550 555 560
 Gly Trp Ala Pro Ala Ser Val Leu Arg Ala Val Ala Lys Arg Glu Tyr
 565 570 575
 Pro Lys Phe Leu Lys Arg Phe Thr Ser Tyr Val Gln Glu Lys Thr Ala
 580 585 590
 Gly Lys Pro Ile Leu Phe
 595

10

20

<210> 11
 <211> 2283
 <212> DNA
 <213> Bos taurus

30

<220>
 <221> CDS
 <222> (421)..(2214)

<400> 11
 eggcaggaag atggcggcct agcggaggtg tgagtggacc tgggtctctg cagctggggt 60
 ttcctctctt cegtctttct cctcttttcc tctccccga ggttggcacc gagggggcca 120
 aattcgggcg gcggcgccgg ggcagcgca ggggtcacia cgacggcgac ggctgacggt 180
 tggaaaggca ggcttctctt gccctcgac ctcttcccc ggtccgcttg gtgtcaggcg 240

cggcggcggc gggcggcggc ggcggcggc eggactccat cccctctccc gctccctcct 300
 gcaccgggagc gggcactcct tccttcgcca tcccccgacc cttcaccocg gggactgggc 360
 gcctccaccg ggcagctca gggagcgggg gccgggtctcc tgctcggtcg tcgcgcctcc 420
 atg tcg gat aac cag agc tgg aac tcg tcg ggc tcg gag gag gat ccg 468
 Met Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro
 1 5 10 15
 gag acg gag tcc ggg ccg ccg gtg gag cgc tgc gga gtc ctc aac aag 516
 Glu Thr Glu Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Asn Lys
 20 25 30
 tgg aca aac tat att cat ggg tgg cag gat cgc tgg gta gtt ttg aaa 564
 Trp Thr Asn Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys
 35 40 45
 aat aac act ctg agt tac tac aaa tct gaa gat gag aca gag tat ggc 612
 Asn Asn Thr Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly
 50 55 60
 tgc aga gga tcc atc tgt ctt agc aag gct gtc atc acg cct cat gat 660
 Cys Arg Gly Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp
 65 70 75 80
 ttt gat gaa tgc cga ttt gat att agt gta aat gat agt gtt tgg tat 708
 Phe Asp Glu Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr
 85 90 95
 ctt cgt gct caa gat cca gat cac aga cag cag tgg ata gat gcc att 756
 Leu Arg Ala Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile
 100 105 110
 gaa cag cac aag act gaa tct gga tat gga tct gaa tcc agc ttg cgt 804
 Glu Gln His Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg
 115 120 125
 cga cat ggc tcc atg gta tca ttg gta tcc gga gca agt ggc tat tct 852
 Arg His Gly Ser Met Val Ser Leu Val Ser Gly Ala Ser Gly Tyr Ser
 130 135 140
 gca aca tcc acc tcc tca ttc aag aag ggc cac agt tta cgt gag aaa 900
 Ala Thr Ser Thr Ser Ser Phe Lys Lys Gly His Ser Leu Arg Glu Lys
 145 150 155 160
 ctg gct gaa atg gaa acc ttt aga gat ata ctg tgt aga caa gtt gat 948
 Leu Ala Glu Met Glu Thr Phe Arg Asp Ile Leu Cys Arg Gln Val Asp
 165 170 175
 acc cta cag aag ttc ttt gat gcc tgt gct gat gct gtc tcc aag gat 996
 Thr Leu Gln Lys Phe Phe Asp Ala Cys Ala Asp Ala Val Ser Lys Asp
 180 185 190
 gaa ttt caa agg gat aaa gtg gta gaa gat gat gaa gat gac ttt cct 1044
 Glu Phe Gln Arg Asp Lys Val Val Glu Asp Asp Glu Asp Asp Phe Pro
 195 200 205
 acg aca cgt tct gat gga gac ttc ttg cat aat acc aat ggc aat aag 1092
 Thr Thr Arg Ser Asp Gly Asp Phe Leu His Asn Thr Asn Gly Asn Lys
 210 215 220

10

20

30

gaa aag gta ttt cca cat gta aca cca aaa gga att aat ggt ata gac 1140
 Glu Lys Val Phe Pro His Val Thr Pro Lys Gly Ile Asn Gly Ile Asp
 225 230 235 240

ttt aaa ggt gag gcg ata act ttt aaa gca act act gcc gga atc ctt 1188
 Phe Lys Gly Glu Ala Ile Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu
 245 250 255

gct aca ctt tct cat tgt att gag ctg atg gta aaa cgt gag gac agc 1236
 Ala Thr Leu Ser His Cys Ile Glu Leu Met Val Lys Arg Glu Asp Ser
 260 265 270

tgg caa aag aga atg gac aag gaa act gag aag aga aga aga gtg gag 1284
 Trp Gln Lys Arg Met Asp Lys Glu Thr Glu Lys Arg Arg Arg Val Glu
 275 280 285

gaa gca tac aaa aat gcc atg aca gaa ctt aag aaa aaa tcc cac ttt 1332
 Glu Ala Tyr Lys Asn Ala Met Thr Glu Leu Lys Lys Lys Ser His Phe
 290 295 300

gga gga cca gat tat gag gaa ggc cca aac agt ttg att aat gaa gag 1380
 Gly Gly Pro Asp Tyr Glu Glu Gly Pro Asn Ser Leu Ile Asn Glu Glu
 305 310 315 320

gag ttc ttt gat gct gtt gaa gct gct ctt gac aga caa gat aaa ata 1428
 Glu Phe Phe Asp Ala Val Glu Ala Ala Leu Asp Arg Gln Asp Lys Ile
 325 330 335

gaa gaa cag tcg cag agt gaa aag gtc agg tta cat tgg tct act tca 1476
 Glu Glu Gln Ser Gln Ser Glu Lys Val Arg Leu His Trp Ser Thr Ser
 340 345 350

atg cca tct gga gat gcc ttt tct tct gtg ggg act cat aga ttt gtc 1524
 Met Pro Ser Gly Asp Ala Phe Ser Ser Val Gly Thr His Arg Phe Val
 355 360 365

caa aag gtt gaa gag atg gtg cag aac cac atg acc tat tca ttg cag 1572
 Gln Lys Val Glu Glu Met Val Gln Asn His Met Thr Tyr Ser Leu Gln
 370 375 380

gat gta ggt ggg gac gcc aac tgg cag ttg gtt gta gaa gaa ggg gag 1620
 Asp Val Gly Gly Asp Ala Asn Trp Gln Leu Val Val Glu Glu Gly Glu
 385 390 395 400

atg aag gta tat aga aga gaa gta gaa gaa aat ggg att gtt ctg gat 1668
 Met Lys Val Tyr Arg Arg Glu Val Glu Glu Asn Gly Ile Val Leu Asp
 405 410 415

cct ttg aaa gct acc cat gca gtt aaa ggc gtt aca gga cac gag gtc 1716
 Pro Leu Lys Ala Thr His Ala Val Lys Gly Val Thr Gly His Glu Val
 420 425 430

tgc aat tac ttc tgg aat gtt gat gtt cgc aat gat tgg gaa aca act 1764
 Cys Asn Tyr Phe Trp Asn Val Asp Val Arg Asn Asp Trp Glu Thr Thr
 435 440 445

ata gaa aac ttt cat gtg gtg gaa aca tta gct gat aat gca atc atc 1812
 Ile Glu Asn Phe His Val Val Glu Thr Leu Ala Asp Asn Ala Ile Ile
 450 455 460

10

20

30

att tat caa acg cac aag aga gtg tgg cca gcc tct cag cgg gat gtc 1860
 Ile Tyr Gln Thr His Lys Arg Val Trp Pro Ala Ser Gln Arg Asp Val
 465 470 475 480
 tta tat ctg tct gcc att cga aag ata cca gct ttg aat gaa aat gac 1908
 Leu Tyr Leu Ser Ala Ile Arg Lys Ile Pro Ala Leu Asn Glu Asn Asp
 485 490 495
 ccg gag act tgg ata gtt tgt aat ttt tct gta gat cac agc agt gct 1956
 Pro Glu Thr Trp Ile Val Cys Asn Phe Ser Val Asp His Ser Ser Ala
 500 505 510
 cct cta aac aat cga tgt gtc cgt gcc aaa ata aac gtt gct atg att 2004
 Pro Leu Asn Asn Arg Cys Val Arg Ala Lys Ile Asn Val Ala Met Ile
 515 520 525
 tgt cag acc ttg gtg agc ccc cca gag gga aac cag gag att agc agg 2052
 Cys Gln Thr Leu Val Ser Pro Pro Glu Gly Asn Gln Glu Ile Ser Arg
 530 535 540
 gac aac att cta tgc aag att aca tac gtg gcc aat gta aac cct gga 2100
 Asp Asn Ile Leu Cys Lys Ile Thr Tyr Val Ala Asn Val Asn Pro Gly
 545 550 555 560
 gga tgg gcc cca gcc tca gtg tta cgg gca gtg gca aag cga gaa tat 2148
 Gly Trp Ala Pro Ala Ser Val Leu Arg Ala Val Ala Lys Arg Glu Tyr
 565 570 575
 cca aag ttt cta aag cgt ttt act tct tac gta caa gaa aaa act gca 2196
 Pro Lys Phe Leu Lys Arg Phe Thr Ser Tyr Val Gln Glu Lys Thr Ala
 580 585 590
 gga aaa cct att ttg ttc tagtattaac agtgactgaa gcaaggctgt 2244
 Gly Lys Pro Ile Leu Phe
 595
 gtgacattcc atggtggagg aaaaaaaaaa aaaaaaaaaa 2283

<210> 12
 <211> 598
 <212> PRT
 <213> Bos taurus

<400> 12
 Met Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro
 1 5 10 15
 Glu Thr Glu Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Asn Lys
 20 25 30
 Trp Thr Asn Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys
 35 40 45
 Asn Asn Thr Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly
 50 55 60
 Cys Arg Gly Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp
 65 70 75 80
 Phe Asp Glu Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr

10

20

30

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

10

20

30

Pro Leu Lys Ala Thr His Ala Val Lys Gly Val Thr Gly His Glu Val
 420 425 430

Cys Asn Tyr Phe Trp Asn Val Asp Val Arg Asn Asp Trp Glu Thr Thr
 435 440 445

Ile Glu Asn Phe His Val Val Glu Thr Leu Ala Asp Asn Ala Ile Ile
 450 455 460

Ile Tyr Gln Thr His Lys Arg Val Trp Pro Ala Ser Gln Arg Asp Val
 465 470 475 480

Leu Tyr Leu Ser Ala Ile Arg Lys Ile Pro Ala Leu Asn Glu Asn Asp
 485 490 495

Pro Glu Thr Trp Ile Val Cys Asn Phe Ser Val Asp His Ser Ser Ala
 500 505 510

Pro Leu Asn Asn Arg Cys Val Arg Ala Lys Ile Asn Val Ala Met Ile
 515 520 525

Cys Gln Thr Leu Val Ser Pro Pro Glu Gly Asn Gln Glu Ile Ser Arg
 530 535 540

Asp Asn Ile Leu Cys Lys Ile Thr Tyr Val Ala Asn Val Asn Pro Gly
 545 550 555 560

Gly Trp Ala Pro Ala Ser Val Leu Arg Ala Val Ala Lys Arg Glu Tyr
 565 570 575

Pro Lys Phe Leu Lys Arg Phe Thr Ser Tyr Val Gln Glu Lys Thr Ala
 580 585 590

Gly Lys Pro Ile Leu Phe
 595

10

20

<210> 13
 <211> 78
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(78)

<400> 13
 ccc tat agt cgc tct tcc tcc atg tct tcc att gat cta gtc agt gcc 48
 Pro Tyr Ser Arg Ser Ser Ser Met Ser Ser Ile Asp Leu Val Ser Ala
 1 5 10 15

tct gat gat gtt cac aga ttc agc tcc cag 78
 Ser Asp Asp Val His Arg Phe Ser Ser Gln
 20 25

30

<210> 14
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 14
 Pro Tyr Ser Arg Ser Ser Ser Met Ser Ser Ile Asp Leu Val Ser Ala
 1 5 10 15
 Ser Asp Asp Val His Arg Phe Ser Ser Gln
 20 25

<210> 15
 <211> 2034
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPBPR3

<220>
 <221> CDS
 <222> (10)..(990)

<400> 15
 gaattcacc atg gcc cca cta gcc gac tac aag gac gac gat gac aag atg 51
 Met Ala Pro Leu Ala Asp Tyr Lys Asp Asp Asp Asp Lys Met
 1 5 10

tcg gat aat cag agc tgg aac tcg tcg ggc tcg gag gag gat cca gag 99
 Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu
 15 20 25 30

acg gag tct ggg ccg cct gtg gag cgc tgc ggg gtc ctc agt aag tgg 147
 Thr Glu Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp
 35 40 45

aca aac tac att cat ggg tgg cag gat cgt tgg gta gtt ttg aaa aat 195
 Thr Asn Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys Asn
 50 55 60

aat gct ctg agt tac tac aaa tct gaa gat gaa aca gag tat ggc tgc 243
 Asn Ala Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys
 65 70 75

aga gga tcc atc tgt ctt agc aag gct gtc atc aca cct cac gat ttt 291
 Arg Gly Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe
 80 85 90

gat gaa tgt cga ttt gat att agt gta aat gat agt gtt tgg tat ctt 339
 Asp Glu Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu
 95 100 105 110

cgt gct cag gat cca gat cat aga cag caa tgg ata gat gcc att gaa 387
 Arg Ala Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu
 115 120 125

cag cac aag act gaa tct gga tat gga tct gaa tcc agc ttg cgt cga 435
 Gln His Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg Arg
 130 135 140

cat ggc tca atg gtg tcc ctg gtg tct gga gca agt ggc tac tct gca 483
 His Gly Ser Met Val Ser Leu Val Ser Gly Ala Ser Gly Tyr Ser Ala
 145 150 155

10

20

30

aca tcc acc tct tca ttc aag aaa ggc cac agt tta cgt gag aag ttg 531
 Thr Ser Thr Ser Ser Phe Lys Lys Gly His Ser Leu Arg Glu Lys Leu
 160 165 170

gct gaa atg gaa aca ttt aga gac atc tta tgt aga caa gtt gac acg 579
 Ala Glu Met Glu Thr Phe Arg Asp Ile Leu Cys Arg Gln Val Asp Thr
 175 180 185 190

cta cag aag tac ttt gat gcc tgt gct gat gct gtc tct aag gat gaa 627
 Leu Gln Lys Tyr Phe Asp Ala Cys Ala Asp Ala Val Ser Lys Asp Glu
 195 200 205

ctt caa agg gat aaa gtg gta gaa gat gat gaa gat gac ttt cct aca 675
 Leu Gln Arg Asp Lys Val Val Glu Asp Asp Glu Asp Asp Phe Pro Thr
 210 215 220

acg cgt tct gat ggt gac ttc ttg cat agt acc aac ggc aat aaa gaa 723
 Thr Arg Ser Asp Gly Asp Phe Leu His Ser Thr Asn Gly Asn Lys Glu
 225 230 235

aag tta ttt cca cat gtg aca cca aaa gga att aat ggt ata gac ttt 771
 Lys Leu Phe Pro His Val Thr Pro Lys Gly Ile Asn Gly Ile Asp Phe
 240 245 250

aaa ggg gaa gcg ata act ttt aaa gca act act gct gga atc ctt gca 819
 Lys Gly Glu Ala Ile Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu Ala
 255 260 265 270

aca ctt tct cat tgt att gaa cta atg gtt aaa cgt gag gac agc tgg 867
 Thr Leu Ser His Cys Ile Glu Leu Met Val Lys Arg Glu Asp Ser Trp
 275 280 285

cag aag aga ctg gat aag gaa act gag aag aaa aga aga aca gag gaa 915
 Gln Lys Arg Leu Asp Lys Glu Thr Glu Lys Lys Arg Arg Thr Glu Glu
 290 295 300

gca tat aaa aat gca atg aca gaa cga aaa aat ccc act ttg gag gac 963
 Ala Tyr Lys Asn Ala Met Thr Glu Arg Lys Asn Pro Thr Leu Glu Asp
 305 310 315

cag att atg aag aag gcc cta aca gtc tgattaatga agaagagttc 1010
 Gln Ile Met Lys Lys Ala Leu Thr Val
 320 325

ttgatgctg ttgaagctgc tcttgacaga caagataaaa tagaagaaca gtcacagagt 1070

gaaaagggtga gattacattg gctacatcc ttgccctctg gagatgcctt ttcttctgtg 1130

gggacacata gatttgcca aaagccctat agtcgctctt cctccatgac ttccattgat 1190

ctagtcaagt cctctgatga tgttcacaga ttcagctccc aggttgaaga gatggtgcag 1250

aaccacatga cttactcatt acaggatgta ggcggagatg ccaattggca gttggttcta 1310

gaagaaggag aaatgaaggt atacagaaga gaagtgaag aaaatgggat tgttctggat 1370

cctttaaag ctacccatgc agttaaaggc gtcacaggac atgaagtctg caattatttc 1430

tggaatgttg acgttcgcaa tgactgggaa acaactatag aaaactttca tgtggtggaa 1490

10

20

30

acattagctg ataatgcaat catcatttat caaacacaca agaggggtgtg gcctgcttct 1550
cagcgagacg tattatatct ttctgtcatt cgaaagatac cagccttgac tgaaaatgac 1610
cctgaaaactt ggatagtttg taatttttct gtggatcatg acagtgtctc tctaaacaac 1670
cgatgtgtcc gtgccaaaaat aaatgttgct atgatttgtc aaaccttggt aagcccacca 1730
gagggaaaacc aggaaattag cagggacaac attctatgca agattacata tgtagctaata 1790
gtgaacctcg gaggatgggc accagcctca gtgttaaggg cagtggcaaaa gcgagagtat 1850
cctaaatttc taaaacgttt tactttctac gtccaagaaa aaactgcagg aaagcctatt 1910
ttgttctagt attaacaggt actagaagat atgttttatac tttttttaac tttatttgac 1970
taatatgact gtcaatacta aaatntagtt gttgaaagta tttactatgt tttttccgga 2030
attc 2034

10

<210> 16
<211> 327
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: GPBPR3

<400> 16
Met Ala Pro Leu Ala Asp Tyr Lys Asp Asp Asp Asp Lys Met Ser Asp
1 5 10 15
Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu Thr Glu
20 25 30
Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp Thr Asn
35 40 45
Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys Asn Asn Ala
50 55 60
Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys Arg Gly
65 70 75 80
Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe Asp Glu
85 90 95
Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu Arg Ala
100 105 110
Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu Gln His
115 120 125
Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg Arg His Gly
130 135 140
Ser Met Val Ser Leu Val Ser Gly Ala Ser Gly Tyr Ser Ala Thr Ser
145 150 155 160
Thr Ser Ser Phe Lys Lys Gly His Ser Leu Arg Glu Lys Leu Ala Glu

20

30

165 170 175

Met Glu Thr Phe Arg Asp Ile Leu Cys Arg Gln Val Asp Thr Leu Gln
180 185 190

Lys Tyr Phe Asp Ala Cys Ala Asp Ala Val Ser Lys Asp Glu Leu Gln
195 200 205

Arg Asp Lys Val Val Glu Asp Asp Glu Asp Asp Phe Pro Thr Thr Arg
210 215 220

Ser Asp Gly Asp Phe Leu His Ser Thr Asn Gly Asn Lys Glu Lys Leu
225 230 235 240

Phe Pro His Val Thr Pro Lys Gly Ile Asn Gly Ile Asp Phe Lys Gly
245 250 255

Glu Ala Ile Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu Ala Thr Leu
260 265 270

Ser His Cys Ile Glu Leu Met Val Lys Arg Glu Asp Ser Trp Gln Lys
275 280 285

Arg Leu Asp Lys Glu Thr Glu Lys Lys Arg Arg Thr Glu Glu Ala Tyr
290 295 300

Lys Asn Ala Met Thr Glu Arg Lys Asn Pro Thr Leu Glu Asp Gln Ile
305 310 315 320

Met Lys Lys Ala Leu Thr Val
325

10

20

<210> 17
 <211> 1978
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: FLAG-GPPPDNLS

<220>
 <221> CDS
 <222> (10)..(1860)

<400> 17
 gaattcacc atg gcc cca cta gcc gac tac aag gac gac gat gac aag atg 51
 Met Ala Pro Leu Ala Asp Tyr Lys Asp Asp Asp Lys Met
 1 5 10

tcg gat aat cag agc tgg aac tcg tcg ggc tcg gag gag gat cca gag 99
 Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu
 15 20 25 30

acg gag tct ggg ccg cct gtg gag cgc tgc ggg gtc ctc agt aag tgg 147
 Thr Glu Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp
 35 40 45

aca aac tac att cat ggg tgg cag gat cgt tgg gta gtt ttg aaa aat 195
 Thr Asn Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys Asn
 50 55 60

30

aat gct ctg agt tac tac aaa tct gaa gat gaa aca gag tat ggc tgc	243
Asn Ala Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys	
65 70 75	
aga gga tcc atc tgt ctt agc aag gct gtc atc aca cct cac gat ttt	291
Arg Gly Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe	
80 85 90	
gat gaa tgt cga ttt gat att agt gta aat gat agt gtt tgg tat ctt	339
Asp Glu Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu	
95 100 105 110	
cgt gct cag gat cca gat cat aga cag caa tgg ata gat gcc att gaa	387
Arg Ala Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu	
115 120 125	
cag cac aag act gaa tct gga tat gga tct gaa tcc agc ttg cgt cga	435
Gln His Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg Arg	
130 135 140	
cat ggc tca atg gtg tcc ctg gtg tct gga gca agt ggc tac tct gca	483
His Gly Ser Met Val Ser Leu Val Ser Gly Ala Ser Gly Tyr Ser Ala	
145 150 155	
aca tcc acc tct tca ttc aag aaa ggc cac agt tta cgt gag aag ttg	531
Thr Ser Thr Ser Ser Phe Lys Lys Gly His Ser Leu Arg Glu Lys Leu	
160 165 170	
gct gaa atg gaa aca ttt aga gac atc tta tgt aga caa gtt gac acg	579
Ala Glu Met Glu Thr Phe Arg Asp Ile Leu Cys Arg Gln Val Asp Thr	
175 180 185 190	
cta cag aag tac ttt gat gcc tgt gct gat gct gtc tct aag gat gaa	627
Leu Gln Lys Tyr Phe Asp Ala Cys Ala Asp Ala Val Ser Lys Asp Glu	
195 200 205	
ctt caa agg gat aaa gtg gta gaa gat gat gaa gat gac ttt cct aca	675
Leu Gln Arg Asp Lys Val Val Glu Asp Asp Glu Asp Asp Phe Pro Thr	
210 215 220	
acg cgt tct gat ggt gac ttc ttg cat agt acc aac ggc aat aaa gaa	723
Thr Arg Ser Asp Gly Asp Phe Leu His Ser Thr Asn Gly Asn Lys Glu	
225 230 235	
aag tta ttt cca cat gtg aca cca aaa gga att aat ggt ata gac ttt	771
Lys Leu Phe Pro His Val Thr Pro Lys Gly Ile Asn Gly Ile Asp Phe	
240 245 250	
aaa ggg gaa gcg ata act ttt aaa gca act act gct gga atc ctt gca	819
Lys Gly Glu Ala Ile Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu Ala	
255 260 265 270	
aca ctt tct cat tgt att gaa cta atg gtt aaa cgt gag gac agc tgg	867
Thr Leu Ser His Cys Ile Glu Leu Met Val Lys Arg Glu Asp Ser Trp	
275 280 285	
cag aag aga ctg gat aag gaa act gag cac ttt gga gga cca gat tat	915
Gln Lys Arg Leu Asp Lys Glu Thr Glu His Phe Gly Gly Pro Asp Tyr	
290 295 300	

10

20

30

gaa gaa ggc cct aac agt ctg att aat gaa gaa gag ttc ttt gat gct 963
 Glu Glu Gly Pro Asn Ser Leu Ile Asn Glu Glu Glu Phe Phe Asp Ala
 305 310 315

gtt gaa gct gct ctt gac aga caa gat aaa ata gaa gaa cag tca cag 1011
 Val Glu Ala Ala Leu Asp Arg Gln Asp Lys Ile Glu Glu Gln Ser Gln
 320 325 330

agt gaa aag gtg aga tta cat tgg cct aca tcc ttg ccc tct gga gat 1059
 Ser Glu Lys Val Arg Leu His Trp Pro Thr Ser Leu Pro Ser Gly Asp
 335 340 345 350

gcc ttt tct tct gtg ggg aca cat aga ttt gtc caa aag ccc tat agt 1107
 Ala Phe Ser Ser Val Gly Thr His Arg Phe Val Gln Lys Pro Tyr Ser
 355 360 365

cgc tct tcc tcc atg tct tcc att gat cta gtc agt gcc tct gat gat 1155
 Arg Ser Ser Ser Met Ser Ser Ile Asp Leu Val Ser Ala Ser Asp Asp
 370 375 380

gtt cac aga ttc agc tcc cag gtt gaa gag atg gtg cag aac cac atg 1203
 Val His Arg Phe Ser Ser Gln Val Glu Glu Met Val Gln Asn His Met
 385 390 395

act tac tca tta cag gat gta ggc gga gat gcc aat tgg cag ttg gtt 1251
 Thr Tyr Ser Leu Gln Asp Val Gly Gly Asp Ala Asn Trp Gln Leu Val
 400 405 410

gta gaa gaa gga gaa atg aag gta tac aga aga gaa gta gaa gaa aat 1299
 Val Glu Glu Gly Glu Met Lys Val Tyr Arg Arg Glu Val Glu Glu Asn
 415 420 425 430

ggg att gtt ctg gat cct tta aaa gct acc cat gca gtt aaa ggc gtc 1347
 Gly Ile Val Leu Asp Pro Leu Lys Ala Thr His Ala Val Lys Gly Val
 435 440 445

aca gga cat gaa gtc tgc aat tat ttc tgg aat gtt gac gtt cgc aat 1395
 Thr Gly His Glu Val Cys Asn Tyr Phe Trp Asn Val Asp Val Arg Asn
 450 455 460

gac tgg gaa aca act ata gaa aac ttt cat gtg gtg gaa aca tta gct 1443
 Asp Trp Glu Thr Thr Ile Glu Asn Phe His Val Val Glu Thr Leu Ala
 465 470 475

gat aat gca atc atc att tat caa aca cac aag agg gtg tgg cct gct 1491
 Asp Asn Ala Ile Ile Ile Tyr Gln Thr His Lys Arg Val Trp Pro Ala
 480 485 490

tct cag cga gac gta tta tat ctt tct gtc att cga aag ata cca gcc 1539
 Ser Gln Arg Asp Val Leu Tyr Leu Ser Val Ile Arg Lys Ile Pro Ala
 495 500 505 510

ttg act gaa aat gac cct gaa act tgg ata gtt tgt aat ttt tct gtg 1587
 Leu Thr Glu Asn Asp Pro Glu Thr Trp Ile Val Cys Asn Phe Ser Val
 515 520 525

gat cat gac agt gct cct cta aac aac cga tgt gtc cgt gcc aaa ata 1635
 Asp His Asp Ser Ala Pro Leu Asn Asn Arg Cys Val Arg Ala Lys Ile
 530 535 540

aat gtt gct atg att tgt caa acc ttg gta agc cca cca gag gga aac 1683

10

20

30

Asn Val Ala Met Ile Cys Gln Thr Leu Val Ser Pro Pro Glu Gly Asn
 545 550 555

cag gaa att agc agg gac aac att cta tgc aag att aca tat gta gct 1731
 Gln Glu Ile Ser Arg Asp Asn Ile Leu Cys Lys Ile Thr Tyr Val Ala
 560 565 570

aat gtg aac cct gga gga tgg gca cca gcc tca gtg tta agg gca gtg 1779
 Asn Val Asn Pro Gly Gly Trp Ala Pro Ala Ser Val Leu Arg Ala Val
 575 580 585 590

gca aag cga gag tat cct aaa ttt cta aaa cgt ttt act tct tac gtc 1827
 Ala Lys Arg Glu Tyr Pro Lys Phe Leu Lys Arg Phe Thr Ser Tyr Val
 595 600 605

caa gaa aaa act gca gga aag cct att ttg ttc tagtattaac aggtactaga 1880
 Gln Glu Lys Thr Ala Gly Lys Pro Ile Leu Phe
 610 615

agatatgttt tatctttttt taactttatt tgactaatat gactgtcaat actaaaattt 1940

agttggtgaa agtatttact atgttttttc cggaattc 1978

10

<210> 18
 <211> 617
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: FLAG-GPBDNLS

20

<400> 18
 Met Ala Pro Leu Ala Asp Tyr Lys Asp Asp Asp Asp Lys Met Ser Asp
 1 5 10 15

Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu Thr Glu
 20 25 30

Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp Thr Asn
 35 40 45

Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys Asn Asn Ala
 50 55 60

Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys Arg Gly
 65 70 75 80

Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe Asp Glu
 85 90 95

Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu Arg Ala
 100 105 110

Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu Gln His
 115 120 125

Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg Arg His Gly
 130 135 140

Ser Met Val Ser Leu Val Ser Gly Ala Ser Gly Tyr Ser Ala Thr Ser

30

Ala Ile Ile Ile Tyr Gln Thr His Lys Arg Val Trp Pro Ala Ser Gln
 485 490 495

Arg Asp Val Leu Tyr Leu Ser Val Ile Arg Lys Ile Pro Ala Leu Thr
 500 505 510

Glu Asn Asp Pro Glu Thr Trp Ile Val Cys Asn Phe Ser Val Asp His
 515 520 525

Asp Ser Ala Pro Leu Asn Asn Arg Cys Val Arg Ala Lys Ile Asn Val
 530 535 540

Ala Met Ile Cys Gln Thr Leu Val Ser Pro Pro Glu Gly Asn Gln Glu
 545 550 555 560

Ile Ser Arg Asp Asn Ile Leu Cys Lys Ile Thr Tyr Val Ala Asn Val
 565 570 575

Asn Pro Gly Gly Trp Ala Pro Ala Ser Val Leu Arg Ala Val Ala Lys
 580 585 590

Arg Glu Tyr Pro Lys Phe Leu Lys Arg Phe Thr Ser Tyr Val Gln Glu
 595 600 605

Lys Thr Ala Gly Lys Pro Ile Leu Phe
 610 615

10

<210> 19
 <211> 1975
 <212> DNA
 <213> Artificial Sequence

20

<220>
 <223> Description of Artificial Sequence: FLAG-GPBPDSDXY

<220>
 <221> CDS
 <222> (10)..(1857)

<400> 19
 gaattcacc atg gcc cca cta gcc gac tac aag gac gac gat gac aag atg 51
 Met Ala Pro Leu Ala Asp Tyr Lys Asp Asp Asp Lys Met
 1 5 10

tcg gat aat cag agc tgg aac tcg tcg ggc tcg gag gag gat cca gag 99
 Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu
 15 20 25 30

acg gag tct ggg ccg cct gtg gag cgc tgc ggg gtc ctc agt aag tgg 147
 Thr Glu Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp
 35 40 45

aca aac tac att cat ggg tgg cag gat cgt tgg gta gtt ttg aaa aat 195
 Thr Asn Tyr Ile His Gly Trp Glu Asp Arg Trp Val Val Leu Lys Asn
 50 55 60

aat gct ctg agt tac tac aaa tct gaa gat gaa aca gag tat ggc tgc 243
 Asn Ala Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys
 65 70 75

30

aga gga tcc atc tgt ctt agc aag gct gtc atc aca cct cac gat ttt	291
Arg Gly Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe	
80 85 90	
gat gaa tgt cga ttt gat att agt gta aat gat agt gtt tgg tat ctt	339
Asp Glu Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu	
95 100 105 110	
cgt gct cag gat cca gat cat aga cag caa tgg ata gat gcc att gaa	387
Arg Ala Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu	
115 120 125	
cag cac aag act gaa tct gga tat gga tct gaa tcc agc ttg cgt cga	435
Gln His Lys Thr Glu Ser Gly Tyr Gln Ser Glu Ser Ser Leu Arg Arg	
130 135 140	
cat ggc aaa ggc cac agt tta cgt gag aag ttg gct gaa atg gaa aca	483
His Gly Lys Gly His Ser Leu Arg Glu Lys Leu Ala Glu Met Glu Thr	
145 150 155	
ttt aga gac atc tta tgt aga caa gtt gac acg cta cag aag tac ttt	531
Phe Arg Asp Ile Leu Cys Arg Gln Val Asp Thr Leu Gln Lys Tyr Phe	
160 165 170	
gat gcc tgt gct gat gct gtc tct aag gat gaa ctt caa agg gat aaa	579
Asp Ala Cys Ala Asp Ala Val Ser Lys Asp Glu Leu Gln Arg Asp Lys	
175 180 185 190	
gtg gta gaa gat gat gaa gat gac ttt cct aca acg cgt tct gat ggt	627
Val Val Glu Asp Asp Glu Asp Asp Phe Pro Thr Thr Arg Ser Asp Gly	
195 200 205	
gac ttc ttg cat agt acc aac ggc aat aaa gaa aag tta ttt cca cat	675
Asp Phe Leu His Ser Thr Asn Gly Asn Lys Glu Lys Leu Phe Pro His	
210 215 220	
gtg aca cca aaa gga att aat ggt ata gac ttt aaa ggg gaa gcg ata	723
Val Thr Pro Lys Gly Ile Asn Gly Ile Asp Phe Lys Gly Glu Ala Ile	
225 230 235	
act ttt aaa gca act act gct gga atc ctt gca aca ctt tct cat tgt	771
Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu Ala Thr Leu Ser His Cys	
240 245 250	
att gaa cta atg gtt aaa cgt gag gac agc tgg cag aag aga ctg gat	819
Ile Glu Leu Met Val Lys Arg Glu Asp Ser Trp Gln Lys Arg Leu Asp	
255 260 265 270	
aag gaa act gag aag aaa aga aga aca gag gaa gca tat aaa aat gca	867
Lys Glu Thr Glu Lys Lys Arg Arg Thr Glu Glu Ala Tyr Lys Asn Ala	
275 280 285	
atg aca gaa ctt aag aaa aaa tcc cac ttt gga gga cca gat tat gaa	915
Met Thr Glu Leu Lys Lys Lys Ser His Phe Gly Gly Pro Asp Tyr Glu	
290 295 300	
gaa ggc cct aac agt ctg att aat gaa gaa gag ttc ttt gat gct gtt	963
Glu Gly Pro Asn Ser Leu Ile Asn Glu Glu Glu Phe Phe Asp Ala Val	
305 310 315	

10

20

30

gaa gct gct ctt gac aga caa gat aaa ata gaa gaa cag tca cag agt 1011
 Glu Ala Ala Leu Asp Arg Gln Asp Lys Ile Glu Glu Gln Ser Gln Ser
 320 325 330

gaa aag gtg aga tta cat tgg cct aca tcc ttg ccc tct gga gat gcc 1059
 Glu Lys Val Arg Leu His Trp Pro Thr Ser Leu Pro Ser Gly Asp Ala
 335 340 345 350

ttt tct tct gtg ggg aca cat aga ttt gtc caa aag ccc tat agt cgc 1107
 Phe Ser Ser Val Gly Thr His Arg Phe Val Gln Lys Pro Tyr Ser Arg
 355 360 365

tct tcc tcc atg tct tcc att gat cta gtc agt gcc tct gat gat gtt 1155
 Ser Ser Ser Met Ser Ser Ile Asp Leu Val Ser Ala Ser Asp Asp Val
 370 375 380

cac aga ttc agc tcc cag gtt gaa gag atg gtg cag aac cac atg act 1203
 His Arg Phe Ser Ser Gln Val Glu Glu Met Val Gln Asn His Met Thr
 385 390 395

tac tca tta cag gat gta ggc gga gat gcc aat tgg cag ttg gtt gta 1251
 Tyr Ser Leu Gln Asp Val Gly Gly Asp Ala Asn Trp Gln Leu Val Val
 400 405 410

gaa gaa gga gaa atg aag gta tac aga aga gaa gta gaa gaa aat ggg 1299
 Glu Glu Gly Glu Met Lys Val Tyr Arg Arg Glu Val Glu Glu Asn Gly
 415 420 425 430

att gtt ctg gat cct tta aaa gct acc cat gca gtt aaa ggc gtc aca 1347
 Ile Val Leu Asp Pro Leu Lys Ala Thr His Ala Val Lys Gly Val Thr
 435 440 445

gga cat gaa gtc tgc aat tat ttc tgg aat gtt gac gtt cgc aat gac 1395
 Gly His Glu Val Cys Asn Tyr Phe Trp Asn Val Asp Val Arg Asn Asp
 450 455 460

tgg gaa aca act ata gaa aac ttt cat gtg gtg gaa aca tta gct gat 1443
 Trp Glu Thr Thr Ile Glu Asn Phe His Val Val Glu Thr Leu Ala Asp
 465 470 475

aat gca atc atc att tat caa aca cac aag agg gtg tgg cct gct tct 1491
 Asn Ala Ile Ile Ile Tyr Gln Thr His Lys Arg Val Trp Pro Ala Ser
 480 485 490

cag cga gac gta tta tat ctt tct gtc att cga aag ata cca gcc ttg 1539
 Gln Arg Asp Val Leu Tyr Leu Ser Val Ile Arg Lys Ile Pro Ala Leu
 495 500 505 510

act gaa aat gac cct gaa act tgg ata gtt tgt aat ttt tct gtg gat 1587
 Thr Glu Asn Asp Pro Glu Thr Trp Ile Val Cys Asn Phe Ser Val Asp
 515 520 525

cat gac agt gct cct cta aac aac cga tgt gtc cgt gcc aaa ata aat 1635
 His Asp Ser Ala Pro Leu Asn Asn Arg Cys Val Arg Ala Lys Ile Asn
 530 535 540

gtt gct atg att tgt caa acc ttg gta agc cca cca gag gga aac cag 1683
 Val Ala Met Ile Cys Gln Thr Leu Val Ser Pro Pro Glu Gly Asn Gln
 545 550 555

gaa att agc agg gac aac att cta tgc aag att aca tat gta gct aat 1731

10

20

30

Glu Ile Ser Arg Asp Asn Ile Leu Cys Lys Ile Thr Tyr Val Ala Asn
 560 565 570

gtg aac cct gga gga tgg gca cca gcc tca gtg tta agg gca gtg gca 1779
 Val Asn Pro Gly Gly Trp Ala Pro Ala Ser Val Leu Arg Ala Val Ala
 575 580 585 590

aag cga gag tat cct aaa ttt cta aaa cgt ttt act tct tac gtc caa 1827
 Lys Arg Glu Tyr Pro Lys Phe Leu Lys Arg Phe Thr Ser Tyr Val Gln
 595 600 605

gaa aaa act gca gga aag cct att ttg ttc tagtattaac aggtactaga 1877
 Glu Lys Thr Ala Gly Lys Pro Ile Leu Phe
 610 615

agatatgttt tatctttttt taactttatt tgactaatat gactgtcaat actasaattt 1937

agttgttgaa agtatttact atgttttttc cggaattc 1975

10

<210> 20
 <211> 616
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: FLAG-GPBPDSDXY

<400> 20
 Met Ala Pro Leu Ala Asp Tyr Lys Asp Asp Asp Lys Met Ser Asp
 1 5 10 15

Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu Thr Glu
 20 25 30

Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp Thr Asn
 35 40 45

Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys Asn Asn Ala
 50 55 60

Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys Arg Gly
 65 70 75 80

Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe Asp Glu
 85 90 95

Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu Arg Ala
 100 105 110

Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu Gln His
 115 120 125

Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg Arg His Gly
 130 135 140

Lys Gly His Ser Leu Arg Glu Lys Leu Ala Glu Met Glu Thr Phe Arg
 145 150 155 160

Asp Ile Leu Cys Arg Gln Val Asp Thr Leu Gln Lys Tyr Phe Asp Ala
 165 170 175

20

30

Cys Ala Asp Ala Val Ser Lys Asp Glu Leu Gln Arg Asp Lys Val Val
 180 185 190
 Glu Asp Asp Glu Asp Asp Phe Pro Thr Thr Arg Ser Asp Gly Asp Phe
 195 200 205
 Leu His Ser Thr Asn Gly Asn Lys Glu Lys Leu Phe Pro His Val Thr
 210 215 220
 Pro Lys Gly Ile Asn Gly Ile Asp Phe Lys Gly Glu Ala Ile Thr Phe
 225 230 235 240
 Lys Ala Thr Thr Ala Gly Ile Leu Ala Thr Leu Ser His Cys Ile Glu
 245 250 255
 Leu Met Val Lys Arg Glu Asp Ser Trp Gln Lys Arg Leu Asp Lys Glu
 260 265 270
 Thr Glu Lys Lys Arg Arg Thr Glu Glu Ala Tyr Lys Asn Ala Met Thr
 275 280 285
 Glu Leu Lys Lys Lys Ser His Phe Gly Gly Pro Asp Tyr Glu Glu Gly
 290 295 300
 Pro Asn Ser Leu Ile Asn Glu Glu Glu Phe Phe Asp Ala Val Glu Ala
 305 310 315 320
 Ala Leu Asp Arg Gln Asp Lys Ile Glu Glu Gln Ser Gln Ser Glu Lys
 325 330 335
 Val Arg Leu His Trp Pro Thr Ser Leu Pro Ser Gly Asp Ala Phe Ser
 340 345 350
 Ser Val Gly Thr His Arg Phe Val Gln Lys Pro Tyr Ser Arg Ser Ser
 355 360 365
 Ser Met Ser Ser Ile Asp Leu Val Ser Ala Ser Asp Asp Val His Arg
 370 375 380
 Phe Ser Ser Gln Val Glu Glu Met Val Gln Asn His Met Thr Tyr Ser
 385 390 395 400
 Leu Gln Asp Val Gly Gly Asp Ala Asn Trp Gln Leu Val Val Glu Glu
 405 410 415
 Gly Glu Met Lys Val Tyr Arg Arg Glu Val Glu Glu Asn Gly Ile Val
 420 425 430
 Leu Asp Pro Leu Lys Ala Thr His Ala Val Lys Gly Val Thr Gly His
 435 440 445
 Glu Val Cys Asn Tyr Phe Trp Asn Val Asp Val Arg Asn Asp Trp Glu
 450 455 460
 Thr Thr Ile Glu Asn Phe His Val Val Glu Thr Leu Ala Asp Asn Ala
 465 470 475 480
 Ile Ile Ile Tyr Gln Thr His Lys Arg Val Trp Pro Ala Ser Gln Arg
 485 490 495

10

20

30

Asp Val Leu Tyr Leu Ser Val Ile Arg Lys Ile Pro Ala Leu Thr Glu
500 505 510

Asn Asp Pro Glu Thr Trp Ile Val Cys Asn Phe Ser Val Asp His Asp
515 520 525

Ser Ala Pro Leu Asn Asn Arg Cys Val Arg Ala Lys Ile Asn Val Ala
530 535 540

Met Ile Cys Gln Thr Leu Val Ser Pro Pro Glu Gly Asn Gln Glu Ile
545 550 555 560

Ser Arg Asp Asn Ile Leu Cys Lys Ile Thr Tyr Val Ala Asn Val Asn
565 570 575

Pro Gly Gly Trp Ala Pro Ala Ser Val Leu Arg Ala Val Ala Lys Arg
580 585 590

Glu Tyr Pro Lys Phe Leu Lys Arg Phe Thr Ser Tyr Val Gln Glu Lys
595 600 605

Thr Ala Gly Lys Pro Ile Leu Phe
610 615

10

<210> 21
<211> 1915
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
FLAG-GPBPSXY/NLS

20

<220>
<221> CDS
<222> (10)..(1797)

<400> 21
gaattcacc atg gcc cca cta gcc gac tac aag gac gac gat gac aag atg 51
Met Ala Pro Leu Ala Asp Tyr Lys Asp Asp Asp Asp Lys Met
1 5 10

tcg gat aat cag agc tgg aac tcg tcg ggc tcg gag gag gat cca gag 99
Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu
15 20 25 30

acg gag tct ggg ccg cct gtg gag cgc tgc ggg gtc ctc agt aag tgg 147
Thr Glu Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp
35 40 45

30

aca aac tac att cat ggg tgg cag gat cgt tgg gta gtt ttg aaa aat 195
Thr Asn Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys Asn
50 55 60

aat gct ctg agt tac tac aaa tct gaa gat gaa aca gag tat ggc tgc 243
Asn Ala Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys
65 70 75

aga gga tcc atc tgt ctt agc aag gct gtc atc aca cct cac gat ttt 291
Arg Gly Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe

80	85	90		
gat gaa tgt cga ttt gat att agt gta aat gat agt gtt tgg tat ctt			339	
Asp Glu Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu				
95	100	105		110
cgt gct cag gat cca gat cat aga cag caa tgg ata gat gcc att gaa			387	
Arg Ala Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu				
115	120	125		
cag cac aag act gaa tct gga tat gga tct gaa tcc agc ttg cgt cga			435	
Gln His Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg Arg				
130	135	140		
cat ggc aaa ggc cac agt tta cgt gag aag ttg gct gaa atg gaa aca			483	
His Gly Lys Gly His Ser Leu Arg Glu Lys Leu Ala Glu Met Glu Thr				
145	150	155		
ttt aga gac atc tta tgt aga caa gtt gac acg cta cag aag tac ttt			531	
Phe Arg Asp Ile Leu Cys Arg Gln Val Asp Thr Leu Gln Lys Tyr Phe				
160	165	170		
gat gcc tgt gct gat gct gtc tct aag gat gaa ctt caa agg gat aaa			579	
Asp Ala Cys Ala Asp Ala Val Ser Lys Asp Glu Leu Gln Arg Asp Lys				
175	180	185		190
gtg gta gaa gat gat gaa gat gac ttt cct aca acg cgt tct gat ggt			627	
Val Val Glu Asp Asp Glu Asp Asp Phe Pro Thr Thr Arg Ser Asp Gly				
195	200	205		
gac ttc ttg cat agt acc aac ggc aat aaa gaa aag tta ttt cca cat			675	
Asp Phe Leu His Ser Thr Asn Gly Asn Lys Glu Lys Leu Phe Pro His				
210	215	220		
gtg aca cca aaa gga att aat ggt ata gac ttt aaa ggg gaa gcg ata			723	
Val Thr Pro Lys Gly Ile Asn Gly Ile Asp Phe Lys Gly Glu Ala Ile				
225	230	235		
act ttt aaa gca act act gct gga atc ctt gca aca ctt tct cat tgt			771	
Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu Ala Thr Leu Ser His Cys				
240	245	250		
att gaa cta atg gtt aaa cgt gag gac agc tgg cag aag aga ctg gat			819	
Ile Glu Leu Met Val Lys Arg Glu Asp Ser Trp Gln Lys Arg Leu Asp				
255	260	265		270
aag gaa act gag cac ttt gga gga cca gat tat gaa gaa ggc cct aac			867	
Lys Glu Thr Glu His Phe Gly Gly Pro Asp Tyr Glu Glu Gly Pro Asn				
275	280	285		
agt ctg att aat gaa gaa gag ttc ttt gat gct gtt gaa gct gct ctt			915	
Ser Leu Ile Asn Glu Glu Phe Phe Asp Ala Val Glu Ala Ala Leu				
290	295	300		
gac aga caa gat aaa ata gaa gaa cag tca cag agt gaa aag gtg aga			963	
Asp Arg Gln Asp Lys Ile Glu Glu Gln Ser Gln Ser Glu Lys Val Arg				
305	310	315		
tta cat tgg cct aca tcc ttg ccc tct gga gat gcc ttt tct tct gtg			1011	
Leu His Trp Pro Thr Ser Leu Pro Ser Gly Asp Ala Phe Ser Ser Val				
320	325	330		

10

20

30

ggg aca cat aga ttt gtc caa aag ccc tat agt cgc tct tcc tcc atg 1059
 Gly Thr His Arg Phe Val Gln Lys Pro Tyr Ser Arg Ser Ser Ser Met
 335 340 345 350

tct tcc att gat cta gtc agt gcc tct gat gat gtt cac aga ttc agc 1107
 Ser Ser Ile Asp Leu Val Ser Ala Ser Asp Asp Val His Arg Phe Ser
 355 360 365

tcc cag gtt gaa gag atg gtg cag aac cac atg act tac tca tta cag 1155
 Ser Gln Val Gln Glu Met Val Gln Asn His Met Thr Tyr Ser Leu Gln
 370 375 380

gat gta ggc gga gat gcc aat tgg cag ttg gtt gta gaa gaa gga gaa 1203
 Asp Val Gly Asp Ala Asn Trp Gln Leu Val Val Glu Glu Gly Glu
 385 390 395

atg aag gta tac aga aga gaa gta gaa gaa aat ggg att gtt ctg gat 1251
 Met Lys Val Tyr Arg Arg Glu Val Glu Glu Asn Gly Ile Val Leu Asp
 400 405 410

cct tta aaa gct acc cat gca gtt aaa ggc gtc aca gga cat gaa gtc 1299
 Pro Leu Lys Ala Thr His Ala Val Lys Gly Val Thr Gly His Glu Val
 415 420 425 430

tgc aat tat ttc tgg aat gtt gac gtt cgc aat gac tgg gaa aca act 1347
 Cys Asn Tyr Phe Trp Asn Val Asp Val Arg Asn Asp Trp Glu Thr Thr
 435 440 445

ata gaa aac ttt cat gtg gtg gaa aca tta gct gat aat gca atc atc 1395
 Ile Glu Asn Phe His Val Val Glu Thr Leu Ala Asp Asn Ala Ile Ile
 450 455 460

att tat caa aca cac aag agg gtg tgg cct gct tct cag cga gac gta 1443
 Ile Tyr Gln Thr His Lys Arg Val Trp Pro Ala Ser Gln Arg Asp Val
 465 470 475

tta tat ctt tct gtc att cga aag ata cca gcc ttg act gaa aat gac 1491
 Leu Tyr Leu Ser Val Ile Arg Lys Ile Pro Ala Leu Thr Glu Asn Asp
 480 485 490

cct gaa act tgg ata gtt tgt aat ttt tct gtg gat cat gac agt gct 1539
 Pro Glu Thr Trp Ile Val Cys Asn Phe Ser Val Asp His Asp Ser Ala
 495 500 505 510

cct cta aac aac cga tgt gtc cgt gcc aaa ata aat gtt gct atg att 1587
 Pro Leu Asn Asn Arg Cys Val Arg Ala Lys Ile Asn Val Ala Met Ile
 515 520 525

tgt caa acc ttg gta agc cca cca gag gga aac cag gaa att agc agg 1635
 Cys Gln Thr Leu Val Ser Pro Pro Glu Gly Asn Gln Glu Ile Ser Arg
 530 535 540

gac aac att cta tgc aag att aca tat gta gct aat gtg aac cct gga 1683
 Asp Asn Ile Leu Cys Lys Ile Thr Tyr Val Ala Asn Val Asn Pro Gly
 545 550 555

gga tgg gca cca gcc tca gtg tta agg gca gtg gca aag cga gag tat 1731
 Gly Trp Ala Pro Ala Ser Val Leu Arg Ala Val Ala Lys Arg Glu Tyr
 560 565 570

10

20

30

cct aaa ttt cta aaa cgt ttt act tct tac gtc caa gaa aaa act gca 1779
 Pro Lys Phe Leu Lys Arg Phe Thr Ser Tyr Val Gln Glu Lys Thr Ala
 575 580 585 590

gga aag cct att ttg ttc tagtattaac aggtactaga agatatgttt 1827
 Gly Lys Pro Ile Leu Phe
 595

tatctttttt taactttatt tgactaatat gactgtcaat actaaaattt agttgttgaa 1887
 agtatttact atgttttttc cggaattc 1915

<210> 22
 <211> 596
 <212> PRT
 <213> Artificial Sequence

10

<220>
 <223> Description of Artificial Sequence:
 FLAG-GPBPDSXY/NLS

<400> 22
 Met Ala Pro Leu Ala Asp Tyr Lys Asp Asp Asp Asp Lys Met Ser Asp
 1 5 10 15
 Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu Thr Glu
 20 25 30
 Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp Thr Asn
 35 40 45
 Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys Asn Asn Ala
 50 55 60
 Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys Arg Gly
 65 70 75 80
 Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe Asp Glu
 85 90 95
 Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu Arg Ala
 100 105 110
 Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu Gln His
 115 120 125
 Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg Arg His Gly
 130 135 140
 Lys Gly His Ser Leu Arg Glu Lys Leu Ala Glu Met Glu Thr Phe Arg
 145 150 155 160
 Asp Ile Leu Cys Arg Gln Val Asp Thr Leu Gln Lys Tyr Phe Asp Ala
 165 170 175
 Cys Ala Asp Ala Val Ser Lys Asp Glu Leu Gln Arg Asp Lys Val Val
 180 185 190
 Glu Asp Asp Glu Asp Asp Phe Pro Thr Thr Arg Ser Asp Gly Asp Phe
 195 200 205

20

30

Leu His Ser Thr Asn Gly Asn Lys Glu Lys Leu Phe Pro His Val Thr
 210 215 220
 Pro Lys Gly Ile Asn Gly Ile Asp Phe Lys Gly Glu Ala Ile Thr Phe
 225 230 235 240
 Lys Ala Thr Thr Ala Gly Ile Leu Ala Thr Leu Ser His Cys Ile Glu
 245 250 255
 Leu Met Val Lys Arg Glu Asp Ser Trp Gln Lys Arg Leu Asp Lys Glu
 260 265 270
 Thr Glu His Phe Gly Gly Pro Asp Tyr Glu Glu Gly Pro Asn Ser Leu
 275 280 285
 Ile Asn Glu Glu Glu Phe Phe Asp Ala Val Glu Ala Ala Leu Asp Arg
 290 295 300
 Gln Asp Lys Ile Glu Glu Gln Ser Gln Ser Glu Lys Val Arg Leu His
 305 310 315 320
 Trp Pro Thr Ser Leu Pro Ser Gly Asp Ala Phe Ser Ser Val Gly Thr
 325 330 335
 His Arg Phe Val Gln Lys Pro Tyr Ser Arg Ser Ser Ser Met Ser Ser
 340 345 350
 Ile Asp Leu Val Ser Ala Ser Asp Asp Val His Arg Phe Ser Ser Gln
 355 360 365
 Val Glu Glu Met Val Gln Asn His Met Thr Tyr Ser Leu Gln Asp Val
 370 375 380
 Gly Gly Asp Ala Asn Trp Gln Leu Val Val Glu Glu Gly Glu Met Lys
 385 390 395 400
 Val Tyr Arg Arg Glu Val Glu Glu Asn Gly Ile Val Leu Asp Pro Leu
 405 410 415
 Lys Ala Thr His Ala Val Lys Gly Val Thr Gly His Glu Val Cys Asn
 420 425 430
 Tyr Phe Trp Asn Val Asp Val Arg Asn Asp Trp Glu Thr Thr Ile Glu
 435 440 445
 Asn Phe His Val Val Glu Thr Leu Ala Asp Asn Ala Ile Ile Ile Tyr
 450 455 460
 Gln Thr His Lys Arg Val Trp Pro Ala Ser Gln Arg Asp Val Leu Tyr
 465 470 475 480
 Leu Ser Val Ile Arg Lys Ile Pro Ala Leu Thr Glu Asn Asp Pro Glu
 485 490 495
 Thr Trp Ile Val Cys Asn Phe Ser Val Asp His Asp Ser Ala Pro Leu
 500 505 510
 Asn Asn Arg Cys Val Arg Ala Lys Ile Asn Val Ala Met Ile Cys Gln
 515 520 525

10

20

30

Thr Leu Val Ser Pro Pro Glu Gly Asn Gln Glu Ile Ser Arg Asp Asn
 530 . 535 540

Ile Leu Cys Lys Ile Thr Tyr Val Ala Asn Val Asn Pro Gly Gly Trp
 545 550 555 560

Ala Pro Ala Ser Val Leu Arg Ala Val Ala Lys Arg Glu Tyr Pro Lys
 565 570 575

Phe Leu Lys Arg Phe Thr Ser Tyr Val Gln Glu Lys Thr Ala Gly Lys
 580 585 590

Pro Ile Leu Phe
 595

<210> 23
 <211> 2038
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPBP-D169A

<220>
 <221> CDS
 <222> (10)..(1920)

<400> 23
 gaattcacc atg gcc cca cta gcc gac tac aag gac gac gat gac aag atg 51
 Met Ala Pro Leu Ala Asp Tyr Lys Asp Asp Asp Lys Met
 1 5 10

tcg gat aat cag agc tgg aac tcg tcg ggc tcg gag gag gat cca gag 99
 Ser Asp Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu
 15 20 25 30

acg gag tct ggg ccg cct gtg gag cgc tgc ggg gtc ctc agt aag tgg 147
 Thr Glu Ser Gly Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp
 35 40 45

aca aac tac att cat ggg tgg cag gat cgt tgg gta gtt ttg aaa aat 195
 Thr Asn Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys Asn
 50 55 60

aat gct ctg agt tac tac aaa tct gaa gat gaa aca gag tat ggc tgc 243
 Asn Ala Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys
 65 70 75

aga gga tcc atc tgt ctt agc aag gct gtc atc aca cct cac gat ttt 291
 Arg Gly Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe
 80 85 90

gat gaa tgt cga ttt gat att agt gta aat gat agt gtt tgg tat ctt 339
 Asp Glu Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu
 95 100 105 110

cgt gct cag gat cca gat cat aga cag caa tgg ata gat gcc att gaa 387
 Arg Ala Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu
 115 120 125

10

20

30

cag cac aag act gaa tct gga tat gga tct gaa tcc agc ttg cgt cga 435
 Gln His Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg Arg
 130 135 140

cat ggc tca atg gtg tcc ctg gtg tct gga gca agt ggc tac tct gca 483
 His Gly Ser Met Val Ser Leu Val Ser Gly Ala Ser Gly Tyr Ser Ala
 145 150 155

aca tcc acc tct tca ttc aag aaa ggc cac agt tta cgt gag aag ttg 531
 Thr Ser Thr Ser Ser Phe Lys Lys Gly His Ser Leu Arg Glu Lys Leu
 160 165 170

gct gaa atg gaa aca ttt aga gcc atc tta tgt aga caa gtt gac acg 579
 Ala Glu Met Glu Thr Phe Arg Ala Ile Leu Cys Arg Gln Val Asp Thr
 175 180 185 190

cta cag aag tac ttt gat gcc tgt gct gat gct gtc tct aag gat gaa 627
 Leu Gln Lys Tyr Phe Asp Ala Cys Ala Asp Ala Val Ser Lys Asp Glu
 195 200 205

ctt caa agg gat aaa gtg gta gaa gat gat gaa gat gac ttt cct aca 675
 Leu Gln Arg Asp Lys Val Val Glu Asp Asp Glu Asp Asp Phe Pro Thr
 210 215 220

acg cgt tct gat ggt gac ttc ttg cat agt acc aac ggc aat aaa gaa 723
 Thr Arg Ser Asp Gly Asp Phe Leu His Ser Thr Asn Gly Asn Lys Glu
 225 230 235

aag tta ttt cca cat gtg aca cca aaa gga att aat ggt ata gac ttt 771
 Lys Leu Phe Pro His Val Thr Pro Lys Gly Ile Asn Gly Ile Asp Phe
 240 245 250

aaa ggg gaa gcg ata act ttt aaa gca act act gct gga atc ctt gca 819
 Lys Gly Glu Ala Ile Thr Phe Lys Ala Thr Thr Ala Gly Ile Leu Ala
 255 260 265 270

aca ctt tct cat tgt att gaa cta atg gtt aaa cgt gag gac agc tgg 867
 Thr Leu Ser His Cys Ile Glu Leu Met Val Lys Arg Glu Asp Ser Trp
 275 280 285

cag aag aga ctg gat aag gaa act gag aag aaa aga aga aca gag gaa 915
 Gln Lys Arg Leu Asp Lys Glu Thr Glu Lys Lys Arg Arg Thr Glu Glu
 290 295 300

gca tat aaa aat gca atg aca gaa ctt aag aaa aaa tcc cac ttt gga 963
 Ala Tyr Lys Asn Ala Met Thr Glu Leu Lys Lys Lys Ser His Phe Gly
 305 310 315

gga cca gat tat gaa gaa ggc cct aac agt ctg att aat gaa gaa gag 1011
 Gly Pro Asp Tyr Glu Glu Gly Pro Asn Ser Leu Ile Asn Glu Glu Glu
 320 325 330

ttc ttt gat gct gtt gaa gct gct ctt gac aga caa gat aaa ata gaa 1059
 Phe Phe Asp Ala Val Glu Ala Ala Leu Asp Arg Gln Asp Lys ile Glu
 335 340 345 350

gaa cag tca cag agt gaa aag gtg aga tta cat tgg cct aca tcc ttg 1107
 Glu Gln Ser Gln Ser Glu Lys Val Arg Leu His Trp Pro Thr Ser Leu
 355 360 365

ccc tct gga gat gcc ttt tct tct gtg ggg aca cat aga ttt gtc caa 1155

10

20

30

Pro	Ser	Gly	Asp	Ala	Phe	Ser	Ser	Val	Gly	Thr	His	Arg	Phe	Val	Gln		
			370					375					380				
aag	ccc	tat	agt	cgc	tct	tcc	tcc	atg	tct	tcc	att	gat	cta	gtc	agt	1203	
Lys	Pro	Tyr	Ser	Arg	Ser	Ser	Ser	Met	Ser	Ser	Ile	Asp	Leu	Val	Ser		
		385					390					395					
gcc	tct	gat	gat	gtt	cac	aga	ttc	agc	tcc	cag	gtt	gaa	gag	atg	gtg	1251	
Ala	Ser	Asp	Asp	Val	His	Arg	Phe	Ser	Ser	Gln	Val	Glu	Glu	Met	Val		
	400				405						410						
cag	aac	cac	atg	act	tac	tca	tta	cag	gat	gta	ggc	gga	gat	gcc	aat	1299	
Gln	Asn	His	Met	Thr	Tyr	Ser	Leu	Gln	Asp	Val	Gly	Gly	Asp	Ala	Asn		
415					420					425					430		
tgg	cag	ttg	ggt	gta	gaa	gaa	gga	gaa	atg	aag	gta	tac	aga	aga	gaa	1347	
Trp	Gln	Leu	Val	Val	Glu	Glu	Gly	Glu	Met	Lys	Val	Tyr	Arg	Arg	Glu		
			435					440						445			
gta	gaa	gaa	aat	ggg	att	gtt	ctg	gat	cct	tta	aaa	gct	acc	cat	gca	1395	
Val	Glu	Glu	Asn	Gly	Ile	Val	Leu	Asp	Pro	Leu	Lys	Ala	Thr	His	Ala		
			450					455						460			
gtt	aaa	ggc	gtc	aca	gga	cat	gaa	gtc	tgc	aat	tat	ttc	tgg	aat	gtt	1443	
Val	Lys	Gly	Val	Thr	Gly	His	Glu	Val	Cys	Asn	Tyr	Phe	Trp	Asn	Val		
		465				470						475					
gac	gtt	cgc	aat	gac	tgg	gaa	aca	act	ata	gaa	aac	ttt	cat	gtg	gtg	1491	
Asp	Val	Arg	Asn	Asp	Trp	Glu	Thr	Thr	Ile	Glu	Asn	Phe	His	Val	Val		
	480				485						490						
gaa	aca	tta	gct	gat	aat	gca	atc	atc	att	tat	caa	aca	cac	aag	agg	1539	
Glu	Thr	Leu	Ala	Asp	Asn	Ala	Ile	Ile	Ile	Tyr	Gln	Thr	His	Lys	Arg		
495					500					505					510		
gtg	tgg	cct	gct	tct	cag	cga	gac	gta	tta	tat	ctt	tct	gtc	att	cga	1587	
Val	Trp	Pro	Ala	Ser	Gln	Arg	Asp	Val	Leu	Tyr	Leu	Ser	Val	Ile	Arg		
			515						520					525			
aag	ata	cca	gcc	ttg	act	gaa	aat	gac	cct	gaa	act	tgg	ata	gtt	tgt	1635	
Lys	Ile	Pro	Ala	Leu	Thr	Glu	Asn	Asp	Pro	Glu	Thr	Trp	Ile	Val	Cys		
			530					535					540				
aat	ttt	tct	gtg	gat	cat	gac	agt	gct	cct	cta	aac	aac	cga	tgt	gtc	1683	
Asn	Phe	Ser	Val	Asp	His	Asp	Ser	Ala	Pro	Leu	Asn	Asn	Arg	Cys	Val		
			545				550						555				
cgt	gcc	aaa	ata	aat	gtt	gct	atg	att	tgt	caa	acc	ttg	gta	agc	cca	1731	
Arg	Ala	Lys	Ile	Asn	Val	Ala	Met	Ile	Cys	Gln	Thr	Leu	Val	Ser	Pro		
	560				565							570					
cca	gag	gga	aac	cag	gaa	att	agc	agg	gac	aac	att	cta	tgc	aag	att	1779	
Pro	Glu	Gly	Asn	Gln	Glu	Ile	Ser	Arg	Asp	Asn	Ile	Leu	Cys	Lys	Ile		
575					580					585				590			
aca	tat	gta	gct	aat	gtg	aac	cct	gga	gga	tgg	gca	cca	gcc	tca	gtg	1827	
Thr	Tyr	Val	Ala	Asn	Val	Asn	Pro	Gly	Gly	Trp	Ala	Pro	Ala	Ser	Val		
			595						600					605			
tta	agg	gca	gtg	gca	aag	cga	gag	tat	cct	aaa	ttt	cta	aaa	cgt	ttt	1875	
Leu	Arg	Ala	Val	Ala	Lys	Arg	Glu	Tyr	Pro	Lys	Phe	Leu	Lys	Arg	Phe		

10

20

30

610 615 620
 act tct tac gtc caa gaa aaa act gca gga aag cct att ttg ttc 1920
 Thr Ser Tyr Val Gln Glu Lys Thr Ala Gly Lys Pro Ile Leu Phe
 625 630 635

tagtattaac aggtactaga agatatgttt tatctttttt taactttatt tgactaatat 1980

gactgtcaat actaaaattt agttgttgaa agtatttact atgttttttc cggaattc 2038

<210> 24

<211> 637

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: GPBP-D169A

<400> 24

Met Ala Pro Leu Ala Asp Tyr Lys Asp Asp Asp Asp Lys Met Ser Asp
 1 5 10 15

Asn Gln Ser Trp Asn Ser Ser Gly Ser Glu Glu Asp Pro Glu Thr Glu
 20 25 30

Ser Gly Pro Pro Val Glu Arg Cys Gly Val Leu Ser Lys Trp Thr Asn
 35 40 45

Tyr Ile His Gly Trp Gln Asp Arg Trp Val Val Leu Lys Asn Asn Ala
 50 55 60

Leu Ser Tyr Tyr Lys Ser Glu Asp Glu Thr Glu Tyr Gly Cys Arg Gly
 65 70 75 80

Ser Ile Cys Leu Ser Lys Ala Val Ile Thr Pro His Asp Phe Asp Glu
 85 90 95

Cys Arg Phe Asp Ile Ser Val Asn Asp Ser Val Trp Tyr Leu Arg Ala
 100 105 110

Gln Asp Pro Asp His Arg Gln Gln Trp Ile Asp Ala Ile Glu Gln His
 115 120 125

Lys Thr Glu Ser Gly Tyr Gly Ser Glu Ser Ser Leu Arg Arg His Gly
 130 135 140

Ser Met Val Ser Leu Val Ser Gly Ala Ser Gly Tyr Ser Ala Thr Ser
 145 150 155 160

Thr Ser Ser Phe Lys Lys Gly His Ser Leu Arg Glu Lys Leu Ala Glu
 165 170 175

Met Glu Thr Phe Arg Ala Ile Leu Cys Arg Gln Val Asp Thr Leu Gln
 180 185 190

Lys Tyr Phe Asp Ala Cys Ala Asp Ala Val Ser Lys Asp Glu Leu Gln
 195 200 205

Arg Asp Lys Val Val Glu Asp Asp Glu Asp Asp Phe Pro Thr Thr Arg
 210 215 220

10

20

30

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

Ser Val Asp His Asp Ser Ala Pro Leu Asn Asn Arg Cys Val Arg Ala
 545 550 555 560

Lys Ile Asn Val Ala Met Ile Cys Gln Thr Leu Val Ser Pro Pro Glu
 565 570 575

Gly Asn Gln Glu Ile Ser Arg Asp Asn Ile Leu Cys Lys Ile Thr Tyr
 580 585 590

Val Ala Asn Val Asn Pro Gly Gly Trp Ala Pro Ala Ser Val Leu Arg
 595 600 605

Ala Val Ala Lys Arg Glu Tyr Pro Lys Phe Leu Lys Arg Phe Thr Ser
 610 615 620

Tyr Val Gln Glu Lys Thr Ala Gly Lys Pro Ile Leu Phe
 625 630 635

10

<210> 25
 <211> 12482
 <212> DNA
 <213> Homo sapiens

<400> 25
 tcgatcattt ccctcttcat attcagtgtt tattgcacag atctctcaac aacacagcca 60
 ttaaatagat attctccaag tgacacttac atcacacatg tttgagttta cgttacttgc 120
 aaacataggg aaagaaagat acatgggata aactggtgca tgagaaatga gatcttagca 180
 gttggtrgaa ataaatgaga acaactgagg caaactaaag aggaagaagg gcaagtggca 240
 gcttaacagg agtaagatga tgagatgaag ggcagaatac ctctcatggag aggaggcaaa 300
 gagatataca tgatatgttc tttagaacat aactgaagca aacaatgata ttatttctaa 360
 ttatatataa acctgtgagt cagccttcca ggggcccctt gctaaggtag aatcattgga 420
 atgatttggc cagggtttgg ataggagaga attggcagca gcgttaagat tgacctatga 480
 taaataatgc tatgcaggtg gcagggagtc tgactaggag caaaatcaac gaacttatcc 540
 ctgcctaac atagtatctg tggagtcaga aagaagaggt taaattggga tatctgaggc 600
 aagtatcagg atttgccatg tctgcccagc agtttcataa ttctaattgg tataagcact 660
 aaggcgttca ctaagtgaat gttggtagtt ccagggtata ttatccattc ttgagttaca 720
 aaatacactt taaaaccttc ccattttaat attatatggt tttttagtca cagagtgaaa 780
 aggtgagatt acattggcct acatccttgc cctctggaga tgccttttct tctgtgggga 840
 cacatagatt tgtccaaaag gtaagctaat gtcagagttt actaaaagta cacottgtat 900
 tgttcttcat tgttgggtgga aatatctttt atttgagacg gagtctcact ctgtcaccag 960
 agtggagtgc agtggcgaga tctcggtcca ctacagcttc cacctcccgg gttcaagaga 1020
 ttctcgtgcc tcagcctccc tggtagctgg gattacaggc atgtaccacc acaccagct 1080

20

30

aatttttcta tttttaatgg agacagtctc accatggcca ggatggctct gatctcctga 1140
ccttgatgac caccacacct agcctcccag agtgcctggga ttacagcgt gagccaccat 1200
gccagccgg aatatccttg tagtatataa gttttctccc cttttcatta atttaagtaa 1260
tgagactgtt tttggtttta tatattgtat tccatataca tccccaataa cagttagaaa 1320
ttttgttctg aaaataaagt tctttcattt ttatttaagg ggaaagtgg ggggtggcaa 1380
ataaggagtg gctagtccaa aatagttaac cagaagtata tccagtata ctaaactctt 1440
ctctctcttg gggttaaatg gtattacttt gtattattgg aagcactaca ttcttttttg 1500
gaatgatttt ggaacataat acataatagg tgcataaagt cagcagtgc tgcctgctt 1560
gtttcatata gtgctttgtt tctctctccc tttatctgtt gtttggagt tggtagtaa 1620
tgctctgttg tgcctttgtt ctgattactt ggttttttct ttgtctgtct ctggtagccc 1680
tatagctgct ctctctccat gtctctcatt gatctagtca gtgcctctga tgatgttcc 1740
agattcagct cccaggtact gtatgaatgt atagagtggc cttgagctct tctgtgctat 1800
atctcagcct gctttcccag ttcctagaaa tcttttgggt aggcactga ttttagtttt 1860
gaattttaaa tagtaacatt aagcattaaa aaggctctcc ttgtctacta aatagttcct 1920
ctgtcaggtt tgcatagtgc ctttactatt cacagcttg aattttgtca tataggaggt 1980
actccagaaa gatcttcaaa ctgaattgaa acaaatagaa gatactgggt tttgtatct 2040
atgtaatact tgtttcttca gtcaggattt agcagttttg atggacgtgg tccatctgat 2100
atgttatagc agaaaagcag atttttacaa gtctcacttt aaagcctaaa gtaccccaaa 2160
ttaatatcca acaaggaat cactttttaa taatatgttt catttccatt ataatactaa 2220
gctctattga gcagattgtg ttttctctat gcaaattacc tttggatatt ataaatgaat 2280
atctctgttc atatgctaaa tctatggaaa tttgttttaa tttttagcat tggtaaggtt 2340
ttaggaattt aagacaggaa gctggatgct tgcggctctt aaagtctgta cctcaaaaat 2400
aaaatcagat taccattgga agaagttttt tttagtgtca gcgttagttc tttttttaat 2460
ttctttaatc ttcacatctt tgcattcaa ctttttatct ttctggatg tgcattttat 2520
tggactagat tatattatgt taatcttata ttaaagacct gagcactctg gtcagaatga 2580
ctcagtttaa acctgggta ggtgtatgat cccagtaagt tttctaactt ttttgtgctt 2640
catttttatg atttagctag aacctgacac ataataagtg ctcaataaat gttacctgtt 2700
attgctatta taacataatt tctttgagct aataaaagt atctacatca ttattttttc 2760
ctctgtgaga gtattgctat aaaagtcttt aaaagtcata gtttaagag atttctatta 2820
tttttatgtt tataaataaa gtttacatta gtttttaacc tgcaatagag aagaatatta 2880
agactttaat tttctgact tgtacagcgt ttttctctt gaatactctt aagaaaaaga 2940

10

20

30

tttagcaatt ctggatcaga aatcatccat aaccaaatat accacagtat attttacctt 3000
 ttgcttgctc atttatgcat ttttttttaa ttttacttat ttattttcga gacagggctt 3060
 tgctctgttg cccaggetgg agtgcagtg gacgatctgg gctcactgca acctccatct 3120
 cccaggttca agcaattctc ctgcctcagc ctcccaagta gctgggatta caggcacgca 3180
 ccactatgcc cagctaattt ttgtattctt agtaaagacg ggttttcacc atggtggcca 3240
 ggctggctca gcaactctga cctcgtgacg tgcccacctc ggcctcccaa agtgctggga 3300
 ttacaggtgt gagccaccat gcccggccct gcgtatgttt ttaaaaagag actcatattc 3360
 ataatgaatc tgtgacaaaa ctacataata ctgggagact ttggtttatt gtgctaagct 3420
 ccacattgca ttaaaatcat atcacagact aatcaaaaat gcaggaatac ataggctata 3480
 aatgaaagaa aatataatga cagcaagaa agaatgtaag ccagtaataa agaatgccta 3540
 agaattaggg gtccagaacc caaaccaggg cctcactgtt agtgctgtag aacagctgaa 3600
 ttgcttttaa gtccaggtaa ctatatcact gagaagcagg tgcctatatt tttacaaaat 3660
 tttgctgaca gcttactctt tcgtaatat aatacccttt tgtaaaactc atgtatgtaa 3720
 cttgagagaa atcttgctgg atttttttct ctaatatatg gtgctcatga ttgatcagat 3780
 cctgttttag cctttgatta tgtactgttt tatatgccag aagaggtaaa aatgaagaaa 3840
 ataacattaa ggcttcaag tatttgttgt ccttgctaaa gcattagtgt tcattagcag 3900
 acgtggactc tagcaattca ctgttgaat taaattgtgt gccttatggt cagcagttcc 3960
 tttataatag atgactaatt cccaattgat aagatttttt gtttcagagg atgttacact 4020
 gccttatcag ccattatcaa aggatctagc aagttgattc tgtatagtca cacttgagaa 4080
 tatagcattg gatgtagatc tggagttaat attagttgag aaacattgtg ttatctggaa 4140
 aactcttcca gttcaacaca gtgtaaaatt atagtagtga ctatacagta gtgttacatt 4200
 ttacagttct cacaccctat agagactttt gtattaacaa aataagaggc tcaaaggtta 4260
 ttcattaaca ttagaaacac ttatgttata ttacattgca tgggtctttt ctggtttttg 4320
 tttttttttt ttttttgaga cggagttctg cttttgttgc ccaggctgga gtgcaatggg 4380
 acgatcttgg ctcaactgac cctctgcccc ctggattcaa gcgattctct tgccctagcc 4440
 acctgagtag ctgggattac aggcacctgc caccacacc agctaatttt ttttcatttt 4500
 tagtagagat ggggtttcac catgttggcc aggctggtct cgaactcctg acctcaggtg 4560
 atctgcccgc ctoggcctcc caaagtgtct ggattacagg catgagccgc cacacctggc 4620
 ctacatcgtt ctttaatacac aatatacat cagttactcc acagcgettg atatgggagg 4680
 taaccaaatt ctttgtttta taatatcttc ataattaatt aaaaaactaa gtcgacattt 4740

10

20

30

ttaatcaact ttaataatTT gccaaaatAT tatataagca taatataatC aattcttact 4800
 tactccaaca aattttaaaa gtccagatac agataccata tctagtttct tgatcattta 4860
 tatcagetcc catacagaag ccttctaaat ctctggtaat ttcactttgc tgtttatata 4920
 agtgttggtc catgactacc ttgttcttct tgaaatgatg ttttatagcc ttgaattggc 4980
 tgaaataatc aagtgtacaa ttgagagatg ccctgaaaac agcttaaaat aaaatatgta 5040
 catctactag gaaattagta ccaacacatg aatctgtctg atgggcagat attaggaatg 5100
 aagtcaactc agatctgaga aattaaagtT gtaaaggact gcaagttctg tgtttttggt 5160
 gttgttgggt ttgttgggtg tgtttgTTTT ttcatttttg ttttttgggt ttttttgaga 5220
 cagagtctca ttctgtcacc caggctgtag tgcagtggca cgatctcaac tcaactgcaac 5280
 ctccgtctcc caggttcaag cgattctcct gtctcagctg ggattacagg cacacgctat 5340
 cacaccagc taatttttgt attttttagta gagacaggtt ttcaccatgt tagccaggct 5400
 ggtctcgaac tcctgacctc aagtgatctg cccgtctcgg cctcccaaag tgctgggatt 5460
 acaggcctga gacaccatgc ccagcatttt tttttttttt tttttttttt gtaaagagac 5520
 aaggtttcac ttgtccaggc caagtgcagt ggcatgatca tagctctgta acctgacctc 5580
 tgacctctga ctccctggac acaagtgatc ctccctgtctc tcagcctccc aagtagctgg 5640
 gactacaggc attccaccac acccaactaa ttgtttttat tttttgtaga gacagggcct 5700
 tgctatggtg cccaggetgg caagttcttg aaataatggc tgtggccaca aactagaaaa 5760
 taattttcag gtgtacagag aatagaaaga atttagattc ataaattgat cattttgttc 5820
 acagttatTT gcataacaca gtccacattt aaagggtgca ccttagaaat caaaggggaa 5880
 gaacatcatc ctctattgaa aaagaaagaa atcaaaggat gtacagtgaa tttgcagctt 5940
 aatctatggg gagcatcatt gcaaaaaatg gttctgtgtg aggctcttcc ccacccttg 6000
 tccataggag cacattattg ttgtagtaat tatttcaccc ctctcccttt ttcagtgtac 6060
 aagtgataca tgctaatttt aacagaactt gaaagtagaa taaaattaaa ataatagttt 6120
 actaatatcc catttatctt ctctcatata tatgagataa atattaaggt gtatgtactt 6180
 atccatattg gcctgatttt ttaaaaatcct tgtatatgca tctttgcacc cttatcctaa 6240
 tatttcctta gaatatatcc ctagaagcat aattgtggga acaaaggcca tgaacatttt 6300
 caagtgttta ttttattatt ttattttatt tttattaatt ttgatacagg gttttgcttt 6360
 gttccccaga ctggagtgca gtggtgagat caccactcac tgcacctga cctcctggac 6420
 tcaagcgatc cacctgcctc agtctctcca gttagcgggg ctaaggacta caggcacatg 6480
 ccatcatgcc cagctaattt ttttatttgt agcagagacg aggtctcact gtgttgccca 6540
 ggctgctatt ttatttattt ttttaagagat agggctctcat tctgtcttcc aggctagaat 6600

10

20

30

gcagtggcac aatcatagct cactgcaacc tcaagcgatc tttgcctcag cctgagtagc 6660
 tgggactaca ggcattggcc accactctca gctaattttt ttttcaattt tttatttttt 6720
 gtagatatgg gggctctact gtgttgccca ggctggctct gaaccocctag cctaaagtga 6780
 tcttcccacc tcagcctccc aaagtgcctag gattacaggg cacaggcctc agccaagttt 6840
 taaaaatttt tactgcccata ctcttcatta gaaaagtga accagcttac attcccaggg 6900
 cagttttcta ttgatatagt agcactgaat attataatc agttaacttt tgtcaatacg 6960
 gtaggctaaa agtgctatgt tcttagccat ctctcttttg ggtaacagt gcactatttt 7020
 gttattaata attattctat ctaacaagcc cctctatgg ttttgtggct ttgtagtaag 7080
 catagttgta tttccttttt tgaggtggag tcttgctatg ttgccagggc tggagtgcag 7140
 tggcgcgac tcggctcact gcaccctccg cctcccgggt tcaagtgatt ctccctgcctc 7200
 agactcctga gtatctggga ctacagggcat gcaccaccac gccagctaa ttttttatat 7260
 ttttagtaga gaggggagtt caccgtgta gccgggatgg tctctatctc ttgacctcgt 7320
 ggtccgcgtg cctcagcctc ccaaaatgct gtgattacag goatgagcca cctgcctgg 7380
 ccaacatttc ttttacatgc ataaaagaga tctgagctgt ttttgagccc ttctagactt 7440
 tctttttttt tttttttttt tttttttttt tttttttttt ttttttttaa gtagatgagg 7500
 tcttgctatg ttgccgagac ttaacctcaa actcctaggg ccaagcaatc ctcccagct 7560
 gctgggacta caggcatgaa ccaccatgcc caacttagac ttttattgta ctatcaaaag 7620
 gcaattttct tttcaaattt ctgggtaata gtgttagaaa aatcctactt ggtaacatcc 7680
 agaaatggca tcatactgag tgattcaaat gtgagatgga agaaaagggt agaattggag 7740
 tgaacgtccc ctcttatctc aaatgtattt tatctccatt ttgtttcata gtttattagt 7800
 ttgaagatgc tttgaatgtc acctaatcat tttcaactct aggtccagaa aaatcaaggg 7860
 catgatttct gaaattacac ttagcctaata taaaacttag aaacactggt caccttcttc 7920
 aatgtttttg actgagctct tttcatttat aagtgcaggg aggtgttact ataacattat 7980
 ttcctagaat gtcaaatttt gagcctaata gcatggtaaa tttggctata tttgttgttt 8040
 tttgtttttg tttttttttt aatgaaactt agtatttctt tgtttccac ttcttttttt 8100
 tttttttttt tttttttttt tgagacggag tctctctctg tcatccaggg tggagtgcaa 8160
 tggcgtgatc ttggctcact gccacctccg cctcgcaggt tcaagctatt ctctttcac 8220
 agcctcctga gtagctggga ctacagggac ccaccaccac gcccgccaa tttttttgta 8280
 ttttttagtag agaoggggtt ttaocatggt aggcaggatg gtctcgaact cctgaccttg 8340
 tgatctgccc gcctcagcct cccaaagtgc tgggattaca ggcgtgagcc accgcacctg 8400

10

20

30

gcctcccact tctttttaat atgtcgtgtc ataactgaac agtaaagtga gcagattatc 8460
aggttaaatc tgaagtgtca gtctggtcac cagtgcocaa gttactgccc ctatggtaat 8520
attggttact ttgtattttc ctacagcaaa cataaaattt gttatagtga gatttttacc 8580
tgtatacctc tcttaacttt aatgttatta cctcaaggaa gatattatca tgaatgaaga 8640
ttccatgatg aaagttttgc agagtttatt gcagtaattt agtacttcat tagaatcttt 8700
agtttttttag gagcacagta ctgaatgttt gtttctttgt tggacctttt gaaaaccggt 8760
ttccattga tgcagtgtag ctgttacagg aatcatcttt ttaaaacggt tttatacagc 8820
atggctgaaa attgaacctg ggcctccctc gtggcctacc attgaaggaa cagcattttt 8880
tgcctatcta gaaagacaat gttaaatgtg ctatctatat attttttaac ttgtgctacc 8940
tactacgcgt ttatattttg ggaatctggt ttcttttggg caaaaccaca aatcaaaaac 9000
acctcatttc ttaggcattt gaaatcccta attcagaata atctcccaaa cagaaacaca 9060
actacctgca ttctttttga caaaagagct aagtagcatt agaaaattat tttaaaccca 9120
attctgtttt ttaacagaat aaaattcttc tgttcttcac atcttctttt cataggtaac 9180
ctattgaaag tagggtttat ttgggggaag catttcttcc tgtctcttat ctcataataa 9240
atacagggtg gcttaactac tagtttccca cctcaaagat atactcaaat ctaaagatgt 9300
ttaagatttt gggatctgaa gagtaaacad ttctccta atcacaatgtga cagagacaaa 9360
tgaatcaagc caatgctact tttattttat cataactaact ggaacttttc tttttggaaa 9420
tcagatacat ttgtatgta ttagtaattt ggaatcctgc attggttatc ctgcctctcc 9480
caaagcagat tctgaaatta taaaggtgca caggttctcc atgcaacacc aaaagttata 9540
ttttccaagg ctttgtaaaa ttgtagaatg tctgtttaa tttctgtcaa atcagtaact 9600
cacactgttt tgagaattat gaataaagga ataaaatatt gttagtgttt atttagtaca 9660
aaagtagatt atagaatctc agcatttttg tcaaaaaatt tctttttgat gattgacaga 9720
tcaggagaca cttaaggcca tacctgcttt cagtaatcaa aaatgcattt aagatccaga 9780
aacttgaggt agcagaacat cactatcaca tataacatat cctttggtat agaaaattat 9840
attcccagag tgagtctctt ttttaaaacc attaatgagg ccaaggtggg aagatcactt 9900
gggaccagga gttcaagacc aagcctgggc cagatggcga gacctgtct ctacaaaaaa 9960
ttaactggat gtggtggtgc actcctgtag tcccacctac tcagaggctg aggcaggagg 10020
atcccttgag cccaggaaat tgtagtggca gtgagctatg atcatactac tgtactgcag 10080
tctgggccac gaagtggagc cgtgtctctt aaaaaaaaaa aaatgttagg catggtggca 10140
caggcatata gttttagcta cttaggaggc tgaggcagga ggatcacttg agcccagaag 10200
ttcaagatta cagtgagtta tgattgtgcc gctgcactcc aacctgggtg acaaaataac 10260

10

20

30

cctgtctctg gcgggtaggg ggggaagtga ttatttactt tgaaatatgt tcaaaactga 10320
ttcctgttct atattcctaa tgaacagaat agactttata taaaacaaat agttaaactt 10380
aaggataaaa ttttaatgga agtataatat atatatcttc cagctcttct gtcttctaata 10440
gtatttatta cagaaaatga aattactttg tttccgcaat ctttgatca cttcagttct 10500
ccaataaatc tgagaattct ggtagtgtga aatattcagc tttctttgct tatttacata 10560
aaatgtataa ggacaatttg tgataattaa gagttacatt taaatatcag gaaaaagtta 10620
taaatttaaa ttaaaaaatt ttaaaaggaa attattagaa attttaaaag aatgaactaa 10680
aaggtgatta tatgtaaatg cttgcatata tgaatattag cattgtcccc aaaataattt 10740
agaacaaaaga aattggaatc aaataaataa aggtttgatt atttttaaag tggtctatat 10800
tccatgataa aagagagggt tatcagtggc ataagaaagg tttttcacct tttttgtatt 10860
gaaatctttg acatatacat atatatcttt gccatcttt gtgtatcttt gctcgtatga 10920
gagcaaagat ataggcaaag atatgctctc tctctctatg tctttgttca taccaagacc 10980
ttcctgatat ctccacataa tottaaatat aggaacatta gactggatga tctctgtgcc 11040
ccctttatct ctactcttcc attattttat actttaacac atcatctctg ttttatgata 11100
taagaatgga atatttcttt tttcctgaaa atgcttattt tggtcacttg atacacatta 11160
ggccaatatg tgttacttga gtgaccatc ttccttcttt tcatttctgt ctctgtctat 11220
taacctggat atctggaatg tggactaaac tcttcaaaca ctatgtaaaa cctactaacc 11280
tttgtgcatt tggttgctca gctactaaga gcaccatttc tgaactgaag ttaactgaag 11340
accattctgt tttagagatt atgacatacc ttttggatc tcatgccttt ttctctcctt 11400
ctcaagggtg aagagatggt gcagaaccac atgacttact cattacagga tgtaggcgga 11460
gatgccaatt ggcagttggt tgtagaagaa ggagaaatga aggtaattcc ccctgaaatg 11520
ttatagattg ccaaaggcgt ctctgtttca gtcataattat cattactatt gatatgaata 11580
aggatagcac tttcaactta cctttaaaac aaattattac atgtgatcaa agcagtacca 11640
tatattgagc aataaaatgt ctttttgctt ttctggcttt gcccttacta aaggttttta 11700
tgattataat ataaatatat gattaaaact ttctgttttg actaggccat gaagaaaata 11760
aaatttagag aattagatat gaccaggta caattagctg atggctctgt atttggatat 11820
ttccttttgt ttgtttttt taacatactg aatgttgtgc ctgatgaca ctttgtttct 11880
ctcccttttt ggtctatacc ctcttctttt tcccttctct tactgcacct ttaattgata 11940
tttggacatt ggtcagttaa tcctggttac atccctaaac acatggacag aaaataagag 12000
cagggactga gagatacaga gatggattga aaagcaaaag caacattgaa ttttggattt 12060

10

20

30

tctcattcct aaggaactat gctaaataaa gatacaaaga taataagaca ctctccaagc 12120
 taaagcttta gtttaaggaaa aagaatattg acatttaaaa gatactattg gccaggcaca 12180
 gtggctatgc ctgtaatccc agcactttta ggaggacatg gcaggcggat tacttgagct 12240
 caggagtcca agtcaaacct gggcaacacg gtgaaacccc gtctctacca aaaatacaaa 12300
 aattagctgg gtgcagtacc acacacttgt agtcccagct acccaggagg ctgggcaaaa 12360
 gattccttga gccagggagg tcaaggctgc aatgagccgc gtttgtgcca ctgcactota 12420
 gcctgggtca caaagtgaga cctgtgtga gatatatata tatatatata tatatatata 12480
 ta 12482

10

<210> 26
 <211> 21
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPpep1

<400> 26
 Lys Gly Lys Arg Gly Asp Ser Gly Ser Pro Ala Thr Trp Thr Thr Arg
 1 5 10 15

Gly Phe Val Phe Thr
 20

20

<210> 27
 <211> 21
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPpep1Ala9

<400> 27
 Lys Gly Lys Arg Gly Asp Ala Gly Ser Pro Ala Thr Trp Thr Thr Arg
 1 5 10 15

Gly Phe Val Phe Thr
 20

30

<210> 28
 <211> 50
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: ON-GPBP-54m

<400> 28
 tcgaattcac catggcccca ctagccgact acaaggacga c gatgacaag 50

<210> 29
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: ON-GPBP-55c

<400> 29
ccgagcccga cgagttccag ctctgattat ccgacatctt gtcacgctcg 50

<210> 30
<211> 32
<212> DNA
<213> Artificial Sequence

10

<220>
<223> Description of Artificial Sequence: ON-HNC-B-N-14m

<400> 30
cgggatccgc tagctaagcc aggcaaggat gg 32

<210> 31
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: ON-HNC-B-N-16c

20

<400> 31
cgggatccat gcataaatag cagttctgct gt 32

<210> 32
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: FLAG peptide

<400> 32
Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

30

<210> 33
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Hypothetical peptide

<400> 33
Pro Arg Ser Ala Arg Cys Gln Ala Arg Arg Arg Arg Gly Gly Arg Thr
1 5 10 15

Ser Ser

<210> 34
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: ON-GPBP-11m

<400> 34
 gcgggactca gcggccggat tttct 25

<210> 35
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: ON-GPBP-15m

<400> 35
 acagctggca gaagagac 18

<210> 36
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: ON-GPBP-20c

<400> 36
 catgggtagc ttttaaag 18

<210> 37
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: ON-GPBP-22m

<400> 37
 tagaagaaca gtcacagagt gaaaagg 27

<210> 38
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: ON-GPBP-53c

10

20

30

<400> 38 gaattcgaac aaaataggct ttc	23	
<210> 39 <211> 17 <212> DNA <213> Artificial Sequence		
<220> <223> Description of Artificial Sequence: ON-GPBP-56m		
<400> 39 ccctatagtc gctcttc	17	
<210> 40 <211> 17 <212> DNA <213> Artificial Sequence		10
<220> <223> Description of Artificial Sequence: ON-GPBP-57c		
<400> 40 ctgggagctg aatctgt	17	
<210> 41 <211> 24 <212> DNA <213> Artificial Sequence		20
<220> <223> Description of Artificial Sequence: ON-GPBP-62c		
<400> 41 gtggttctgc accatctctt caac	24	
<210> 42 <211> 41 <212> DNA <213> Artificial Sequence		
<220> <223> Description of Artificial Sequence: ON-GPBP-26		30
<400> 42 cacatagatt tgtccaaaag gttgaagaga tggcgcagaa c	41	
<210> 43 <211> 19 <212> PRT <213> Artificial Sequence		
<220> <223> Description of Artificial Sequence: GPIII derived peptide		

<400> 43
 Gln Arg Ala His Gly Gln Asp Leu Asp Ala Leu Phe Val Lys Val Leu
 1 5 10 15

Arg Ser Pro

<210> 44
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPIII-IV-V
 derived peptide

10

<400> 44
 Gln Arg Ala His Gly Gln Asp Leu Glu Ser Leu Phe His Gln
 1 5 10

<210> 45
 <211> 685
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPDV

<220>
 <221> CDS
 <222> (1)..(633)

20

<400> 45
 ggt ttg aaa gga aaa cgt gga gac agt gga tca cct gca acc tgg aca 48
 Gly Leu Lys Gly Lys Arg Gly Asp Ser Gly Ser Pro Ala Thr Trp Thr
 1 5 10 15
 acg aga ggc ttt gtc ttc acc cga cac agt caa acc aca gca att cct 96
 Thr Arg Gly Phe Val Phe Thr Arg His Ser Gln Thr Thr Ala Ile Pro
 20 25 30
 tca tgt cca gag ggg aca gtg cca ctc tac agt ggg ttt tct ttt ctt 144
 Ser Cys Pro Glu Gly Thr Val Pro Leu Tyr Ser Gly Phe Ser Phe Leu
 35 40 45
 ttt gta caa gga aat caa cga gcc cac gga caa gac ctt gga act ctt 192
 Phe Val Gln Gly Asn Gln Arg Ala His Gly Gln Asp Leu Gly Thr Leu
 50 55 60
 ggc agc tgc ctg cag cga ttt acc aca atg cca ttc tta ttc tgc aat 240
 Gly Ser Cys Leu Gln Arg Phe Thr Thr Met Pro Phe Leu Phe Cys Asn
 65 70 75 80
 gtc aat gat gta tgt aat ttt gca tct cga aat gat tat tca tac tgg 288
 Val Asn Asp Val Cys Asn Phe Ala Ser Arg Asn Asp Tyr Ser Tyr Trp
 85 90 95
 ctg tca aca cca gct ctg atg cca atg aac atg gct ccc att act ggc 336
 Leu Ser Thr Pro Ala Leu Met Pro Met Asn Met Ala Pro Ile Thr Gly

30

```

      100              105              110
aga gcc ctt gag cct tat ata agc aga tgc act gtt tgt gaa ggt cct 384
Arg Ala Leu Glu Pro Tyr Ile Ser Arg Cys Thr Val Cys Glu Gly Pro
      115              120              125

gcg atc gcc ata gcc gtt cac agc caa acc act gac att cct cca tgt 432
Ala Ile Ala Ile Ala Val His Ser Gln Thr Thr Asp Ile Pro Pro Cys
      130              135              140

cct cac gcc tgg att tct ctc tgg aaa gga ttt tca ttc atc atg aaa 480
Pro His Gly Trp Ile Ser Leu Trp Lys Gly Phe Ser Phe Ile Met Lys
      145              150              155              160

gcc tat tcc atc aac tgt gaa agc tgg gga att aga aaa aat aat aag 528
Ala Tyr Ser Ile Asn Cys Glu Ser Trp Gly Ile Arg Lys Asn Asn Lys
      165              170              175

tcg ctg tca ggt gtg cat gaa gaa aag aca ctg aag cta aaa aag aca 576
Ser Leu Ser Gly Val His Glu Glu Lys Thr Leu Lys Leu Lys Lys Thr
      180              185              190

gca gaa ctg cta ttt ttc atc cta aag aac aaa gta atg aca gaa cat 624
Ala Glu Leu Leu Phe Phe Ile Leu Lys Asn Lys Val Met Thr Glu His
      195              200              205

gct gtt att taggtatctt tctttaacca aacaatattg ctccatgatg 673
Ala Val Ile
      210

acttagtaca aa 685

```

```

<210> 46
<211> 211
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: GPDV

```

```

<400> 46
Gly Leu Lys Gly Lys Arg Gly Asp Ser Gly Ser Pro Ala Thr Trp Thr
  1              5              10              15

Thr Arg Gly Phe Val Phe Thr Arg His Ser Gln Thr Thr Ala Ile Pro
      20              25              30

Ser Cys Pro Glu Gly Thr Val Pro Leu Tyr Ser Gly Phe Ser Phe Leu
      35              40              45

Phe Val Gln Gly Asn Gln Arg Ala His Gly Gln Asp Leu Gly Thr Leu
      50              55              60

Gly Ser Cys Leu Gln Arg Phe Thr Thr Met Pro Phe Leu Phe Cys Asn
      65              70              75              80

Val Asn Asp Val Cys Asn Phe Ala Ser Arg Asn Asp Tyr Ser Tyr Trp
      85              90              95

Leu Ser Thr Pro Ala Leu Met Pro Met Asn Met Ala Pro Ile Thr Gly

```

10

20

30

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

10

<210> 47
 <211> 680
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPDIII

20

<220>
 <221> CDS
 <222> (1)..(216)

<400> 47
 ggt ttg aaa gga aaa cgt gga gac agt gga tca cct gca acc tgg aca 48
 Gly Leu Lys Gly Lys Arg Gly Asp Ser Gly Ser Pro Ala Thr Trp Thr
 1 5 10 15
 acg aga ggc ttt gtc ttc acc cga cac agt caa acc aca gca att cct 96
 Thr Arg Gly Phe Val Phe Thr Arg His Ser Gln Thr Thr Ala Ile Pro
 20 25 30
 tca tgt cca gag ggg aca gtg cca ctc tac agt ggg ttt tct ttt ctt 144
 Ser Cys Pro Glu Gly Thr Val Pro Leu Tyr Ser Gly Phe Ser Phe Leu
 35 40 45
 ttt gta caa gga aat caa cga gcc cac gga caa gac ctt gat gca ctg 192
 Phe Val Gln Gly Asn Gln Arg Ala His Gly Gln Asp Leu Asp Ala Leu
 50 55 60
 ttt gtg aag gtc ctg cga tcg cca tagcogttca cagccaaacc actgacatcc 246
 Phe Val Lys Val Leu Arg Ser Pro
 65 70
 ctccatgtcc tcaoggetgg atttctctct ggaaggatt ttcattcatc atgttcacaa 306
 gtgcaggttc tgagggcacc gggcaagcac tggcctcccc tggctcctgc ctggaagaat 366

30

tccgagccag cccatttcta gaatgtcatg gaagaggaac gtgcaactac tattcaaatt 426
 cctacagttt ctggctggct tcattaacc cagaaagaat gttcagaaag cctattccat 486
 caactgtgaa agctggggaa tttagaaaaa taataagtcg ctgtcagggtg tgcataaaga 546
 aaagacactg aagctaaaaa agacagcaga actgctatatt ttcacocaa agaacaaaagt 606
 aatgacagaa catgctgtta tttaggtatt tttctttaac caaacaatat tgctccatga 666
 tgacttagta caaa 680

<210> 48
 <211> 72
 <212> PRT
 <213> Artificial Sequence

10

<220>
 <223> Description of Artificial Sequence: GPDIII

<400> 48
 Gly Leu Lys Gly Lys Arg Gly Asp Ser Gly Ser Pro Ala Thr Trp Thr
 1 5 10 15
 Thr Arg Gly Phe Val Phe Thr Arg His Ser Gln Thr Thr Ala Ile Pro
 20 25 30
 Ser Cys Pro Glu Gly Thr Val Pro Leu Tyr Ser Gly Phe Ser Phe Leu
 35 40 45
 Phe Val Gln Gly Asn Gln Arg Ala His Gly Gln Asp Leu Asp Ala Leu
 50 55 60
 Phe Val Lys Val Leu Arg Ser Pro
 65 70

20

<210> 49
 <211> 392
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPDIII-IV-V

<220>
 <221> CDS
 <222> (1)..(207)

30

<400> 49
 ggt ttg aaa gga aaa cgt gga gac agt gga tca cct gca acc tgg aca 48
 Gly Leu Lys Gly Lys Arg Gly Asp Ser Gly Ser Pro Ala Thr Trp Thr
 1 5 10 15
 acg aga ggc ttt gtc ttc acc cga cac agt caa acc aca gca att cct 96
 Thr Arg Gly Phe Val Phe Thr Arg His Ser Gln Thr Thr Ala Ile Pro
 20 25 30
 tca tgt cca gag ggg aca gtg cca ctc tac agt ggg ttt tct ttt ctt 144
 Ser Cys Pro Glu Gly Thr Val Pro Leu Tyr Ser Gly Phe Ser Phe Leu

35 40 45
 ttt gta caa gga aat caa cga gcc cac gga caa gac ott gaa agc cta 192
 Phe Val Gln Gly Asn Gln Arg Ala His Gly Gln Asp Leu Glu Ser Leu
 50 55 60
 ttc cat caa ctg tga aagctgggga attagaaaaa ataataagtc gctgtcaggt 247
 Phe His Gln Leu
 65
 gtgcatgaag aaaagacact gaagctaaaa aagacagcag aactgctatt tttcatccta 307
 aagaacaaaag taatgacaga acatgctggt atttaggtat ttttctttaa ccaaacaata 367
 ttgctccatg atgacttagt acaaa 392

10

<210> 50
 <211> 68
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPDIII-IV-V

<400> 50
 Gly Leu Lys Gly Lys Arg Gly Asp Ser Gly Ser Pro Ala Thr Trp Thr
 1 5 10 15
 Thr Arg Gly Phe Val Phe Thr Arg His Ser Gln Thr Thr Ala Ile Pro
 20 25 30
 Ser Cys Pro Glu Gly Thr Val Pro Leu Tyr Ser Gly Phe Ser Phe Leu
 35 40 45

20

Phe Val Gln Gly Asn Gln Arg Ala His Gly Gln Asp Leu Glu Ser Leu
 50 55 60

Phe His Gln Leu
 65

<210> 51
 <211> 507
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPDIII-V

30

<220>
 <221> CDS
 <222> (1)..(216)

<400> 51
 ggt ttg aaa gga aaa cgt gga gac agt gga tca cct gca acc tgg aca 48
 Gly Leu Lys Gly Lys Arg Gly Asp Ser Gly Ser Pro Ala Thr Trp Thr
 1 5 10 15
 acg aga ggc ttt gtc ttc acc cga cac agt caa acc aca gca att cct 96
 Thr Arg Gly Phe Val Phe Thr Arg His Ser Gln Thr Thr Ala Ile Pro

```

      20              25              30
tca tgt cca gag ggg aca gtg cca ctc tac agt ggg ttt tct ttt ctt 144
Ser Cys Pro Glu Gly Thr Val Pro Leu Tyr Ser Gly Phe Ser Phe Leu
      35              40              45

ttt gta caa gga aat caa cga gcc cac gga caa gac ctt gat gca ctg 192
Phe Val Gln Gly Asn Gln Arg Ala His Gly Gln Asp Leu Asp Ala Leu
      50              55              60

ttt gtg aag gtc ctg cga tcg cca tagccgttca cagccaaacc actgacattc 246
Phe Val Lys Val Leu Arg Ser Pro
      65              70

ctccatgtcc tcacggctgg atttctctct ggaaaggatt ttcattcatc atgaaagcct 306

attccatcaa ctgtgaaagc tggggaatta gaaaaaataa taagtcgctg tcaggtgtgc 366

atgaagaaaa gacactgaag ctaaaaaaga cagcagaact gctatTTTTc atcctaaaga 426

acaaagtaat gacagaacat gctgttattt aggtattttt cttaaaccAA acaatattgc 486

tccatgatga cttagtACAA a 507

```

<210> 52
 <211> 72
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: GPDIII-V

```

<400> 52
Gly Leu Lys Gly Lys Arg Gly Asp Ser Gly Ser Pro Ala Thr Trp Thr
  1              5              10              15

Thr Arg Gly Phe Val Phe Thr Arg His Ser Gln Thr Thr Ala Ile Pro
      20              25              30

Ser Cys Pro Glu Gly Thr Val Pro Leu Tyr Ser Gly Phe Ser Phe Leu
      35              40              45

Phe Val Gln Gly Asn Gln Arg Ala His Gly Gln Asp Leu Asp Ala Leu
      50              55              60

Phe Val Lys Val Leu Arg Ser Pro
      65              70

```

<210> 53
 <211> 659
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: HMBP-21

<220>
 <221> CDS
 <222> (37)..(627)

10

20

30

<220>

<223> Description of Artificial Sequence: HMBP-21

<400> 54

```

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

```

10

20

【 図 1 】

GCAGGAAGATGGCGCGGTAGCGGAGGTGTGAGTGGACCGCGGACTACGCGGGCGGATTTCTCTTCCT 70
 TCTTTTCCCTTTTCTCCCTATTTGAAATGGCATCGAGGGGGCTAAGTTCGGGTGGCAGCGCGGGG 140
 CAACGCGGGGTACCGCGCACCGCGCGCGGCTGACGGCTGGAAGGGTAGGCTTCATTCCACGCGCTGTC 210
 CTCCTTCCTCGCTCCGCTCGGTGTCAGGCGCGCGCGCGCGCGCGCGCGGACTTCTGCTCCCTCTCTGC 280
 TCCCGCCACACCGAGCGGCACTTCTGCTGCGCATCCCGCCACCTTCCACCCCGGAGACTGGCGCC 350
 CTCCTCCGCGCAGCTGAGGAGCGGGGCGCGCTCTCTGCTGGTTCGAGCCCTCATGTCGGATAAT 420
 M...S...D...M 4
 CAGAGCTGGAATCCTCGGGCTCGGAGGAGTCCAGAGCGGAGTCTCGGCGCGCTGTGGAGCGCTGCG 490
 Q...S...W...N...S...S...E...E...D...P...E...T...R...S...G...P...V...E...R...C 27
 GGGTCTCAGTAACTGACAACTCATTCATGGGTGGCAGGATCTGGGTAGTTTGAATAAATAATGC 560
 G...V...L...S...K...M...T...H...Y...E...H...S...W...Q...D...H...W...V...Y...L...K...N...H...A 51
 TCTGAGTTACTACAAATCTGAAGATGAACAAGAGTATGGCTCGAGAGGATCCATCTGCTTAGCAAGGCT 630
 L...S...Y...Y...K...S...E...E...D...E...T...H...Y...S...C...R...G...S...I...C...L...S...K...A 74
 GTCATCAGCTCCAGGATTTGATGATCTGCGATTTGATATAGTGTAAATGATAGTCTTTGGTATCTTC 700
 V...I...T...E...H...D...F...D...E...C...R...F...D...I...S...V...N...D...S...V...W...Y...L 97
 GTGCTCAGGATCCAGATCAGAGCAGCAAGTATGATAGATGATGATGATGATGATGATGATGATGATG 770
 R...A...Q...D...P...D...H...R...Q...W...I...D...A...I...E...Q...H...K...E...S...G...Y 121
 TGGATCTGAATCAGCTTGGCTCGACATGCTCAATGGTGTCCCTGGTGTCTGGAGCAAGTGGCTACTCT 840
 G...S...E...S...L...R...R...H...G...S...M...V...S...L...V...S...G...A...S...G...Y...S 144
 CCAACATCCACTCTTCATTCAGAAAGCCACAGTTTACGTGAGAGTTGCTGAAATGGAACATTTA 910
 A...T...S...T...S...S...F...K...K...G...H...S...L...R...E...K...L...A...E...M...E...T...F 167
 GAGACATCTTATGTAGACAGTGTACACGCTACAGCAAGTACTTTGATGCTGTGCTGATGCTGTCTCAA 980
 R...D...I...L...C...R...Q...V...D...T...L...Q...K...Y...F...D...A...C...A...D...A...V...S...K 191
 GGATGACTTCAAGGAGTAAAGTGTAGAGATGATGAGAGTACTTCTCACAACCGGCTCTGATGGT 1050
 D...E...L...Q...R...D...K...V...E...D...D...E...D...D...F...P...T...T...R...S...D...G 214
 GACTTCTTGATGATACACCGGCAATAAGAAAGTATTCCACATGTGACACCAAAAGGATTAATG 1120
 D...F...L...H...S...T...N...G...N...K...B...K...L...P...P...H...V...T...P...K...G...I...M 237
 GTATGACTTTAAAGGGGAGGATAACTTTTAAAGCAACTCTGCTGAACTCCTTCCACACTTCTCTCA 1190
 G...I...D...F...K...G...B...A...I...T...F...K...A...T...T...A...G...I...L...A...T...L...S...H 261
 TTTATGAACTAATGTTAAACGTGAGGACAGCTGGCGAAGAGCTGGATAAGGAACTGGAAGAGAA 1260
 C...I...E...L...M...V...K...R...E...D...S...W...Q...K...B...L...D...K...E...T...E...K...K 284
 AGAAGACGAGCGAGCATATRAAATGCAATGACAGAACTTAAGAAAAATCCCACTTTGGAGACCCG 1330
 R...E...T...E...E...A...Y...K...N...A...M...T...E...L...K...K...E...S...H...E...G...G...P 307
 ATTATGAGAAGGCCCTAACAGTCTGATTAAATGAGAAGAGTCTTTGATGCTGTTGAAAGCTGCTTGA 1400
 D...Y...E...E...G...P...N...S...L...I...N...E...E...E...F...F...D...A...V...E...A...A...L...D 331

FIG. 1

CGACAGATFAAATAGAAAGACACTCAGAGGTGAAAGGTGAGATTAGATTGGCTACTCCTTGCCC 1470
 R...Q...D...K...I...E...B...E...Q...S...Q...S...E...K...V...R...L...H...W...P...T...S...L...P 354
 TCTGAGATGCCCTTTCTCTGTGGGACACATAGATTGTCCAAAAGCCCTATAGTCGCTCTTCCCTCA 1540
 S...G...D...A...F...S...S...V...G...T...H...R...F...V...Q...K...P...Y...S...R...B...S...S 377
 TGCTTCCATTGATCTAGTCACTGCTGATGATGATTCACAGATTCCAGTCCAGGTTGAAGAGAGGT 1610
 M...S...S...Y...D...L...V...S...A...S...D...D...V...H...R...F...S...S...Q...V...E...E...M...V 401
 GCAGAACCACTGACTACTCATTACAGGATGTAGCGGAGATGCCAATGGCAGTTGGTGTAGAGAA 1680
 Q...N...H...W...Y...V...S...L...Q...D...V...G...O...D...A...H...W...Q...L...V...V...E 424
 GGAGAAATGAGGTATACAGAGAGAGATGAAAGAAAATGGGATTTCTGATCCCTTTAAAGCTACCC 1750
 G...E...M...K...V...Y...R...R...E...V...E...E...N...G...I...V...L...D...P...L...K...A...T 447
 ATCAGTTAAAGGCGCTCAGAGACATGAAGTCTCAATTAATTTCTGAAATGTGACGTTCCCAATGACTG 1830
 H...A...V...K...G...V...T...G...H...E...V...C...N...Y...F...W...N...V...D...V...R...N...D...W 471
 GGAACACTATAGAAAATTTTCATGTGTGGAACATTAAGTGTGATGCAATCATCAATTTATCAACA 1890
 E...T...T...I...E...H...F...H...V...V...S...T...L...A...D...H...A...I...I...Y...Q...T 494
 CACAAGAGGTTGGCTCTCTCAGCGAGAGTATTATATCTTTCTGATCCCTGAAAGATACCAAGCT 1960
 H...K...R...V...W...P...A...S...O...R...D...V...L...Y...L...S...V...I...R...K...I...P...A 517
 TGACTGAAATGACCTGAACTTGAATGATGTTTAAATTTTCTGGATCATGACAGTCTCTCTTAA 2030
 L...T...E...N...D...P...E...T...W...I...V...C...N...F...S...V...D...H...D...S...A...P...L...N 541
 CACCGATGTCTCCGTGCCAAAATAATGTTGCTATGATTTGTCAACCTTGGTAGCCACCAGAGGGA 2100
 N...R...C...V...R...A...K...I...M...V...A...M...I...C...Q...T...L...V...S...P...P...E...G 564
 AACCGAAATGACGAGGCAACATCTATGCAAGATTACATATAGCTAAATGGAACCTGGAGAT 2170
 N...Q...E...I...S...R...D...N...I...L...C...K...I...T...Y...V...A...N...V...N...P...G...G 587
 GGCACCGCTCAGTGTAAAGGCGAGTGCAGAGGATGATCTAAATTTCTAAACGTTTACTTC 2240
 W...A...P...A...S...V...L...R...A...V...A...K...R...E...Y...P...K...P...L...K...R...P...T...S 611
 TTACTCCAGAAAAAATCCAGGAAAGCTATTGTTGCTAGTATTACAGGACTAGAAAGATATGTT 2310
 Y...V...Q...E...K...T...A...G...K...P...J...L...F 624
 TACTTTTTTAACTTTTGTGACTAATATGACTGCTCAACTAAAATTTAGTTTGTGAAAGATTTACT 2380
 ATGTTTTT 2389

FIG. 1

【 図 2 】

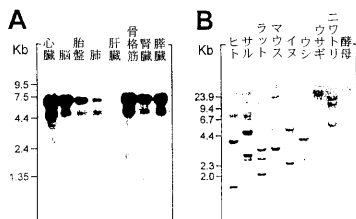


FIG. 2

【 図 3 】

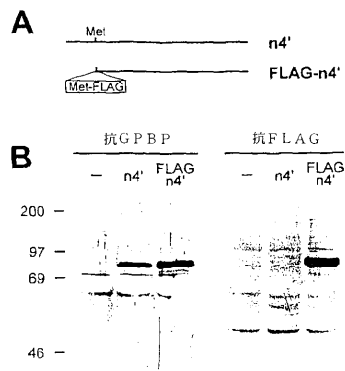


FIG. 3

【 図 4 】

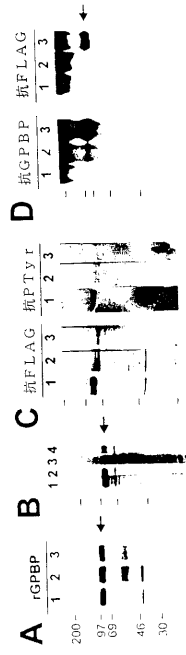


FIG. 4

【 図 5 】

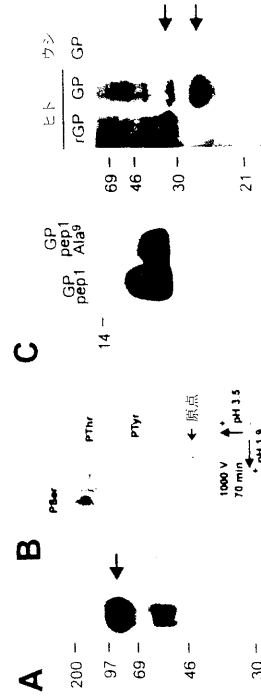


FIG. 5

【 図 6 】

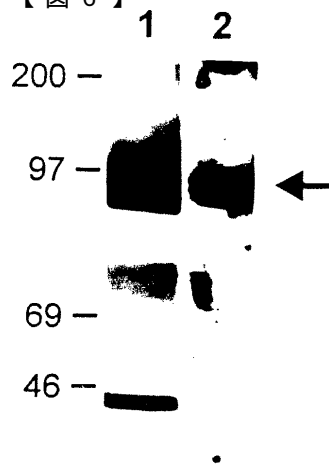


FIG. 6

【 図 7 】

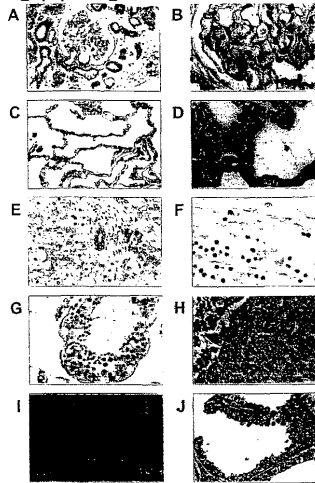


FIG. 7

【 図 8 】

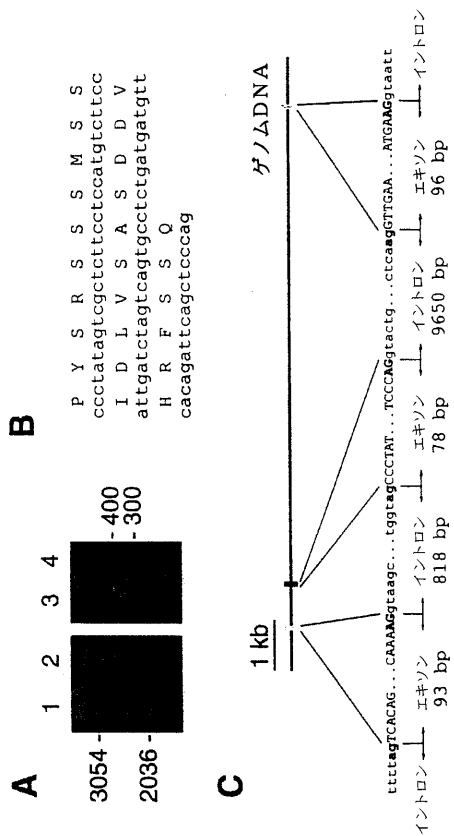


FIG. 8

【 図 9 】

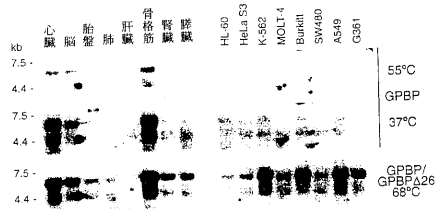


FIG. 9

【 図 10 】

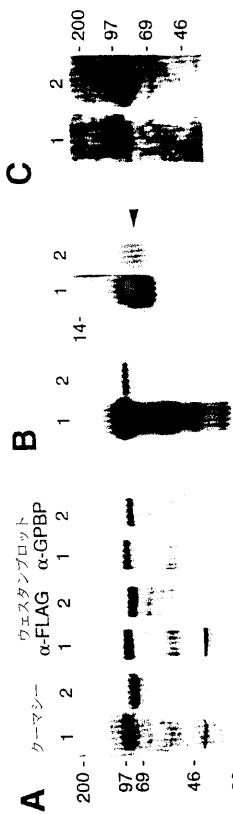


FIG. 10

【 図 11 】

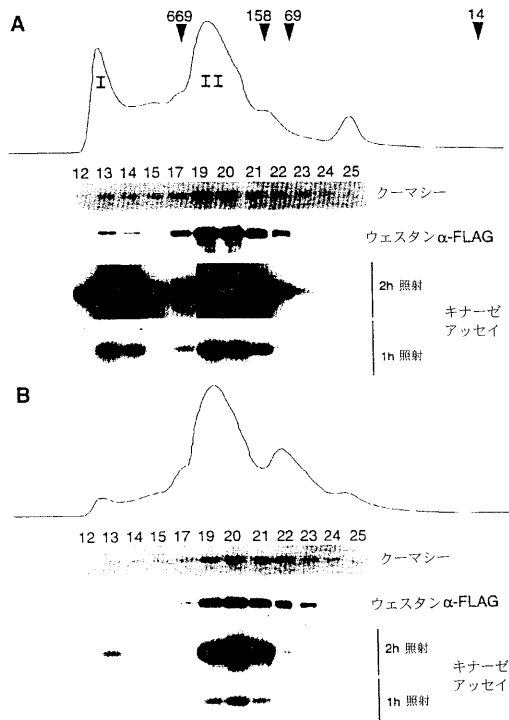


FIG. 11

【 図 1 2 】



FIG. 12

【 図 1 3 】

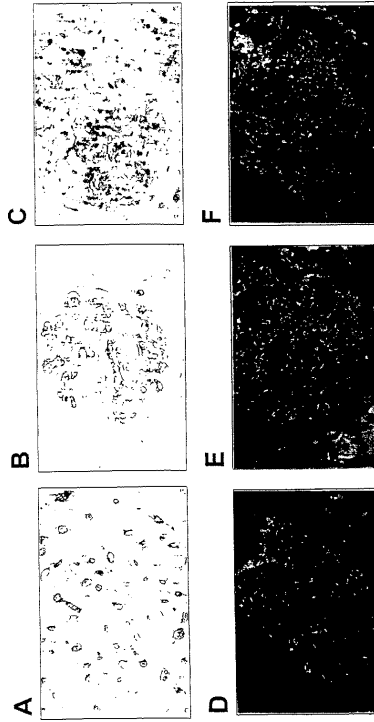


FIG. 13

【 図 1 4 】

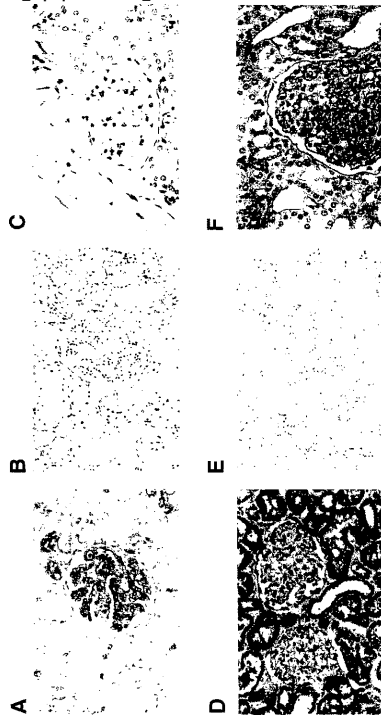


FIG. 14

【 図 1 5 】

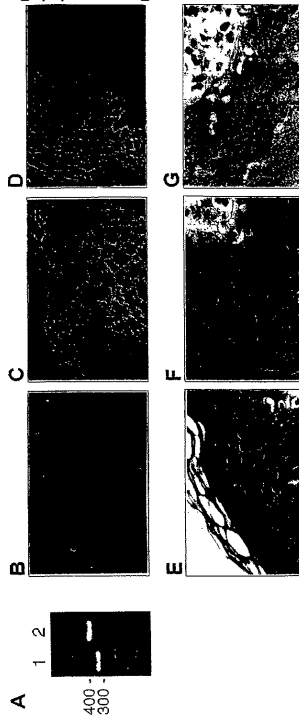


FIG. 15

【 図 16 】

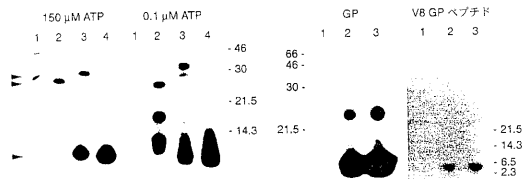


FIG. 16

【 図 17 】

GPΔIII GLKGRGDSGSPATWTRGFVTRHSQTTAI
 MBP MASQRRP-SQRHGSKYLATASTMDHARHGFL
 GPΔIII PSCPRGFPVPLYSGFSFLVQGNQRAHGQDL
 MBP PRHRDTGILDSIGRFFGGDRGAPKRGSGK--
 GPΔIII ALFVKVLRSP
 MBP VFWLKPGRSP

FIG. 17

【 図 18 】

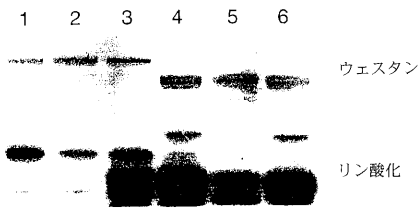


FIG. 18

【 図 19 】



FIG. 19

【 図 2 0 】

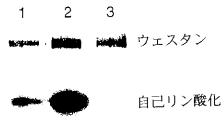


FIG. 20

フロントページの続き

(51) Int.Cl.			F I		
C 1 2 N	1/19	(2006.01)	C 1 2 N	1/19	
C 1 2 N	1/21	(2006.01)	C 1 2 N	1/21	
C 1 2 N	9/12	(2006.01)	C 1 2 N	9/12	
C 1 2 Q	1/02	(2006.01)	C 1 2 Q	1/02	
C 1 2 Q	1/68	(2006.01)	C 1 2 Q	1/68	A
G 0 1 N	33/53	(2006.01)	G 0 1 N	33/53	D
G 0 1 N	33/564	(2006.01)	G 0 1 N	33/53	M
C 1 2 N	5/10	(2006.01)	G 0 1 N	33/564	Z
C 1 2 P	21/08	(2006.01)	C 1 2 N	5/00	A
			C 1 2 P	21/08	

(56) 参考文献 米国特許第 5 4 2 4 4 0 8 (U S , A)

Eur.J.Biochem., 1995, Vol.229, pp.754-760

Proc.Natl.Acad.Sci.USA, 1986, Vol.83, pp.4962-4966

The Journal of Biological Chemistry, 1995, Vol.270, No.22, pp.13254-13261

The Journal of Biological Chemistry, April 1999, Vol.274, No.18, pp.12642-12649

(58) 調査した分野(Int.Cl. , D B 名)

C12N 15/09

C12N 15/54

C07K 14/47

C07K 16/18

C07K 16/40

C12N 9/12

C12Q 1/68

BIOSIS/WPI (DIALOG)

PubMed

SwissProt/PIR/GeneSeq

GenBank/EMBL/DDBJ/GeneSeq

专利名称(译)	Goodpasture抗原结合蛋白		
公开(公告)号	JP3761784B2	公开(公告)日	2006-03-29
申请号	JP2000601171	申请日	2000-02-24
[标]申请(专利权)人(译)	Sausufuan		
申请(专利权)人(译)	南, 胡安		
当前申请(专利权)人(译)	南, 胡安		
[标]发明人	サウスフアン		
发明人	サウス、フアン		
IPC分类号	C12N15/09 C07K14/47 C07K16/18 C07K16/40 C12N1/15 C12N1/19 C12N1/21 C12N9/12 C12Q1/02 C12Q1/68 G01N33/53 G01N33/564 C12N5/10 C12P21/08 G01N33/50 A61K38/00 A61K39/00 A61P35/00 A61P37/00 A61P37/06 C12N15/54 G01N33/15		
CPC分类号	C12N9/1205 A61K38/00 C07K16/40 C12Y207/01037 G01N33/564		
FI分类号	C12N15/00.ZNAA C07K14/47 C07K16/18 C07K16/40 C12N1/15 C12N1/19 C12N1/21 C12N9/12 C12Q1/02 C12Q1/68.A G01N33/53.D G01N33/53.M G01N33/564.Z C12N5/00.A C12P21/08		
代理人(译)	三好秀		
优先权	60/121483 1999-02-24 US		
其他公开文献	JP2003525023A		
外部链接	Espacenet		

摘要(译)

本发明提供了编码Goodpasture抗原结合蛋白 (GPBP) 的分离的核酸序列和表达载体, 基本上纯化的GPBP, 针对GPBP的抗体以及检测GPBP的方法。

