(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)特許出願公表番号
特表2004-507201
(P2004-507201A)
(43)公表日　平成16年3月11日(2004.3.11)

(51)Int.Cl.[7]	F I		テーマコード（参考）
C12N 15/09	C12N 15/00	ZNAA	4B024
C07K 14/47	C07K 14/47		4B064
C07K 16/18	C07K 16/18		4B065
C07K 19/00	C07K 19/00		4H045
C12N 1/19	C12N 1/19		

審査請求 有　予備審査請求 有　（全 605 頁）　最終頁に続く

(21)出願番号　特願2000-570323(P2000-570323)
(86)(22)出願日　平成11年9月15日(1999.9.15)
(85)翻訳文提出日　平成13年3月16日(2001.3.16)
(86)国際出願番号　PCT/US1999/021090
(87)国際公開番号　WO2000/015796
(87)国際公開日　平成12年3月23日(2000.3.23)
(31)優先権主張番号　US9819330
(32)優先日　平成10年9月16日(1998.9.16)
(33)優先権主張国　米国(US)

(71)出願人　596168317
ジェネンテック・インコーポレーテッド
ＧＥＮＥＮＴＥＣＨ，ＩＮＣ．
アメリカ合衆国カリフオルニア・９４０８０－４９９０・サウス・サン・フランシスコ・ディーエヌエー・ウェイ・１
(74)代理人　100109726
弁理士　園田　吉隆
(74)代理人　100101199
弁理士　小林　義教
(72)発明者　チェン，ジアン
アメリカ合衆国　ニュージャージ　０８５４０，プリンストン，ヨーク　ドライブ　１２１

最終頁に続く

(54)【発明の名称】分泌及び膜貫通ポリペプチドとそれをコードしている核酸

(57)【要約】
本発明は新規なポリペプチド及びこれらポリペプチドをコードしている核酸分子に関する。ここにまた提供されるものはら核酸配列を含んでなるベクターと宿主細胞、異種性ポリペプチド配列に融合した本発明のポリペプチドを含んでなるキメラポリペプチド分子、本発明のポリペプチドに結合する抗体及び本発明のポリペプチドを生産する方法に関する。

【特許請求の範囲】
【請求項１】
Ｆｉｇ２（配列番号：２）、Ｆｉｇ４（配列番号：４）、Ｆｉｇ６（配列番号：１２）、Ｆｉｇ９（配列番号：１８）、Ｆｉｇ１１（配列番号：２３）、Ｆｉｇ１３（配列番号：２８）、Ｆｉｇ１５（配列番号：３４）、Ｆｉｇ１７（配列番号：３９）、Ｆｉｇ１９（配列番号：４９）、Ｆｉｇ２２（配列番号：５９）、Ｆｉｇ２４（配列番号：６４）、Ｆｉｇ２６（配列番号：６９）、Ｆｉｇ２８（配列番号：７１）、Ｆｉｇ３０（配列番号：７３）、Ｆｉｇ３２（配列番号：８４）、Ｆｉｇ３４（配列番号：９１）、Ｆｉｇ３６（配列番号：９６）、Ｆｉｇ３８（配列番号：１０４）、Ｆｉｇ４０（配列番号：１０９）、Ｆｉｇ４２（配列番号：１１４）、Ｆｉｇ４４（配列番号：１１９）、Ｆｉｇ４６（配列番号：１２７）、Ｆｉｇ４８（配列番号：１３２）、Ｆｉｇ５０（配列番号：１３７）、Ｆｉｇ５２（配列番号：１４２）、Ｆｉｇ５４（配列番号：１４８）、Ｆｉｇ５６（配列番号：１５３）、Ｆｉｇ５８（配列番号：１５９）、Ｆｉｇ６０（配列番号：１６４）、Ｆｉｇ６２（配列番号：１７０）、Ｆｉｇ６４（配列番号：１７５）、Ｆｉｇ６６（配列番号：１７７）、Ｆｉｇ６８（配列番号：１８５）、Ｆｉｇ７０（配列番号：１９０）、Ｆｉｇ７２（配列番号：１９５）、Ｆｉｇ７４（配列番号：２０１）、Ｆｉｇ７６（配列番号：２０７）、Ｆｉｇ７８（配列番号：２１３）、Ｆｉｇ８０（配列番号：２２１）、Ｆｉｇ８２（配列番号：２２７）、Ｆｉｇ８４（配列番号：２３６）、Ｆｉｇ８６（配列番号：２４５）、Ｆｉｇ８８（配列番号：２５０）、Ｆｉｇ９０（配列番号：２５５）、Ｆｉｇ９２（配列番号：２５７）、Ｆｉｇ９４（配列番号：２５９）、Ｆｉｇ９６（配列番号：２６１）、Ｆｉｇ９８（配列番号：２６３）、Ｆｉｇ１００（配列番号：２８５）、Ｆｉｇ１０２（配列番号：２９０）、Ｆｉｇ１０４（配列番号：２９２）、Ｆｉｇ１０６（配列番号：２９４）、Ｆｉｇ１０８（配列番号：３１０）、Ｆｉｇ１１０（配列番号：３１５）、Ｆｉｇ１１２（配列番号：３２０）、Ｆｉｇ１１４（配列番号：３２５）、Ｆｉｇ１１６（配列番号：３３２）、Ｆｉｇ１１８（配列番号：３３９）、Ｆｉｇ１２０（配列番号：３４１）及びＦｉｇ１２２（配列番号：３７７）に示したアミノ酸配列からなる群から選択されるアミノ酸配列を含んでなるポリペプチドをコードするヌクレオチド配列に対して少なくとも８０％の配列同一性を有する単離された核酸。
【請求項２】
前記ヌクレオチド配列が、Ｆｉｇ１（配列番号：１）、Ｆｉｇ３（配列番号：３）、Ｆｉｇ５（配列番号：１１）、Ｆｉｇ８（配列番号：１７）、Ｆｉｇ１０（配列番号：２２）、Ｆｉｇ１２（配列番号：２７）、Ｆｉｇ１４（配列番号：３３）、Ｆｉｇ１６（配列番号：３８）、Ｆｉｇ１８（配列番号：４８）、Ｆｉｇ２１（配列番号：５８）、Ｆｉｇ２３（配列番号：６３）、Ｆｉｇ２５（配列番号：６８）、Ｆｉｇ２７（配列番号：７０）、Ｆｉｇ２９（配列番号：７２）、Ｆｉｇ３１（配列番号：８３）、Ｆｉｇ３３（配列番号：９０）、Ｆｉｇ３５（配列番号：９５）、Ｆｉｇ３７（配列番号：１０３）、Ｆｉｇ３９（配列番号：１０８）、Ｆｉｇ４１（配列番号：１１３）、Ｆｉｇ４３（配列番号：１１８）、Ｆｉｇ４５（配列番号：１２６）、Ｆｉｇ４７（配列番号：１３１）、Ｆｉｇ４９（配列番号：１３６）、Ｆｉｇ５１（配列番号：１４１）、Ｆｉｇ５３（配列番号：１４７）、Ｆｉｇ５５（配列番号：１５２）、Ｆｉｇ５７（配列番号：１５８）、Ｆｉｇ５９（配列番号：１６３）、Ｆｉｇ６１（配列番号：１６９）、Ｆｉｇ６３（配列番号：１７４）、Ｆｉｇ６５（配列番号：１７６）、Ｆｉｇ６７（配列番号：１８４）、Ｆｉｇ６９（配列番号：１８９）、Ｆｉｇ７１（配列番号：１９４）、Ｆｉｇ７３（配列番号：２００）、Ｆｉｇ７５（配列番号：２０６）、Ｆｉｇ７７（配列番号：２１２）、Ｆｉｇ７９（配列番号：２２０）、Ｆｉｇ８１（配列番号：２２６）、Ｆｉｇ８３（配列番号：２３５）、Ｆｉｇ８５（配列番号：２４４）、Ｆｉｇ８７（配列番号：２４９）、Ｆｉｇ８９（配列番号：２５４）、Ｆｉｇ９１（配列番号：２５６）、Ｆｉｇ９３（配列番号：２５８）、Ｆｉｇ９５（配列番号：２６０）、Ｆｉｇ９７（配列番号：２６２）、Ｆｉｇ９９（配列番号：２８４）、Ｆｉｇ１０１（配列番号：２８９）、Ｆｉｇ１０３（配列番号：２９１）、Ｆｉｇ１０５（配列番号：２９３）、Ｆｉｇ１０７（配列番号：３０９）

、Ｆｉｇ１０９（配列番号：３１４）、Ｆｉｇ１１１（配列番号：３１９）、Ｆｉｇ１１３（配列番号：３２４）、Ｆｉｇ１１５（配列番号：３３１）、Ｆｉｇ１１７（配列番号：３３８）、Ｆｉｇ１１９（配列番号：３４０）及びＦｉｇ１２１（配列番号：３７６）に示した配列からなる群から選択されるヌクレオチド配列又はその補体を含んでなる請求項１に記載の核酸。

【請求項３】

前記ヌクレオチド配列が、Ｆｉｇ１（配列番号：１）、Ｆｉｇ３（配列番号：３）、Ｆｉｇ５（配列番号：１１）、Ｆｉｇ８（配列番号：１７）、Ｆｉｇ１０（配列番号：２２）、Ｆｉｇ１２（配列番号：２７）、Ｆｉｇ１４（配列番号：３３）、Ｆｉｇ１６（配列番号：３８）、Ｆｉｇ１８（配列番号：４８）、Ｆｉｇ２１（配列番号：５８）、Ｆｉｇ２３（配列番号：６３）、Ｆｉｇ２５（配列番号：６８）、Ｆｉｇ２７（配列番号：７０）、Ｆｉｇ２９（配列番号：７２）、Ｆｉｇ３１（配列番号：８３）、Ｆｉｇ３３（配列番号：９０）、Ｆｉｇ３５（配列番号：９５）、Ｆｉｇ３７（配列番号：１０３）、Ｆｉｇ３９（配列番号：１０８）、Ｆｉｇ４１（配列番号：１１３）、Ｆｉｇ４３（配列番号：１１８）、Ｆｉｇ４５（配列番号：１２６）、Ｆｉｇ４７（配列番号：１３１）、Ｆｉｇ４９（配列番号：１３６）、Ｆｉｇ５１（配列番号：１４１）、Ｆｉｇ５３（配列番号：１４７）、Ｆｉｇ５５（配列番号：１５２）、Ｆｉｇ５７（配列番号：１５８）、Ｆｉｇ５９（配列番号：１６３）、Ｆｉｇ６１（配列番号：１６９）、Ｆｉｇ６３（配列番号：１７４）、Ｆｉｇ６５（配列番号：１７６）、Ｆｉｇ６７（配列番号：１８４）、Ｆｉｇ６９（配列番号：１８９）、Ｆｉｇ７１（配列番号：１９４）、Ｆｉｇ７３（配列番号：２００）、Ｆｉｇ７５（配列番号：２０６）、Ｆｉｇ７７（配列番号：２１２）、Ｆｉｇ７９（配列番号：２２０）、Ｆｉｇ８１（配列番号：２２６）、Ｆｉｇ８３（配列番号：２３５）、Ｆｉｇ８５（配列番号：２４４）、Ｆｉｇ８７（配列番号：２４９）、Ｆｉｇ８９（配列番号：２５４）、Ｆｉｇ９１（配列番号：２５６）、Ｆｉｇ９３（配列番号：２５８）、Ｆｉｇ９５（配列番号：２６０）、Ｆｉｇ９７（配列番号：２６２）、Ｆｉｇ９９（配列番号：２８４）、Ｆｉｇ１０１（配列番号：２８９）、Ｆｉｇ１０３（配列番号：２９１）、Ｆｉｇ１０５（配列番号：２９３）、Ｆｉｇ１０７（配列番号：３０９）、Ｆｉｇ１０９（配列番号：３１４）、Ｆｉｇ１１１（配列番号：３１９）、Ｆｉｇ１１３（配列番号：３２４）、Ｆｉｇ１１５（配列番号：３３１）、Ｆｉｇ１１７（配列番号：３３８）、Ｆｉｇ１１９（配列番号：３４０）及びＦｉｇ１２１（配列番号：３７６）に示した配列の全長コード化配列からなる群から選択されるヌクレオチド配列又はその補体を含んでなる請求項１に記載の核酸。

【請求項４】

登録番号ＡＴＣＣ２０９２５８、ＡＴＣＣ２０９２５６、ＡＴＣＣ２０９２６４、ＡＴＣＣ２０９２５０、ＡＴＣＣ２０９３７５、ＡＴＣＣ２０９３７８、ＡＴＣＣ２０９３８４、ＡＴＣＣ２０９３９６、ＡＴＣＣ２０９４２０、ＡＴＣＣ２０９４８０、ＡＴＣＣ２０９２６５、ＡＴＣＣ２０９２５７、ＡＴＣＣ２０９２６２、ＡＴＣＣ２０９２５３、ＡＴＣＣ２０９４０２、ＡＴＣＣ２０９４０１、ＡＴＣＣ２０９３９７、ＡＴＣＣ２０９４００、ＡＴＣＣ２０９３８５、ＡＴＣＣ２０９３６７、ＡＴＣＣ２０９４３２、ＡＴＣＣ２０９２６３、ＡＴＣＣ２０９２５１、ＡＴＣＣ２０９２５５、ＡＴＣＣ２０９２５２、ＡＴＣＣ２０９３７３、ＡＴＣＣ２０９３７０、ＡＴＣＣ２０９５２３、ＡＴＣＣ２０９３７２、ＡＴＣＣ２０９３７４、ＡＴＣＣ２０９３７３、ＡＴＣＣ２０９３８２、ＡＴＣＣ２０９３８３、ＡＴＣＣ２０９４０３、ＡＴＣＣ２０９３９８、ＡＴＣＣ２０９３９９、ＡＴＣＣ２０９３９２、ＡＴＣＣ２０９３８７、ＡＴＣＣ２０９３８８、ＡＴＣＣ２０９３９４、ＡＴＣＣ２０９４２１、ＡＴＣＣ２０９３９３、ＡＴＣＣ２０９４１８、ＡＴＣＣ２０９４８５、ＡＴＣＣ２０９４８３、ＡＴＣＣ２０９４８２、ＡＴＣＣ２０９４９１、ＡＴＣＣ２０９４８１、ＡＴＣＣ２０９４３８、ＡＴＣＣ２０９９２７、ＡＴＣＣ２０９４３９、ＡＴＣＣ２０９４８９、ＡＴＣＣ２０９４３３、ＡＴＣＣ２０９４８８、ＡＴＣＣ２０９４３４、ＡＴＣＣ２０９３９５、ＡＴＣＣ２０９４８６、ＡＴＣＣ２０９４９０、ＡＴＣＣ

２０９４８４又はＡＴＣＣ ２０９３７１で寄託されたＤＮＡの全長コード化配列を含む単離された核酸。
【請求項５】
請求項１に記載の核酸を含んでなるベクター。
【請求項６】
ベクターで形質転換された宿主細胞に認識されるコントロール配列に作用可能に結合した請求項５に記載のベクター。
【請求項７】
請求項５に記載のベクターを含んでなる宿主細胞。
【請求項８】
前記細胞がＣＨＯ細胞である請求項７に記載の宿主細胞。
【請求項９】
前記細胞が大腸菌である請求項７に記載の宿主細胞。
【請求項１０】
前記細胞が酵母細胞である請求項７に記載の宿主細胞。
【請求項１１】
請求項７に記載の宿主細胞を、前記ＰＲＯポリペプチドの発現に適した条件下で培養し、細胞培養から前記ＰＲＯポリペプチドを回収することを含んでなるＰＲＯポリペプチドの製造方法。
【請求項１２】
Ｆｉｇ２（配列番号：２）、Ｆｉｇ４（配列番号：４）、Ｆｉｇ６（配列番号：１２）、Ｆｉｇ９（配列番号：１８）、Ｆｉｇ１１（配列番号：２３）、Ｆｉｇ１３（配列番号：２８）、Ｆｉｇ１５（配列番号：３４）、Ｆｉｇ１７（配列番号：３９）、Ｆｉｇ１９（配列番号：４９）、Ｆｉｇ２２（配列番号：５９）、Ｆｉｇ２４（配列番号：６４）、Ｆｉｇ２６（配列番号：６９）、Ｆｉｇ２８（配列番号：７１）、Ｆｉｇ３０（配列番号：７３）、Ｆｉｇ３２（配列番号：８４）、Ｆｉｇ３４（配列番号：９１）、Ｆｉｇ３６（配列番号：９６）、Ｆｉｇ３８（配列番号：１０４）、Ｆｉｇ４０（配列番号：１０９）、Ｆｉｇ４２（配列番号：１１４）、Ｆｉｇ４４（配列番号：１１９）、Ｆｉｇ４６（配列番号：１２７）、Ｆｉｇ４８（配列番号：１３２）、Ｆｉｇ５０（配列番号：１３７）、Ｆｉｇ５２（配列番号：１４２）、Ｆｉｇ５４（配列番号：１４８）、Ｆｉｇ５６（配列番号：１５３）、Ｆｉｇ５８（配列番号：１５９）、Ｆｉｇ６０（配列番号：１６４）、Ｆｉｇ６２（配列番号：１７０）、Ｆｉｇ６４（配列番号：１７５）、Ｆｉｇ６６（配列番号：１７７）、Ｆｉｇ６８（配列番号：１８５）、Ｆｉｇ７０（配列番号：１９０）、Ｆｉｇ７２（配列番号：１９５）、Ｆｉｇ７４（配列番号：２０１）、Ｆｉｇ７６（配列番号：２０７）、Ｆｉｇ７８（配列番号：２１３）、Ｆｉｇ８０（配列番号：２２１）、Ｆｉｇ８２（配列番号：２２７）、Ｆｉｇ８４（配列番号：２３６）、Ｆｉｇ８６（配列番号：２４５）、Ｆｉｇ８８（配列番号：２５０）、Ｆｉｇ９０（配列番号：２５５）、Ｆｉｇ９２（配列番号：２５７）、Ｆｉｇ９４（配列番号：２５９）、Ｆｉｇ９６（配列番号：２６１）、Ｆｉｇ９８（配列番号：２６３）、Ｆｉｇ１００（配列番号：２８５）、Ｆｉｇ１０２（配列番号：２９０）、Ｆｉｇ１０４（配列番号：２９２）、Ｆｉｇ１０６（配列番号：２９４）、Ｆｉｇ１０８（配列番号：３１０）、Ｆｉｇ１１０（配列番号：３１５）、Ｆｉｇ１１２（配列番号：３２０）、Ｆｉｇ１１４（配列番号：３２５）、Ｆｉｇ１１６（配列番号：３３２）、Ｆｉｇ１１８（配列番号：３３９）、Ｆｉｇ１２０（配列番号：３４１）及びＦｉｇ１２２（配列番号：３７７）に示したアミノ酸配列からなる群から選択されるアミノ酸配列に対して少なくとも８０％の配列同一性を有する単離された天然配列ＰＲＯポリペプチド。
【請求項１３】
登録番号ＡＴＣＣ ２０９２５８、ＡＴＣＣ ２０９２５６、ＡＴＣＣ ２０９２６４、ＡＴＣＣ ２０９２５０、ＡＴＣＣ ２０９３７５、ＡＴＣＣ ２０９３７８、ＡＴＣＣ ２０９３８４、ＡＴＣＣ ２０９３９６、ＡＴＣＣ ２０９４２０、ＡＴＣＣ ２０９４８０、

ＡＴＣＣ ２０９２６５、ＡＴＣＣ ２０９２５７、ＡＴＣＣ ２０９２６２、ＡＴＣＣ ２０９２５３、ＡＴＣＣ ２０９４０２、ＡＴＣＣ ２０９４０１、ＡＴＣＣ ２０９３９７、ＡＴＣＣ ２０９４００、ＡＴＣＣ ２０９３８５、ＡＴＣＣ ２０９３６７、ＡＴＣＣ ２０９４３２、ＡＴＣＣ ２０９２６３、ＡＴＣＣ ２０９２５１、ＡＴＣＣ ２０９２５５、ＡＴＣＣ ２０９２５２、ＡＴＣＣ ２０９３７３、ＡＴＣＣ ２０９３７０、ＡＴＣＣ ２０９５２３、ＡＴＣＣ ２０９３７２、ＡＴＣＣ ２０９３７４、ＡＴＣＣ ２０９３７３、ＡＴＣＣ ２０９３８２、ＡＴＣＣ ２０９３８３、ＡＴＣＣ ２０９４０３、ＡＴＣＣ ２０９３９８、ＡＴＣＣ ２０９３９９、ＡＴＣＣ ２０９３９２、ＡＴＣＣ ２０９３８７、ＡＴＣＣ ２０９３８８、ＡＴＣＣ ２０９３９４、ＡＴＣＣ ２０９４２１、ＡＴＣＣ ２０９３９３、ＡＴＣＣ ２０９４１８、ＡＴＣＣ ２０９４８５、ＡＴＣＣ ２０９４８３、ＡＴＣＣ ２０９４８２、ＡＴＣＣ ２０９４９１、ＡＴＣＣ ２０９４８１、ＡＴＣＣ ２０９４３８、ＡＴＣＣ ２０９９２７、ＡＴＣＣ ２０９４３９、ＡＴＣＣ ２０９４８９、ＡＴＣＣ ２０９４３３、ＡＴＣＣ ２０９４８８、ＡＴＣＣ ２０９４３４、ＡＴＣＣ ２０９３９５、ＡＴＣＣ ２０９４８６、ＡＴＣＣ ２０９４９０、ＡＴＣＣ ２０９４８４又はＡＴＣＣ ２０９３７１で寄託されたヌクレオチドによってコードされるアミノ酸配列に対して少なくとも８０％の配列同一性を有する単離されたＰＲＯポリペプチド。

【請求項１４】

異種アミノ酸配列に融合した請求項１２に記載のポリペプチドを含んでなるキメラ分子。

【請求項１５】

前記異種アミノ酸配列がエピトープタグ配列である請求項１４に記載のキメラ分子。

【請求項１６】

前記異種アミノ酸配列が免疫グロブリンのＦｃ領域である請求項１４に記載のキメラ分子。

【請求項１７】

請求項１２に記載のＰＲＯポリペプチドに特異的に結合する抗体。

【請求項１８】

前記抗体がモノクローナル抗体である請求項１７に記載の抗体。

【発明の詳細な説明】

【０００１】

（発明の分野）

本発明は、一般的に、新規なＤＮＡの同定及び単離、及び新規なポリペプチドの組換え生産に関する。

【０００２】

（発明の背景）

細胞外タンパク質は、特に、多細胞生物の形成、分化及び維持において重要な役割を担っている。多くの個々の細胞の運命、例えば増殖、泳動、分化又は他の細胞との相互作用は、典型的には、他の細胞及び／又は直近の環境から受け取る情報に支配される。この情報は、しばしば分泌ポリペプチド（例えば、分裂促進因子、生存因子、細胞障害性因子、分化因子、神経ペプチド、及びホルモン）により伝達され、これが、翻って多様な細胞レセプター又は膜結合タンパク質により受け取られ解釈される。これらの分泌ポリペプチド又はシグナル分子は、通常は細胞分泌経路を通過して、細胞外環境におけるその作用部位に到達する。

分泌タンパク質は、製薬、診断、バイオセンサー及びバイオリアクターを含む、様々な産業上の利用性を有している。血栓溶解剤、インターフェロン、インターロイキン、エリスロポエチン、コロニー刺激因子、及び種々の他のサイトカインのような、現在入手可能な大抵のタンパク質薬物は分泌タンパク質である。新規な未変性分泌タンパク質を同定する努力が産業界及び学術界の両方によってなされている。多くの努力が新規な分泌タンパク質のコード配列を同定するために哺乳類組換えＤＮＡライブラリーのスクリーニングに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅ

ｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．９３；７１０８－７１１３（１９９６）；米国特許第５５３６６３７号を参照されたい］。
膜結合タンパク質及びレセプターは、多細胞生物の形成、分化及び維持において重要な役割を担っている。多くの個々の細胞の運命、例えば増殖、遊走、分化又は他の細胞との相互作用は、典型的には他の細胞及び／又は直近の環境から受け取られる情報に支配される。この情報は、しばしば分泌ポリペプチド（例えば、分裂促進因子、生存因子、細胞障害性因子、分化因子、神経ペプチド、及びホルモン）により伝達され、これが次に多様な細胞レセプター又は膜結合タンパク質により受け取られ解釈される。このような膜結合タンパク質及び細胞レセプターは、これらに限定されるものではないが、サイトカインレセプター、レセプターキナーゼ、レセプターホスファターゼ、細胞－細胞間相互作用に関与するレセプター、及びセレクチン及びインテグリンのような細胞接着分子を含む。例えば、細胞の増殖及び分化を調節するシグナルの伝達は、様々な細胞タンパク質のリン酸化により部分的に調節される。そのプロセスを触媒する酵素であるプロテインチロシンキナーゼはまた成長因子レセプターとしても作用する。具体例には、繊維芽細胞増殖因子及び神経成長因子レセプターが含まれる。
膜結合タンパク質及びレセプター分子は、製薬及び診断薬を含む、様々な産業上の利用性を有している。例えば、レセプターイムノアドヘシンはレセプター－リガンド間相互作用を阻止する治療薬として使用することができる。膜結合タンパク質はまた、関連するレセプター／リガンド間相互作用の可能性のあるペプチド又は小分子インヒビターをスクリーニングするために使用することもできる。
新規な天然のレセプター又は膜結合タンパク質を同定するための努力が産業界と学術界の双方によってなされている。多くの努力が、新規なレセプター又は膜結合タンパク質のコード配列を同定するために、哺乳動物の組換えＤＮＡライブラリーのスクリーニングに注がれている。
【０００３】
１．ＰＲＯ２１１及びＰＲＯ２１７
上皮成長因子（ＥＧＦ）は上皮細胞及び線維芽細胞を含む様々なタイプの細胞の増殖を刺激する一般的な分裂促進因子である。ＥＧＦはＥＧＦレセプター（ＥＧＦＲ）に結合してそれを活性化させ、それが細胞内シグナル伝達及びそれに引き続く効果を開始させる。ＥＧＦＲは中枢神経系（ＣＮＳ）の他の領域に加えて、大脳皮質、小脳、及び海馬のニューロン中で発現される。また、ＥＧＦはＣＮＳの様々な領域において発現される。従って、ＥＧＦは有糸分裂細胞に対してばかりでなく分裂終了細胞に対しても作用する。実際、多くの研究において、ＥＧＦがＣＮＳ中の様々なタイプのニューロンに神経栄養性及び神経調節性効果を有していることが示されている。例えば、ＥＧＦは培養された大脳皮質及び小脳ニューロンに直接作用し、神経突起成長と生存を亢進する。他方、ＥＧＦはまた中隔コリン作用性及び中脳ドーパミン作用性ニューロンを含む他の細胞型にもグリア細胞を通して間接的に作用する。ＣＮＳ中のニューロンに対するＥＧＦの効果の証拠は蓄積されつつあるが、作用の機構は本質的に知られていないままである。有糸分裂細胞におけるＥＧＦ－－誘発シグナル伝達は細胞分裂終了ニューロンにおいてよりもよりよく理解されている。クローン化されたクロム親和細胞腫ＰＣ１２細胞と培養された大脳皮質ニューロンの研究により、ＥＧＦ誘発神経栄養性作用がＥＧＦに応答してのＥＧＦＲとマイトジェン活性化プロテインキナーゼ（ＭＡＰＫ）の持続性活性化により媒介されることが示唆されている。持続性細胞内シグナル伝達はＥＧＦＲ下方制御の減少速度と相関し、これはＥＧＦに対するニューロン細胞の応答を決定するかもしれない。ＥＧＦは有糸細胞分裂細胞及び細胞分裂終了ニューロンを含む細胞の様々なタイプに作用する多分化能成長因子である可能性が高い。
ＥＧＦは、腎臓、膵臓、甲状腺、脳下垂体、神経系及び胃腸系の唾液腺及びブルンアー腺によりつくられ、体液、例えば唾液、血液、脳脊髄液（ＣＳＦ）、尿、羊水、前立腺液、膵液、及び母乳に見いだされる。Ｐｌａｔａ－Ｓａｌａｍａｎ，Ｐｅｐｔｉｄｅｓ １２：６５３－６６３（１９９１）。

ＥＧＦは、内因性チロシンキナーゼを含むその膜特異的レセプターにより媒介される。Ｓｔｏｓｃｈｅｃｋ等，Ｊ．Ｃｅｌｌ Ｂｉｏｃｈｅｍ．３１：１３５－１５２（１９８６）。ＥＧＦは内因性チロシンキナーゼを活性化する膜貫通シグナルを誘発するそのレセプターの細胞外部分に結合することにより機能すると信じられている。

ＥＧＦ様ドメインの精製と配列分析により、交差結合して３つのペプチドループをつくりだす６個の保存システイン残基の存在が明らかになった。Ｓａｖａｇｅ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２４８：７６６９－７６７２（１９７９）。幾つかの他のペプチドは、同じ一般化されたモチーフ$X_nCX_7CX_{4/5}CX_{10}CXCX_5GX_2CX_n$（ここで、Ｘは任意の非システインアミノ酸を表し、ｎは可変の繰り返し数である）を共有するＥＧＦレセプターと反応しうることが今は一般的に知られている。このモチーフを有する非単離ペプチドは、ＴＧＦ－α、アンフィレギュリン、シュワン細胞腫由来成長因子（ＳＤＧＦ）、ヘパリン結合ＥＧＦ様成長因子及びある種のウィルス性コード化ペプチド（例えばワクシニアウィルス、Ｒｅｉｓｎｅｒ，Ｎａｔｕｒｅ ３１３：８０１－８０３（１９８５），ショープ線維腫ウィルス、Ｃｈａｎｇ等，Ｍｏｌ Ｃｅｌｌ Ｂｉｏｌ．７：５３５－５４０（１９８７），伝染性軟属腫、Ｐｏｒｔｅｒ及びＡｒｃｈａｒｄ，Ｊ．Ｇｅｎ．Ｖｉｒｏｌ．６８：６７３－６８２（１９８７）、及び粘液腫ウィルス、Ｕｐｔｏｎ等，Ｊ．Ｖｉｒｏｌ．６１：１２７１－１２７５（１９８７）、Ｐｒｉｇｅｎｔ及びＬｅｍｏｉｎｅ，Ｐｒｏｇ．Ｇｒｏｗｔｈ Ｆａｃｔｏｒ Ｒｅｓ．．４：１－２４（１９９２）を含む。

ＥＧＦ様ドメインは成長因子に制限されず、細胞接着、タンパク質－タンパク質相互作用と発達において興味有る性質を有している様々な細胞表面及び細胞外タンパク質において観察されている。Ｌａｕｒｅｎｃｅ及びＧｕｓｔｅｒｓｏｎ，Ｔｕｍｏｒ Ｂｉｏｌ．１１：２２９－２６１（１９９０）。これらのタンパク質は血液凝固因子（因子ＶＩ、ＩＸ、Ｘ、ＸＩＩ、プロテインＣ、プロテインＳ、プロテインＺ、組織プラスミノゲンアクチベーター、ウロキナーゼ）、細胞外マトリックス成分（ラミニン、サイトタクチン、エンタクチン）、細胞表面レセプター（ＬＤＬレセプター、トロンボモジュリンレセプター）及び免疫関連タンパク質（補体Ｃ１ｒ、ウロモジュリン）を含む。

更により興味深いことには、ＥＧＦ様前駆体の一般的構造パターンは下等生物並びに哺乳動物細胞を通して保存されている。発達上の優位性を持つ多くの遺伝子はＥＧＦ様繰り返しを持つ無脊椎動物において同定されている。例えば、ショウジョウバエのｎｏｔｃｈ遺伝子は、ＥＧＦとの相同性を示す３６の直列に配置された４０のアミノ酸反復をコードしている。Ｗｈａｒｔｏｎ等，Ｃｅｌｌ ４３：５５７－５８１（１９８５）。ヒドロパシープロットは推定膜貫通ドメインを示し、そのＥＤＦ関連配列は膜の細胞外側に位置している。ＥＧＦ様反復を持つ他のホメオティック遺伝子は、Ｎｏｔｃｈに基づくプローブを使用して同定されたデルタ、９５Ｆ及び５ＺＤ、及び２つの特定された細胞の間に伝達される発生的シグナルに対する推定上レセプターをコードする線虫遺伝子Ｌｉｎ－１２を含む。

特に、ＥＧＦは胃腸粘膜の保存と維持及び急性及び慢性粘膜病状の修復、Ｋｏｎｔｕｒｅｋ等，Ｅｕｒ．Ｊ．Ｇａｓｔｒｏｅｎｔｅｒｏｌ Ｈｅｐａｔｏｌ．７（１０），９３３－３７（１９９５）で、壊死性腸炎、ゾリンジャー・エリソン症候群、胃腸潰瘍形成及び先天性微絨毛萎縮症、Ｇｕｇｌｉｅｔｔａ及び Ｓｕｌｌｉｖａｎ，Ｅｕｒ．Ｊ．Ｇａｓｔｒｏｅｎｔｅｒｏｌ Ｈｅｐａｔｏｌ，７（１０），９４５－５０（１９９５）に可能性を有していることが示された。また、ＥＧＦは、毛包分化；ｄｕ Ｃｒｏｓ，Ｊ．Ｉｎｖｅｓｔ．Ｄｅｒｍａｔｏｌ．１０１（１ Ｓｕｐｐｌ．），１０６Ｓ－１１３Ｓ（１９９３），Ｈｉｌｌｉｅｒ，Ｃｌｉｎ．Ｅｎｄｏｃｒｉｎｏｌ．３３（４），４２７－２８（１９９０）；腎機能、Ｈａｍｍ等，Ｓｅｍｉｎ．Ｎｅｐｈｒｏｌ．１３（１）：１０９－１５（１９９３），Ｈａｒｒｉｓ，Ａｍ．Ｊ．Ｋｉｄｎｅｙ Ｄｉｓ．１７（６）：６２７－３０（１９９１）；涙液、ｖａｎ Ｓｅｔｔｅｎ，Ｉｎｔ．Ｏｐｈｔｈａｌｍｏｌ １５（６）；３５９－６２（１９９１）；ビタミンＫ媒介血液凝固 Ｓｔｅｎｆｌｏ等，Ｂｌｏｏｄ ７８（７）：１６３７－

５１（１９９１）に関与していた。ＥＧＦはまた異常なケラチノサイト分化により特徴づけられる様々な皮膚疾患、例えば乾癬、上皮ガン、例えば肺の扁平上皮ガン、外陰部の類表皮ガン及び神経膠腫に関連している。Ｋｉｎｇ等，Ａｍ．Ｊ．Ｍｅｄ．Ｓｃｉ．２９６：１５４－１５８（１９９８）。

非常に興味深いのは、成長因子シグナル伝達経路における遺伝的変更が発育異常とガンを含む慢性疾患に密接に関連しているという増大している証拠である。Ａａｒｏｎｓｏｎ，Ｓｃｉｅｎｃｅ２５４：１１４６－１１５３（１９９１）。例えば、ＥＧＦレセプタータンパク質と密接な構造類似性を持つプロトガン遺伝子ｃ－ｅｒｂ－２（またＨＥＲ－２としても知られている）がヒト乳ガンにおいて過剰発現されている。Ｋｉｎｇ等，Ｓｃｉｅｎｃｅ２２９：９７４－９７６（１９８５）；Ｇｕｌｌｉｃｋ，Ｈｏｒｍｏｎｅｓ ａｎｄ ｔｈｅｉｒ ａｃｔｉｏｎｓ，Ｃｏｏｋｅ等，ｅｄｓ，Ａｍｓｔｅｒｄａｍ，Ｅｌｓｅｖｉｅｒ，ｐｐ３４９－３６０（１９８６）。

我々はここにＥＧＦに対して相同性を有する新規なポリペプチドの同定と特徴付けを記載し、そのポリペプチドはここでＰＲＯ２１１及びＰＲＯ２１７と命名する。

【０００４】

２．ＰＲＯ２３０

腎炎は腎臓の構造と正常な機能に影響を及ぼす腎臓の炎症により特徴付けられる病状である。この病状は慢性的の場合も急性の場合もあり、一般には感染、変性プロセス又は血管疾患により引き起こされる。全ての場合において、腎炎を持つ患者が病状の治療を始めることができるように早期の検出が望ましい。

腎炎を検出するためのアプローチ法は自然に関連する抗原とそれに対する抗体を決定することである。ウサギにおいて、尿細管間質性腎炎抗原（ＴＩＮ－ａｇ）が、Ｎｅｌｓｏｎ，Ｔ．Ｒ．等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２７０（２７）：１６２６５－７０（１９９５ ７月）に報告されている（ＧＥＮＢＡＮＫ／Ｕ２４２７０）。この研究は、システインプロテイナーゼファミリーのタンパク質のメンバーであるヒトプレプロカテプシンＢと３０％相同性を示すカルボキシ末端領域を有する予想アミノ酸配列を有する基底膜糖タンパク質であるということを報告している。ウサギＴＩＮ－ａｇがラミニンＡ及びＳ鎖、Ｉ型コラーゲンのアルファ１鎖、フォンウィルブランド因子及びムチンと相同性を共有する上皮成長因子様モチーフを含むアミノ末端領域にドメインを有しており、構造的及び機能的類似性を示していることがまた報告されている。マウスにおいてもまた研究がなされている。しかし、ヒトにおける尿細管間質性腎炎抗原を同定して、腎炎の早期検出方法と治療の開発を補助することが望ましい。

尿細管間質性腎炎抗原と相同性を有するタンパク質は医学及び産業界にとって特に興味深い。しばしば、互いに相同性を有するタンパク質は類似した機能を有している。相同性を有するタンパク質がいつ類似した機能を有していないことはまた興味深く、ある構造的モチーフが機能以外の情報、例えば機能の局所性を特定することを示している。

我々はここに尿細管間質性腎炎抗原に対して相同性を有する、ここでＰＲＯ２３０と命名した新規なポリペプチドの同定と特徴付けを記載する。

【０００５】

３．ＰＲＯ２３２

幹細胞は（ａ）増殖、（ｂ）自己維持、（ｃ）多数の分化した機能性子孫の生産、（ｄ）損傷後の組織の再生及び／又は（ｅ）これらの選択肢の使用の柔軟性が可能な未分化細胞である。幹細胞はしばしば幹細胞を同定するために使用できる細胞特異的マーカーとなりうる細胞表面抗原を発現し、それによって特異的幹細胞集団を同定し単離する手段を提供する。

異なった幹細胞集団を所有することにより複数の重要な応用を可能にする。例えば、特異的幹細胞集団を所有していることは、その増殖と分化に関与している成長因子と他のタンパク質の同定を可能にする。また、（１）幹細胞の特定の系列への提供の早期の段階、（２）そのような提供の防止、及び（３）幹細胞増殖の負の制御に関連するまだ発見されていないタンパク質がありえ、その全ては幹細胞集団を所有すれば同定できる。更に、幹細

胞は、幹細胞が移植された個人の健康を挿入された遺伝子が促進する遺伝子療法に対して重要で理想的な標的である。最後に、幹細胞は器官又は組織の移植術、例えば肝臓再生及び皮膚移植において重要な役割を果たしうる。

様々な異なった応用における幹細胞の重要性が分かっているので、新規な天然の幹細胞抗原タンパク質を同定し、幹細胞集団を同定するため並びに細胞増殖及び分化において幹細胞が果たす機能的役割の洞察手段を提供するための特異的細胞表面マーカーを提供するために産業界と学術界の両方により努力が現在なされている。我々はここに幹細胞抗原に対して相同性を有する新規なポリペプチドの同定と特徴付けを記載し、そのポリペプチドをここでＰＲＯ２３２ポリペプチドと命名する。

【０００６】

４．ＰＲＯ１８７

成長因子は、単独で又は協力して細胞成長又は増殖を亢進する分子シグナル又はメディエータである。しかし、成長因子への発現時には成長のみ以外の細胞性反応がある。その結果、成長因子は多機能性で強力な細胞レギュレーターとしてよりよく特徴付けられる。その生物学的効果は増殖、走化性及び細胞外マトリックス生産の刺激を含む。成長因子は刺激性及び阻害成功化の双方を有しうる。例えば、トランスフォーミング成長因子（ＴＧＦ－α）は高度に多面的で、ある細胞、特に結合組織では増殖を刺激する一方、他のもの、例えばリンパ球及び上皮細胞では増殖の強力な阻害剤である。

成長因子による成長刺激又は阻害の生理学的効果は標的組織の発達と分化の状態に依存する。古典的な内分泌分子による局所的細胞調節はオートクライン（同じ細胞）、ジャクスタクリン（近くの細胞）、及びパラクリン（隣接細胞）経路の理解を含む。ペプチド成長因子は複雑な生物学的言語の要素であり、細胞間情報交換の基礎を提供する。それらは細胞が互いに情報を運ぶことを可能にし、細胞間の相互作用を媒介し、遺伝子発現を変化させる。これらの多機能及び多面的因子の効果は他のペプチドの有無に依存している。

ＦＧＦ－８は、正常な二倍体線維芽細胞と樹立株化細胞の双方に対するヘパリン結合性の強力なマイトジェンのファミリーである線維芽細胞成長因子（ＦＧＦｓ）のメンバーである。Ｇｏｓｐｏｄａｒｏｗｉｃｚ等，（１９８４），Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８１：６９６３。ＦＧＦファミリーはとりわけ、酸性ＦＧＦ（ＦＧＦ－１）、塩基性ＦＧＦ（ＦＧＦ－２）、ＩＮＴ－２（ＦＧＦ－３）、Ｋ－ＦＧＦ／ＨＳＴ（ＦＧＦ－４）、ＦＧＦ－５、ＦＧＦ－６、ＫＧＦ（ＦＧＦ－７）、ＡＩＧＦ（ＦＧＦ－８）を含む。全てのＦＧＦｓは２つの保存されたシステイン残基を持ち、アミノ酸レベルで３０－５０％の配列相同性を共有する。これらの因子は、顆粒膜細胞、副腎皮質細胞、軟骨細胞、筋芽細胞、角膜及び血管内皮細胞（ウシ又はヒト）、血管平滑筋細胞、水晶体、網膜及び前立腺上皮細胞、オリゴデンドロサイト、アストロサイト、クロンドサイト（ｃｈｒｏｎｄｏｃｙｔｅｓ）、筋芽細胞及び骨芽細胞を含む広範囲の正常な二倍体中胚葉由来及び神経冠由来細胞に対して分裂促進的である。

線維芽細胞成長因子はまた非分裂促進的な形で多数の細胞型を刺激することができる。これらの活動は、線維芽細胞成長因子はまた創傷領域への細胞遊走の促進（走化作用）、新しい血管形成の開始（血管新生）、神経再生と生存の変調（ニューロトロフィズム（ｎｅｕｒｏｔｒｏｐｈｉｓｍ））、内分泌機能の変調、及び特異的細胞性タンパク質発現、細胞外マトリックス生産及び細胞生存の刺激又は抑制を含む。Ｂａｉｒｄ ＆ Ｂｏｈｌｅｎ，Ｈａｎｄｂｏｏｋ ｏｆ Ｅｘｐ．Ｐｈａｒｍａｃｏｌ．９５（１）：３６９－４１８，Ｓｐｒｉｎｇｅｒ，（１９９０）。これらの性質は、創傷治癒、神経修復、側枝血管形成等々を加速させる治療的アプローチにおいて線維芽細胞成長因子を使用するための基礎を提供する。例えば、線維芽細胞成長因子が心臓疾患及び手術において心筋損傷を最小にすることが示唆されている（Ｕ．Ｓ．Ｐ．４，３７８，３４７）。

アンドロゲン誘発成長因子（ＡＩＧＦ）としても知られているＥＧＦ－８はＦＧＦファミリーの他のメンバーと３０－４０％の配列相同性を共有する２１５のアミノ酸タンパク質である。ＦＧＦ－８はマウス乳ガン株化細胞ＳＣ３においてアンドロゲン調節及び誘導下にあることが提案されている。Ｔａｎａｋａ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓ

ｃｉ．ＵＳＡ８９：８９２８－８９３２（１９９２）；Ｓａｔｏ等，Ｊ．Ｓｔｅｒｏｉｄ Ｂｉｏｃｈｅｍ．Ｍｏｌｅｃ．Ｂｉｏｌ．４７：９１－９８（１９９３）。その結果、ＦＧＦ－８は、アンドロゲン応答性器官であることが知られている前立腺において局所的な役割を有しうる。ＦＧＦ－８がまたＮＩＨ－３Ｔ３線維芽細胞内に形質移入されるとトランスフォーミング活性を示すので発ガン性でありうる。Ｋｏｕｈａｒａ等，Ｏｎｃｏｇｅｎｅ ９ ４５５－４６２（１９９４）。ＦＧＦ－８は心臓、脳、肺、腎臓、精巣、前立腺及び卵巣において検出されている一方、発現はまた外因性アンドロゲンのない場合にもまた検出された。Ｓｃｈｍｉｔｔ等，Ｊ．Ｓｔｅｒｏｉｄ Ｂｉｏｃｈｅｍ．Ｍｏｌ．Ｂｉｏｌ．５７（３－４）：１７３－７８（１９９６）。

ＦＧＦ－８は、マウス胚形成の様々な段階で発現されるという性質を幾つかの他のＦＧＦｓと共有し、これが、様々なＥＧＦｓが分化と胚形成において多重的でおそらくは協調的な役割を有するという理論を裏付けている。更に、ＦＧＦ－８はまた乳房腫瘍形成の過程においてＷｎｔ－１と共同するプロトオンコジーンとして同定されている（Ｓｈａｃｋｌｅｆｏｒｄ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９０，７４０－７４４（１９９３）；Ｈｅｉｋｉｎｈｅｉｍｏ等，Ｍｅｃｈ．Ｄｅｖ．４８：１２９－１３８（１９９４））。

他のＦＧＦｓとは対照的に、ＦＧＦ－８は一次転写物の選択的スプライシングの結果として、３種のタンパク質アイソフォームとして存在している、上掲のＴａｎａｋａ等。ＦＧＦ－８の正常な成人の発現は弱く生殖腺組織に限られているが、ノーザンブロット分析は、ＦＧＦ－８ｍＲＮＡがマウス妊娠期間の１０日から１２日存在していることを示しており、ＦＧＦ－８が正常な発育に重要であることを示唆している。Ｈｅｉｋｉｎｈｅｉｍｏ等，Ｍｅｃｈ Ｄｅｖ．４８（２）：１２９－３８（１９９４）。妊娠の８から１６日の間の更なるインサイツハイブリッド形成アッセイは、第一気管支アーチの表面外胚葉、前線鼻突起、前脳及び中脳－後脳移行部において初期の発現を示した。１０－１２日で、ＦＧＦ－８は前肢及び後肢芽の表面外胚葉、鼻及び上咽頭、漏斗及び終脳、間脳及び後脳に発現された。発現は妊娠の１３日で発育中の後肢において継続するがその後は検出できない。この結果は、ＦＧＦ－８が胚形成において独特な時間的かつ空間的パターンを有していることを示唆しており、嚢胚形成後の胚における外胚葉分化の多重領域におけるこの成長因子の役割を示唆している。

ここで我々は、ＦＧＦ－８と相同性を有する新規なポリペプチドの同定を記載し、これらポリペプチドをここでＰＲＯ１８７ポリペプチドと命名した。

【０００７】

５．ＰＲＯ２６５

タンパク質－タンパク質相互作用はレセプター及び抗原複合体及びシグナル伝達機構に関連するものを含む。タンパク質－タンパク質相互作用の基となる構造的及び機能的機構についてよく知られているように、タンパク質－タンパク質相互作用は、タンパク質－タンパク質相互作用の特定の結果を調節するように更に容易に操作することができる。従って、タンパク質－タンパク質相互作用の基となる機構は科学及び医学界にとって興味深い。ロイシンリッチ反復を含む全てのタンパク質はタンパク質－タンパク質相互作用に関与していると考えられている。ロイシンリッチ反復は多様な機能を持つ多くのタンパク質と細胞位置に存在する短い配列モチーフである。リボ核酸インヒビタータンパク質の結晶構造から、ロイシンリッチ反復はβ－α構造単位に対応することが明らかになった。これらの単位は、一面が溶媒に暴露された平行なβシートを形成するように配置され、タンパク質は珍しい非球状の形状を獲得する。これらの二つの特徴はロイシンリッチ反復を含むタンパク質のタンパク質結合機能の原因であることが示されている。Ｋｏｂｅ及びＤｅｉｓｅｎｈｏｆｅｒ，Ｔｒｅｎｄｓ Ｂｉｏｃｈｅｍ．Ｓｃｉ．，１９（１０）：４１５－４２１（Ｏｃｔ．１９９４）を参照されたい。

個体発生の間にコラーゲン原線維を配向し指令する組織オーガナイザーとなり、創傷治癒、組織修復、及び腫瘍ストローマ形成のような病理プロセスに関与するロイシンリッチプロテオグリカンについて研究が報告されている。Ｉｏｚｚｏ，Ｒ．Ｖ．，Ｃｒｉｔ．

Ｒｅｖ．Ｂｉｏｃｈｅｍ．Ｍｏｌ．Ｂｉｏｌ．，３２（２）：１４１－１７４（１９９７）。創傷治癒、組織修復におけるロイシンリッチタンパク質の関与を示す他の研究は、Ｄｅ Ｌａ Ｓａｌｌｅ，Ｃ．等，Ｖｏｕｖ．Ｒｅｖ．Ｆｒ．Ｈｅｍａｔｏｌ．（Ｇｅｒｍａｎｙ），３７（４）：２１５－２２２（１９９５）であり、出血疾患ベルナール－スーリエ症候群に伴う複合体におけるロイシンリッチモチーフの突然変異を報告しており、Ｃｈｌｅｍｅｔｓｏｎ，Ｋ．Ｊ．，Ｔｈｒｏｍｂ．Ｈａｅｍｏｓｔ．（Ｇｅｒｍａｎｙ），７４（１）：１１１－１１６（Ｊｕｌｙ １９９５）は、血小板がロイシンリッチ反復を有していることを報告している。ロイシンリッチ反復を有することが報告されている特に興味深い他のタンパク質は、アルツハイマー病のような神経変性疾患、パーキンソン病のような神経損傷の治療において、及びガンの診断に対して有用であることが報告されているＳＬＩＴタンパク質である。エール大学のＡｒｔａｖａｎｉｓｔｓａｋｏｎａｓ，Ｓ．及びＲｏｔｈｂｅｒｇ，Ｊ．Ｍ．のＷＯ９２１０５１８－Ａ１を参照されたい。ロイシンリッチ反復を有するタンパク質の生物学的機能を報告している他の研究には：Ｔａｙａｒ，Ｎ．等，Ｍｏｌ．Ｃｅｌｌ Ｅｎｄｏｃｒｉｎｏｌ．，（Ｉｒｅｌａｎｄ），１２５（１－２）：６５－７０（Ｄｅｃ．１９９６）（ゴナドトロピンレセプター関連）；Ｍｉｕｒａ，Ｙ．等，Ｎｉｐｐｏｎ Ｒｉｎｓｈｏ（Ｊａｐａｎ），５４（７）：１７８４－１７８９（Ｊｕｌｙ １９９６）（アポトーシス関連）；Ｈａｒｒｉｓ，Ｐ．Ｃ．等，Ｊ．Ａｍ．Ｓｏｃ．Ｎｅｐｈｒｏｌ．，６（４）：１１２５－１１３３（Ｏｃｔ．１９９５）（腎疾患関連）；及びＬａ Ｊｏｌｌａ Ｃａｎｃｅｒ Ｒｅｓｅａｒｃｈ ＦｏｕｎｄａｔｉｏｎのＲｕｏｓｌａｈｔｉ，Ｅ．Ｉ．等のＷＯ９１１０７２７－Ａ（トランスフォーミング成長因子βに結合するデコリンの、ガンの治療、創傷治癒及び瘢痕化への関与）が含まれる。また興味深いのはフィブロモジュリンと皮膚瘢痕化を防止又は低減するためのその使用である。フィブロモジュリンの研究はＲｏｕｓｌａｈｔｉ等の米国特許番号５，６５４，２７０に見いだされる。

よって、タンパク質－タンパク質相互作用をよりよく理解するためにロイシンリッチ反復を有する新規なタンパク質を同定するための努力が産業界と学術界の両方によってなされている。特に興味があるものはロイシンリッチ反復とフィブロモジュリン、ＳＬＩＴタンパク質及び血小板糖タンパク質Ｖのようなロイシンリッチ反復を有する既知のタンパク質と相同性を有するタンパク質である。多くの努力は、ロイシンリッチ反復を有する新規な分泌及び膜貫通タンパク質についてのコード化配列を同定するための哺乳動物組換えＤＮＡライブラリのスクリーニングに集中している。ここで我々は、ここにＰＲＯ２６５と命名する、フィブロモジュリンと相同性を有する新規なポリペプチドの同定及び特徴づけについて記載する。

【０００８】

６．ＰＲＯ２１９

ヒトマトリリン－２ポリペプチドはフォ・ンウィルブランド因子Ａ型様モジュールスーパーファミリーのメンバーである。フォン・ウィルブランド因子は止血の維持に重要な役割を果たすタンパク質である。より詳細には、フォンウィルブランド因子はＶＩＩＩ因子と相互作用して複合体を形成する能力を介して血管損傷の部位で血小板－血管壁相互作用に関与することが知られているタンパク質である。血液中にフォン・ウィルブランド因子がないと血管損傷の部位での血管壁への血小板の接着を防止する血液血小板の異常を引き起こす。その結果はフォン・ウィルブランド病を含むｂｒｕｓｉｎｇ、鼻血、腸出血、性向である。

血液凝固因子の生理学的な重要性が分かっているので、凝固プロセスに関連している新規な天然のタンパク質を同定するために産業界と学術界の両方により努力が現在なされている。我々はここにヒトマトリン－２前駆体ポリペプチドと相同性を有する新規な全長ポリペプチドの同定を記載する。

【０００９】

７．ＰＲＯ２４６

細胞表面タンパク質ＨＣＡＲはアデノウィルスの亜群Ｃ及びコクサッキーウィルスの亜群

Ｂのためのレセプターとして作用する膜結合タンパク質である。よって、ＨＣＡＲは細胞表面上のＨＣＡＲレセプターの存在がウィルス粒子に対する結合部位を提供する点において細胞のウィルス感染を媒介する手段を提供しえ、よってウィルス感染を容易にする。
膜結合タンパク質及び特にウィルスに対する細胞表面レセプターを提供するものの生理学的な重要性に鑑みて、新規な膜結合レセプタータンパク質を同定するための努力が産業界と学術界の双方によって現在なされている。これらの努力の多くは、新規なレセプタータンパク質に対するコード化配列を同定するために哺乳動物組換えＤＮＡライブラリのスクリーニングに注がれている。我々はここに細胞表面タンパク質ＨＣＡＲとＡ３３及びガン胎児性抗原を含む様々な腫瘍抗原に対して相同性を有する新規な膜結合ポリペプチド（ここでＰＲＯ２４６と命名）を記載し、ここでこのポリペプチドは新規な細胞表面ウィルスレセプター又は腫瘍抗原でありうる。

【０００１０】

８．ＰＲＯ２２８

多数の既知の７種の膜貫通タンパク質があり、このファミリーの中にはＣＤ９７とＥＭＲ１を含む群がある。ＣＤ９７は細胞性リガンド、ＣＤ５５、ＤＡＦを有する７スパン膜貫通レセプターである。Ｈａｍａｎｎ等，Ｊ．Ｅｘｐ．Ｍｅｄ．（Ｕ．Ｓ．），１８４（３）：１１８９（１９９６）。また、ＣＤ９７はヒト甲状腺ガンにおける脱分化マーカーであり、炎症に関連していると報告されている。Ａｕｓｔ等，Ｃａｎｃｅｒ Ｒｅｓ．（Ｕ．Ｓ．），５７（９）：１７９８（１９９７）；Ｇｒａｙ等，Ｊ．Ｉｍｍｕｎｏｌ．（Ｕ．Ｓ．），１５７（１２）：５４３８（１９９６）。ＣＤ９７はまたセクレチンレセプタースーパーファミリーに関連していると報告されているが、そのファミリーの既知のメンバーとは異なり、ＣＤ９７とＥＭＲ１はＮ末端に幾つかのＥＧＦドメインを有する細胞外領域を伸展している。Ｈａｍａｎｎ等，Ｇｅｎｏｍｉｃｓ，３２（１）：１４４（１９９６）；Ｈａｒｍａｎｎ等，Ｊ．Ｉｍｍｕｎｏｌ．，１５５（４）：１９４２（１９９５）。ＥＭＲ１は更にＬｉｎ等，Ｇｅｎｏｍｉｃｓ，４１（３）：３０１（１９９７）及びＢａｕｄ等，Ｇｅｎｏｍｉｃｓ，２６（２）：３３４（１９９５）に記載されている。ＣＤ９７とＥＭＲ１はセクレチンレセプターと関連しているように思われるが、Ｇプロテイン結合レセプターのセクレチンファミリーの既知のメンバーは、α－ラトロキシンレセプター、ラトロフィリンを含み、これはカルシウム依存性でニューロン組織中に豊富にあると記載されている。Ｌｅｌｉａｎｏｖａ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２７２（３４），２１５０４（１９９７）；Ｄａｖｌｅｔｏｖ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．（Ｕ．Ｓ．），２７１（３８）：２３２３９（１９９６）。セクレチンレセプタースーパーファミリーのメンバーとセクレチンレセプタースーパーファミリー、あるいはＣＲＦ及びカルシトニンレセプターに関連している非メンバーの両方が興味深い。特に、既知のタンパク質との相同性により同定されるこれらのファミリーの新しいメンバーは興味深い。
新規な膜結合レセプタータンパク質、特にＣＤ９７とＥＭＲ１がメンバーである７つの膜貫通タンパク質のファミリーに属しうるＥＧＦ反復と大きなＮ末端を持つ膜貫通タンパク質を同定するための努力が産業界と学術界の双方によって現在なされている。我々はここにＰＲＯ２２８ポリペプチドと命名した、ＣＤ９７とＥＭＲ１に対して相同性を有する新規なポリペプチドの同定と特徴付けを記載する。

【００１１】

９．ＰＲＯ５３３

成長因子は、単独で又は協力して細胞成長又は増殖を亢進する分子シグナル又はメディエータである。しかし、成長因子への発現時には成長のみ以外の細胞性反応がある。その結果、成長因子は多機能性で強力な細胞レギュレーターとしてよりよく特徴付けられる。その生物学的効果は増殖、走化性及び細胞外マトリックス生産の刺激を含む。成長因子は刺激性及び阻害成功化の双方を有しうる。例えば、トランスフォーミング成長因子（ＴＧＦ－β）は高度に多面的で、ある細胞、特に結合組織では増殖を刺激する一方、他のもの、例えばリンパ球及び上皮細胞では増殖の強力な阻害剤である。

成長因子による成長刺激又は阻害の生理学的効果は標的組織の発達と分化の状態に依存する。古典的な内分泌分子による局所的細胞調節はオートクライン（同じ細胞）、ジャクスタクリン（近くの細胞）、及びパラクリン（隣接細胞）経路の理解を含む。ペプチド成長因子は複雑な生物学的言語の要素であり、細胞間情報交換の基礎を提供する。それらは細胞が互いに情報を運ぶことを可能にし、細胞間の相互作用を媒介し、遺伝子発現を変化させる。これらの多機能及び多面的因子の効果は他のペプチドの有無に依存している。

繊維芽細胞成長因子（ＦＧＦｓ）は、正常な二倍体線維芽細胞と樹立株化細胞の双方に対するヘパリン結合性の強力なマイトジェンのファミリーである。Ｇｏｓｐｏｄａｒｏｗｉｃｚ等，（１９８４），Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ８１：６９６３。ＦＧＦファミリーはとりわけ、酸性ＦＧＦ（ＦＧＦ－１）、塩基性ＦＧＦ（ＦＧＦ－２）、ＩＮＴ－２（ＦＧＦ－３）、Ｋ－ＦＧＦ／ＨＳＴ（ＦＧＦ－４）、ＦＧＦ－５、ＦＧＦ－６、ＫＧＦ（ＦＧＦ－７）、ＡＩＧＦ（ＦＧＦ－８）を含む。全てのＦＧＦｓは２つの保存されたシステイン残基を持ち、アミノ酸レベルで３０－５０％の配列相同性を共有する。これらの因子は、顆粒膜細胞、副腎皮質細胞、軟骨細胞、筋芽細胞、角膜及び血管内皮細胞（ウシ又はヒト）、血管平滑筋細胞、水晶体、網膜及び前立腺上皮細胞、オリゴデンドロサイト、アストロサイト、クロンドサイト（ｃｈｒｏｎｄｏｃｙｔｅｓ）、筋芽細胞及び骨芽細胞を含む広範囲の正常な二倍体中胚葉由来及び神経冠由来細胞に対して分裂促進的である。

線維芽細胞成長因子はまた非分裂促進的な形で多数の細胞型を刺激することができる。これらの活動は、線維芽細胞成長因子はまた創傷領域への細胞遊走の促進（走化作用）、新しい血管形成の開始（血管新生）、神経再生と生存の変調（ニューロトロフィズム（ｎｅｕｒｏｔｒｏｐｈｉｓｍ））、内分泌機能の変調、及び特異的細胞性タンパク質発現、細胞外マトリックス生産及び細胞生存の刺激又は抑制を含む。Ｂａｉｒｄ ＆ Ｂｏｈｌｅｎ，Ｈａｎｄｂｏｏｋ ｏｆ Ｅｘｐ．Ｐｈａｒｍａｃｏｌ．９５（１）：３６９－４１８，Ｓｐｒｉｎｇｅｒ，（１９９０）。これらの性質は、創傷治癒、神経修復、側枝血管形成等々を加速させる治療的アプローチにおいて線維芽細胞成長因子を使用するための基礎を提供する。例えば、線維芽細胞成長因子が心臓疾患及び手術において心筋損傷を最小にすることが示唆されている（Ｕ．Ｓ．Ｐ．４，３７８，３４７）。

ここで我々は、ここでＰＲＯ５３３ポリペプチドと命名した、ＦＧＦと相同性を有する新規なポリペプチドの同定と特徴付けを記載する。

【００１２】

１０．ＰＲＯ２４５

細胞性プロセスの上述した調節と変調に最も重要なタンパク質の幾つかは細胞中におけるタンパク質のリン酸化のレベルを調節する酵素である。例えば、細胞成長と分化を調節するシグナルの伝達は様々な細胞性タンパク質のリン酸化と脱リン酸化により少なくとも部分的に調節されることが知られている。これらのプロセスを触媒する酵素は、様々な細胞性タンパク質をリン酸化するように機能するプロテインキナーゼと、様々な細胞性タンパク質からリン酸残基を除去するように機能するプロテインホスファターゼを含む。細胞中でのタンパク質リン酸化のレベルのバランスはよってこれらの２種の酵素の相対活動により媒介される。

多くのプロテインキナーゼ酵素が同定されているが、これらの触媒性タンパク質の多くにより果たされる役割はまだ明らかにされなければならない。しかし、多数の既知のプロテインキナーゼがタンパク質中のチロシン残基をリン酸化するように機能し、よって様々な異なった効果を生じることがよく知られている。おそらく最も重要なこととしては、多くの発ガン産物と成長因子が固有のチロシンキナーゼ活性を有しているという発見のために、チロシンキナーゼに非常に大きな興味が存在している。従って、チロシンキナーゼファミリーの新しいメンバーを同定することが望まれている。

プロテインキナーゼの生理学的な重要性が分かっているので、新規な天然のキナーゼタンパク質を同定するために産業界と学術界の両方により努力がなされている。これらの努力の多くは新規なキナーゼタンパク質に対するコード化配列を同定するために哺乳動物の組

換えＤＮＡライブラリのスクリーニングに集中している。我々はここにＰＲＯ２４５ポリペプチドと命名した、チロシンキナーゼタンパク質と相同性を有する新規なポリペプチドの同定と特徴付けを記載する。

【００１３】

１１．ＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７

タンパク質－タンパク質相互作用はレセプター及び抗原複合体及びシグナル伝達機構に関連するものを含む。タンパク質－タンパク質相互作用の基となる構造的及び機能的機構についてよく知られているように、タンパク質－タンパク質相互作用は、タンパク質－タンパク質相互作用の特定の結果を調節するように更に容易に操作することができる。従って、タンパク質－タンパク質相互作用の基となる機構は科学及び医学界にとって興味深い。ロイシンリッチ反復を含む全てのタンパク質はタンパク質－タンパク質相互作用に関与していると考えられている。ロイシンリッチ反復は多様な機能を持つ多くのタンパク質と細胞位置に存在する短い配列モチーフである。リボ核酸インヒビタータンパク質の結晶構造から、ロイシンリッチ反復はβ－α構造単位に対応することが明らかになった。これらの単位は、一面が溶媒に暴露された平行なβシートを形成するように配置され、タンパク質は珍しい非球状の形状を獲得する。これらの二つの特徴はロイシンリッチ反復を含むタンパク質のタンパク質結合機能の原因であることが示されている。Ｋｏｂｅ及びＤｅｉｓｅｎｈｏｆｅｒ，Ｔｒｅｎｄｓ Ｂｉｏｃｈｅｍ．Ｓｃｉ．，１９（１０）：４１５－４２１（Ｏｃｔ．１９９４）を参照されたい。

個体発生の間にコラーゲン原線維を配向し指令する組織オーガナイザーとなり、創傷治癒、組織修復、及び腫瘍ストローマ形成のような病理プロセスに関与するロイシンリッチプロテオグリカンについて研究が報告されている。Ｉｏｚｚｏ，Ｒ．Ｖ．，Ｃｒｉｔ．Ｒｅｖ．Ｂｉｏｃｈｅｍ．Ｍｏｌ．Ｂｉｏｌ．，３２（２）：１４１－１７４（１９９７）。創傷治癒、組織修復におけるロイシンリッチタンパク質の関与を示す他の研究は、Ｄｅ Ｌａ Ｓａｌｌｅ，Ｃ．等，Ｖｏｕｖ．Ｒｅｖ．Ｆｒ．Ｈｅｍａｔｏｌ．（Ｇｅｒｍａｎｙ），３７（４）：２１５－２２２（１９９５）であり、出血疾患ベルナール－スーリエ症候群に伴う複合体におけるロイシンリッチモチーフの突然変異を報告しており、Ｃｈｌｅｍｅｔｓｏｎ，Ｋ．Ｊ．，Ｔｈｒｏｍｂ．Ｈａｅｍｏｓｔ．（Ｇｅｒｍａｎｙ），７４（１）：１１１－１１６（Ｊｕｌｙ １９９５）は、血小板がロイシンリッチ反復を有していることを報告している。ロイシンリッチ反復を有することが報告されている特に興味深い他のタンパク質は、アルツハイマー病のような神経変性疾患、パーキンソン病のような神経損傷の治療において、及びガンの診断に対して有用であることが報告されているＳＬＩＴタンパク質である。エール大学のＡｒｔａｖａｎｉｓｔｓａｋｏｎａｓ，Ｓ．及びＲｏｔｈｂｅｒｇ，Ｊ．Ｍ．のＷＯ９２１０５１８－Ａ１を参照されたい。ロイシンリッチ反復を有するタンパク質の生物学的機能を報告している他の研究には：Ｔａｙａｒ，Ｎ．等，Ｍｏｌ．Ｃｅｌｌ Ｅｎｄｏｃｒｉｎｏｌ．，（Ｉｒｅｌａｎｄ），１２５（１－２）：６５－７０（Ｄｅｃ．１９９６）（ゴナドトロピンレセプター関連）；Ｍｉｕｒａ，Ｙ．等，Ｎｉｐｐｏｎ Ｒｉｎｓｈｏ（Ｊａｐａｎ），５４（７）：１７８４－１７８９（Ｊｕｌｙ １９９６）（アポトーシス関連）；Ｈａｒｒｉｓ，Ｐ．Ｃ．等，Ｊ．Ａｍ．Ｓｏｃ．Ｎｅｐｈｒｏｌ．，６（４）：１１２５－１１３３（Ｏｃｔ．１９９５）（腎疾患関連）；及びＬａ Ｊｏｌｌａ Ｃａｎｃｅｒ Ｒｅｓｅａｒｃｈ ＦｏｕｎｄａｔｉｏｎのＲｕｏｓｌａｈｔｉ，Ｅ．Ｉ．等のＷＯ９１１０７２７－Ａ（トランスフォーミング成長因子βに結合するデコリンの、ガンの治療、創傷治癒及び瘢痕化への関与）が含まれる。

よって、タンパク質－タンパク質相互作用をよりよく理解するためにロイシンリッチ反復を有する新規なタンパク質を同定するための努力が産業界と学術界の両方によってなされている。特に興味があるものはロイシンリッチ反復とＳＬＩＴタンパク質及び血小板糖タンパク質Ⅴのようなロイシンリッチ反復を有する既知のタンパク質と相同性を有するタンパク質である。

【００１４】

１２．ＰＲＯ２５８
免疫グロブリンは、Ｂ細胞膜上の抗原に対するレセプターと、形質細胞の分泌産物との双方として機能するタンパク質である抗体分子である。全ての抗体分子と同様に、免疫グロブリンは２つの主要な機能を遂行する：それらは抗原に特異的に結合し、限られた数の生物学的エフェクター機能に関与する。従って、Ｉｇスーパーファミリーの新しいメンバーは常に興味深い。様々なウィルスによりレセプターとして作用する分子と免疫機能を調節する様に作用するものが特に興味深い。また特に興味深いものはウィルスレセプターとして作用するか免疫機能を調節する既知のＩｇファミリーメンバーと相同性を有する分子である。よって、ポリオウィルスレセプター、ＣＲＴＡＭ及びＣＤ１６６（リンパ球抗原ＣＤ６のリガンド）と相道営を有する分子は特に興味深い。
細胞外及び膜結合タンパク質は多細胞生物の形成、分化及び維持において重要な役割を果たしている。多くの個々の細胞の運命、例えば増殖、遊走、分化、又は他の細胞との相互作用は典型的には他の細胞及び／又は直近の環境から受け取る情報により支配される。この情報はしばしば分泌ポリペプチド（例えば分裂促進分子、生存因子、細胞障害性因子、分化因子、神経ペプチド、及びホルモン）により伝達され、これがまた多様な細胞レセプター又は膜結合タンパク質によって受け取られ解釈される。これらの分泌ポリペプチド又はシグナル伝達分子は通常は細胞性分泌経路を通過して、通常は膜結合レセプタータンパク質にある細胞外環境のその作用部位に達する。
ここで我々は、ここにＰＲＯ２５８と命名した、ＣＲＴＡＭと相同性を有する新規なポリペプチドの同定及び特徴づけについて記載する。
【００１５】
１３．ＰＲＯ２６６
タンパク質－タンパク質相互作用はレセプター及び抗原複合体及びシグナル伝達機構に関連するものを含む。タンパク質－タンパク質相互作用の基となる構造的及び機能的機構についてよく知られているように、タンパク質－タンパク質相互作用は、タンパク質－タンパク質相互作用の特定の結果を調節するように更に容易に操作することができる。従って、タンパク質－タンパク質相互作用の基となる機構は科学及び医学界にとって興味深い。ロイシンリッチ反復を含む全てのタンパク質はタンパク質－タンパク質相互作用に関与していると考えられている。ロイシンリッチ反復は多様な機能を持つ多くのタンパク質と細胞位置に存在する短い配列モチーフである。リボ核酸インヒビタータンパク質の結晶構造から、ロイシンリッチ反復はβ－α構造単位に対応することが明らかになった。これらの単位は、一面が溶媒に暴露された平行なβシートを形成するように配置され、タンパク質は珍しい非球状の形状を獲得する。これらの二つの特徴はロイシンリッチ反復を含むタンパク質のタンパク質結合機能の原因であることが示されている。Ｋｏｂｅ及びＤｅｉｓｅｎｈｏｆｅｒ，Ｔｒｅｎｄｓ Ｂｉｏｃｈｅｍ．Ｓｃｉ．，１９（１０）：４１５－４２１（Ｏｃｔ．１９９４）を参照されたい。
個体発生の間にコラーゲン原線維を配向し指令する組織オーガナイザーとなり、創傷治癒、組織修復、及び腫瘍ストローマ形成のような病理プロセスに関与するロイシンリッチプロテオグリカンについて研究が報告されている。Ｉｏｚｚｏ，Ｒ．Ｖ．，Ｃｒｉｔ．Ｒｅｖ．Ｂｉｏｃｈｅｍ．Ｍｏｌ．Ｂｉｏｌ．，３２（２）：１４１－１７４（１９９７）。創傷治癒、組織修復におけるロイシンリッチタンパク質の関与を示す他の研究は、Ｄｅ Ｌａ Ｓａｌｌｅ，Ｃ．等，Ｖｏｕｖ．Ｒｅｖ．Ｆｒ．Ｈｅｍａｔｏｌ．（Ｇｅｒｍａｎｙ），３７（４）：２１５－２２２（１９９５）であり、出血疾患ベルナール－スーリエ症候群に伴う複合体におけるロイシンリッチモチーフの突然変異を報告しており、Ｃｈｌｅｍｅｔｓｏｎ，Ｋ．Ｊ．，Ｔｈｒｏｍｂ．Ｈａｅｍｏｓｔ．（Ｇｅｒｍａｎｙ），７４（１）：１１１－１１６（Ｊｕｌｙ １９９５）は、血小板がロイシンリッチ反復を有していることを報告している。ロイシンリッチ反復を有することが報告されている特に興味深い他のタンパク質は、アルツハイマー病のような神経変性疾患、パーキンソン病のような神経損傷の治療において、及びガンの診断に対して有用であることが報告されているＳＬＩＴタンパク質である。エール大学のＡｒｔａｖａｎｉｓｔｓａ

ｋｏｎａｓ，Ｓ．及びＲｏｔｈｂｅｒｇ，Ｊ．Ｍ．のＷＯ９２１０５１８－Ａ１を参照されたい。ロイシンリッチ反復を有するタンパク質の生物学的機能を報告している他の研究には：Ｔａｙａｒ，Ｎ．等，Ｍｏｌ．Ｃｅｌｌ Ｅｎｄｏｃｒｉｎｏｌ．，（Ｉｒｅｌａｎｄ），１２５（１－２）：６５－７０（Ｄｅｃ．１９９６）（ゴナドトロピンレセプター関連）；Ｍｉｕｒａ，Ｙ．等，Ｎｉｐｐｏｎ Ｒｉｎｓｈｏ（Ｊａｐａｎ），５４（７）：１７８４－１７８９（Ｊｕｌｙ １９９６）（アポトーシス関連）；Ｈａｒｒｉｓ，Ｐ．Ｃ．等，Ｊ．Ａｍ．Ｓｏｃ．Ｎｅｐｈｒｏｌ．，６（４）：１１２５－１１３３（Ｏｃｔ．１９９５）（腎疾患関連）；及びＬａ Ｊｏｌｌａ Ｃａｎｃｅｒ Ｒｅｓｅａｒｃｈ ＦｏｕｎｄａｔｉｏｎのＲｕｏｓｌａｈｔｉ，Ｅ．Ｉ．等のＷＯ９１１０７２７－Ａ（トランスフォーミング成長因子βに結合するデコリンの、ガンの治療、創傷治癒及び瘢痕化への関与）が含まれる。

よって、タンパク質－タンパク質相互作用をよりよく理解するためにロイシンリッチ反復を有する新規なタンパク質を同定するための努力が産業界と学術界の両方によってなされている。特に興味があるものはロイシンリッチ反復と、ＳＬＩＴタンパク質のようなロイシンリッチ反復を有する既知のタンパク質と相同性を有するタンパク質である。ここで我々は、ここにＰＲＯ２６６ポリペプチドと命名する、ＳＬＩＴタンパク質に相同性を持つ新規なポリペプチドについて記載する。

【００１６】

１４．ＰＲＯ２６９

トロンボモジュリンはトロンビンに結合しその活性を調節する。それは血液凝固の制御において重要である。トロンボモジュリンはトロンビンによるプロテインＣの活性化を加速することにより天然の抗凝血剤として機能する。可溶型トロンボモジュリンは、ヘパリンと比較して出血の危険が低くなった抗血栓剤としての治療用途を有しうる。トロンボモジュリンは内皮細胞及び血小板に存在する細胞表面膜貫通糖タンパク質である。トロンボモジュリンのより小さい機能的に活性な型は血漿中を循環し、尿にもまた見いだされる。（Ｈａｅｂｅｒｌｉ，Ａ．，Ｈｕｍａｎ Ｐｒｏｔｅｉｎ Ｄａｔａ，ＶＣＨ Ｏｕｂ．，Ｎ．Ｙ．，１９９２）。トロンボモジュリンと相同性を有するポリペプチドが特に望ましい。

ここで我々は、ここにＰＲＯ２６９ポリペプチドと命名する、トロンボモジュリンと相同性を持つ新規なポリペプチドの同定と特徴付けを記載する。

【００１７】

１５．ＰＲＯ２８７

プロコラーゲンＣ－プロテイナーゼエンハンサータンパク質は結合して骨形態形成タンパク質「ＢＭＰ１」／プロコラーゲンＣ－プロテイナーゼ（ＰＣＰ）の活性を亢進する。それは、細胞外マトリックスの沈着においてある役割を果たす。ＢＭＰ１タンパク質は骨及び／又は軟骨形成を誘導するために、また創傷治癒と組織修復において使用することができる。従って、プロコラーゲンＣ－プロテイナーゼエンハンサータンパク質ＢＭＰ１とそれと相同性を有するタンパク質は学術界と医学界にとって興味深い。

ここで我々は、ここにＰＲＯ２８７ポリペプチドと命名する、プロコラーゲンＣ－プロテイナーゼエンハンサータンパク質前駆体とプロコラーゲンＣ－プロテイナーゼエンハンサータンパク質と相同性を持つ新規なポリペプチドの同定と特徴付けを記載する。

【００１８】

１６．ＰＲＯ２１４

成長因子は、単独で又は協力して細胞成長又は増殖を亢進する分子シグナル又はメディエータである。しかし、成長因子への発現時には成長のみ以外の細胞性反応がある。その結果、成長因子は多機能性で強力な細胞レギュレーターとしてよりよく特徴付けられる。その生物学的効果は増殖、走化性及び細胞外マトリックス生産の刺激を含む。成長因子は刺激性及び阻害成功化の双方を有しうる。例えば、トランスフォーミング成長因子（ＴＧＦ－α）は高度に多面的で、ある細胞、特に結合組織では増殖を刺激する一方、他のもの、例えばリンパ球及び上皮細胞では増殖の強力な阻害剤である。

成長因子による成長刺激又は阻害の生理学的効果は標的組織の発達と分化の状態に依存する。古典的な内分泌分子による局所的細胞調節はオートクライン（同じ細胞）、ジャクスタクリン（近くの細胞）、及びパラクリン（隣接細胞）経路の理解を含む。ペプチド成長因子は複雑な生物学的言語の要素であり、細胞間情報交換の基礎を提供する。それらは細胞が互いに情報を運ぶことを可能にし、細胞間の相互作用を媒介し、遺伝子発現を変化させる。これらの多機能及び多面的因子の効果は他のペプチドの有無に依存している。

上皮成長因子（ＥＧＦ）は、上皮細胞と線維芽細胞を含む様々な型の細胞の増殖を刺激する一般的な分裂促進因子である。ＥＧＦはＥＧＦレセプター（ＥＧＦＲ）に結合してこれを活性化させ、これが細胞内シグナル伝達及びそれに引き続く効果を開始させる。ＥＧＦＲは中枢神経系（ＣＮＳ）の他の領域に加えて、大脳皮質、小脳、及び海馬のニューロン中で発現される。また、ＥＧＦはＣＮＳの様々な領域において発現される。従って、ＥＧＦは有糸分裂細胞に対してばかりでなく分裂終了細胞に対しても作用する。実際、多くの研究において、ＥＧＦがＣＮＳ中の様々なタイプのニューロンに神経栄養性及び神経調節性効果を有していることが示されている。例えば、ＥＧＦは培養された大脳皮質及び小脳ニューロンに直接作用し、神経突起成長と生存を亢進する。他方、ＥＧＦはまた中隔コリン作用性及び中脳ドーパミン作用性ニューロンを含む他の細胞型にもグリア細胞を通して間接的に作用する。ＣＮＳ中のニューロンに対するＥＧＦの効果の証拠は蓄積されつつあるが、作用の機構は本質的に知られていないままである。有糸分裂細胞におけるＥＧＦ－－誘発シグナル伝達は細胞分裂終了ニューロンにおいてよりもよりよく理解されている。クローン化されたクロム親和細胞腫ＰＣ１２細胞と培養された大脳皮質ニューロンの研究により、ＥＧＦ誘発神経栄養性作用がＥＧＦに応答してのＥＧＦＲとマイトジェン活性化プロテインキナーゼ（ＭＡＰＫ）の持続性活性化により媒介されることが示唆されている。持続性細胞内シグナル伝達はＥＧＦＲ下方制御の減少速度と相関し、これはＥＧＦに対するニューロン細胞の応答を決定するかもしれない。ＥＧＦは有糸細胞分裂細胞及び細胞分裂終了ニューロンを含む細胞の様々なタイプに作用する多分化能成長因子である可能性が高い。

ＥＧＦは、腎臓、膵臓、甲状腺、脳下垂体、神経系及び胃腸系の唾液腺及びブルンアー腺によりつくられ、体液、例えば唾液、血液、脳脊髄液（ＣＳＦ）、尿、羊水、前立腺液、膵液、及び母乳に見いだされる。Ｐｌａｔａ－Ｓａｌａｍａｎ，Ｐｅｐｔｉｄｅｓ １２：６５３－６６３（１９９１）。

ＥＧＦは、内因性チロシンキナーゼを含むその膜特異的レセプターにより媒介される。Ｓｔｏｓｃｈｅｃｋ等，Ｊ．Ｃｅｌｌ Ｂｉｏｃｈｅｍ．３１：１３５－１５２（１９８６）。ＥＧＦは内因性チロシンキナーゼを活性化する膜貫通シグナルを誘発するそのレセプターの細胞外部分に結合することにより機能すると信じられている。

ＥＧＦ様ドメインの精製と配列分析により、交差結合して３つのペプチドループをつくりだす６個の保存システイン残基の存在が明らかになった。Ｓａｖａｇｅ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２４８：７６６９－７６７２（１９７９）。幾つかの他のペプチドは、同じ一般化されたモチーフ$X_nCX_7CX_{4/5}CX_{10}CXCX_5GX_2CX_n$（ここで、Ｘは任意の非システインアミノ酸を表し、ｎは可変の繰り返し数である）を共有するＥＧＦレセプターと反応しうることが今は一般的に知られている。このモチーフを有する非単離ペプチドは、ＴＧＦ－α、アンフィレギュリン、シュワン細胞腫由来成長因子（ＳＤＧＦ）、ヘパリン結合ＥＧＦ様成長因子及びある種のウィルス性コード化ペプチド（例えばワクシニアウィルス、Ｒｅｉｓｎｅｒ，Ｎａｔｕｒｅ ３１３：８０１－８０３（１９８５），ショープ線維腫ウィルス、Ｃｈａｎｇ等，Ｍｏｌ Ｃｅｌｌ Ｂｉｏｌ．７：５３５－５４０（１９８７），伝染性軟属腫、Ｐｏｒｔｅｒ及びＡｒｃｈａｒｄ，Ｊ．Ｇｅｎ．Ｖｉｒｏｌ．６８：６７３－６８２（１９８７）、及び粘液腫ウィルス、Ｕｐｔｏｎ等，Ｊ．Ｖｉｒｏｌ．６１：１２７１－１２７５（１９８７）、Ｐｒｉｇｅｎｔ及びＬｅｍｏｉｎｅ，Ｐｒｏｇ．Ｇｒｏｗｔｈ Ｆａｃｔｏｒ Ｒｅｓ．．４：１－２４（１９９２）を含む。

ＥＧＦ様ドメインは成長因子に制限されず、細胞接着、タンパク質－タンパク質相互作用

と発達において興味有る性質を有している様々な細胞表面及び細胞外タンパク質において観察されている。Ｌａｕｒｅｎｃｅ及びＧｕｓｔｅｒｓｏｎ，Ｔｕｍｏｒ Ｂｉｏｌ．１１：２２９－２６１（１９９０）。これらのタンパク質は血液凝固因子（因子ＶＩ、ＩＸ、Ｘ、ＸＩＩ、プロテインＣ、プロテインＳ、プロテインＺ、組織プラスミノゲンアクチベーター、ウロキナーゼ）、細胞外マトリックス成分（ラミニン、サイトタクチン、エンタクチン）、細胞表面レセプター（ＬＤＬレセプター、トロンボモジュリンレセプター）及び免疫関連タンパク質（補体Ｃ１ｒ、ウロモジュリン）を含む。

更により興味深いことには、ＥＧＦ様前駆体の一般的構造パターンは下等生物並びに哺乳動物細胞を通して保存されている。発達上の優位性を持つ多くの遺伝子はＥＧＦ様繰り返しを持つ無脊椎動物において同定されている。例えば、ショウジョウバエのｎｏｔｃｈ遺伝子は、ＥＧＦとの相同性を示す３６の直列に配置された４０のアミノ酸反復をコードしている。Ｗｈａｒｔｏｎ等，Ｃｅｌｌ ４３：５５７－５８１（１９８５）。ヒドロパシープロットは推定膜貫通ドメインを示し、そのＥＤＦ関連配列は膜の細胞外側に位置している。ＥＧＦ様反復を持つ他のホメオティック遺伝子は、Ｎｏｔｃｈに基づくプローブを使用して同定されたデルタ、９５Ｆ及び５ＺＤ、及び２つの特定された細胞の間に伝達される発生的シグナルに対する推定上レセプターをコードする線虫遺伝子Ｌｉｎ－１２を含む。

特に、ＥＧＦは胃腸粘膜の保存と維持及び急性及び慢性粘膜病状の修復、Ｋｏｎｔｕｒｅｋ等，Ｅｕｒ．Ｊ．Ｇａｓｔｒｏｅｎｔｅｒｏｌ Ｈｅｐａｔｏｌ．７（１０），９３３－３７（１９９５）で、壊死性腸炎、ゾリンジャー・エリソン症候群、胃腸潰瘍形成及び先天性微絨毛萎縮症、Ｇｕｇｌｉｅｔｔａ及びＳｕｌｌｉｖａｎ，Ｅｕｒ．Ｊ．Ｇａｓｔｒｏｅｎｔｅｒｏｌ Ｈｅｐａｔｏｌ，７（１０），９４５－５０（１９９５）に可能性を有していることが示された。また、ＥＧＦは、毛包分化；ｄｕ Ｃｒｏｓ，Ｊ．Ｉｎｖｅｓｔ．Ｄｅｒｍａｔｏｌ．１０１（１ Ｓｕｐｐｌ．），１０６Ｓ－１１３Ｓ（１９９３），Ｈｉｌｌｉｅｒ，Ｃｌｉｎ．Ｅｎｄｏｃｒｉｎｏｌ．３３（４），４２７－２８（１９９０）；腎機能、Ｈａｍｍ等，Ｓｅｍｉｎ．Ｎｅｐｈｒｏｌ．１３（１）：１０９－１５（１９９３），Ｈａｒｒｉｓ，Ａｍ．Ｊ．Ｋｉｄｎｅｙ Ｄｉｓ．１７（６）：６２７－３０（１９９１）；涙液、ｖａｎ Ｓｅｔｔｅｎ，Ｉｎｔ．Ｏｐｈｔｈａｌｍｏｌ １５（６）；３５９－６２（１９９１）；ビタミンＫ媒介血液凝固 Ｓｔｅｎｆｌｏ等，Ｂｌｏｏｄ ７８（７）：１６３７－５１（１９９１）に関与していた。ＥＧＦはまた異常なケラチノサイト分化により特徴づけられる様々な皮膚疾患、例えば乾癬、上皮ガン、例えば肺の扁平上皮ガン、外陰部の類表皮ガン及び神経膠腫に関連している。Ｋｉｎｇ等，Ａｍ．Ｊ．Ｍｅｄ．Ｓｃｉ．２９６：１５４－１５８（１９９８）。

非常に興味深いのは、成長因子シグナル伝達経路における遺伝的変更が発育異常とガンを含む慢性疾患に密接に関連しているという増大している証拠である。Ａａｒｏｎｓｏｎ，Ｓｃｉｅｎｃｅ ２５４：１１４６－１１５３（１９９１）。例えば、ＥＧＦレセプタータンパク質と密接な構造類似性を持つプロトガン遺伝子ｃ－ｅｒｂ－２（またＨＥＲ－２としても知られている）がヒト乳ガンにおいて過剰発現されている。Ｋｉｎｇ等，Ｓｃｉｅｎｃｅ ２２９：９７４－９７６（１９８５）；Ｇｕｌｌｉｃｋ，Ｈｏｒｍｏｎｅｓ ａｎｄ ｔｈｅｉｒ ａｃｔｉｏｎｓ，Ｃｏｏｋｅ等，ｅｄｓ，Ａｍｓｔｅｒｄａｍ，Ｅｌｓｅｖｉｅｒ，ｐｐ３４９－３６０（１９８６）。

【００１９】

１７．ＰＲＯ３１７

ＴＧＦ－βスーパー遺伝子ファミリー、あるいは簡単にＴＧＦ－βスーパーファミリーは、分泌タンパク質の一群であり、実質的に全ての門において発現される多数の関連した成長及び分化因子を含む。スーパーファミリーメンバーはシグナル伝達機構を活性化させてその多機能サイトカイン効果を顕在化させる特異的細胞表面レセプターに結合する。Ｋｏｌｏｄｚｉｅｊｃｚｙｋ及びＨａｌｌ，Ｂｉｏｃｈｅｍ．Ｃｅｌｌ．Ｂｉｏｌ．，７４：２９９－３１４（１９９６）；Ａｔｔｉｓａｎｏ及びＷｒａｎａ，Ｃｙｔｏｋｉ

ｎｅ Ｇｒｗｏｔｈ Ｆａｃｔｏｒ Ｒｅｖ．，７：３２７－３３９（１９９６）；及びＨｉｌｌ，Ｃｅｌｌｕｌａｒ Ｓｉｇｎａｌｉｎｇ，８：５３３－５４４（１９９６）。このファミリーのメンバーはＴＧＦ－βの５つの区別できる型（Ｓｐｏｒｎ及びＲｏｂｅｒｔｓ，Ｐｅｐｔｉｄｅ Ｇｒｏｗｔｈ Ｆａｃｔｏｒｓ ａｎｄ Ｔｈｅｉｒ ｒｅｃｅｐｔｏｒｓ，Ｓｐｏｒｎ及びＲｏｂｅｒｔｓ，ｅｄｓ．（Ｓｐｒｉｎｇｅｒ－Ｖｅｒｌａｇ：Ｂｅｒｌｉｎ，１９９０ ｐｐ．４１９－４７２）、並びに分化因子ｖｇ１（Ｗｅｅｋｓ及びＭｅｌｔｏｎ，Ｃｅｌｌ，５１：８６１－８６７（１９８７））及びＤＰＰ－Ｃポリペプチド（Ｐａｄｇｅｔｔ等，Ｎａｔｕｒｅ，３２５：８１－８４（１９８７））、ホルモンのアクチンとインヒビン（Ｍａｓｏｎ等，Ｎａｔｕｒｅ，３１８：６５９－６６３（１９８５）；Ｎａｓｏｎ等，Ｇｒｏｗｔｈ Ｆａｃｔｏｒｓ，１：７７－８８（１９８７））、ミュラー阻害物質（ＭＩＳ）（Ｃａｔｅ等，Ｃｅｌｌ，４５：６８５－６９８（１９８６））、骨形成タンパク質（ＢＭＰｓ）（Ｗｏｚｎｅｙ等，Ｓｃｉｅｎｃｅ，２４２：１５２８－１５３４（１９８８）；１９８８年１月１４日公開のＰＣＴ ＷＯ ８８／００２０５；１９８９年１０月３１日発行のＵ．Ｓ．４，８７７，８６４）、発生的制御タンパク質Ｖｇｒ－１（Ｌｙｏｎｓ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ．８６：４５５４－４５５８（１９８９））及びＶｇｒ－２（Ｊｏｎｅｓ等，Ｍｏｌｅｃ．Ｅｎｄｏｃｒｉｎｏｌ．，６：１９６１－１９６８（１９９２））、マウス成長分化因子（ＧＤＦ）、例えばＧＤＦ－３及びＧＤＦ－９（Ｋｉｎｇｓｌｅｙ，Ｇｅｎｅｓ Ｄｅｖ．，８：１３３－１４６（１９９４）；ＭｃＰｈｅｒｒｏｎ及びＬｅｅ，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２６８：３４４４－３４４９（１９９３））、マウスｌｅｆｔｙ／Ｓｔｒａ１（Ｍｅｎｏ等，Ｎａｔｕｒｅ，３８１：１５１－１５５（１９９６）；Ｂｏｕｉｌｌｅｔ等，Ｄｅｖ．Ｂｉｏｌ．，１７０：４２０－４３３（１９９５））、グリア株化細胞由来神経栄養因子（ＧＤＮＦ）（Ｌｉｎ等，Ｓｃｉｅｎｃｅ，２６０：１１３０－１１３２（１９９３））、ニュールツリン（ｎｅｕｒｔｕｒｉｎ）（Ｋｏｔｓｂａｕｅｒ等，Ｎａｔｕｒｅ，３８４：４６７－４７０（１９９６））、及び子宮内膜出血関連因子（ＥＢＡＦ）（Ｋｏｔｈａｐａｌｌｉ等，Ｊ．Ｃｌｉｎ．Ｉｎｖｅｓｔ．，９９：２３４２－２３５０（１９９７））を含んでいる。サブセットＢＭＰ－２Ａ及びＢＭＰ－２ＢはＤＰＰ－Ｃとおよそ７５％の配列相同性があり、そのタンパク質の哺乳動物の対応物を表しうる。

ＴＧＦ－βスーパーファミリーのタンパク質は、疎水性シグナル配列、数百のアミノ酸の長く相対的に不十分に保存されたＮ末端プロ領域、切断部位（通常は多塩基性）、及びより短くより高度に保存されたＣ末端領域を含むより大きな前駆体ポリペプチド鎖によりコードされるジスルフィド結合ホモ又はヘテロダイマーである。このＣ末端領域はプロセシングされた成熟タンパク質に対応し、特徴的なシステインモチーフ、つまり全ての既知のファミリーメンバーの中のＴＧＦ－βの９つのシステイン残基の７つの保存を持つおよそ１００アミノ酸を含む。成熟及びプロ領域の間の切断部位の位置はファミリーメンバーの間で変わるが、全てのタンパク質のＣ末端は同一の位置にあり、配列Ｃｙｓ－Ｘ－Ｃｙｓ－Ｘで終了するが、Ｃｙｓ－Ｌｙｓ－Ｃｙｓ－ＳｅｒのＴＧＦ－βコンセンサスＣ末端とは全ての場合において異なっている。上掲のＳｐｏｒｎ及びＲｏｂｅｒｔｓ，１９９０。

現在同定されている少なくとも５つの型のＴＧＦ－β、つまりＴＧＦ－β１、ＴＧＦ－β２ＴＧＦ－β３、ＴＧＦ－β４、及びＴＧＦ－β５がある。ＴＧＦ－β１の活性化型は３９０のアミノ酸前駆体のカルボキシ末端の１１２のアミノ酸の二量体化により形成されたホモダイマーである。組換えＴＧＦ－β１はクローン化され（Ｄｅｒｙｎｃｋ等，Ｎａｔｕｒｅ，３１６：７０１－７０５（１９８５））、チャイニーズハムスター卵巣細胞に発現されている（Ｇｅｎｔｒｙ等，Ｍｏｌ．Ｃｅｌｌ．Ｂｉｏｌ．，７：３４１８－３４２７（１９８７））。また、組換えヒトＴＧＦ－β２（ｄｅＭａｒｔｉｎ等，ＥＭＢＯ Ｊ．，６：３６７３（１９８７））、並びにヒト及びブタＴＧＦ－β３（Ｄｅｒｙｎｃｋ等，ＥＭＢＯ Ｊ．，７：３７３７－３７４３（１９８８）；ｔｅｎ Ｄｉ

ｊｋｅ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，８５：４７１５（１９８８））がクローン化されている。ＴＧＦ－β２は、４１４のアミノ酸の前駆体型を有しており、またＴＧＦ－β１の活性型とおよそ７０％の相同性を共有するカルボキシ末端の１１２のアミノ酸からホモダイマーにプロセシングされる（Ｍａｒｑｕａｒｄｔ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２６２：１２１２７（１９８７））。またＥＰ２００，３４１；１６９，０１６；２６８，５６１；及び２６７，４６３；Ｕ．Ｓ．特許番号４，７７４，３２２；Ｃｈｅｉｆｅｔｚ等，Ｃｅｌｌ，４８：４０９－４１５（１９８７）；Ｊａｋｏｗｌｅｗ等，Ｍｏｌｅｃｕｌａｒ Ｅｎｄｏｃｒｉｎ．，２：７４７－７５５（１９８８）；Ｄｅｒｔｎｃｋ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２６１：４３７７－４３７９（１９８６）；Ｓｈａｒｐｌｅｓ等，ＤＮＡ，６：２３９－２４４（１９８７）；Ｄｅｒｙｎｃｋ等，Ｎｕｃｌ．Ａｃｉｄｓ．Ｒｅｓ．，１５：３１８８－３１８９（１９８７）；Ｄｅｒｙｎｃｋ等，Ｎｕｃｌ．Ａｃｉｄｓ．Ｒｅｓ．，１５：３１８７（１９８７）；Ｓｅｙｅｄｉｎ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２６１：５６９３－５６９５（１９８６）；Ｍａｄｉｓｅｎ等，ＤＮＡ，７：１－８（１９８８）；及びＨａｎｋ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．（Ｕ．Ｓ．Ａ．），８５：７９－８２（１９８８）を参照されたい。

ＴＧＦ－β４及びＴＧＦ－β５はニワトリ軟骨細胞ｃＤＮＡライブラリ（Ｊａｋｏｗｌｅｗ等，Ｍｏｌｅｃ．Ｅｎｄｏｃｒｉｎｏｌ．，２：１１８６－１１９５（１９８８））とカエル卵母細胞からそれぞれクローン化されている。

ＴＧＦ－βのプロ領域は成熟ＴＧＦ－βダイマーと非共有結合的に結合し（Ｗａｋｅｆｉｅｌｄ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２６３：７６４６－７６５４（１９８８）；Ｗａｋｅｆｉｅｌｄ等，Ｇｒｏｗｔｈ Ｆａｃｔｏｒｓ，１：２０３－２１８（１９８９））、プロ領域は、ＴＧＦ－βとアクチビンの両方の活性な成熟ダイマーの正しい折り畳みと分泌に必要であることが見いだされている（Ｇｒａｙ及びＭａｓｏｎ，Ｓｃｉｅｎｃｅ，２４７：１３２８－１３３０（１９９０））。ＴＧＦ－βの成熟及びプロ領域の間の結合は成熟ダイマーの生物活性を遮蔽し、不活性な潜伏型を形成する。潜伏期は、プロ領域の存在がアクチビン又はインヒビンの生物活性に効果を持たないので、ＴＧＦ－βスーパーファミリーの定数ではない。

ＴＧＦ－βスーパーファミリーからのタンパク質の生物学上の一貫した特徴は、その発生プロセスを調節する能力である。ＴＧＦ－βは非常に広範な正常な細胞と新生物細胞の両方に数多くの調節性作用を有することが知られている。ＴＧＦ－βは多機能であり、細胞増殖、分化、及び細胞機能におけるその他の重要なプロセスを刺激あるいは阻害することができる（上掲のＳｐｏｒｎ及びＲｏｂｅｒｓ）。

ＴＧＦ－βスーパーファミリーの一メンバーであるＥＢＡＦは、子宮内膜において後期分泌期と異常な子宮内膜出血の間のみに発現される。Ｋｏｔｈａｐａｌｌｉ等，Ｊ．Ｃｌｉｎ．Ｉｎｖｅｓｔ．，９９：２３４２－２３５０（１９９７）。ヒト子宮内膜は、規則的な感覚で出血する体内の唯一の組織であるという点で独特である。また、異常な子宮内膜出血は婦人科疾患の最もよくある徴候であり、子宮摘出の重要な兆しである。インサイツハイブリッド形成は、ＥＢＡＦのｍＲＮＡが子宮内膜腺又は内皮細胞に有意にｍＲＮＡを発現させることなくストローマに発現されることを示している。

ＥＢＡＦの予想されるタンパク質配列はＴＧＦ－βスーパーファミリーのマウスｌｅｆｔｙ／ｓｔｒａ３によりコードされたタンパク質と強い相同性を示した。モチーフ検索により、予想されたＥＢＡＦタンパク質はＴＧＦ－β関連タンパク質の中で保存されており、システインノット構造の形成に必要であるシステイン残基の殆どを含むことが明らかにされた。ＥＢＡＦ列は最初の保存されたシステイン残基から１２のアミノ酸だけ上流の更なるシステイン残基を含む。更なるシステイン残基を含むことが知られている唯一の他のファミリーメンバーはＴＧＦ－βｓ、インヒビン、及びＧＤＦ－３である。ＬＥＦＴＹ、ＧＤＦ－３／Ｖｇｒ２、及びＧＤＦ－９と同様にＥＢＡＦは分子間ジスルフィド結合を形成することが知られているシステイン残基を欠く。従って、ＥＢＡＦはダイマーとしては存在していない可能性もある不対システイン残基を持つＴＧＦ－βスーパーファミリーの更

なるメンバーであると思われる。しかし、２つのモノマーサブユニット間の疎水的接触はダイマー形成を促進しうる。蛍光インサイツハイブリッド形成は、ｅｂａｆ遺伝子がヒト染色体１のバンドｑ４２．１に位置していることを示している。
ＥＢＡＦに関連しているもののようなＴＧＦ－βスーパーファミリーの更なるメンバーが産業界と学術界によって探されている。
ここで我々は、ここにＰＲＯ３１７ポリペプチドと命名する、ＥＢＡＦと相同性を持つ新規なポリペプチドの同定と特徴付けを記載する。
【００２０】
１８．ＰＲＯ３０１
ガンの広範な発生はかなりの資源を新しい治療法を発見することに捧げることを促した。特定の一方法は腫瘍抗原に特異的である腫瘍又はガン特異的モノクローナル抗体（ｍＡｂｓ）を創造を含む。正常細胞とガン細胞を区別できるそのようなｍＡｂｓは疾患の診断、予防及び治療に有用である。特定の抗原は、結腸直腸ガンのような新生物疾患に関連していることが知られている。
一つの特定の抗原であるＡ３３抗原は原発性又は転移性大腸ガン並びに正常な大腸上皮の９０％以上において発現される。大腸ガンは広範性疾患であるので、早期の診断と治療が重要な医療目標である。大腸ガンの診断と治療は蛍光、核磁気又は放射性タグを持つそれに特異的なモノクローナル抗体（ｍＡｂｓ）を使用して実施することができる。放射性遺伝子、毒素及び／又は薬物タグｍＡＢｓを最小の患者の処方でインサイツでの治療に対して使用できる。ｍＡｂｓはまた大腸ガンの診断と治療における診断に使用することができる。例えば、Ａ３３抗原の血清レベルがある患者において上昇している場合、手術後のレベルの降下は腫瘍切除が成功したことを示している。他方、手術後の血清Ａ３３抗原レベルの引き続く上昇は元の腫瘍の転移が生じたかあるいは新しい原発性腫瘍が現れたことを示している。そのようなモノクローナル抗体は手術及び／又は化学療法の代わりに又はこれと同時に使用することができる。例えばＵ．Ｓ．Ｐ．４，５７９，８２７とＵ．Ｓ．Ｓ．Ｎ．４２４，９９１（Ｅ．Ｐ．１９９，１４１）はモノクローナル抗体の治療的投与に関しており、その後者は抗Ａ３３ｍＡｂの適用に関している。
上皮由来の多くのガンはアデノウィルスレセプターを有している。実際、アデノウィルス由来ベクターが腫瘍中にアンチセンス核酸を挿入する手段として提案されている（Ｕ．Ｓ．Ｐ．５，５１８，８８５）。よって、新生物腫瘍とのウィルスレセプターの結合は期待できないものではない。
我々は、ここでＰＲＯ３０１と命名した、ある種のガン関連抗原に対して相同性を有する新規なポリペプチドの同定及び特徴づけをここに記載する。
【００２１】
１９．ＰＲＯ２２４
コレステロール摂取は人の健康に深刻な意味を持ちうる。コレステロール摂取は膜合成に必要なコレステロールの殆どを細胞に提供する。この摂取が阻害された場合には、コレステロールは血液中に蓄積し血管壁におけるアテローム斑の形成に寄与しうる。殆どのコレステロールは低密度リポタンパク質（ＬＤＬｓ）として知られている複合体の形態のタンパク質に結合して血液中を輸送される。ＬＤＬｓはＬＤＬレセプタータンパク質を介して細胞中に飲食運動される。従って、ＬＤＬレセプタータンパク質とそれと相同性を有するタンパク質は科学界及び医学界にとって興味深い。
膜結合タンパク質とレセプターは多細胞生物の精製、分化及び維持において重要な役割を果たしうる。ＬＤＬレセプターは細胞膜の合成と形成に関与する膜結合タンパク質の例であり、個人の健康はその機能によって直接かつ間接に影響を受ける。多くの膜結合タンパク質はＬＤＬレセプターのようなレセプターとして作用する。これらのレセプターは基質を飲食運動するように機能するか、チャンネルのためのレセプターとして機能しうる。他の膜結合タンパク質はシグナル又は抗原として機能する。
膜結合タンパク質とレセプター分子は、製薬的及び診断用薬剤を含む様々な産業上の利用性を有している。膜結合タンパク質はまた関連レセプター／リガンド相互作用の強力なペ

プチド又は小分子制御因子をスクリーニングするために使用することもできる。ＬＤＬレセプターの場合には、血液中のコレステロール量とプラーク形成を低下させるように飲食運動を亢進する分子を見いだすことが望ましい。また、これらの分子が高い血液コレステロールを有する個人により避けられ又は制御されるように飲食運動を阻害する分子を同定することが望ましい。リポタンパク質レセプターと相同であるがリポタンパク質レセプターとしては機能しないポリペプチドはまた相同性を示す断片の機能の決定において興味深い。

次の研究はアポリポタンパク質を含むこれまでに既知の低密度リポタンパク質レセプターと関連タンパク質について報告している：Ｓａｗａｍｕｒａ等，Ｎｉｐｐｏｎ Ｃｈｅｍｉｐｈａｒ Ｃｏ，日本特許出願Ｊ０９０９８７８７；Ｎｏｖａｋ，Ｓ．等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２７１：（２０）１１７３２－６（１９９６）；Ｂｌａｓｓ，Ｄ．，Ｊ．Ｖｉｒｏｌ．，６９（１１）７２４４－７（Ｎｏｖ．１９９５）；Ｓｃｏｔｔ，Ｊ．，Ｊ．Ｉｎｈｅｒｔ．Ｍｅｔａｂ．Ｄｉｓ．（ＵＫ），９／Ｓｕｐｐ．１（３－１６）（１９８６）；Ｙａｍａｍｏｔｏ等，Ｃｅｌｌ，３９：２７－３８（１９８４）；Ｒｅｂｅｃｅ等，Ｎｅｕｒｏｂｉｏｌ．Ａｇｉｎｇ，１５：５１１７（１９９４）；Ｎｏｖａｋ，Ｓ．等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍｉｓｔｒｙ，２７１：１１７３２－１１７３６（１９９６）；及びＳｅｓｔａｖｅｌ及びＦｒｕｃｈａｒｔ，Ｃｅｌｌ Ｍｏｌ．Ｂｉｏｌ．，４０（４）：４６１－８１（Ｊｕｎｅ １９９４）。これらの刊行物とこの出願の提出の前に刊行された他のものは当該分野で既に知られている更なる背景を提供する。

新規な天然の膜結合レセプタータンパク質、特にリポタンパク質レセプターと相同性を有するものを同定する努力が産業界と学術界の両方によってなされている。我々は、ここにＰＲＯ２２４ポリペプチドと命名する、リポタンパク質と相同性を持つ新規なポリペプチドの同定と特徴付けをここに記載する。

【００２２】

２０．ＰＲＯ２２２

補体は体液性免疫と炎症において重要である血液中に見いだされる一群のタンパク質である。補体タンパク質は連続的に抗原－抗体複合体真野はタンパク分解性酵素により活性化される。活性化されると、補体タンパク質は細菌及び他の微生物を殺し、血管透過性に影響を及ぼし、ヒスタミンを放出し白血球を誘引する。補体はまた標的細胞に結合するとファゴサイトーシスを亢進する。自家細胞への害を防止するために、補体活性化経路は厳重に調節される。

補体活性化の調節又は補体タンパク質自体における欠陥は免疫複合体疾患、例えば全身性エリテマトーデスを生じ、また細菌感染を受けやすくなる結果になる。全ての場合、患者が治療を開示できる様に補体欠陥の早期の検出が望ましい。よって、補体活性化を制御する可溶型及び膜タンパク質の同定に研究努力が現在向けられている。

ヒトにおける補体活性化を制御する際に重要であることが知られているタンパク質はＨ因子と補体レセプター１型（ＣＲ１）を含む。Ｈ因子は補体タンパク質Ｃ３ｂと相互作用して、Ｃ３転換酵素の崩壊を加速し、補体タンパク質Ｃ４ｂのＩ因子媒介切断のための補助因子として作用する１５０ｋＤの可溶性血清タンパク質である。補体レセプター１型は肥満細胞と殆どの血球に見いだされる１９０－２８０ｋＤの膜結合タンパク質である。ＣＲ１は補体タンパク質Ｃ３ｂ、Ｃ４ｂ、及びｉＣ３ｂ相互作用してＣ３転換酵素の解離を加速し、Ｃ３ｂとＣ４ｂのＩ因子媒介切断のための補助因子として作用し、免疫複合体を結合し、その分解とファーゴサイトーシスを促進する。

補体タンパク質と相同性を有するタンパク質は特に医学界と産業界にとって興味深い。しばしば、互いに相同性を有するタンパク質は類似した機能を有している。相同性を有するタンパク質がいつ類似した機能を有していないことはまた興味深く、ある構造的モチーフが機能以外の情報、例えば機能の局所性を特定することを示している。

新規な天然の分泌及び膜結合タンパク質、特に補体経路に関連する既知のタンパク質と相同性を有するものを同定する努力が産業界と学術界の両方によってなされている。補体経

路に関連するタンパク質は、Ｂｉｒｍｉｎｇｈａｍ ＤＪ （１９９５）, Ｃｒｉｔｉｃａｌ Ｒｅｖｉｅｗｓ ｉｎ Ｉｍｍｕｎｏｌｏｇｙ, １５（２）：１３３－１５４及Ａｂｂａｓ ＡＫ等 （１９９４） Ｃｅｌｌｕｌａｒ ａｎｄ Ｍｏｌｅｃｕｌａｒ Ｉｍｍｕｎｏｌｏｇｙ, ２ｎｄ Ｅｄ. Ｗ. Ｂ. Ｓａｕｎｄｅｒｓ Ｃｏｍｐａｎｙ, Ｐｈｉｌａｄｅｌｐｈｉａ, ｐｐ ２９５－３１５において概説されている。

我々は、ここでＰＲＯ２２２と命名した、補体レセプターと相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。

【００２３】

２１．ＰＲＯ２３４

多細胞生物内の多くの系の成功裡の機能は細胞－細胞間相互作用に依存する。そのような相互作用はリガンド－レセプター結合、そして細胞－細胞接着を可能にする形で特定のレセプターに特定のリガンドをアラインメントさせることにより影響を受ける。細胞認識におけるタンパク質－タンパク質間相互作用が認識されてからしばらく経つが、生理的に関連した認識における炭水化物の役割は最近になって広く検討されるようになった（Ｂ. Ｋ. Ｂｒａｂｄｌｅｙ等, Ｊ. Ｌｅｕｋ. Ｂｉｏｌ. ４０：９７ （１９８６）及びＮ. Ｓｈａｒｏｎ等, Ｓｃｉｅｎｃｅ ２４６： ２２７ （１９８９）を参照されたい）。オリゴ糖はその細胞表面位置と構造の多様性のために新規な認識レクチンとして作用するようにうまく位置させられている。多くのオリゴ糖構造は少数のグリコシルトランスフェラーゼの差次的活性を通してつくられうる。オリゴ糖の多様な構造は比較的少ない遺伝子産物の転写によりつくられ、これが、オリゴ糖が広範囲の細胞－細胞間相互作用が生じるもっともらしい機構であることを示唆している。相互作用する細胞への細胞表面炭水化物と推定炭水化物結合タンパク質（レクチン）の差次的発現の例は記載されている（Ｊ. Ｄｏｄｄ及びＴ. Ｍ. Ｊｅｓｓｅｌ, Ｊ. Ｎｅｕｒｏｓｃｉ. ５： ３２７８ （１９８５）；Ｌ. Ｊ. Ｒｅｇａｎ等, Ｐｒｏｃ. Ｎａｔｌ. Ａｃａｄ. Ｓｃｉ. ＵＳＡ ８３： ２２４８ （１９８６）；Ｍ. Ｃｏｎｓｔａｎｔｉｎｅ－Ｐａｔｏｎ等, Ｎａｔｕｒｅ ３２４： ４５９ （１９８６）；及びＭ. Ｔｉｅｍｅｙｅｒ等, Ｊ. Ｂｉｏｌ. Ｃｈｅｍ. ２６３： １６７１ （１９８９）。レクチンファミリーの一つの興味深いメンバーはセレクチンである。

急性又は慢性の炎症部位までの白血球の遊走はこれらの細胞と内皮との間の接着相互作用を含む。この特異的接着は炎症性侵襲により開始されるカスケードの最初の事象であり、よって生物の調節された防御にとって極めて重要なものである。

炎症性応答の間に白血球と内皮の間の相互作用に関連している細胞接着分子のタイプは現在は４種に達する：（１）セレクチン；（２）セレクチンに対する（炭水化物及び糖タンパク質）リガンド；（３）インテグリン；及び（４）インテグリンリガンドで、免疫グロブリン遺伝子スーパーファミリーのメンバーであるものである。

セレクチンは構造的にも機能的にも統一化された細胞接着分子である。構造的には、セレクチンはカルシウム依存性レクチン（Ｃ－レクチン）と相同性を持つドメイン、上皮成長因子（ｅｇｆ）様ドメイン及び幾つかの補体結合様ドメインを含むことにより特徴づけられる。Ｂｅｖｉｌａｃｑｕａ, Ｍ. Ｐ. 等, Ｓｃｉｅｎｃｅ ２４３：１１６０－１１６５ （１９８９）；Ｊｏｈｎｓｔｏｎ等, Ｃｅｌｌ ５６： １０３３－１０４４ （１９８９）；Ｌａｓｋｙ等, Ｃｅｌｌ ５６： １０４５－１０５５ （１９８９）；Ｓｉｅｇａｌｍａｎ, Ｍ. 等, Ｓｃｉｅｎｃｅ ２４３： １１６５－１１７２ （１９８９）；Ｓｔｏｏｌｍａｎ, Ｌ. Ｍ., Ｃｅｌｌ ５６： ９０７－９１０ （１９８９）。

機能的には、セレクチンはそのレクチンドメインと細胞表面炭水化物リガンドの間の相互作用を通して細胞結合を媒介する能力という共通の性質を共有する（Ｂｒａｎｄｌｅｙ, Ｂ等, Ｃｅｌｌ ６３, ８６１－８６３ （１９９０）；Ｓｐｒｉｎｇｅｒ, Ｔ. 及びＬａｓｋｙ, Ｌ. Ａ., Ｎａｔｕｒｅ ３４９, １９－１９７ （１９９１）；Ｂｅｖｉｌａｃｑｕａ, Ｍ. Ｐ. 及びＮｅｌｓｏｎ, Ｒ. Ｗ., Ｊ. Ｃｌｉｎ. Ｉｎｖｅｓｔ. ９１ ３７９－３８７ （１９９３）及びＴｅｄｄｅｒ等, Ｊ. Ｅｘｐ. Ｍｅｄ. １７０： １２３－１３３ （１９８９））。

細胞接着分子のセレクチンファミリーにはこれまで３つのメンバーが同定されている：Ｌ－セレクチン（末梢リンパ節ホーミングレセプター（ｐｎＨＲ）とも呼ばれる，ＬＥＣ－ＣＡＭ－１，ＬＡＭ－１，ｇｐ９０^{ＭＥＬ}，ｇｐ１００^{ＭＥＬ}，ｇｐ１１０^{ＭＥＬ}，ＭＥＬ－１４抗原，Ｌｅｕ－８抗原，ＴＱ－１抗原，ＤＲＥＧ抗原）、Ｅ－セレクチン（ＬＥＣ－ＣＡＭ－２，ＬＥＣＡＭ－２，ＥＬＡＭ－１）及びＰ－セレクチン（ＬＥＣ－ＣＡＭ－３，ＬＥＣＡＭ－３，ＧＭＰ－１４０，ＰＡＤＧＥＭ）である。
Ｃ－レクチンドメインの同定はそのようなドメインを含むタンパク質に対する炭水化物結合リガンドを定める強い努力に至った。Ｅ－セレクチンは炭水化物配列ＮｅｕＮＡｃ．２－３Ｇａｌ．１－４（Ｆｕｃ．１－３）ＧｌｃＮＡｃ（シアリルールイスｘ、又はｓＬｅ^{x}）と関連オリゴ糖を認識すると考えられる。Ｂｅｒｇ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２６５：１４８６９－１４８７２（１９９１）；Ｌｏｗｅ等，Ｃｅｌｌ ６３：４７５－４８４（１９９０）；Ｐｈｉｌｌｉｐｓ等，Ｓｃｉｅｎｃｅ ２５０：１１３０－１１３２（１９９０）；Ｔｉｅｍｅｙｅｒ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８８：１１３８－１１４２（１９９１）。
Ｌ－セレクチンは、レクチンドメインを含むが、内皮細胞上の炭水化物含有リガンドを認識することによりその接着機能を遂行する。Ｌ－セレクチンは白血球、例えばリンパ球、好中球、単球及び好酸球の表面に発現され、末梢リンパ組織へのリンパ球の輸送（Ｇａｌｌａｔｉｎ等，Ｎａｔｕｒｅ ３０３：３０－３４（１９８３））と急性好中球媒介炎症性応答（Ｗａｔｓｏｎ，Ｓ．Ｒ．，Ｎａｔｕｒｅ ３４９：１６４－１６７（１９９１））に関与している。Ｌ－セレクチンのアミノ酸配列とコード核酸配列は例えば１９９２年3月２４日に発行の米国特許５，０９８，８３３に開示されている。
Ｌ－セレクチン（ＬＥＣＡＭ－１）は好中球流入を阻止するその能力のために特に興味深い（Ｗａｔｓｏｎ等，Ｎａｔｕｒｅ ３４９：１６４－１６７（１９９１））。それはＨＥＶに結合する慢性リンパ性白血病細胞に発現される（Ｓｐｅｒｔｉｎｉ等，Ｎａｔｕｒｅ ３４９：６９１－６９４（１９９１））。また、慢性炎症部位におけるＨＥＶ構造は例えばリウマチ関節炎、乾癬及び多発性硬化症のような疾患の徴候と関連している。
Ｅ－セレクチン（ＥＬＡＭ－１）は、ＩＬ－１又はＴＮＦへ応答して内皮細胞上で一過性に発現するため特に興味深い。Ｂｅｖｉｌａｃｑｕａ等，Ｓｃｉｅｎｃｅ ２４３：１１６０（１９８９）。この誘導された発現の時間過程（２－８ｈ）は感染と傷害に応答しての初期の好中球誘導血管外遊走におけるこのレセプターの役割を示唆している。抗ＥＬＡＭ－１抗体が霊長類喘息モデルにおいて好中球流入を阻止し、よって炎症性応答から生じる気道閉塞を防止するのに有用であることが更に報告されている。Ｇｕｎｄｅｌ等，Ｊ．Ｃｌｉｎ．Ｉｎｖｅｓｔ．８８：１４０７（１９９１）。
刺激された血管内皮への循環好中球の接着は炎症性応答の主要な事象である。Ｐ－セレクチンはルイスｘ構造（Ｇａｌ．１－４（Ｆｕｃ．１－３）ＧｌｃＮＡｃ）を認識すると報告されている。Ｌａｒｓｅｎ等，Ｃｅｌｌ ６３：４６７－４７４（１９９０）。更なる末端結合シアル酸が高い親和性結合に必要とされる。Ｍｏｏｒｅ等，Ｊ．Ｃｅｌｌ．Ｂｉｏｌ．１１２：４９１－４９９（１９９１）。Ｐ－セレクチンは急性肺傷害において重要であることが示されている。抗Ｐ－セレクチン抗体は齧歯類肺傷害モデルにおいて強い保護効果を有していることが示されている。Ｍ．Ｓ．Ｍｕｌｌｉｇａｎ等，Ｊ．Ｃｌｉｎ．Ｉｎｖｅｓｔ．９０：１６００（１９９１）。
我々はここにＰＲＯ２３４ポリペプチドと命名した、レクチンタンパク質と相同性を有する新規なポリペプチドの同定と特徴付けを記載する。
【００２４】
２２．ＰＲＯ２３１
細胞性プロセスの上述した調節と変調に最も重要なタンパク質の幾つかは細胞中におけるタンパク質のリン酸化のレベルを調節する酵素である。例えば、細胞成長と分化を調節するシグナルの伝達は様々な細胞性タンパク質のリン酸化と脱リン酸化により少なくとも部分的に調節されることが知られている。これらのプロセスを触媒する酵素は、様々な細胞

性タンパク質をリン酸化するように機能するプロテインキナーゼと、様々な細胞性タンパク質からリン酸残基を除去するように機能するプロテインホスファターゼを含む。細胞中でのタンパク質リン酸化のレベルのバランスはよってこれらの２種の酵素の相対活動により媒介される。
プロテインホスファターゼは、膜結合及び可溶型の双方を含む多くの多様な形態で見いだされる増大しているファミリーを表す。多くのプロテインホスファターゼが記載されているが、非常に僅かなものの機能が理解され始めた状態である（Ｔｏｎｋｓ，Ｓｅｍｉｎ．Ｃｅｌｌ Ｂｉｏｌ．４：３７３－４５３（１９９３）及びＤｉｘｏｎ，Ｒｅｃｅｎｔ Ｐｒｏｇ．Ｈｏｒｍ．Ｒｅｓ．５１：４０５－４１４（１９９６））。しかし、一般には多くのプロテインホスファターゼ様々なプロテインキナーゼにより誘導される正又は負のシグナルを変調するように機能するようである。従って、プロテインホスファターゼは数多くの多様な細胞性プロセスにおいて重要な役割を果たしている可能性が高い。
プロテインホスファターゼの生理学的な重要性が分かっているので、新規な天然のホスファターゼタンパク質を同定するために産業界と学術界の両方により努力がなされている。これらの努力の多くは新規なホスファターゼタンパク質に対するコード化配列を同定するために哺乳動物の組換えＤＮＡライブラリのスクリーニングに集中している。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．９３；７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。
我々はここにＰＲＯ２３１ポリペプチドと命名した、酸性ホスファターゼと相同性を有する新規なポリペプチドの同定と特徴付けを記載する。
【００２５】
２３．ＰＲＯ２２９
スカベンジャーレセプターは異化分解からＩｇＧ分子を保護することが知られている。Ｒｉｅｃｈｍａｎｎ及びＨｏｌｌｉｎｇｅｒ，Ｎａｔｕｒｅ Ｂｉｏｔｅｃｈｎｏｌｏｇｙ，１５：６１７（１９９７）。特に、ＣＨ２及びＣＨ３ドメインの研究は、これらのドメインの特異的配列が抗体の半減期を決定する際に重要であることを示している。Ｅｌｌｅｒｓｏｎ等，Ｊ．Ｉｍｍｕｎｏｌ．，１１６：５１０（１９７６）；Ｙａｓｍｅｅｎ等，Ｊ．Ｉｍｍｕｎｏｌ．１１６：５１８（１９７６；Ｐｏｌｌｏｃｋ等，Ｅｕｒ．Ｊ．Ｉｍｍｕｎｏｌ．，２０：２０２１（１９９０））。スカベンジャーレセプタータンパク質とそれに対する抗体は更にＫｒｉｅｇｅｒ等の米国特許５，５１０，４６６に報告されている。ポリペプチドの半減期を増大させるスカベンジャーレセプターの能力と免疫機能へのその関与のために、スカベンジャーレセプターと相同性を有する分子は科学界及び医学界にとって重要である。
新規な天然の分泌及び膜結合タンパク質、特にスカベンジャーレセプターと相同性を有するものを同定する努力が産業界と学術界の両方によってなされている。これらの努力の多くは新規な分泌及び膜結合レセプタータンパク質に対するコード化配列を同定するために哺乳動物の組換えＤＮＡライブラリのスクリーニングに集中している。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．９３；７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。
我々は、ここでＰＲＯ２２９ポリペプチドと命名した、スカベンジャーレセプターと相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。
【００２６】
２４．ＰＲＯ２３８
酸素フリーラジカルと酸化防止剤は脳虚血及び再灌流後における中枢神経系に重要な役割を果たしていると思われる。更に、虚血及び再灌流に関連した心臓損傷がフリーラジカルの作用によって引き起こされることが報告されている。また、細胞のレドックス状態は細胞運命の中心的な決定要因であることが研究により報告されている。更に、反応性酸素種

が、細胞毒性であり、組織壊死、臓器不全、アテローム性動脈硬化症、不妊症、出生欠損、早老、突然変異、悪性腫瘍を含む、炎症性疾患を引き起こすことが報告されている。しかして、酸化と還元の制御は、脳卒中、心臓発作、酸化ストレス、高血圧症の制御と防止を含む多くの理由から重要である。この点において、レダクターゼ、特にオキシドレダクターゼは興味深い。この主題事項を更に記述している刊行物には、Ｋｅｌｓｅｙ等，Ｂｒ．Ｊ．Ｃａｎｃｅｒ，７６（７）：８５２－８５４（１９９７）；Ｆｒｉｅｄｒｉｃｈ及びＷｅｉｓｓ，Ｊ．Ｔｈｅｏｒ．Ｂｉｏｌ．，１８７（４）：５２９－４０（１９９７）及びＰｉｅｕｌｌｅ等，Ｊ．Ｂａｃｔｅｒｉｏｌ．，１７９（１８）：５６８４－９２（１９９７）が含まれる。

新規な天然の分泌及び膜結合レセプタータンパク質、特にレダクターゼと相同性を有する分泌タンパク質を同定する努力が産業界と学術界の双方によってなされている。多くの努力が、新規な分泌及び膜結合レセプタータンパク質のコード配列を同定するために哺乳類組換えＤＮＡライブラリーのスクリーニングに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，９３：７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。

我々は、ここでＰＲＯ２３８ポリペプチドと命名した、レダクターゼと相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。

【００２７】

２５．ＰＲＯ２３３

細胞のレドックス状態は細胞の運命の重要な決定因子であることが、研究において報告されている。更に、反応性酸素種は、組織壊死、器官機能不全、アテローム性動脈硬化、不妊症、出生欠損症、時期尚早の加齢、変異及び悪性腫瘍を含む炎症性疾患の原因となる細胞毒性であると報告されている。よって、酸化と還元の制御は、脳卒中、心臓発作、酸化ストレス、高血圧の制御及び予防を含む、多くの理由から重要である。酸素フリーラジカルと酸化防止剤は脳虚血及び再灌流後における中枢神経系に重要な役割を果たしていると思われる。更に、虚血及び再灌流に関連した心臓損傷がフリーラジカルの作用によって引き起こされることが報告されている。この点に関して、レダクターゼ、特にオキシドレダクターゼが興味深い。加えて、転写因子、ＮＦ－ｋａｐｐａ Ｂ及びＡＰ－１は酸化還元状態により調節され、エイズ、ガン、アテローム性動脈硬化及び糖尿病の合併症の病因に関与すると考えられる種々の遺伝子の発現に影響を及ぼすことが知られている。さらに、この主題事項について記載した出版物には、Ｋｅｌｓｅｙ等，Ｂｒ．Ｊ．Ｃａｎｃｅｒ，７６（７）：８５２－４（１９９７）；Ｆｒｉｅｄｒｉｃｈ及びＷｅｉｓｓ．Ｊ．Ｔｈｅｏｒ．Ｂｉｏｌ．，１８７（４）：５２９－４０（１９９７）及びＰｉｅｕｌｌｅ等，Ｊ．Ｂａｃｔｅｒｉｏｌ．，１７９（１８）：５６８４－９２（１９９７）が含まれる。インビボにおけるレドックス反応の生理学的な重要性が分かっているため、レドックス反応に関与している新規な天然のタンパク質を同定するための努力がなされている。我々は、ここでＰＲＯ２３３ポリペプチドと命名した、レダクターゼと相同性を有する新規なポリペプチドの同定についてここに記載する。

【００２８】

２６．ＰＲＯ２２３

エキソペプチダーゼのカルボキシペプチダーゼファミリーはポリペプチドにおけるカルボキシ末端アミド結合を加水分解する酵素の多様な群を構成し、非常に多くの哺乳動物組織がこれらの酵素をつくりだす。これまでに同定されているカルボキシペプチダーゼの多くはポリペプチドのあるアミノ酸に対してかなり強い切断特異性を示す。例えば、リジン、アルギニン、セリン又はポリペプチドのカルボキシ末端に基質として芳香族か分岐状脂肪族側鎖の何れかを持つアミノ酸を好むカルボキシペプチダーゼ酵素が同定されている。

セリンカルボキシペプチダーゼに関しては、そのようなアミノ酸特異的酵素は様々な異なったほ乳類及び非哺乳類生物から同定されている。哺乳類セリンカルボキシペプチダーゼ酵素は、例えばタンパク消化、活性化、不活性化、又はペプチドホルモン活性の変調、及

びタンパク質と酵素の物理特性の変更を含む多くの異なった生物学的プロセスにおいて重要な役割を果たしている。
セリンカルボキシペプチダーゼの生理学的な重要性に鑑みて、新規な天然の分泌及び膜結合レセプタータンパク質、特に新規なカルボキシペプチダーゼを同定する努力が産業界と学術界の双方によってなされている。多くの努力が、新規な分泌及び膜結合レセプタータンパク質のコード配列を同定するために哺乳類組換えＤＮＡライブラリーのスクリーニングに注がれている。我々は、ここでＰＲＯ２２３ポリペプチドと命名した、一又は複数のセリンカルボキシペプチダーゼポリペプチドと相同性を有する新規なポリペプチドをここに記載する。
【００２９】
２７．ＰＲＯ２３５
プレキシンは、カルシウムイオンの存在下で同種親和性結合機構を介して細胞接着を媒介することが知られている膜糖タンパク質としてアフリカツメガエルオタマジャクシの神経系中において最初に同定された。アフリカツメガエル、マウス及びヒト相同体の間でのプレキシンの強い進化的保存が観察されている。［Ｋａｎｅｙａｍａ等，Ｂｉｏｃｈｅｍ．Ａｎｄ Ｂｉｏｐｈｙｓ．Ｒｅｓ．Ｃｏｍｍ．２２６：５２４－５２９（１９９６）］。インビボでの細胞接着機構の生理学的重要性が分かっているので、細胞接着に関与する新規な天然のタンパク質を同定する努力が現在なされている。我々は、ここでＰＲＯ２３５ポリペプチドと命名した、プレキシンと相同性を有する新規なポリペプチドの同定をここに記載する。
【００３０】
２８．ＰＲＯ２３６及びＰＲＯ２６２
β－ガラクトシダーゼは、β－ガラクトシド分子を加水分解するように機能するよく知られた酵素タンパク質である。β－ガラクトシダーゼはインビトロとインビボの両方で様々な異なった用途に対して用いられており、極めて有用な研究ツールであることが分かっている。しかして、β－ガラクトシダーゼポリペプチドと相同性を示す新規なポリペプチドを得ることには興味がある。
β－ガラクトシダーゼと相同性を有する新規なポリペプチドを得ることに強い興味があるので、新規な天然のβ－ガラクトシダーゼ相同体タンパク質を同定するための努力が産業界と学術界の双方によってなされている。これらの努力の多くは、新規なβ－ガラクトシダーゼ様タンパク質のコード配列を同定するために哺乳類組換えＤＮＡライブラリーのスクリーニングに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，９３：７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。我々は、ここでＰＲＯ２３６及びＰＲＯ２６２ポリペプチドと命名した、β－ガラクトシダーゼ酵素と有意な相同性を有する新規なポリペプチドをここに記載する。
【００３１】
２９．ＰＲＯ２３９
デンシンは、接着分子の全ての特徴を有する脳から単離された糖タンパク質である。それは、脳のシナプス部位に非常に集中しており、発達中のニューロン中の樹枝状突起において顕著に発現している。デンシンはＯ－結合シアロ糖タンパク質のメンバーとして特徴づけられている。デンシンは再生のような医学的に重要なプロセスと関連性を有している。シナプス突起とインビボでの細胞接着機構の生理学的重要性が分かっているので、シナプス機構と細胞接着に関与する新規な天然のタンパク質を同定する努力が現在なされている。我々は、ここでＰＲＯ２３９ポリペプチドと命名した、デンシンと相同性を有する新規なポリペプチドの同定をここに記載する。
【００３２】
３０．ＰＲＯ２５７
エブネリン（Ｅｂｎｅｒｉｎ）は、哺乳動物のフォン・エブナー腺に関連する細胞表面タンパク質である。新規で天然の細胞表面レセプタータンパク質、特にエブネリンのような

細胞表面タンパク質と配列相同性を有するものを同定する努力が産業界と学術界の双方によってなされている。これらの努力の多くは新規なレセプタータンパク質に対するコード化配列を同定するために哺乳動物の組換えＤＮＡライブラリをスクリーニングすることに注がれている。我々は、ここでＰＲＯ２５７ポリペプチドと命名した、フォン・エブナー腺関連タンパク質エブネリンと相同性を有する新規なポリペプチドの同定をここに記載する。

【００３３】

３１．ＰＲＯ２６０

フコシダーゼはプロテオグリカンを含むフコースからフコース残基を取り除く酵素である。例えばガン、関節リウマチ、及び糖尿病のようなある種の病理状態では、血清タンパク質の異常なフコシル化がある。従って、フコシダーゼ及びフコシダーゼと相同性を有するタンパク質はこれらの症状の研究と抑止に対して重要である。特に、α－Ｉ－フコシダーゼ前駆体と相同性を有するタンパク質が興味深い。フコシダーゼとフコシダーゼ阻害剤は更に米国特許５，６３７，４９０、５，３８２，７０９、５，２４０，７０７、５，１５３，３２５、５，１００，７９７、５，０９６，９０９及び５，０１７，７０４に記載されている。研究はまたＶａｌｋ等，Ｊ．Ｖｉｒｏｌ．，７１（９）：６７９６（１９９７），Ａｋｔｏｇｕ等，Ｍｏｎａｌｄｉ．Ａｒｃｈ．Ｃｈｅｓｔ Ｄｉｓ．（Ｉｔａｌｙ），５２（２）：１１８（１９９７）及びＦｏｃａｒｅｌｌｉ等，Ｂｉｏｃｈｅｍ．Ｂｉｏｐｈｙｓ．Ｒｅｓ．Ｃｏｍｍｕｎ．（Ｕ．Ｓ．），２３４（１）：５４（１９９７）においても報告されている。

新規な天然の分泌及び膜結合レセプタータンパク質を同定する努力が産業界と学術界の双方によってなされている。多くの努力が新規な分泌及び膜結合レセプタータンパク質に対するコード化配列を同定するために哺乳動物の組換えＤＮＡライブラリをスクリーニングすることに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，９３：７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。

我々は、ここでＰＲＯ２６０ポリペプチドと命名した、フコシダーゼと相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。

【００３４】

３２．ＰＲＯ２６３

ＣＤ４４は細胞－細胞及び細胞－マトリックス相互作用に関与する細胞表面接着分子である。細胞外真トリックスの成分であるヒアルロン酸は主要なリガンドである。他のリガンドにはコラーゲン、フィブロネクチン、ラミニン、コンドロイチン硫酸、粘膜アドレシン、セルグリチン及びオステオポニンが含まれる。ＣＤ４４はまた細胞輸送の調節、リンパ節ホーミング、成長シグナルの伝達、及び移動細胞への成長因子及びケモカインの提示において重要である。ＣＤ４４表面タンパク質は転移性腫瘍に関連し、ＣＤ４４はＨＩＶ感染に対するマーカーとして使用されている。ある種のスプライスバリアントはガン患者の転移と不十分な予後に関連している。違って、ＣＤ４４と相同性を有する分子は、その相同性が、それらがＣＤ４４の機能に関連した機能を有しているであろうということを示しているので、特に興味深い。ＣＤ４４は更に米国特許５，５０６，１１９、５，５０４，１９４及び５，１０８，９０４；Ｇｅｒｂｅｒｉｃｋ等，Ｔｏｘｉｃｏｌ．Ａｐｐｌ．Ｐｈａｒｍａｃｏｌ．，１４６（１）：１（１９９７）；Ｗｉｔｔｉｇ等，Ｉｍｍｕｎｏｌ．Ｌｅｔｔｅｒｓ（Ｎｅｔｈｅｒｌａｎｄｓ），５７（１－３）：２１７（１９９７）；及びＯｌｉｖｅｉｒａとＯｄｅｌｌ，Ｏｒａｌ Ｏｎｃｏｌ．（Ｅｎｇｌａｎｄ），３３（４）：２６０（１９９７）に記載されている。

新規な天然の分泌及び膜結合レセプタータンパク質、特にＣＤ４４抗原と相同性を有する膜貫通タンパク質を同定するための努力が産業界と学術界の両方によってなされている。多くの努力が新規な分泌及び膜結合レセプタータンパク質に対するコード化配列を同定するために哺乳動物の組換えＤＮＡライブラリをスクリーニングすることに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐ

ｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，９３：７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。
我々は、ここでＰＲＯ２６３ポリペプチドと命名した、ＣＤ４４抗原と相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。
【００３５】
３３．ＰＲＯ２７０
チオレドキシンは還元－酸化（レドックス）状態をもたらす。多くの疾患は潜在的にレドックス状態に関連しており、反応性酸素種が多くの重要な生物学的プロセスにおいて役割を担っている。転写因子であるＮＫ－ｋａｐｐａ Ｂ及びＡＰ－１はレドックス状態により調節され、ＡＩＤＳ、ガン、アテローム性動脈硬化症及び糖尿病合併症の病状に関与していると思われる非常に多彩な遺伝子の発現に影響を及ぼすことが知られている。そのようなタンパク質は、アポトーシスにおいて役割を担っているのに加えて、細胞性抗酸化防御と、酸化ストレスに関与する病状、例えば脳卒中及び炎症においても役割を担っている。従って、チオレドキシンとそれに対して相同性を有するタンパク質は科学界と医学界にとって興味深い。
我々は、ここでＰＲＯ２７０ポリペプチドと命名した、チオレドキシンと相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。
【００３６】
３４．ＰＲＯ２７１
プロテオグリカン連結タンパク質は、様々な細胞外マトリックスタンパク質、より特定的にはコラーゲンのようなタンパク質と密接に関連しているタンパク質である。例えば、コラーゲンの一つの主要な成分はアグレカン（ａｇｇｒｅｃａｎ）と呼ばれる大きなプロテオグリカンである。この分子はコアタンパク質のアミノ末端Ｇ１球状ドメインを通してグリコサミノグリカンヒアルロナンに結合することにより保持される。この結合はベルシカンのようなヒアルロナン結合プロテオグリカンを含む他の組織ともまた関連しているタンパク質であるプロテオグリカン関連タンパク質により安定化される。
連結タンパク質は若年性関節リウマチを患っている個人の自己免疫抗体に対する潜在的な標的として同定された（Ｇｕｅｒａｓｓｉｍｏｖ等，Ｊ．Ｒｈｅｕｍａｔｏｌｏｇｙ ２４（５）：９５９－９６４（１９９７）を参照）。しかして、連結タンパク質と相同性を有する新規なタンパク質を同定する点には強い興味がある。我々は、ここでＰＲＯ２７１ポリペプチドと命名した、かかる相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。
【００３７】
３５．ＰＲＯ２７２
レティキュロカルビン（ｒｅｔｉｃｕｌｏｃａｌｂｉｎ）はタンパク質輸送及び管腔タンパク質プロセシングに関与する小胞体タンパク質である。レティキュロカルビンは小胞体管腔に存在し、カルシウムを結合することが知られ、小胞体の管腔保持機構に関与している。それは高親和性カルシウム結合と関連したＥＦ－ハンドモチーフの６つのドメインを含んでいる。我々は、ここでＰＲＯ２７２と命名した、レティキュロカルビンと相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。
【００３８】
３６．ＰＲＯ２９４
天然に生じるタンパク質であるコラーゲンは産業において広い利用性を有している。化学的に加水分解された天然のコラーゲンは加熱及び冷却により変性させまた再生させ、他の用途のなかでも写真及び医学に使用されるゼラチンをつくることができる。コラーゲンは三重螺旋体と命名されている高次構造を持つ鎖間凝集体を形成する能力のような重要な性質を有している。我々は、ここでＰＲＯ２９４と命名した、コラーゲン分子の一部と相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。
【００３９】
３７．ＰＲＯ２９５

インテグリンは細胞接着を促進する細胞表面糖タンパク質レセプターのスーパーファミリーからなる。各細胞はその細胞接着能力を形成する数多くのレセプターを有している。インテグリンは細胞と他の細胞又はマトリックス成分との間の広範な相互作用に関与する。インテグリンは免疫系細胞の移動と機能の調節において特に重要である。血小板ＩＩｂ／ＩＩＩＡインテグリン複合体は血小板凝集の調節において特に重要である。インテグリンファミリーのメンバーであるインテグリンβ－６は、上皮細胞上に発現され上皮炎症を変調する。他のインテグリンである白血球随伴抗原－１（ＬＦＡ－１）は免疫応答の間のリンパ球の接着において重要である。インテグリンは非共有結合的に結合したα及びβサブユニットのヘテロダイマーとして発現される。インビボでの細胞接着機構の生理学的な重要性が分かっているので、細胞接着に関与する新規な天然のタンパク質を同定する努力が現在なされている。我々はここでＰＲＯ２９５と命名する、インテグリンと相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。

【００４０】

３８．ＰＲＯ２９３

タンパク質－タンパク質相互作用はレセプター及び抗原複合体及びシグナル伝達機構に関連するものを含む。タンパク質－タンパク質相互作用の基となる構造的及び機能的機構についてよく知られているように、タンパク質－タンパク質相互作用は、タンパク質－タンパク質相互作用の特定の結果を調節するように更に容易に操作することができる。従って、タンパク質－タンパク質相互作用の基となる機構は科学及び医学界にとって興味深い。ロイシンリッチ反復を含む全てのタンパク質はタンパク質－タンパク質相互作用に関与していると考えられている。ロイシンリッチ反復は多様な機能を持つ多くのタンパク質と細胞位置に存在する短い配列モチーフである。リボ核酸インヒビタータンパク質の結晶構造から、ロイシンリッチ反復はβ－α構造単位に対応することが明らかになった。これらの単位は、一面が溶媒に暴露された平行なβシートを形成するように配置され、タンパク質は珍しい非球状の形状を獲得する。これらの二つの特徴はロイシンリッチ反復を含むタンパク質のタンパク質結合機能の原因であることが示されている。Ｋｏｂｅ及びＤｅｉｓｅｎｈｏｆｅｒ，Ｔｒｅｎｄｓ Ｂｉｏｃｈｅｍ．Ｓｃｉ．，１９（１０）：４１５－４２１（Ｏｃｔ．１９９４）を参照されたい。

個体発生の間にコラーゲン原線維を配向し指令する組織オーガナイザーとなり、創傷治癒、組織修復、及び腫瘍ストローマ形成のような病理プロセスに関与するロイシンリッチプロテオグリカンについて研究が報告されている。Ｉｏｚｚｏ，Ｒ．Ｖ．，Ｃｒｉｔ．Ｒｅｖ．Ｂｉｏｃｈｅｍ．Ｍｏｌ．Ｂｉｏｌ．，３２（２）：１４１－１７４（１９９７）。創傷治癒、組織修復におけるロイシンリッチタンパク質の関与を示す他の研究は、Ｄｅ Ｌａ Ｓａｌｌｅ，Ｃ．等，Ｖｏｕｖ．Ｒｅｖ．Ｆｒ．Ｈｅｍａｔｏｌ．（Ｇｅｒｍａｎｙ），３７（４）：２１５－２２２（１９９５）であり、出血疾患ベルナール－スーリエ症候群に伴う複合体におけるロイシンリッチモチーフの突然変異を報告しており、Ｃｈｌｅｍｅｔｓｏｎ，Ｋ．Ｊ．，Ｔｈｒｏｍｂ．Ｈａｅｍｏｓｔ．（Ｇｅｒｍａｎｙ），７４（１）：１１１－１１６（Ｊｕｌｙ １９９５）は、血小板がロイシンリッチ反復を有していることを報告している。ロイシンリッチ反復を有することが報告されている特に興味深い他のタンパク質は、アルツハイマー病のような神経変性疾患、パーキンソン病のような神経損傷の治療において、及びガンの診断に対して有用であることが報告されているＳＬＩＴタンパク質である。エール大学のＡｒｔａｖａｎｉｓｔｓａｋｏｎａｓ，Ｓ．及びＲｏｔｈｂｅｒｇ，Ｊ．Ｍ．のＷＯ９２１０５１８－Ａ１を参照されたい。ロイシンリッチ反復を有するタンパク質の生物学的機能を報告している他の研究には：Ｔａｙａｒ，Ｎ．等，Ｍｏｌ．Ｃｅｌｌ Ｅｎｄｏｃｒｉｎｏｌ．，（Ｉｒｅｌａｎｄ），１２５（１－２）：６５－７０（Ｄｅｃ．１９９６）（ゴナドトロピンレセプター関連）；Ｍｉｕｒａ，Ｙ．等，Ｎｉｐｐｏｎ Ｒｉｎｓｈｏ（Ｊａｐａｎ），５４（７）：１７８４－１７８９（Ｊｕｌｙ １９９６）（アポトーシス関連）；Ｈａｒｒｉｓ，Ｐ．Ｃ．等，Ｊ．Ａｍ．Ｓｏｃ．Ｎｅｐｈｒｏｌ．，６（４）：１１２５－１１３３（Ｏｃｔ．１９９５）（腎疾患関連）；及びＬａ Ｊｏｌｌａ Ｃａｎｃｅｒ Ｒ

ｅｓｅａｒｃｈ ＦｏｕｎｄａｔｉｏｎのＲｕｏｓｌａｈｔｉ，Ｅ．Ｉ．等のＷＯ９１０７２７－Ａ（トランスフォーミング成長因子βに結合するデコリンの、ガンの治療、創傷治癒及び瘢痕化への関与）が含まれる。
よって、タンパク質－タンパク質相互作用をよりよく理解するためにロイシンリッチ反復を有する新規なタンパク質を同定するための努力が産業界と学術界の両方によってなされている。特に興味があるものはロイシンリッチ反復と、既知のニューロンのロイシンリッチ反復タンパク質と相同性を有するタンパク質である。多くの努力がロイシンリッチ反復を有する新規な分泌及び膜結合レセプタータンパク質に対するコード化配列を同定するために哺乳動物の組換えＤＮＡライブラリをスクリーニングすることに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，９３：７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。
我々は、ここでＰＲＯ２９３ポリペプチドと命名した、ロイシンリッチ反復タンパク質と相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。
【００４１】
３９．ＰＲＯ２４７
タンパク質－タンパク質相互作用はレセプター及び抗原複合体及びシグナル伝達機構に関連するものを含む。タンパク質－タンパク質相互作用の基となる構造的及び機能的機構についてよく知られているように、タンパク質－タンパク質相互作用は、タンパク質－タンパク質相互作用の特定の結果を調節するように更に容易に操作することができる。従って、タンパク質－タンパク質相互作用の基となる機構は科学及び医学界にとって興味深い。ロイシンリッチ反復を含む全てのタンパク質はタンパク質－タンパク質相互作用に関与していると考えられている。ロイシンリッチ反復は多様な機能を持つ多くのタンパク質と細胞位置に存在する短い配列モチーフである。リボ核酸インヒビタータンパク質の結晶構造から、ロイシンリッチ反復はβ－α構造単位に対応することが明らかになった。これらの単位は、一面が溶媒に暴露された平行なβシートを形成するように配置され、タンパク質は珍しい非球状の形状を獲得する。これらの二つの特徴はロイシンリッチ反復を含むタンパク質のタンパク質結合機能の原因であることが示されている。Ｋｏｂｅ及びＤｅｉｓｅｎｈｏｆｅｒ，Ｔｒｅｎｄｓ Ｂｉｏｃｈｅｍ．Ｓｃｉ．，１９（１０）：４１５－４２１（Ｏｃｔ．１９９４）を参照されたい。
個体発生の間にコラーゲン原線維を配向し指令する組織オーガナイザーとなり、創傷治癒、組織修復、及び腫瘍ストローマ形成のような病理プロセスに関与するロイシンリッチプロテオグリカンについて研究が報告されている。Ｉｏｚｚｏ，Ｒ．Ｖ．，Ｃｒｉｔ．Ｒｅｖ．Ｂｉｏｃｈｅｍ．Ｍｏｌ．Ｂｉｏｌ．，３２（２）：１４１－１７４（１９９７）。創傷治癒、組織修復におけるロイシンリッチタンパク質の関与を示す他の研究は、Ｄｅ Ｌａ Ｓａｌｌｅ，Ｃ．等，Ｖｏｕｖ．Ｒｅｖ．Ｆｒ．Ｈｅｍａｔｏｌ．（Ｇｅｒｍａｎｙ），３７（４）：２１５－２２２（１９９５）であり、出血疾患ベルナール－スーリエ症候群に伴う複合体におけるロイシンリッチモチーフの突然変異を報告しており、Ｃｈｌｅｍｅｔｓｏｎ，Ｋ．Ｊ．，Ｔｈｒｏｍｂ．Ｈａｅｍｏｓｔ．（Ｇｅｒｍａｎｙ），７４（１）：１１１－１１６（Ｊｕｌｙ １９９５）は、血小板がロイシンリッチ反復を有していることを報告している。ロイシンリッチ反復を有することが報告されている特に興味深い他のタンパク質は、アルツハイマー病のような神経変性疾患、パーキンソン病のような神経損傷の治療において、及びガンの診断に対して有用であることが報告されているＳＬＩＴタンパク質である。エール大学のＡｒｔａｖａｎｉｓｔｓａｋｏｎａｓ，Ｓ．及びＲｏｔｈｂｅｒｇ，Ｊ．Ｍ．のＷＯ９２１０５１８－Ａ１を参照されたい。ロイシンリッチ反復を有するタンパク質の生物学的機能を報告している他の研究には：Ｔａｙａｒ，Ｎ．等，Ｍｏｌ．Ｃｅｌｌ Ｅｎｄｏｃｒｉｎｏｌ．，（Ｉｒｅｌａｎｄ），１２５（１－２）：６５－７０（Ｄｅｃ．１９９６）（ゴナドトロピンレセプター関連）；Ｍｉｕｒａ，Ｙ．等，Ｎｉｐｐｏｎ Ｒｉｎｓｈｏ（Ｊａｐａｎ），５４（７）：１７８４－１７８９（Ｊｕｌｙ １９９６）（アポトーシス関連）；Ｈａｒｒ

ｉｓ，Ｐ．Ｃ．等，Ｊ．Ａｍ．Ｓｏｃ．Ｎｅｐｈｒｏｌ．，６（４）：１１２５－１１３３（Ｏｃｔ．１９９５）（腎疾患関連）；及びＬａ Ｊｏｌｌａ Ｃａｎｃｅｒ Ｒｅｓｅａｒｃｈ ＦｏｕｎｄａｔｉｏｎのＲｕｏｓｌａｈｔｉ，Ｅ．Ｉ．等のＷＯ９１０７２７－Ａ（トランスフォーミング成長因子βに結合するデコリンの、ガンの治療、創傷治癒及び瘢痕化への関与）が含まれる。

デンシンは、接着分子の全ての特徴を有する脳から単離された糖タンパク質である。それは、脳のシナプス部位に非常に集中しており、発達中のニューロン中の樹枝状突起において顕著に発現している。デンシンはＯ－結合シアロ糖タンパク質のメンバーとして特徴づけられている。デンシンは再生のような医学的に重要なプロセスと関連性を有している。シナプス突起とインビボでの細胞接着機構の生理学的重要性が分かっているので、シナプス機構と細胞接着に関与する新規な天然のタンパク質を同定する努力が現在なされている。デンシンは更にＫｅｎｎｅｄｙ，Ｍ．Ｂ，Ｔｒｅｎｄｓ Ｎｅｕｒｏｓｃｉ．（Ｅｎｇｌａｎｄ），２０（６）：２６４（１９９７）及びＡｐｐｅｒｓｏｎ等，Ｊ．Ｎｅｕｒｏｓｃｉ．，１６（２１）：６８３９（１９９６）に記載されている。

よって、タンパク質－タンパク質相互作用をよりよく理解するためにロイシンリッチ反復を有する新規なタンパク質を同定するための努力が産業界と学術界の両方によってなされている。特に興味があるものはロイシンリッチ反復と、ＫＩＡＡ０２３１及びデンシンのようなロイシンリッチ反復を有する既知のタンパク質と相同性を有するタンパク質である。多くの努力がロイシンリッチ反復を有する新規な分泌及び膜結合レセプタータンパク質に対するコード化配列を同定するために哺乳動物の組換えＤＮＡライブラリをスクリーニングすることに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，９３：７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。

我々は、ここでＰＲＯ２４７と命名した、ロイシンリッチ反復タンパク質と相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。

【００４２】

４０．ＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３

プロテアーゼは哺乳類及び非哺乳類生物における多数の非常に重要な生物学的プロセスに関与している酵素タンパク質である。様々な異なる哺乳類及び非哺乳類生物からの多くの異なるプロテアーゼ酵素が同定され、特徴づけられている。哺乳類のプロテアーゼ酵素は、例えばタンパク質の消化、活性化、不活性化、又はペプチドホルモン活性の変調、及びタンパク質と酵素の物理特性の改変を含む多くの異なった生物学的プロセスにおいて重要な役割を担っている。

プロテアーゼ酵素が担っている重要な生理学的役割に鑑みて、新規な天然のプロテアーゼ相同体を同定するための努力が、産業界と学術界の両方によって現在なされている。これらの努力の多くは、新規な分泌及び膜結合レセプタータンパク質のコード配列を同定するための哺乳類の組換えＤＮＡライブラリーのスクリーニングに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，９３；７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照］。我々は、ここでＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３ポリペプチドと命名した、様々なプロテアーゼ酵素に対して相同性を有する新規なポリペプチドの同定についてここに記載する。

【００４３】

４１．ＰＲＯ３２８

ＧＬＩＰタンパク質ファミリーは胚形成において重要な役割を担うジンクフィンガータンパク質を含むものとして特徴づけられている。これらのタンパク質は転写制御タンパク質として機能しえ、ヒト腫瘍のサブセットにおいて増幅されることが知られている。神経膠腫病原性タンパク質は一群の植物病原性関連タンパク質と構造的に関連している。それは膠芽細胞腫において非常に発現される。米国特許５，５８２，９８１（１９９６年１２月１０日発行）及び５，３２２，８０１（１９９６年６月２１日発行）、Ｅｌｌｉｎｇｔｏ

ｎ，Ａ．Ｄ．等，Ｎａｔｕｒｅ，３４６：８１８（１９９０），Ｇｒｉｎｄｌｅｙ，Ｊ．Ｃ．等，Ｄｅｖ．Ｂｉｏｌ．，１８８（２）：３３７（１９９７），Ｍａｒｉｎｅ，Ｊ．Ｃ．等，Ｍｅｃｈ．Ｄｅｖ．，６３（２）：２１１（１９９７）を参照されたい。ＣＲＩＳＰ又はシステインリッチ分泌タ性ンパク質ファミリーもまた一群の植物病原性関連タンパク質と構造的に関連している一群のタンパク質である。［Ｓｃｈｗｉｄｅｔｚｋｙ，Ｕ．，Ｂｉｏｃｈｅｍ．Ｊ．，３２１：３２５（１９９７），Ｐｆｉｓｔｅｒｅｒ，Ｐ．，Ｍｏｌ．Ｃｅｌｌ Ｂｉｏｌ．，１６（１１）：６１６０（１９９６），Ｋｒａｔｚｓｃｈｍａｒ，Ｊ．，Ｅｕｒ．Ｊ．Ｂｉｏｃｈｅｍ．，２３６（３）：８２７（１９９６）］。我々は、ここでＰＲＯ３２８ポリペプチドと命名した、ＧＬＩＰ及びＣＲＩＳＰと相同性を有する新規なポリペプチドの同定についてここに記載する。

【００４４】

４２．ＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６

タンパク質－タンパク質相互作用はレセプター及び抗原複合体及びシグナル伝達機構に関連するものを含む。タンパク質－タンパク質相互作用の基となる構造的及び機能的機構についてよく知られているように、タンパク質－タンパク質相互作用は、タンパク質－タンパク質相互作用の特定の結果を調節するように更に容易に操作することができる。従って、タンパク質－タンパク質相互作用の基となる機構は科学及び医学界にとって興味深い。ロイシンリッチ反復を含む全てのタンパク質はタンパク質－タンパク質相互作用に関与していると考えられている。ロイシンリッチ反復は多様な機能を持つ多くのタンパク質と細胞位置に存在する短い配列モチーフである。リボ核酸インヒビタータンパク質の結晶構造から、ロイシンリッチ反復はβ－α構造単位に対応することが明らかになった。これらの単位は、一面が溶媒に暴露された平行なβシートを形成するように配置され、タンパク質は珍しい非球状の形状を獲得する。これらの二つの特徴はロイシンリッチ反復を含むタンパク質のタンパク質結合機能の原因であることが示されている。Ｋｏｂｅ及びＤｅｉｓｅｎｈｏｆｅｒ，Ｔｒｅｎｄｓ Ｂｉｏｃｈｅｍ．Ｓｃｉ．，１９（１０）：４１５－４２１（Ｏｃｔ．１９９４）を参照されたい。

個体発生の間にコラーゲン原線維を配向し指令する組織オーガナイザーとなり、創傷治癒、組織修復、及び腫瘍ストローマ形成のような病理プロセスに関与するロイシンリッチプロテオグリカンについて研究が報告されている。Ｉｏｚｚｏ，Ｒ．Ｖ．，Ｃｒｉｔ．Ｒｅｖ．Ｂｉｏｃｈｅｍ．Ｍｏｌ．Ｂｉｏｌ．，３２（２）：１４１－１７４（１９９７）。創傷治癒及び組織修復におけるロイシンリッチタンパク質の関与を示す他の研究は、Ｄｅ Ｌａ Ｓａｌｌｅ，Ｃ．等，Ｖｏｕｖ．Ｒｅｖ．Ｆｒ．Ｈｅｍａｔｏｌ．（Ｇｅｒｍａｎｙ），３７（４）：２１５－２２２（１９９５）であり、出血疾患ベルナール－スーリエ症候群に伴う複合体におけるロイシンリッチモチーフの突然変異を報告しており、Ｃｈｌｅｍｅｔｓｏｎ，Ｋ．Ｊ．，Ｔｈｒｏｍｂ．Ｈａｅｍｏｓｔ．（Ｇｅｒｍａｎｙ），７４（１）：１１１－１１６（Ｊｕｌｙ １９９５）は、血小板がロイシンリッチ反復を有していることを報告しており、そしてＬａ Ｊｏｌｌａ Ｃａｎｃｅｒ Ｒｅｓｅａｒｃｈ ＦｏｕｎｄａｔｉｏｎのＲｕｏｓｌａｈｔｉ，Ｅ．Ｉ．等のＷＯ９１１０７２７－Ａは、トランスフォーミング成長因子βに結合するデコリンが、ガンの治療、創傷治癒及び瘢痕化へ関与していることを報告している。創傷治癒と結合組織のような組織、皮膚及び骨の再成長と関連した付随療法において；ヒトと動物における体の成長において；及び他の成長関連プロセスを刺激するのに有用である点において、この群のタンパク質に機能が関連しているものはインスリン様成長因子（ＩＧＦ）である。ＩＧＦの酸不安定サブユニット（ＡＬＳ）もまたＩＧＦの半減期を増加させインビボでのＩＧＦ複合体の一部である点において興味深い。ロイシンリッチ反復を有することが報告されている他のタンパク質は、アルツハイマー病のような神経変性疾患、パーキンソン病のような神経損傷の治療において、またガンの診断に対して有用であることが報告されているＳＬＩＴタンパク質である。エール大学のＡｒｔａｖａｎｉｓｔｓａｋｏｎａｓ，Ｓ．及びＲｏｔｈｂｅｒｇ，Ｊ．Ｍ．のＷＯ９２１０５１８－Ａ１を参照されたい。特に興

味深いのはＬＩＧ－１、つまりマウス脳中のグリア細胞において特異的に発現される膜糖タンパク質であり、ロイシンリッチ反復と免疫グロブリン様ドメインを有している。Ｓｕｚｕｋｉ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．（Ｕ．Ｓ．），２７１（３７）：２２５２２（１９９６）。ロイシンリッチ反復を有するタンパク質の生物学的機能を報告している他の研究には：Ｔａｙａｒ，Ｎ．等，Ｍｏｌ．Ｃｅｌｌ Ｅｎｄｏｃｒｉｎｏｌ．，（Ｉｒｅｌａｎｄ），１２５（１－２）：６５－７０（Ｄｅｃ．１９９６）（ゴナドトロピンレセプター関連）；Ｍｉｕｒａ，Ｙ．等，Ｎｉｐｐｏｎ Ｒｉｎｓｈｏ（Ｊａｐａｎ），５４（７）：１７８４－１７８９（Ｊｕｌｙ １９９６）（アポトーシス関連）；Ｈａｒｒｉｓ，Ｐ．Ｃ．等，Ｊ．Ａｍ．Ｓｏｃ．Ｎｅｐｈｒｏｌ．，６（４）：１１２５－１１３３（Ｏｃｔ．１９９５）（腎疾患関連）が含まれる。

よって、タンパク質－タンパク質相互作用をよりよく理解するためにロイシンリッチ反復を有する新規なタンパク質を同定するための努力が産業界と学術界の両方によってなされている。特に興味があるものはロイシンリッチ反復と、ＬＩＧ－１、ＡＬＳ及びデコリンのようなロイシンリッチ反復を有する既知のタンパク質と相同性を有するタンパク質である。多くの努力がロイシンリッチ反復を有する新規な分泌及び膜結合レセプタータンパク質に対するコード化配列を同定するために哺乳動物の組換えＤＮＡライブラリをスクリーニングすることに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，９３：７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。

我々は、ここでＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６ポリペプチドと命名した、ロイシンリッチ反復スーパーファミリーのタンパク質と相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。

【００４５】

４３．ＰＲＯ３３２

ある配列の保存されたロイシン残基（ロイシンリッチ反復モチーフ）により特徴づけられる反復を含んでなる分泌タンパク質は多様な生物学的役割を持っている。ある種のプロテオグリカン、例えばをビグリカン、フィブロモジュリン及びデコリンは例えば約２４個のアミノ酸のロイシンリッチ反復が存在することを特徴としている［Ｒｕｏｓｌａｈｔｉ，Ａｎｎ．Ｒｅｖ．Ｃｅｌｌ．Ｂｉｏｌ．４２２９－２５５（１９８８）；Ｏｌｄｂｅｒｇ等，ＥＭＢＯ Ｊ．８，２６０１－２６０４（１９８９）］。一般に、プロテオグリカンは細胞外マトリックス、軟骨又は骨機能の調節において役割を担っていると考えられている。プロテオグリカンのデコリンはコラーゲンＩ型及びＩＩ型に結合し、原繊維の形成速度に影響を及ぼす。フィブロモジュリンはまたコラーゲンに結合して原繊維形成を遅延させる。フィブロモジュリンとデコリンの両方がトランスフォーミング成長因子β（ＴＧＦ－β）の活性を阻害する（１９９６年１２月１０日発行の米国特許５，５８３，１０３）。ＴＧＦ－βは細胞外マトリックスの誘導に重要な役割を担うことが知られており、ガンや糸球体腎炎のような線維性疾患の発生に関与している。従って、プロテオグリカンは、ガン細胞に対するＴＧＦ－βの成長刺激活性を阻害する能力に基づいて、線維性ガンの治療に対して提案されている。プロテオグリカンはまた、関節リウマチ、動脈硬化症、成人呼吸窮迫症候群、肝硬変、肺線維症、心筋梗塞後症、心臓線維症、血管形成後再狭窄、腎臓間質線維症及びある種の皮膚線維性症状、例えば火傷、他の浸潤性皮膚傷害、又は美容もしくは再構成手術により生じるかもしれないケロイド及び瘢痕化を含む他の増殖性病状の治療に潜在的に有用であると記載されている（１９９７年８月５日発行の米国特許５，６５４，２７０）。

我々は、ここでＰＲＯ３３２ポリペプチドと命名した、ロイシンリッチ反復スーパーファミリーのタンパク質と相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。

【００４６】

４４．ＰＲＯ３３４

ミクロフィブリル束とこれらの束に付随して見いだされるタンパク質、特に付着分子は、

皮膚科学の分野、特に老化、傷害又は日光によりダメージを受けた皮膚の研究において興味深い。フィブリリンミクロフィブリルは皮膚の連続する弾性網目構造を形成し、真皮－上皮接合部から延びる測定可能なエラスチンを欠くミクロフィブリル束として真皮中に存在し、深い網状真皮中に厚い弾性繊維の成分として存在する。更に、マルファン症候群はフィルリリンモノマー又は他の結合組織要素の多量体化と干渉する変異に関連している。フィブリン－１はアミノ末端アナフィラトキシン様分子に９つの表皮成長因子（ＥＧＦ）様分子が続き、選択的スプライシングに応じて４つの可能なカルボキシル末端を有するモジュラー糖タンパク質である。フィブリン－２は、フィブロネクチン又はフィブリリンを含むミクロフィブリルと密に結合して頻繁に見出される新規な細胞外マトリックスタンパク質である。よって、フィブリリン、フィブリン、及びこれらに関連した分子は、特に、加齢、傷害又は日光から皮膚がダメージを受けるのを防止し、又はこれらからダメージを受けた皮膚を回復させるための用途に対して興味深い。更に、これらの分子は一般に結合組織と付着分子及び関連機構の研究において興味深い。フィブリリン、フィブリン及び関連分子は更にＡｄａｍｓ等，Ｊ．Ｍｏｌ．Ｂｉｏｌ．，２７２（２）：２２６－３６（１９９７）；Ｋｉｅｌｔｙ及びＳｈｕｔｔｌｅｗｏｒｔｈ，Ｍｉｃｒｏｓｃ．Ｒｅｓ．Ｔｅｃｈ．，３８（４）：４１３－２７（１９９７）；及びＣｈｉｌｄ，Ｊ．Ｃａｒｄ．Ｓｕｒｇ．，１２（２Ｓｕｐｐ．）：１３１－５（１９９７）に記載されている。

現在、新規な分泌及び膜結合レセプタータンパク質、特にフィブリンとフィブリリンと相同性を有する分泌タンパク質を同定するための努力が産業界と学術界の両方によってなされている。多くの努力が新規な分泌及び膜結合レセプタータンパク質に対するコード化配列を同定するために哺乳動物の組換えＤＮＡライブラリをスクリーニングすることに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，９３：７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。

我々は、ここでＰＲＯ３３４ポリペプチドと命名した、フィブリン及びフィブリリンと相同性を有する新規なポリペプチドの同定と特徴付けをここに記載する。

【００４７】

４５．ＰＲＯ３４６

ガンの広範な発生はかなりの資源を新しい治療法を発見することに捧げることを促した。特定の一方法は腫瘍抗原に特異的である腫瘍又はガン特異的モノクローナル抗体（ｍＡｂｓ）を創造を含む。正常細胞とガン細胞を区別できるそのようなｍＡｂｓは疾患の診断、予防及び治療に有用である。特定の抗原は、結腸直腸ガンのような新生物疾患に関連していることが知られている。大腸ガンは広範性疾患であるので、早期の診断と治療が重要な医療目標である。大腸ガンの診断と治療は蛍光、核磁気又は放射性タグを持つそれに特異的なモノクローナル抗体（ｍＡｂｓ）を使用して実施することができる。放射性遺伝子、毒素及び／又は薬物タグｍＡＢｓを最小の患者の処方でインサイツでの治療に対して使用できる。

ガン胎児性抗原（ＣＥＡ）はヒト大腸ガンと２－６月齢のヒト胎児の消化器に見いだされる糖タンパク質である。ＣＥＡは既知のヒト腫瘍マーカーであり、新生物疾患、例えば大腸ガンの診断において広く使用されている。例えば、ＣＥＡの血清レベルがある患者において上昇している場合、手術後のＣＥＡレベルの降下は腫瘍切除が成功したことを示している。他方、手術後の血清ＣＥＡレベルの引き続く上昇は元の腫瘍の転移が生じたかあるいは新しい原発性腫瘍が現れたことを示している。ＣＥＡはまたｍＡｂに対する標的であるアンチセンスヌクレオチドとなりうる。

【００４８】

４６．ＰＲＯ２６８

プロテインジスルフィドイソメラーゼは、正しいジスルフィド結合の形成の確立を通してタンパク質の正しいリフォールディングの促進に関与する酵素タンパク質である。プロテインジスルフィドイソメラーゼは当初は還元された変性ＲＮＡ分解酵素の再生を触媒する

その能力に基づいて同定された（Ｇｏｌｄｂｅｒｇｅｒ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２３９：１４０６－１４１０（１９６４）及びＥｐｓｔｅｉｎ等，Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｓｙｍｐ．Ｑｕａｎｔ．Ｂｉｏｌ．２８：４３９－４４９（１９６３））。プロテインジスルフィドイソメラーゼは、そのＣ末端に－ＫＤＥＬ又は－ＨＤＥＬアミノ酸配列を介して小胞体に保持されている小胞体の常在性酵素であることが示されている。

ジスルフィド結合形成酵素の重要性と多くの異なった応用、例えば組み換え的に産生されたタンパク質の正しいリフォールディングの収量を増大させる際の、その潜在的な用途が分かったので、現在は産業界と学術界の双方により、プロテインジスルフィドイソメラーゼに対して相同性を有する新規で天然のタンパク質を同定するための努力がなされている。これらの努力の多くは、新規なジスルフィドイソメラーゼ相同体に対するコード配列を同定するための哺乳動物組み換えＤＮＡライブラリーのスクリーニングに照準が合わされている。我々は、ここでＰＲＯ２６８と命名した、プロテインジスルフィドイソメラーゼと相同性を有する新規なポリペプチドをここに記載する。

【００４９】

４７．ＰＲＯ３３０

プロリル４－ヒロキシラーゼは、コラーゲンとプロコラーゲンの双方に見いだされた繰り返しの３個のアミノ酸の配列であるアミノ酸配列Ｇｌｙ－Ｘ－ＹのＹ位でプロリン残基を翻訳後にヒドロキシル化するように機能する酵素である。Ｇｌｙ－Ｘ－Ｙアミノ酸トリプレットのＹ位でのプロリン残基のヒドロキシル化によるそれらの位置での４－ヒドロキシプロリン残基の形成は、新たに合成されたコラーゲンポリペプチド鎖がその正しい三次元の三重螺旋高次構造に折り畳まれる前に必要となる。ヒドロキシル化が生じない場合は、合成されたコラーゲンポリペプチドは非螺旋のままで、細胞によりあまり分泌されず、安定な機能性のコラーゲン原繊維には構築されない。Ｖｕｏｒｉｏ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８９：７４６７－７４７０（１９９２）。プロリル４－ヒドロキシラーゼは少なくとも２つの異なったポリペプチドサブユニットであるαとβから構成されている。

新規な天然の分泌及び膜結合レセプタータンパク質を同定するための努力が産業界と学術界の両方によってなされている。多くの努力が新規な分泌及び膜結合レセプタータンパク質に対するコード化配列を同定するために哺乳動物の組換えＤＮＡライブラリをスクリーニングすることに注がれている。スクリーニング方法及び技術の例は文献に記載されている［例えば、Ｋｌｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，９３：７１０８－７１１３（１９９６）；米国特許５，５３６，６３７を参照されたい］。これらの努力に基づいて、出願人は、ここでＰＲＯ３３０と命名した、プロリル４－ヒドロキシラーゼのαサブユニットと相同性を有する新規なポリペプチドをここに同定し記載した。

【００５０】

４８．ＰＲＯ３３９及びＰＲＯ３１０

フリンジ（ｆｒｉｎｇｅ）はショウジョウバエの羽の成虫原基の背側区画におけるノッチのセレート（ｓｅｒｒａｔｅ）媒介活性化を特異的に阻止するタンパク質である。Ｆｌｅｍｉｎｇ等，Ｄｅｖｅｌｏｐｍｅｎｔ，１２４（１５）：２９７３－８１（１９９７）。従って、フリンジは発達におけるその役割並びにセレート、特にセレートのシグナル伝達能を制御するその能力のために興味深い。また興味深いものはセレート様分子の調節及び／又は発達に役割を担っている新規なポリペプチドである。特に興味深いものは、ここに同定し記載し、ここでＰＲＯ３３９及びＰＲＯ３１０ポリペプチドと命名した、フリンジと相同性を有する新規なポリペプチドである。

【００５１】

４９．ＰＲＯ２４４

レクチンは炭水化物を特異的かつ非供給結合的に結合する領域を有するタンパク質のクラスである。膜結合と可溶型の双方の数多くのレクチンが高等動物において同定されており、様々な細胞認識現象と腫瘍転移に関連している。

殆どのレクチンはＣ型（カルシウム依存性）又はＳ型（チオール依存性）の何れかに分類できる。
レクチンは原形質膜のレベルで開始される細胞性事象を調節する役割を担っていると考えられる。例えば、原形質膜関連タンパク質はリンパ系細胞の様々なサブセット、例えばＴリンパ球の活性化に関与しており、細胞表面分子はこれらの細胞の活性化、従って免疫反応の間はその応答の原因であることが知られている。
細胞接着分子の特定のグループであるセレクチンは、Ｃ型レクチンのスーパーファミリーに属している。このグループはＬ－セレクチン（末梢リンパ節ホーミングレセプター（ｐｎＨＲ），ＬＥＣ－ＣＡＭ－１，ＬＡＭ－１，ｇｐ９０^{ＭＥＬ}，ｇｐ１００^{ＭＥＬ}，ｇｐ１１０^{ＭＥＬ}，ＭＥＬ－１４抗原，Ｌｅｕ－８抗原，ＴＱ－１抗原，ＤＲＥＧ抗原）、Ｅ－セレクチン（ＬＥＣ－ＣＡＭ－２，ＬＥＣＡＭ－２，ＥＬＡＭ－１）及びＰ－セレクチン（ＬＥＣ－ＣＡＭ－３，ＬＥＣＡＭ－３，ＧＭＰ－１４０，ＰＡＤＧＥＭ）を含む。セレクチンの構造はＣ型レクチン（炭水化物結合）ドメイン、上皮成長因子様（ＥＧＦ様）モチーフ、及び補体調節性（ＣＲ）モチーフの可変メンバーからなる。セレクチンは白血球接着、例えば炎症に隣接した小静脈内皮細胞に対する好中球の付着（Ｅ－セレクチン）、あるいは血液から二次リンパ系器官、例えばリンパ節及びバイエル板へのリンパ球の輸送（Ｌ－セレクチン）に関連している。
他の例示的なレクチンは、細胞接着と免疫応答に関連していると考えられる細胞関連マクロファージ抗原Ｍａｃ－２である。マクロファージはまたヒトガン関連エピトープのＴｎＡｇを認識するレクチンを発現する。
他のＣ型レクチンは免疫学的に関連した制御された又は往路グラムされた細胞死（アポトーシス）の重要なメディエーターであるＣＤ９５（Ｆａｓ抗原／ＡＰＯ－１）である。「アポトーシス」は内因性細胞自殺プログラムの活性化に続いて後生動物において生じる非壊死性細胞死である。Ｆａｓ抗原のクローニングはＰＣＴ公報ＷＯ９１／１０４４８と欧州特許出願ＥＰ５１０６９１に記載されている。成熟Ｆａｓ分子は３１９のアミノ酸からなり、その１５７は細胞外で、１７が膜貫通ドメインを構成し、１４５が細胞内である。Ｔ細胞表面でのＦａｓ発現のレベルの増加は腫瘍細胞及びＨＩＶ感染細胞に付随している。ＣＤ９５の連結はインターロイキン－１（ＩＬ－２）の存在下でアポトーシスを惹起する。Ｃ型レクチンはまた酸化された低密度リポタンパク質（ＬＤＬ）に対するレセプターを含む。これは、アテローム性動脈硬化症の病状における可能な役割を示唆している。
我々はここにＰＲＯ２４４ポリペプチドと命名した、Ｃ型レクチンと相同性を有する新規なポリペプチドの同定と特徴付けを記載する。
【００５２】
（発明の概要）
１．　ＰＲＯ２１１及びＰＲＯ２１７
出願人は「ＰＲＯ２１１」及び「ＰＲＯ２１７」ポリペプチドと命名され、ＥＧＦに対して相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定した。
一実施態様では、本発明はＰＲＯ２１１又はＰＲＯ２１７ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１（配列番号：１）及び／又はＦｉｇ３（配列番号：３）に示されたＦｉｇ２（配列番号：２）及び／又は４（配列番号：４）のＥＧＦ様相同体ＰＲＯ２１１及びＰＲＯ２１７ポリペプチドをコードするＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２１１及びＰＲＯ２１７ＥＧＦ様相同体ＰＲＯ２１１及びＰＲＯ２１７ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２１１及びＰＲＯ２１７ＥＧＦ様相同体ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ２（配列番号：２）の残基：１から３５３又は（２）Ｆｉｇ４（配列番号：４）の１から３７９を含んでなるアミノ酸配列を含む。
【００５３】

２．ＰＲＯ２３０
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２３０」として命名した。
一実施態様では、本発明はＰＲＯ２３０ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ６（配列番号：１２）のアミノ酸残基１から４６７を有するＰＲＯ２３０ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２３０ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２３０ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ６（配列番号：１２）の残基１から４６７を含んでなるアミノ酸配列を含む。
他の実施態様では、本発明はここでＤＮＡ２００８８と命名する配列番号：１３（Ｆｉｇ７）のヌクレオチド配列を含んでなる発現配列タグ（ＥＳＴ）を提供する。
【００５４】
３．ＰＲＯ２３２
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２３２」として命名した。
一実施態様では、本発明はＰＲＯ２３２ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ９（配列番号：１８）のアミノ酸残基１から１１４を有するＰＲＯ２３２ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２３２ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２３２ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ９（配列番号：１８）の残基１から１１４を含んでなるアミノ酸配列を含む。
【００５５】
４．ＰＲＯ１８７
出願人は、本出願において「ＰＲＯ１８７」と命名する新規なポリペプチドをコードするｃＤＮＡクローンを同定した。
一実施態様では、本発明はＰＲＯ１８７ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１１（配列番号：２３）のＰＲＯ１８７ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。他の態様では、本発明はＦｉｇ１０（配列番号：２２）のコード化配列を含んでなる核酸又はその補体を提供する。他の態様では、本発明は寄託番号ＡＴＣＣ２０９３７５でＡＴＣＣに寄託されたクローンＤＮＡ２７８６４－１１５５の全長タンパク質の核酸、あるいは寄託番号ＡＴＣＣ２０９３７５で寄託されたクローンＤＮＡ２７８６４－１１５５のコード化配列を提供する。
更に他の実施態様では、本発明は、単離されたＰＲＯ１８７ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ１８７ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１１（配列番号：２３）の残基１から２０５を含んでなるアミノ酸配列を含む。あるいは、本発明は寄託番号ＡＴＣＣ２０９３７５で寄託された核酸によりコードされるポリペプチドを提供する。
【００５６】
５．ＰＲＯ２６５
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２６５」として命名した。
一実施態様では、本発明はＰＲＯ２６５ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１３（配列番号：２８）のアミノ酸残基１から６６０を有するＰＲＯ２６５ポリペプチドをコードしている

ＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２６５ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２６５ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１３（配列番号：２８）の残基１から６６０を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ２６５ポリペプチドの単離された細胞外ドメインに対するものである。

【００５７】

６．ＰＲＯ２１９

出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２１９」として命名した。
一実施態様では、本発明はＰＲＯ２１９ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１５（配列番号：３４）のアミノ酸残基１から９１５を有するＰＲＯ２１９ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２１９ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２１９ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１５（配列番号：３４）の残基１から９１５を含んでなるアミノ酸配列を含む。

【００５８】

７．ＰＲＯ２４６

出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２４６」として命名した。
一実施態様では、本発明はＰＲＯ２４６ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１７（配列番号：３９）のアミノ酸残基１から３９０を有するＰＲＯ２４６ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２４６ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２４６ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１７（配列番号：３９）の残基１から３９０を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ２４６ポリペプチドの単離された細胞外ドメインに対するものである。

【００５９】

８．ＰＲＯ２２８

出願人はＣＤ９７、ＥＭＲ１及びラトロフィリンと相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２２８」として命名した。
一実施態様では、本発明はＰＲＯ２２８ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１９（配列番号：４９）のアミノ酸残基１から６９０を有するＰＲＯ２２８ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２２８ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２２８ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１９（配列番号：４９）の残基１から６９０を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ２２８ポリペプチドの単離された細胞外ドメインに対するものである。
他の実施態様では、本発明は、ここでＤＮＡ２１９５１と命名された、配列番号：５０のヌクレオチド配列を含んでなる発現配列タグ（ＥＳＴ）を提供する。

【００６０】
９．　ＰＲＯ５３３
本出願人は本出願においてＰＲＯ５３３と命名された新規な分泌ポリペプチドをコードするｃＤＮＡクローン（ＤＮＡ４９４３５－１２１９）を同定した。
一実施態様では、本発明は、（ａ）Ｆｉｇ２２（配列番号：５９）のアミノ酸残基２３から２１６の配列を有するＰＲＯ５３３ポリペプチドをコードするＤＮＡ分子、又は（ｂ）（ａ）のＤＮＡ分子の補体に対して、少なくとも約８０％の配列同一性を有する単離された核酸分子を提供する。配列同一性は好ましくは約８５％、より好ましくは約９０％、最も好ましくは約９５％である。一態様では、単離された核酸は、Ｆｉｇ２２（配列番号：５９）のアミノ酸残基２３から２１６を有するポリペプチドと少なくとも約８０％、好ましくは少なくとも約８５％、より好ましくは少なくとも約９０％、そして最も好ましくは少なくとも約９５％の配列同一性を有する。好ましくは、配列同一性の最も高い度合いは分泌タンパク質において生じる（Ｆｉｇ２２のアミノ酸２３から２１６、配列番号：５９）。更なる実施態様では、単離された核酸は、Ｆｉｇ２２（配列番号：５９）のアミノ酸残基１から２１６を有するＰＲＯ５３３ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。他の態様では、本発明は寄託番号ＡＴＣＣ２０９４８０でＡＴＣＣに寄託されたクローンＤＮＡ４９４３５－１２１９の全長タンパク質の核酸を提供する。
更に他の実施態様では、本発明は、単離されたＰＲＯ５３３ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ５３３ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ２２（配列番号：５９）の残基２３から２１６を含んでなるアミノ酸配列を含む。天然シグナル配列（Ｆｉｇ２２（配列番号：５９）のアミノ酸１から２２）を伴うか伴わない、かつ開始メチオニンを伴うか伴わない天然ＰＲＯ５３３ポリペプチドが特に含まれる。あるいは、本発明は寄託番号ＡＴＣＣ２０９４８０で寄託された核酸によりコードされるＰＲＯ５３３ポリペプチドを提供する。
【００６１】
１０．ＰＲＯ２４５
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２４５」として命名した。
一実施態様では、本発明はＰＲＯ２４５ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ２４（配列番号：６４）のアミノ酸残基１から３１２を有するＰＲＯ２４５ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２４５ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２４５ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ２４（配列番号：６４）の残基１から３１２を含んでなるアミノ酸配列を含む。
【００６２】
１１．ＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７
出願人は、全てロイシンリッチ反復を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定した。これらのポリペプチドは本出願においてＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７と命名する。
一実施態様では、本発明はＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７をそれぞれコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、ここで提供されるものは、Ｆｉｇ２６（配列番号：６９）のアミノ酸残基１から７０８を有するＰＲＯ２２０ポリペプチドをコードしているＤＮＡを含有する単離された核酸か、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。またここで提供されるものは、Ｆｉｇ２８（配列番号：７１）のアミノ酸残基１から２５９を有するＰＲＯ２２１ポリペプチドをコードしている

ＤＮＡを含有する単離された核酸か、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。更にまたここで提供されるものは、Ｆｉｇ３０（配列番号：７３）のアミノ酸残基１から６２０を有するＰＲＯ２２７ポリペプチドをコードしているＤＮＡを含有する単離された核酸か、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７ポリペプチドを提供する。特に、ここで提供されるものはＰＲＯ２２０ポリペプチドの単離された天然配列であり、それは一実施態様では、Ｆｉｇ２６（配列番号：６９）の残基１から７０８を含んでなるアミノ酸配列を含む。また、ここで提供されるものはＰＲＯ２２１ポリペプチドの単離された天然配列であり、それは一実施態様では、Ｆｉｇ２８（配列番号：７１）の残基１から２５９を含んでなるアミノ酸配列を含む。更にここで提供されるものはＰＲＯ２２７ポリペプチドの単離された天然配列であり、それは一実施態様では、Ｆｉｇ３０（配列番号：７３）の残基１から６２０を含んでなるアミノ酸配列を含む。

【００６３】

１２．ＰＲＯ２５８

出願人はＣＲＴＡＭ及びポリオウィルスレセプター前駆体と相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２５８」として命名した。

一実施態様では、本発明はＰＲＯ２５８ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ３２（配列番号：８４）のアミノ酸残基１から３９８を有するＰＲＯ２５８ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ２５８ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２５８ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ３２（配列番号：８４）の残基１から３９８を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ２５８ポリペプチドの単離された細胞外ドメインに対するものである。

【００６４】

１３．ＰＲＯ２６６

出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２６６」として命名した。

一実施態様では、本発明はＰＲＯ２６６ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ３４（配列番号：９１）のアミノ酸残基１から６９６を有するＰＲＯ２６６ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ２６６ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２６６ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ３４（配列番号：９１）の残基１から６９６を含んでなるアミノ酸配列を含む。

【００６５】

１４．ＰＲＯ２６９

出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２６９」として命名した。

一実施態様では、本発明はＰＲＯ２６９ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ３６（配列番号：９６）のアミノ酸残基１から４９０を有するＰＲＯ２６９ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、

場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２６９ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２６９ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ３６（配列番号：９６）の残基１から４９０を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ２６９ポリペプチドの単離された細胞外ドメインに対するものである。
【００６６】
１５．ＰＲＯ２８７
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２８７」として命名した。
一実施態様では、本発明はＰＲＯ２８７ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ３８（配列番号：１０４）のアミノ酸残基１から４１５を有するＰＲＯ２８７ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２８７ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２８７ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ３８（配列番号：１０４）の残基１から４１５を含んでなるアミノ酸配列を含む。
【００６７】
１６．ＰＲＯ２１４
出願人は、本出願において「ＰＲＯ２１４」と命名する新規なポリペプチドをコードするｃＤＮＡクローンを同定した。
一実施態様では、本発明はＰＲＯ２１４ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ４０（配列番号：１０９）のＰＲＯ２１４ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。他の態様では、本発明はＦｉｇ３９（配列番号：１０８）のコード化配列を含んでなる核酸又はその補体を提供する。他の態様では、本発明は寄託番号ＡＴＣＣ２０９３８５でＡＴＣＣに寄託されたクローンＤＮＡ３２２８６－１１９１の全長タンパク質の核酸を提供する。
更に他の実施態様では、本発明は、単離されたＰＲＯ２１４ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２１４ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ４０（配列番号：１０９）の残基を含んでなるアミノ酸配列を含む。あるいは、本発明は寄託番号ＡＴＣＣ２０９３８５で寄託された核酸によりコードされるポリペプチドを提供する。
【００６８】
１７．ＰＲＯ３１７
出願人は、本出願において「ＰＲＯ３１７」と命名する新規なポリペプチドをコードするｃＤＮＡクローンを同定した。
一実施態様では、本発明はＰＲＯ３１７ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ４２のアミノ酸残基１から３６６を有するＰＲＯ３１７ポリペプチドをコードしているＤＮＡ（配列番号：１１３）を含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ３１７ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ３１７ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ４２（配列番号：１１４）の残基１から３６６を含んでなるアミノ酸配列を含む。
更に他の実施態様では、本発明は、試料中のＰＲＯ３１７の存在を検出する方法を提供し、該方法は、

ａ）ＰＲＯ３１７を含んでいることが疑われている試料に検出可能な抗ＰＲＯ３１７抗体を接触させ；そして
ｂ）試料への抗体の結合を検出することを含み；ここで、試料は体液、組織試料、細胞抽出物、及び細胞培地からなる群から選択される。
また更なる実施態様では、試料中のＰＲＯ３１７ｍＲＮＡの存在を検出する方法が提供され、該方法は、
ａ）ＰＲＯ３１７ｍＲＮＡを含んでいることが疑われている試料を、中程度から緊縮性条件下でＰＲＯ３１７ｍＲＮＡとハイブリッド形成する検出可能な核酸プローブに接触させ；そして
ｂ）試料へのプローブのハイブリッド形成を検出することを含む。
好ましくは、この方法において、試料は組織試料であり、検出工程はインサイツハイブリッド形成により、又は試料は細胞抽出物であり検出はノーザン分析法である。
更に、本発明は、有効量のＰＲＯ３１７アゴニスト又はＰＲＯ３１７アンタゴニスト、例えばＰＲＯ３１７に特異的に結合する抗体で、又は有効量のＰＲＯ３１７ポリペプチド又は担体を含むその組成物を哺乳動物に投与することを含んでなるＰＲＯ３１７関連疾患を治療する方法を提供する。

【００６９】

１８．ＰＲＯ３０１

本出願人は本出願において「ＰＲＯ３０１」と命名された新規な分泌ポリペプチドをコードするｃＤＮＡクローン（ＤＮＡ４０６２８－１２１６）を同定した。
一実施態様では、本発明は、（ａ）Ｆｉｇ４４（配列番号：１１９）のアミノ酸残基２８から２５８の配列を有するＰＲＯ３０１ポリペプチドをコードするＤＮＡ分子、又は（ｂ）（ａ）のＤＮＡ分子の補体に対して、少なくとも約８０％の配列同一性を有する単離された核酸分子を提供する。配列同一性は好ましくは約８５％、より好ましくは約９０％、最も好ましくは約９５％である。一態様では、単離された核酸は、Ｆｉｇ４４（配列番号：１１９）のアミノ酸残基２８から２５８を有するポリペプチドと少なくとも約８０％、好ましくは少なくとも約８５％、より好ましくは少なくとも約９０％、そして最も好ましくは少なくとも約９５％の配列同一性を有する。好ましくは、配列同一性の最も高い度合いは細胞外ドメインにおいて生じる（Ｆｉｇ４４のアミノ酸２８から２５８、配列番号：１１９）。更なる実施態様では、単離された核酸は、Ｆｉｇ４４（配列番号：１１９）のアミノ酸残基２８から２９９を有するＰＲＯ３０１ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。他の態様では、本発明は寄託番号ＡＴＣＣ２０９４３２でＡＴＣＣに寄託されたクローンＤＮＡ４０６２８－１２１６の全長タンパク質の核酸、あるいは寄託番号ＡＴＣＣ２０９４３２で寄託されたクローンＤＮＡ４０６２８－１２１６のコード化配列を提供する。
更に他の実施態様では、本発明は、単離されたＰＲＯ３０１ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ３０１ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ４４（配列番号：１１９）の細胞外ドメイン残基２８から２５８を含んでなるアミノ酸配列を含む。天然シグナル配列（Ｆｉｇ４４（配列番号：１１９）のアミノ酸１から２７）を伴うか伴わない、かつ開始メチオニンを伴うか伴わない天然ＰＲＯ３０１ポリペプチドが特に含まれる。また、本発明の配列はまた膜貫通ドメイン（Ｆｉｇ４４の残基２３６から約２５８；配列番号：１１９）及び／又は細胞内ドメイン（Ｆｉｇ４４の残基約２５９から２９９；配列番号：１１９）を含んでもよい。あるいは、本発明は寄託番号ＡＴＣＣ２０９４３２で寄託された核酸によりコードされるＰＲＯ３０１ポリペプチドを提供する。

【００７０】

１９．ＰＲＯ２２４

出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２２４」として命名した。

一実施態様では、本発明はＰＲＯ２２４ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ４６（配列番号：１２７）のアミノ酸残基１から２８２を有するＰＲＯ２２４ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２２４ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２２４ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ４６（配列番号：１２７）の残基１から２８２を含んでなるアミノ酸配列を含む。
【００７１】
２０．ＰＲＯ２２２
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２２２」として命名した。
一実施態様では、本発明はＰＲＯ２２２ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ４８（配列番号：１３２）のアミノ酸残基１から４９０を有するＰＲＯ２２２ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２２２ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２２２ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ４８（配列番号：１３２）の残基１から４９０を含んでなるアミノ酸配列を含む。
【００７２】
２１．ＰＲＯ２３４
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、本出願において「ＰＲＯ２３４」として命名した。
一実施態様では、本発明はＰＲＯ２３４ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ５０（配列番号：１３７）のアミノ酸残基１から３８２を有するＰＲＯ２３４ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。他の態様では、本発明はＦｉｇ４９（配列番号：１３６）のヌクレオチド配列を含んでなる単離された核酸分子を提供する。
他の実施態様では、本発明は、単離された新規なＰＲＯ２３４ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２３４ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ５０（配列番号：１３７）の残基１から３８２を含んでなるアミノ酸配列を含む。
更に他の実施態様では、本発明はゲノム及びｃＤＮＡヌクレオチド配列を単離するために有用なオリゴヌクレオチドプローブを提供する。
【００７３】
２２．ＰＲＯ２３１
出願人は推定酸性フォスファターゼと相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２３１」と命名した。
一実施態様では、本発明はＰＲＯ２３１ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ５２（配列番号：１４２）のアミノ酸残基１から４２８を有するＰＲＯ２３１ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２３１ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２３１ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ５２（配列番号：１４２）の残基１から４２８を含んでなるアミノ酸配列を含む。
【００７４】

２３．ＰＲＯ２２９
出願人はスカベンジャーレセプターと相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２２９」と命名した。
一実施態様では、本発明はＰＲＯ２２９ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ５４（配列番号：１４８）のアミノ酸残基１から３４７を有するＰＲＯ２２９ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２２９ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２２９ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ５４（配列番号：１４８）の残基１から３４７を含んでなるアミノ酸配列を含む。
【００７５】
２４．ＰＲＯ２３８
出願人はレダクターゼと相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２３８」と命名した。
一実施態様では、本発明はＰＲＯ２３８ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ５６（配列番号：１５３）のアミノ酸残基１から３１０を有するＰＲＯ２３８ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２３８ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２３８ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ５６（配列番号：１５３）の残基１から３１０を含んでなるアミノ酸配列を含む。
【００７６】
２５．ＰＲＯ２３３
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２３３」と命名した。
一実施態様では、本発明はＰＲＯ２３３ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ５８（配列番号：１５９）のアミノ酸残基１から３００を有するＰＲＯ２３３ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２３３ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２３３ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ５８（配列番号：１５９）の残基１から３００を含んでなるアミノ酸配列を含む。
【００７７】
２６．ＰＲＯ２２３
出願人はセリンカルボキシペプチダーゼポリペプチドと相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２２３」と命名した。
一実施態様では、本発明はＰＲＯ２２３ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ６０（配列番号：１６４）のアミノ酸残基１から４７６を有するＰＲＯ２２３ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２２３ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２２３ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ６０（配列番号：１６４）の残基１から４７６を含んでなるアミノ酸配列を含む。
【００７８】

２７．ＰＲＯ２３５
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２３５」と命名した。
一実施態様では、本発明はＰＲＯ２３５ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ６２（配列番号：１７０）のアミノ酸残基１から５５２を有するＰＲＯ２３５ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２３５ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ２３５ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ６２（配列番号：１７０）の残基１から５５２を含んでなるアミノ酸配列を含む。

【００７９】
２８．ＰＲＯ２３６及びＰＲＯ２６２
出願人は、β－ガラクトシダーゼと相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、これらのポリペプチドを本出願において「ＰＲＯ２３６」及び「ＰＲＯ２６２」と命名した。
一実施態様では、本発明はＰＲＯ２３６をコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸分子は、Ｆｉｇ６４（配列番号：１７５）のアミノ酸残基１から６３６を有するＰＲＯ２３６ポリペプチドをコードしているＤＮＡを含有する単離された核酸か、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明はＰＲＯ２６２をコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸分子は、Ｆｉｇ６６（配列番号：１７７）のアミノ酸残基１から６５４を有するＰＲＯ２６２ポリペプチドをコードしているＤＮＡを含有する単離された核酸か、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２３６ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２３６ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ６４（配列番号：１７５）の残基１から６３６を含んでなるアミノ酸配列を含む。
他の実施態様では、本発明は、単離されたＰＲＯ２６２ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２６２ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ６６（配列番号：１７７）の残基１から６５４を含んでなるアミノ酸配列を含む。

【００８０】
２９．ＰＲＯ２３９
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２３９」と命名した。
一実施態様では、本発明はＰＲＯ２３９ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ６８（配列番号：１８５）のアミノ酸残基１から５０１を有するＰＲＯ２３９ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２３９ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２３９ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ６８（配列番号：１８５）の残基１から５０１を含んでなるアミノ酸配列を含む。

【００８１】
３０．ＰＲＯ２５７
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを

本出願において「ＰＲＯ２５７」と命名した。
一実施態様では、本発明はＰＲＯ２５７ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ７０（配列番号：１９０）のアミノ酸残基１から６０７を有するＰＲＯ２５７ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２５７ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２５７ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ７０（配列番号：１９０）の残基１から６０７を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ２５７ポリペプチドの単離された細胞外ドメインに対するものである。

【００８２】

３１．ＰＲＯ２６０

出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２６０」と命名した。
一実施態様では、本発明はＰＲＯ２６０ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ７２（配列番号：１９５）のアミノ酸残基１から４６７を有するＰＲＯ２６０ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２６０ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２６０ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ７２（配列番号：１９５）の残基１から４６７を含んでなるアミノ酸配列を含む。

【００８３】

３２．ＰＲＯ２６３

出願人はＣＤ４４抗原と相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２６３」と命名した。
一実施態様では、本発明はＰＲＯ２６３ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ７４（配列番号：２０１）のアミノ酸残基１から３２２を有するＰＲＯ２６３ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２６３ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２６３ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ７４（配列番号：２０１）の残基１から３２２を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ２６３ポリペプチドの単離された細胞外ドメインに対するものである。

【００８４】

３３．ＰＲＯ２７０

出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２７０」と命名した。
一実施態様では、本発明はＰＲＯ２７０ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ７６（配列番号：２０７）のアミノ酸残基１から２９６を有するＰＲＯ２７０ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２７０ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２７０ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ７６（配列番号：２０７）の残基１から２９６を含んでなるアミノ酸配列を含む

。
【００８５】
３４．ＰＲＯ２７１
出願人はプロテオグリカン連結タンパク質と相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２７１」と命名した。
一実施態様では、本発明はＰＲＯ２７１ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ７８（配列番号：２１３）のアミノ酸残基１から３６０を有するＰＲＯ２７１ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２７１ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２７１ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ７８（配列番号：２１３）の残基１から３６０を含んでなるアミノ酸配列を含む。
【００８６】
３５．ＰＲＯ２７２
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２７２」と命名した。
一実施態様では、本発明はＰＲＯ２７２ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ８０（配列番号：２２１）のアミノ酸残基１から３２８を有するＰＲＯ２７２ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２７２ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２７２ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ８０（配列番号：２１１）の残基１から３２８を含んでなるアミノ酸配列を含む。
【００８７】
３６．ＰＲＯ２９４
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２９４」と命名した。
一実施態様では、本発明はＰＲＯ２９４ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ８２（配列番号：２２７）のアミノ酸残基１から５５０を有するＰＲＯ２９４ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２９４ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２９４ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ８２（配列番号：２２７）の残基１から５５０を含んでなるアミノ酸配列を含む。
【００８８】
３７．ＰＲＯ２９５
出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２９５」と命名した。
一実施態様では、本発明はＰＲＯ２９５ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ８４（配列番号：２３６）のアミノ酸残基１から３５０を有するＰＲＯ２９５ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ２９５ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２９５ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ８４（配列番号：２３６）の残基１から３５０を含んでなるアミノ酸配列を含む。

【００８９】

３８．ＰＲＯ２９３

出願人は新規なヒトニューロンのロイシンリッチ反復ポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２９３」と命名した。

一実施態様では、本発明はＰＲＯ２９３ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ８６（配列番号：２４５）のアミノ酸残基１から７１３を有するＰＲＯ２９３ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ２９３ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２９３ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ８６（配列番号：２４５）の残基１から７１３を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ２９３ポリペプチドの単離された細胞外ドメインに対するものである。

【００９０】

３９．ＰＲＯ２４７

出願人はロイシンリッチ反復を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２４７」と命名した。

一実施態様では、本発明はＰＲＯ２４７ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ８８（配列番号：２５０）のアミノ酸残基１から５４６を有するＰＲＯ２４７ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ２４７ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２４７ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ８８（配列番号：２５０）の残基１から５４６を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ２４７ポリペプチドの単離された細胞外ドメインに対するものである。

【００９１】

４０．ＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３

出願人は様々なプロテアーゼと相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、それらポリペプチドを本出願において「ＰＲＯ３０２」、「ＰＲＯ３０３」、「ＰＲＯ３０４」、「ＰＲＯ３０７」及び「ＰＲＯ３４３」と命名した。

一実施態様では、本発明はＰＲＯ３０２ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ９０（配列番号：２５５）のアミノ酸残基１から４５２を有するＰＲＯ３０２ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明はＰＲＯ３０３ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ９２（配列番号：２５７）のアミノ酸残基１から３１４を有するＰＲＯ３０３ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

更に他の実施態様では、本発明はＰＲＯ３０４ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ９４（配列番号：２５９）のアミノ酸残基１から５５６を有するＰＲＯ３０４ポリペプチドをコード

しているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明はＰＲＯ３０７ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ９６（配列番号：２６１）のアミノ酸残基１から３８３を有するＰＲＯ３０７ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
更に他の実施態様では、本発明はＰＲＯ３４３ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ９８（配列番号：２６３）のアミノ酸残基１から３１７を有するＰＲＯ３４３ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ３０２ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ３０２ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ９０（配列番号：２５５）の残基１から４５２を含んでなるアミノ酸配列を含む。
他の実施態様では、本発明は、単離されたＰＲＯ３０３ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ３０４ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ９４（配列番号：２５９）の残基１から５５６を含んでなるアミノ酸配列を含む。
他の実施態様では、本発明は、単離されたＰＲＯ３０７ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ３０７ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ９６（配列番号：２６１）の残基１から３８３を含んでなるアミノ酸配列を含む。
他の実施態様では、本発明は、単離されたＰＲＯ３４３ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ３４４ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ９８（配列番号：２６３）の残基１から３１７を含んでなるアミノ酸配列を含む。

【００９２】

４１．ＰＲＯ３２８

出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ３２８」と命名した。
一実施態様では、本発明はＰＲＯ３２８ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１００（配列番号：２８５）のアミノ酸残基１から４６３を有するＰＲＯ３２８ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ３２８ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ３２８ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１００（配列番号：２８５）の残基１から４６３を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ３０６ポリペプチドの単離された細胞外ドメインに対するものである。

【００９３】

４２．ＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６

出願人は、ロイシンリッチ反復とＬＩＧ－１及びＡＬＳとの相同性を有する３種の新規なポリペプチドをコードする３種のｃＤＮＡクローンを同定した。これらのポリペプチドは本出願においてＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６と命名する。
一実施態様では、本発明はＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６をそれぞれコードするＤＮＡを含んでなる３種の単離された核酸分子を提供する。一態様では、ここで提供されるものは、Ｆｉｇ１０２（配列番号：２９０）のアミノ酸残基１から１０５９を有

するＰＲＯ３３５ポリペプチドをコードしているＤＮＡを含有する単離された核酸か、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。またここで提供されるものは、Ｆｉｇ１０４（配列番号：２９２）のアミノ酸残基１から６４０を有するＰＲＯ３３１ポリペプチドをコードしているＤＮＡを含有する単離された核酸か、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。更にまたここで提供されるものは、Ｆｉｇ１０６（配列番号：２９４）のアミノ酸残基１から１１１９を有するＰＲＯ３２６ポリペプチドをコードしているＤＮＡを含有する単離された核酸か、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６ポリペプチドを提供する。特に、本発明はＰＲＯ３３５ポリペプチドに対する単離された天然配列を提供し、それは一実施態様では、Ｆｉｇ１０２（配列番号：２９０）の残基１から１０５９を含んでなるアミノ酸配列を含む。また、ここで提供されるものはＰＲＯ３３１ポリペプチドの単離された天然配列であり、それは一実施態様では、Ｆｉｇ１０４（配列番号：２９２）の残基１から６４０を含んでなるアミノ酸配列を含む。更にここで提供されるものはＰＲＯ３２６ポリペプチドの単離された天然配列であり、それは一実施態様では、Ｆｉｇ１０６（配列番号：２９４）の残基１から１１１９を含んでなるアミノ酸配列を含む。

【００９４】

４３．ＰＲＯ３３２

本出願人は本出願において「ＰＲＯ３３２」と命名された新規な分泌ポリペプチドをコードするｃＤＮＡクローン（ＤＮＡ４０９８２－１２３５）を同定した。

一実施態様では、本発明は、（ａ）Ｆｉｇ１０８（配列番号：３１０）のアミノ酸残基４９から６４２の配列を有するＰＲＯ３５８ポリペプチドをコードするＤＮＡ分子、又は（ｂ）（ａ）のＤＮＡ分子の補体に対して、少なくとも約８０％の配列同一性を有する単離された核酸分子を提供する。配列同一性は好ましくは約８５％、より好ましくは約９０％、最も好ましくは約９５％である。一態様では、単離された核酸は、Ｆｉｇ１０８（配列番号：３１０）のアミノ酸残基１から６４２を有するポリペプチドと少なくとも約８０％、好ましくは少なくとも約８５％、より好ましくは少なくとも約９０％、そして最も好ましくは少なくとも約９５％の配列同一性を有する。好ましくは、配列同一性の最も高い度合いはロイシンリッチ反復ドメインにおいて生じる（Ｆｉｇ３１０のアミノ酸１１６から６２４、配列番号：３１０）。更なる実施態様では、単離された核酸は、Ｆｉｇ１０８（配列番号：３１０）のアミノ酸残基４９から６４２を有するＰＲＯ３３２ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ３３２ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ３３２ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１０８（配列番号：３１０）の残基４９から６２４を含んでなるアミノ酸配列を含む。天然シグナル配列（Ｆｉｇ１０８（配列番号：３１０）のアミノ酸１から４８）を伴うか伴わない、かつ開始メチオニンを伴うか伴わない天然ＰＲＯ３３２ポリペプチドが特に含まれる。

【００９５】

４４．ＰＲＯ３３４

出願人はフィブリン及びフィブリリンと相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ３３４」と命名した。

一実施態様では、本発明はＰＲＯ３３４ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１１０（配列番号：３１５）のアミノ酸残基１から５０９を有するＰＲＯ３３４ポリペプチドをコードして

いるＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ３３４ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ３３４ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１１０（配列番号：３１５）の残基１から５０９を含んでなるアミノ酸配列を含む。
【００９６】
４５．ＰＲＯ３４６
本出願人は本出願において「ＰＲＯ３４６」と命名された新規な分泌ポリペプチドをコードするｃＤＮＡクローン（ＤＮＡ４４１６７－１２４３）を同定した。
一実施態様では、本発明は、（ａ）Ｆｉｇ１１２（配列番号：３２０）のアミノ酸残基１９から３３９の配列を有するＰＲＯ３４６ポリペプチドをコードするＤＮＡ分子、又は（ｂ）（ａ）のＤＮＡ分子の補体に対して、少なくとも約８０％の配列同一性を有する単離された核酸分子を提供する。配列同一性は好ましくは約８５％、より好ましくは約９０％、最も好ましくは約９５％である。一態様では、単離された核酸は、Ｆｉｇ１１２（配列番号：３２０）のアミノ酸残基１９から３３９を有するポリペプチドと少なくとも約８０％、好ましくは少なくとも約８５％、より好ましくは少なくとも約９０％、そして最も好ましくは少なくとも約９５％の配列同一性を有する。好ましくは、配列同一性の最も高い度合いは細胞外ドメイン内で生じる（Ｆｉｇ１１２、配列番号：３２０のアミノ酸１９から３３９）。別の実施態様では、相同性が測定されるポリペプチドはＦｉｇ１１２、配列番号：３２０の残基１－３３９、１９－３６０又は１９－４５０を含む。更なる実施態様では、単離された核酸は、Ｆｉｇ１１２（配列番号：３２０）のアミノ酸残基１９から３３９、あるいはＦｉｇ１１２（配列番号：３２０）の残基１－３３９、１９－３６０又は１９－４５０を有するＰＲＯ３４６ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。他の態様では、本発明は寄託番号ＡＴＣＣ２０９４３４でＡＴＣＣに寄託されたクローンＤＮＡ４４１６７－１２４３の全長タンパク質の核酸、あるいは寄託番号ＡＴＣＣ２０９４３４で寄託されたクローンＤＮＡ４４１６７－１２４３のコード化配列を提供する。
更に他の実施態様では、本発明は、単離されたＰＲＯ３４６ポリペプチドを提供する。特に、本発明は天然配列ＰＲＯ３４６ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１１２（配列番号：３２０）の残基１９から３３９を含んでなるアミノ酸配列を含む。天然シグナル配列（Ｆｉｇ１１２（配列番号：３２０）のアミノ酸１から１８）を伴うか伴わない、開始メチオニンを伴うか伴わない、膜貫通ドメイン（残基３４０から３６０）を伴うか伴わない、そして細胞内ドメイン（残基３６１から４５０）を伴うかともなわない天然ＰＲＯ３４６ポリペプチドが特に含まれる。あるいは、本発明は寄託番号ＡＴＣＣ２０９４３４で寄託された核酸によりコードされるＰＲＯ３４６ポリペプチドを提供する。
【００９７】
４６．ＰＲＯ２６８
出願人はタンパク質ジスルフィドイソメラーゼと相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２６８」と命名した。
一実施態様では、本発明はＰＲＯ２６８ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１１４（配列番号：３２５）のアミノ酸残基１から２８０を有するＰＲＯ２６８ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。
他の実施態様では、本発明は、単離されたＰＲＯ２６８ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ２６８ポリペプチドを提供し、それは一実施態様では

、Ｆｉｇ１１４（配列番号：３２５）の残基１から２８０を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ２６８ポリペプチドの単離された細胞外ドメインに対するものである。

【００９８】

４７．ＰＲＯ３３０

出願人はプロリル４－ヒドロキシラーゼのαサブユニットと相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ３３０」と命名した。

一実施態様では、本発明はＰＲＯ３３０ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１１６（配列番号：３３２）のアミノ酸残基１から５３３を有するＰＲＯ３３０ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ３３０ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ３３０ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１１６（配列番号：３３２）の残基１から５３３を含んでなるアミノ酸配列を含む。本発明の更なる実施態様はＰＲＯ３３０ポリペプチドの単離された細胞外ドメインに対するものである。

【００９９】

４８．ＰＲＯ３３９及びＰＲＯ３１０

出願人はフリンジ（ｆｒｉｎｇｅ）と相同性を有する新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ３３９」及び「ＰＲＯ３１０」と命名した。

一実施態様では、本発明はＰＲＯ３３９及び／又はＰＲＯ３１０ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１１８（配列番号：３３９）のアミノ酸残基１から７７２を有するＰＲＯ３３９ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ３３９ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ３３９ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１１８（配列番号：３３９）の残基１から７７２を含んでなるアミノ酸配列を含む。本発明は更に単離された天然配列ＰＲＯ３１０ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１２０（配列番号：３４１）の残基１から３１８を含んでなるアミノ酸配列を含む。

【０１００】

４９．ＰＲＯ２４４

出願人は新規なポリペプチドをコードするｃＤＮＡクローンを同定し、該ポリペプチドを本出願において「ＰＲＯ２４４」と命名した。

一実施態様では、本発明はＰＲＯ２４４ポリペプチドをコードするＤＮＡを含んでなる単離された核酸分子を提供する。一態様では、単離された核酸は、Ｆｉｇ１２２（配列番号：３７７）のアミノ酸残基１から２１９を有するＰＲＯ２４４ポリペプチドをコードしているＤＮＡを含有するか、そのようなコード化核酸配列に相補的であり、少なくとも中程度、場合によっては高緊縮性条件下でそれに安定して結合したままである。

他の実施態様では、本発明は、単離されたＰＲＯ２４４ポリペプチドを提供する。特に、本発明は単離された天然配列ＰＲＯ３３０ポリペプチドを提供し、それは一実施態様では、Ｆｉｇ１２２（配列番号：３７７）の残基１から２１９を含んでなるアミノ酸配列を含む。

【０１０１】

５０．更なる実施態様

本発明の他の実施態様では、本発明はここに記載したポリペプチドの任意のものをコードするＤＮＡを含んでなるベクターを提供する。そのようなベクターを含んでなる宿主細胞もまた提供される。例を挙げると、宿主細胞はＣＨＯ細胞、大腸菌、又は酵母でありうる。ここに記載のポリペプチドの任意のものを生産するための方法が更に提供され、これは、所望のポリペプチドの発現に適した条件下で宿主細胞を培養し、細胞培養から所望のポリペプチドを回収することを含んでなる。

他の実施態様では、本発明は、異種性ポリペプチド又はアミノ酸配列に融合したここに記載のポリペプチドの任意のものを含んでなるキメラ分子を提供する。そのようなキメラ分子の例は、免疫グロブリンのＦｃ領域又はエピトープタグ配列に融合したここに記載のポリペプチドの任意のものが含まれる。

他の実施態様では、本発明は上述の又は下記のポリペプチドの任意のものに特異的に結合する抗体を提供する。場合によっては、抗体はモノクローナル抗体、ヒト化抗体、抗体断片又は一本鎖抗体である。

さらに他の実施態様では、本発明は、ゲノム及びｃＤＮＡヌクレオチド配列を単離するために有用なオリゴヌクレオチドプローブを提供し、ここでこれらプローブは上述の又は下記のヌクレオチド配列の任意のものから誘導されうる。

他の実施態様では、本発明はＰＲＯポリペプチドをコードするヌクレオチド配列を含んでなる単離された核酸分子を提供する。

一態様では、単離された核酸分子は、（ａ）ここに開示された全長アミノ酸配列、ここに開示されたシグナルペプチドを欠くアミノ酸配列、又はシグナルペプチド有無のここに開示した膜貫通タンパク質の細胞外ドメインを有するＰＲＯポリペプチドをコードするＤＮＡ分子、又は（ｂ）（ａ）のＤＮＡ分子の補体に対して、少なくとも約８０％の配列同一性、好ましくは少なくとも約８１％の配列同一性、更により好ましくは少なくとも約８２％の配列同一性、更により好ましくは少なくとも約８３％の配列同一性、更により好ましくは少なくとも約８４％の配列同一性、更により好ましくは少なくとも約８５％の配列同一性、更により好ましくは少なくとも約８６％の配列同一性、更により好ましくは少なくとも約８７％の配列同一性、更により好ましくは少なくとも約８８％の配列同一性、更により好ましくは少なくとも約８９％の配列同一性、更により好ましくは少なくとも約９０％の酸配列同一性、更により好ましくは少なくとも約９１％の配列同一性、更により好ましくは少なくとも約９２％の配列同一性、更により好ましくは少なくとも約９３％の配列同一性、更により好ましくは少なくとも約９４％の配列同一性、更により好ましくは少なくとも約９５％の配列同一性、更により好ましくは少なくとも約９６％の配列同一性、更により好ましくは少なくとも約９７％の配列同一性、更により好ましくは少なくとも約９８％の配列同一性、更により好ましくは少なくとも約９９％の配列同一性を有しているヌクレオチド配列を含んでなる。

【０１０２】

他の態様では、単離された核酸分子は、（ａ）ここに開示された全長ＰＲＯポリペプチドｃＤＮＡのコード化配列、ここに開示されたシグナルペプチドを欠くＰＲＯポリペプチドｃＤＮＡのコード化配列、又はシグナルペプチド有無のここに開示した膜貫通ＰＲＯポリペプチドの細胞外ドメインのコード化配列、又は（ｂ）（ａ）のＤＮＡ分子の補体に対して、少なくとも約８０％の配列同一性、好ましくは少なくとも約８１％の配列同一性、更により好ましくは少なくとも約８２％の配列同一性、更により好ましくは少なくとも約８３％の配列同一性、更により好ましくは少なくとも約８４％の配列同一性、更により好ましくは少なくとも約８５％の配列同一性、更により好ましくは少なくとも約８６％の配列同一性、更により好ましくは少なくとも約８７％の配列同一性、更により好ましくは少なくとも約８８％の配列同一性、更により好ましくは少なくとも約８９％の配列同一性、更により好ましくは少なくとも約９０％の酸配列同一性、更により好ましくは少なくとも約９１％の配列同一性、更により好ましくは少なくとも約９２％の配列同一性、更により好ましくは少なくとも約９３％の配列同一性、更により好ましくは少なくとも約９４％の配列同一性、更により好ましくは少なくとも約９５％の配列同一性、更により好ましくは少

なくとも約９６％の配列同一性、更により好ましくは少なくとも約９７％の配列同一性、更により好ましくは少なくとも約９８％の配列同一性、更により好ましくは少なくとも約９９％の配列同一性を有しているヌクレオチド配列を含んでなる。
更なる態様では、本発明は、（ａ）ここに開示されたＡＴＴＣに寄託されたヒトタンパク質ｃＤＮＡの任意のものによりコードされる同じ成熟ポリペプチドをコードするＤＮＡ分子、又は（ｂ）（ａ）のＤＮＡ分子の補体に対して、少なくとも約８０％の配列同一性、好ましくは少なくとも約８１％の配列同一性、更により好ましくは少なくとも約８２％の配列同一性、更により好ましくは少なくとも約８３％の配列同一性、更により好ましくは少なくとも約８４％の配列同一性、更により好ましくは少なくとも約８５％の配列同一性、更により好ましくは少なくとも約８６％の配列同一性、更により好ましくは少なくとも約８７％の配列同一性、更により好ましくは少なくとも約８８％の配列同一性、更により好ましくは少なくとも約８９％の配列同一性、更により好ましくは少なくとも約９０％の酸配列同一性、更により好ましくは少なくとも約９１％の配列同一性、更により好ましくは少なくとも約９２％の配列同一性、更により好ましくは少なくとも約９３％の配列同一性、更により好ましくは少なくとも約９４％の配列同一性、更により好ましくは少なくとも約９５％の配列同一性、更により好ましくは少なくとも約９６％の配列同一性、更により好ましくは少なくとも約９７％の配列同一性、更により好ましくは少なくとも約９８％の配列同一性、更により好ましくは少なくとも約９９％の配列同一性を有しているヌクレオチド配列を含んでなる単離された核酸分子に関する。
本発明の他の態様は、膜貫通ドメインが欠失されているか膜貫通ドメインが不活性化されているＰＲＯポリペプチドをコードするヌクレオチド配列を含んでなる単離された核酸分子を提供し、またはそのようなコード化ヌクレオチド配列に相補的であり、ここでそのようなポリペプチドの膜貫通ドメインがここに開示される。従って、ここに記載されたＰＲＯポリペプチドの可溶性細胞外ドメインが考慮される。
【０１０３】
他の実施態様は、例えばハイブリッド形成プローブとしての又は抗－ＰＲＯ抗体に対する結合部位を含んでなるポリペプチドを場合によってはコードするＰＲＯポリペプチドの断片をコードするための用途が見出されるＰＲＯポリペプチドコード配列の断片に関する。そのような核酸断片は通常少なくとも約２０のヌクレオチド長さ、好ましくは少なくとも約３０のヌクレオチド長さ、より好ましくは少なくとも約４０のヌクレオチド長さ、更により好ましくは少なくとも約５０のヌクレオチド長さ、更により好ましくは少なくとも約６０のヌクレオチド長さ、更により好ましくは少なくとも約７０のヌクレオチド長さ、更により好ましくは少なくとも約８０のヌクレオチド長さ、更により好ましくは少なくとも約９０のヌクレオチド長さ、更により好ましくは少なくとも約１００のヌクレオチド長さ、更により好ましくは少なくとも約１１０のヌクレオチド長さ、更により好ましくは少なくとも約１２０のヌクレオチド長さ、更により好ましくは少なくとも約１３０のヌクレオチド長さ、更により好ましくは少なくとも約１４０のヌクレオチド長さ、更により好ましくは少なくとも約１５０のヌクレオチド長さ、更により好ましくは少なくとも約１６０のヌクレオチド長さ、更により好ましくは少なくとも約１７０のヌクレオチド長さ、更により好ましくは少なくとも約１８０のヌクレオチド長さ、更により好ましくは少なくとも約１９０のヌクレオチド長さ、更により好ましくは少なくとも約２００のヌクレオチド長さ、更により好ましくは少なくとも約２５０のヌクレオチド長さ、更により好ましくは少なくとも約３００のヌクレオチド長さ、更により好ましくは少なくとも約３５０のヌクレオチド長さ、更により好ましくは少なくとも約４００のヌクレオチド長さ、更により好ましくは少なくとも約４５０のヌクレオチド長さ、更により好ましくは少なくとも約５００のヌクレオチド長さ、更により好ましくは少なくとも約６００のヌクレオチド長さ、更により好ましくは少なくとも約７００のヌクレオチド長さ、更により好ましくは少なくとも約８００のヌクレオチド長さ、更により好ましくは少なくとも約９００のヌクレオチド長さ、更により好ましくは少なくとも約１０００のヌクレオチド長さであり、ここで、「約」という用語は、その参照長さのプラス又はマイナス１０％の参照ヌクレオチド配列長さを

意味する。ＰＲＯポリペプチドコード化ヌクレオチド配列の新規な断片は、多くのよく知られた配列アラインメントプログラムの任意のものを使用してＰＲＯポリペプチドコードヌクレオチド配列を他の既知のヌクレオチド配列にアラインメントさせ、どのＰＲＯポリペプチドコード化ヌクレオチド配列断片が新規であるかを決定することにより常套的に決定することができることに留意される。そのようなＰＲＯポリペプチドコード化ヌクレオチド配列の全てがここで考慮される。また考慮されるものは、これらのヌクレオチド分子断片によりコードされるＰＲＯポリペプチド断片、好ましくは抗－ＰＲＯ抗体に対する結合部位を含んでなるＰＲＯポリペプチド断片である。

他の実施態様では、本発明は上記において特定した単離核酸配列の任意のものによりコードされる単離されたＰＲＯポリペプチドを提供する。

【０１０４】

ある態様では、本発明は、ここに開示された全長アミノ酸配列、ここに開示されたシグナルペプチドを欠く全長アミノ酸配列、又はシグナルペプチド有無のここに開示した膜貫通タンパク質の細胞外ドメインを有するＰＲＯポリペプチドに対して、少なくとも約８０％の配列同一性、好ましくは少なくとも約８１％の配列同一性、更により好ましくは少なくとも約８２％の配列同一性、更により好ましくは少なくとも約８３％の配列同一性、更により好ましくは少なくとも約８４％の配列同一性、更により好ましくは少なくとも約８５％の配列同一性、更により好ましくは少なくとも約８６％の配列同一性、更により好ましくは少なくとも約８７％の配列同一性、更により好ましくは少なくとも約８８％の配列同一性、更により好ましくは少なくとも約８９％の配列同一性、更により好ましくは少なくとも約９０％の酸配列同一性、更により好ましくは少なくとも約９１％の配列同一性、更により好ましくは少なくとも約９２％の配列同一性、更により好ましくは少なくとも約９３％の配列同一性、更により好ましくは少なくとも約９４％の配列同一性、更により好ましくは少なくとも約９５％の配列同一性、更により好ましくは少なくとも約９６％の配列同一性、更により好ましくは少なくとも約９７％の配列同一性、更により好ましくは少なくとも約９８％の配列同一性、更により好ましくは少なくとも約９９％の配列同一性を有するアミノ酸配列を含んでなる、単離されたＰＲＯポリペプチドに関する。

更なる態様では、本発明は、ここに開示されたＡＴＴＣに寄託されたヒトタンパク質ｃＤＮＡの任意のものによりコードされるアミノ酸配列に対して、少なくとも約８０％の配列同一性、好ましくは少なくとも約８１％の配列同一性、更により好ましくは少なくとも約８２％の配列同一性、更により好ましくは少なくとも約８３％の配列同一性、更により好ましくは少なくとも約８４％の配列同一性、更により好ましくは少なくとも約８５％の配列同一性、更により好ましくは少なくとも約８６％の配列同一性、更により好ましくは少なくとも約８７％の配列同一性、更により好ましくは少なくとも約８８％の配列同一性、更により好ましくは少なくとも約８９％の配列同一性、更により好ましくは少なくとも約９０％の酸配列同一性、更により好ましくは少なくとも約９１％の配列同一性、更により好ましくは少なくとも約９２％の配列同一性、更により好ましくは少なくとも約９３％の配列同一性、更により好ましくは少なくとも約９４％の配列同一性、更により好ましくは少なくとも約９５％の配列同一性、更により好ましくは少なくとも約９６％の配列同一性、更により好ましくは少なくとも約９７％の配列同一性、更により好ましくは少なくとも約９８％の配列同一性、更により好ましくは少なくとも約９９％の配列同一性を有するアミノ酸配列を含んでなる単離されたＰＲＯポリペプチドに関する。

【０１０５】

さらなる態様では、本発明は、ここに開示された全長アミノ酸配列、ここに開示されたシグナルペプチドを欠く全長アミノ酸配列、又はシグナルペプチド有無のここに開示された膜貫通タンパク質の細胞外ドメインを有するＰＲＯポリペプチドのアミノ酸配列と比較したとき、少なくとも約８０％のポジティブ、好ましくは少なくとも約８１％のポジティブ、より好ましくは少なくとも約８２％のポジティブ、更により好ましくは少なくとも約８３％のポジティブ、更により好ましくは少なくとも約８４％のポジティブ、更により好ましくは少なくとも約８５％のポジティブ、更により好ましくは少なくとも約８６％のポジ

ティブ、更により好ましくは少なくとも約８７％のポジティブ、更により好ましくは少なくとも約８８％のポジティブ、更により好ましくは少なくとも約８９％のポジティブ、更により好ましくは少なくとも約９０％のポジティブ、更により好ましくは少なくとも約９１％のポジティブ、更により好ましくは少なくとも約９２％のポジティブ、更により好ましくは少なくとも約９３％のポジティブ、更により好ましくは少なくとも約９４％のポジティブ、更により好ましくは少なくとも約９５％のポジティブ、更により好ましくは少なくとも約９６％のポジティブ、更により好ましくは少なくとも約９７％のポジティブ、更により好ましくは少なくとも約９８％のポジティブ、更により好ましくは少なくとも約９９％のポジティブのスコアを示すアミノ酸配列を含んでなる単離されたＰＲＯポリペプチドに関する。

特定の態様では、本発明は、Ｎ末端シグナル配列及び／又は開始メチオニンを持たない単離されたＰＲＯポリペプチドを提供し、それは上述したそのようなアミノ酸配列をコードするヌクレオチド配列によりコードされる。これを生産する方法もまたここに記載され、ここで、これらの方法はＰＲＯポリペプチドの発現に適した条件下で適当なコード化核酸分子を含むベクターを含んでなる宿主細胞を培養し、細胞培地からＰＲＯポリペプチドを回収することを含んでなる。

【０１０６】

本発明の他の態様は、膜貫通ドメインが欠失されたか膜貫通ドメインが不活性化された単離されたＰＲＯポリペプチドを提供する。これを生産する方法もまたここに記載され、ここで、これらの方法はＰＲＯポリペプチドの発現に適した条件下で適当なコード化核酸分子を含むベクターを含んでなる宿主細胞を培養し、細胞培地からＰＲＯポリペプチドを回収することを含んでなる。

さらに他の実施態様では、本発明は、ここで特定した天然ＰＲＯポリペプチドのアゴニスト及びアンタゴニストに関する。

さらなる実施態様では、本発明は、ＰＲＯポリペプチドを候補分子と接触させ、前記ＰＲＯポリペプチドによって媒介される生物学的活性を監視することを含んでなる、ＰＲＯポリペプチドに対するアゴニスト又はアンタゴニストの同定方法に関する。

またさらなる実施態様では、本発明は、ＰＲＯポリペプチド、又はここに記載のアゴニスト又はアンタゴニスト、又は抗－ＰＲＯ抗体を担体とともに含んでなる物質の組成物に関する。場合によっては、担体は製薬的に許容される担体である。

本発明の他の実施態様は、上述のＰＲＯポリペプチド、そのアゴニスト又はアンタゴニスト又は抗－ＰＲＯ抗体に応答性である症状の治療に有用な医薬の調製のための、ＰＲＯポリペプチド、またはそのアゴニスト又はアンタゴニスト又は抗－ＰＲＯ抗体の使用に関する。

【０１０７】

（好適な実施態様の詳細な説明）

Ｉ．定義

ここで使用される際の「ＰＲＯポリペプチド」及び「ＰＲＯ」又は「ＵＣＰ」という用語は、直後に数値符号がある場合に種々のポリペプチドを指し、完全な符号（例えば、ＰＲＯ／数字）は、ここに記載する特定のポリペプチド配列を意味する。ここで使用される「ＰＲＯ／数字ポリペプチド」及び「ＰＲＯ／数字」であって、「数字」がここで使用される実際の数値符号として与えられる用語は、天然配列ポリペプチド及び変異体（ここで更に詳細に定義する）を含む。ここに記載されるＰＲＯポリペプチドは、ヒト組織型又は他の供給源といった種々の供給源から単離してもよく、組換え又は合成方法によって調製してもよい。

「天然配列ＰＲＯポリペプチド」又は「ＵＣＰ」は、天然由来の対応するＰＲＯポリペプチドと同一のアミノ酸配列を有するポリペプチドを含んでいる。このような天然配列ＰＲＯポリペプチドは、自然から単離することもできるし、組換え又は合成手段により生産することもできる。「天然配列ＰＲＯポリペプチド」という用語には、特に、特定のＰＲＯポリペプチドの自然に生じる切断又は分泌形態（例えば、細胞外ドメイン配列）、自然に

生じる変異形態（例えば、選択的にスプライシングされた形態）及びそのポリペプチドの自然に生じる対立遺伝子変異体が含まれる。本発明の種々の実施態様において、天然配列ＰＲＯポリペプチドは、添付の図面に示される全長アミノ酸配列を含む成熟又は全長天然配列ポリペプチドである。開始及び停止コドンは、図において太字又は下線で示した。しかし、添付の図面に開示したＰＲＯポリペプチドは、図面におけるアミノ酸１としてここに命名されるメチオニン残基で始まるように示されているが、図面におけるアミノ酸位置１の上流又は下流に位置する他のメチオニン残基をＰＲＯポリペプチドの開始アミノ酸残基として用いることも考えられるし可能でもある。

ＰＲＯポリペプチド「細胞外ドメイン」又は「ＥＣＤ」は、膜貫通及び細胞質ドメインを実質的に有しないＰＲＯポリペプチドの形態を意味する。通常、ＰＲＯポリペプチドＥＣＤは、それらの膜貫通及び／又は細胞質ドメインを１％未満、好ましくはそのようなドメインを０．５％未満しか持たない。本発明のＰＲＯポリペプチドについて同定された任意の膜貫通ドメインは、疎水性ドメインのその型を同定するために当該分野において日常的に使用される基準に従い同定されることが理解されるであろう。膜貫通ドメインの厳密な境界は変わり得るが、最初に同定されたドメインのいずれかの末端から約５アミノ酸を越えない可能性が高い。従って、ＰＲＯポリペプチド細胞外ドメインは、場合によっては、実施例又は明細書で同定されるように膜貫通ドメイン及び／又は細胞外ドメインの境界のいずれかの側から約５を越えないアミノ酸を含んでもよく、シグナルペプチドを伴う又は伴わない、それらのポリペプチド及びそれらをコードする核酸は、本発明で考慮される。

【０１０８】

ここに開示する種々のＰＲＯポリペプチドの「シグナルペプチド」の適切な位置は、添付の図面に示す。しかし、注記するように、シグナルペプチドのＣ－末端境界は変化しうるが、ここで最初に定義したようにシグナルペプチドＣ－末端境界のいずれかの側で約５アミノ酸未満である可能性が最も高く、シグナルペプチドのＣ－末端境界は、そのような型のアミノ酸配列成分を同定するのに日常的に使用される基準に従って同定しうる（例えば、Ｎｉｅｌｓｅｎ等，Ｐｒｏｔ．Ｅｎｇ．１０：１－６（１９９７）及びｖｏｎ Ｈｅｉｎｊｅ等，Ｎｕｃｌ．Ａｃｉｄｓ．Ｒｅｓ．１４：４６８３－４６９０（１９８６））。さらに、幾つかの場合には、分泌ポリペプチドからのシグナルペプチドの切断は完全に均一ではなく、一以上の分泌種をもたらすことも認められる。シグナルペプチドがここに定義されるシグナルペプチドのＣ－末端境界の何れかの側の約５アミノ酸未満内で切断されるこれらの成熟ポリペプチド、及びそれらをコードするポリヌクレオチドは、本発明で考慮される。

「ＰＲＯポリペプチド変異体」とは、上記又は下記に定義されるように、ここに開示される全長天然配列ＰＲＯポリペプチド、ここに開示されたシグナルペプチドを欠く全長天然配列ＰＲＯポリペプチド配列、シグナルペプチド有無のここに開示されたＰＲＯの細胞外ドメイン又はここに開示された全長ＰＲＯポリペプチドの他の断片と、少なくとも約８０％のアミノ酸配列同一性を有する活性ＰＲＯポリペプチドを意味する。このようなＰＲＯポリペプチド変異体には、例えば、全長天然アミノ酸配列のＮ－又はＣ－末端において一又は複数のアミノ酸残基が付加、もしくは欠失されたＰＲＯポリペプチドが含まれる。通常、ＰＲＯポリペプチド変異体は、ここに開示される全長天然アミノ酸配列、ここに開示されたシグナルペプチドを欠く全長天然配列ＰＲＯポリペプチド配列、シグナルペプチド有無のここに開示されたＰＲＯの細胞外ドメイン又はここに開示された全長ＰＲＯポリペプチドの他の断片と、少なくとも約８０％のアミノ酸配列同一性、好ましくは少なくとも約８１％のアミノ酸配列同一性、より好ましくは少なくとも約８２％のアミノ酸配列同一性、より好ましくは少なくとも約８３％のアミノ酸配列同一性、より好ましくは少なくとも約８４％のアミノ酸配列同一性、より好ましくは少なくとも約８５％のアミノ酸配列同一性、より好ましくは少なくとも約８６％のアミノ酸配列同一性、より好ましくは少なくとも約８７％のアミノ酸配列同一性、より好ましくは少なくとも約８８％のアミノ酸配列同一性、より好ましくは少なくとも約８９％のアミノ酸配列同一性、より好ましくは少なくとも約９０％のアミノ酸配列同一性、より好ましくは少なくとも約９１％のアミノ酸配

列同一性、より好ましくは少なくとも約９２％のアミノ酸配列同一性、より好ましくは少なくとも約９３％のアミノ酸配列同一性、より好ましくは少なくとも約９４％のアミノ酸配列同一性、より好ましくは少なくとも約９５％のアミノ酸配列同一性、より好ましくは少なくとも約９６％のアミノ酸配列同一性、より好ましくは少なくとも約９７％のアミノ酸配列同一性、より好ましくは少なくとも約９８％のアミノ酸配列同一性、そして、より好ましくは少なくとも約９９％のアミノ酸配列同一性を有している。通常は、ＰＲＯ変異体ポリペプチドは、少なくとも約１０アミノ酸長、より多くは少なくとも約２０アミノ酸長、より多くは少なくとも約３０アミノ酸長、より多くは少なくとも約４０アミノ酸長、より多くは少なくとも約５０アミノ酸長、より多くは少なくとも約６０アミノ酸長、より多くは少なくとも約７０アミノ酸長、より多くは少なくとも約８０アミノ酸長、より多くは少なくとも約９０アミノ酸長、より多くは少なくとも約１００アミノ酸長、より多くは少なくとも約１５０アミノ酸長、より多くは少なくとも約２００アミノ酸長、より多くは少なくとも約３００アミノ酸長、又はそれ以上である。

【０１０９】

ここに定義されるＰＲＯポリペプチドに対してここで同定されている「パーセント（％）アミノ酸配列同一性」は、配列を整列させ、最大のパーセント配列同一性を得るために必要ならば間隙を導入し、如何なる保存的置換も配列同一性の一部と考えないとした、ＰＲＯポリペプチドのアミノ酸残基と同一である候補配列中のアミノ酸残基のパーセントとして定義される。パーセントアミノ酸配列同一性を決定する目的のためのアラインメントは、当業者の技量の範囲にある種々の方法、例えばＢＬＡＳＴ、ＢＬＡＳＴ－２、ＡＬＩＧＮ又はＭｅｇａｌｉｇｎ（ＤＮＡＳＴＡＲ）ソフトウエアのような公に入手可能なコンピュータソフトウエアを使用することにより達成可能である。当業者であれば、比較される配列の全長に対して最大のアラインメントを達成するために必要な任意のアルゴリズムを含む、アラインメントを測定するための適切なパラメータを決定することができる。しかし、ここでの目的のためには、％アミノ酸配列同一性値は、配列比較コンピュータプログラムＡＬＩＧＮ－２を用いて発生させ、ここでＡＬＩＧＮ－２プログラムに対する完全なソースコードは以下の表１０に提供される。ＡＬＩＧＮ－２配列比較コンピュータプログラムはジェネンテク社により作成され、以下の表１０に示されたソースコードは、米国著作権庁、Ｗａｓｈｉｎｇｔｏｎ Ｄ．Ｃ．, ２０５５９にユーザー資料と共に提出されており、米国著作権登録番号ＴＸＵ５１００８７で登録されている。ＡＬＩＧＮ－２プログラムはジェネンテク社（Ｓｏｕｔｈ Ｓａｎ Ｆｒａｎｃｉｓｃｏ, Ｃａｌｉｆｏｒｎｉａ）を通じて公的に利用でき、以下の表１０に提供されるソースコードからコンパイルすることができる。ＡＬＩＧＮ－２プログラムはＵＮＩＸオペレーティングシステム、好ましくはデジタルＵＮＩＸ Ｖ４．０Ｄでの使用のためにコンパイルされなければならない。全ての配列比較パラメータはＡＬＩＧＮ－２プログラムにより設定され、変動しない。

【０１１０】

アミノ酸配列比較にＡＬＩＧＮ－２が用いられる状況では、与えられたアミノ酸配列Ａの、与えられたアミノ酸配列Ｂとの、又はそれに対する％アミノ酸配列同一性（あるいは、与えられたアミノ酸配列Ｂと、又はそれに対して或る程度の％アミノ酸配列同一性を持つ又は含む与えられたアミノ酸配列Ａと言うこともできる）は次のように計算される：

分率Ｘ／Ｙの１００倍

ここで、Ｘは配列アラインメントプログラムＡＬＩＧＮ－２のＡ及びＢのアラインメントによって同一であると一致したスコアのアミノ酸残基の数であり、ＹはＢの全アミノ酸残基数である。アミノ酸配列Ａの長さがアミノ酸配列Ｂの長さと異なる場合、ＡのＢに対する％アミノ酸配列同一性は、ＢのＡに対する％アミノ酸配列同一性とは異なることは理解されるであろう。この方法を用いた％アミノ酸配列同一性の計算の例として、以下の表１１及び１２は、「比較タンパク質」と称されるアミノ酸配列の「ＰＲＯ」と称されるアミノ酸配列に対する％アミノ酸配列同一性を如何に計算するかを証明する。

特に断らない限り、ここで使用される全ての％アミノ酸配列同一性値は、ＡＬＩＧＮ－２コンピュータプログラムを使用してすぐ前の段落に記載したようにして得られる。しかし

、％アミノ酸配列同一性は、ＷＵ－ＢＬＡＳＴ－２コンピュータプログラム（Ａｌｔｓｃｈｕｌ等，Ｍｅｔｈｏｄｓ ｉｎ Ｅｎｚｙｍｏｌｏｇｙ ２６６：４６０－４８０（１９９６））を用いて以下に記載されているようにして得ることもまたできる。ＷＵ－ＢＬＡＳＴ－２検索パラメータの殆どはデフォルト値に設定される。デフォルト値に設定されないもの、すなわち調節できるパラメータは、次の値に設定される：オーバーラップスパン＝１、オーバーラップフラクション＝０．１２５、ワード閾値（Ｔ）＝１１、及びスコアリングマトリックス＝ＢＬＯＳＵＭ６２。ＷＵ－ＢＬＡＳＴ－２を用いる場合、％アミノ酸配列同一性値は、（ａ）天然ＰＲＯポリペプチドから誘導される配列を有する対象のＰＲＯポリペプチドのアミノ酸配列と対象の比較アミノ酸配列（すなわち、ＰＲＯ変異体ポリペプチドであってもよい対象のＰＲＯポリペプチドが比較されている配列）の間の合致する同一のアミノ酸残基の数を、（ｂ）対象のＰＲＯポリペプチドのアミノ酸残基の全数により割ることにより決定される。例えば、「アミノ酸配列Ｂと少なくとも８０％のアミノ酸配列同一性を有する又は有しているアミノ酸配列Ａを含むポリペプチド」という記載において、アミノ酸配列Ａは対象の比較アミノ酸配列であり、アミノ酸配列Ｂは対象のＰＲＯポリペプチドのアミノ酸配列である。

また、％アミノ酸配列同一性は、配列比較プログラムＮＣＢ１－ＢＬＡＳＴ－２（Ａｌｔｓｃｈｕｌ等，Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．２５：３３８９－３４０２（１９９７））を用いて決定することもできる。ＮＣＢ１－ＢＬＡＳＴ２配列比較プログラムはｈｔｔｐ：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｏｖからダウンロードすることができる。ＮＣＢ１－ＢＬＡＳＴ２は幾つかの検索パラメータを使用し、それら検索パラメータの全ては、例えば、ｕｎｍａｓｋ＝可、鎖＝全て、予測される発生＝１０、最小低複合長＝１５／５、マルチパスｅ－値＝０．０１、マルチパスの定数＝２５、最終ギャップアラインメントのドロップオフ＝２５、及びスコアリングマトリクス＝ＢＬＯＳＵＭ６２を含むデフォルト値に設定される。

アミノ酸配列比較にＮＣＢ１－ＢＬＡＳＴ２が用いれれる状況では、与えられたアミノ酸配列Ａの、与えられたアミノ酸配列Ｂとの、又はそれに対する％アミノ酸配列同一性（あるいは、与えられたアミノ酸配列Ｂと、又はそれに対して或る程度の％アミノ酸配列同一性を持つ又は含む与えられたアミノ酸配列Ａと言うこともできる）は次のように計算される：

分率Ｘ／Ｙの１００倍

ここで、Ｘは配列アラインメントプログラムＮＣＢ１－ＢＬＡＳＴ２のＡ及びＢのアラインメントによって同一であると一致したスコアのアミノ酸残基の数であり、ＹはＢの全アミノ酸残基数である。アミノ酸配列Ａの長さがアミノ酸配列Ｂの長さと異なる場合、ＡのＢに対する％アミノ酸配列同一性は、ＢのＡに対する％アミノ酸配列同一性とは異なることは理解されるであろう。

【０１１１】

「ＰＲＯ変異体ポリヌクレオチド」又は「ＰＲＯ変異体核酸配列」とは、下記に定義されるように、活性ＰＲＯポリペプチドをコードする核酸分子であり、ここに開示する全長天然配列ＰＲＯポリペプチド配列、ここに開示するシグナルペプチドを欠いた全長天然配列ＰＲＯポリペプチド配列、シグナルペプチド有無のここに開示するＰＲＯポリペプチドの細胞外ドメイン、又はここに開示する全長ＰＲＯポリペプチド配列の他の任意の断片をコードする核酸配列と少なくとも８０％の配列同一性を有する。通常は、ＰＲＯ変異体ポリペプチドヌクレオチドは、ここに開示する全長天然配列ＰＲＯポリペプチド配列、ここに開示するシグナルペプチドを欠いた全長天然配列ＰＲＯポリペプチド配列、シグナルペプチド有無のここに開示するＰＲＯポリペプチドの細胞外ドメイン、又はここに開示する全長ＰＲＯポリペプチドの他の任意の断片をコードする核酸配列と、少なくとも約８０％の核酸配列同一性、好ましくは少なくとも約８１％の核酸配列同一性、より好ましくは少なくとも約８２％の核酸配列同一性、より好ましくは少なくとも約８３％の核酸配列同一性、より好ましくは少なくとも約８４％の核酸配列同一性、より好ましくは少なくとも約８５％の核酸配列同一性、より好ましくは少なくとも約８６％の核酸配列同一性、より好ま

しくは少なくとも約８７％の核酸配列同一性、より好ましくは少なくとも約８８％の核酸配列同一性、より好ましくは少なくとも約８９％の核酸配列同一性、より好ましくは少なくとも約９０％の核酸配列同一性、より好ましくは少なくとも約９１％の核酸配列同一性、より好ましくは少なくとも約９２％の核酸配列同一性、より好ましくは少なくとも約９３％の核酸配列同一性、より好ましくは少なくとも約９４％の核酸配列同一性、より好ましくは少なくとも約９５％の核酸配列同一性、より好ましくは少なくとも約９６％の核酸配列同一性、より好ましくは少なくとも約９７％の核酸配列同一性、より好ましくは少なくとも約９８％の核酸配列同一性、そして、より好ましくは少なくとも約９９％の核酸配列同一性を有している。変異体は、天然ヌクレオチド配列を包含しない。

通常は、ＰＲＯ変異体ポリヌクレオチドは、少なくとも約３０ヌクレオチド長、より多くは少なくとも約６０ヌクレオチド長、より多くは少なくとも約９０ヌクレオチド長、より多くは少なくとも約１２０ヌクレオチド長、より多くは少なくとも約１５０ヌクレオチド長、より多くは少なくとも約１８０ヌクレオチド長、より多くは少なくとも約２１０ヌクレオチド長、より多くは少なくとも約２４０ヌクレオチド長、より多くは少なくとも約２７０ヌクレオチド長、より多くは少なくとも約３００ヌクレオチド長、より多くは少なくとも約４５０ヌクレオチド長、より多くは少なくとも約６００ヌクレオチド長、より多くは少なくとも約９００ヌクレオチド長、又はそれ以上である。

【０１１２】

ここで同定されるＰＲＯコード化核酸配列に対する「パーセント（％）核酸配列同一性」は、配列を整列させ、最大のパーセント配列同一性を得るために必要ならば間隙を導入し、ＰＲＯ配列のヌクレオチドと同一である候補配列中のヌクレオチドのパーセントとして定義される。パーセント核酸配列同一性を決定する目的のためのアラインメントは、当業者の知る範囲にある種々の方法、例えばＢＬＡＳＴ、ＢＬＡＳＴ－２、ＡＬＩＧＮ又はＭｅｇａｌｉｇｎ（ＤＮＡＳＴＡＲ）ソフトウエアのような公に入手可能なコンピュータソフトウエアを使用することにより達成可能である。しかし、ここでの目的のためには、％核酸配列同一性値は、配列比較コンピュータプログラムＡＬＩＧＮ－２を用いて発生され、ここでＡＬＩＧＮ－２プログラムに対する完全なソースコードは以下の表１０に提供される。ＡＬＩＧＮ－２配列比較コンピュータプログラムはジェネンテク社により作成され、以下の表１０に示されたソースコードは、米国著作権庁、Ｗａｓｈｉｎｇｔｏｎ Ｄ．Ｃ．，２０５５９にユーザー資料と共に提出されており、米国著作権登録番号ＴＸＵ５１００８７で登録されている。ＡＬＩＧＮ－２プログラムはジェネンテク社（Ｓｏｕｔｈ Ｓａｎ Ｆｒａｎｃｉｓｃｏ，Ｃａｌｉｆｏｒｎｉａ）を通じて公的に利用でき、以下の表１０に提供されるソースコードからコンパイルすることができる。ＡＬＩＧＮ－２プログラムはＵＮＩＸオペレーティングシステム、好ましくはデジタルＵＮＩＸ Ｖ４．０Ｄでの使用のためにコンパイルされなければならない。全ての配列比較パラメータはＡＬＩＧＮ－２プログラムにより設定され、変動しない。

【０１１３】

核酸配列比較にＡＬＩＧＮ－２が用いれられる状況では、与えられた核酸配列Ｃの、与えられた核酸配列Ｄとの、又はそれに対する％核酸配列同一性（あるいは、与えられた核酸配列Ｄと、又はそれに対して或る程度の％核酸配列同一性を持つ又は含む与えられた核酸配列Ｃと言うこともできる）は次のように計算される：

分率Ｗ／Ｚの１００倍

ここで、Ｗは配列アラインメントプログラムＡＬＩＧＮ－２のＣ及びＤのアラインメントによって同一であると一致したスコアのヌクレオチドの数であり、ＺはＤのヌクレオチドの全数である。核酸配列Ｃの長さが核酸配列Ｄの長さと異なる場合、ＣのＤに対する％核酸配列同一性は、ＤのＣに対する％核酸配列同一性とは異なることは理解されるであろう。％核酸配列同一性の計算の例として、以下の表１３及び１４は、「比較ＤＮＡ」と称される核酸配列の「ＰＲＯ－ＤＮＡ」と称される核酸配列に対する％核酸配列同一性を如何に計算するかを証明する。

特に断らない限り、ここで使用される全ての％核酸配列同一性値は、ＡＬＩＧＮ－２コン

ピュータプログラムを使用してすぐ前の段落に記載したようにして得られる。しかし、％核酸配列同一性は、ＷＵ－ＢＬＡＳＴ－２コンピュータプログラム（Ａｌｔｓｃｈｕｌ等，Ｍｅｔｈｏｄｓ ｉｎ Ｅｎｚｙｍｏｌｏｇｙ ２６６：４６０－４８０ （１９９６））を用いて以下に記載されているようにして得ることもまたできる。ＷＵ－ＢＬＡＳＴ－２検索パラメータの殆どはデフォルト値に設定される。デフォルト値に設定されないもの、すなわち調節できるパラメータは、次の値に設定される：オーバーラップスパン＝１、オーバーラップフラクション＝０．１２５、ワード閾値（Ｔ）＝１１、及びスコアリングマトリックス＝ＢＬＯＳＵＭ６２。ＷＵ－ＢＬＡＳＴ－２を用いる場合、％核酸配列同一性値は、（ａ）天然配列ＰＲＯポリペプチドコード化核酸から誘導される配列を有する対象のＰＲＯポリペプチドコード化核酸分子と対象の比較核酸分子（すなわち、変異体ＰＲＯポリヌクレオチドであってもよい対象のＰＲＯポリペプチドーコード化核酸分子が比較されている配列）の間の合致する同一のヌクレオチドの数を、（ｂ）対象のＰＲＯポリペプチドコード化核酸分子のヌクレオチドの全数により割ることにより決定される。例えば、「核酸配列Ｂと少なくとも８０％の核酸配列同一性を有する又は有している核酸配列Ａを含む単離された核酸分子」という記載において、核酸配列Ａは対象の比較核酸配列であり、核酸配列Ｂは対象のＰＲＯポリペプチドコード化核酸分子の核酸配列である。

また、％核酸配列同一性は、配列比較プログラムＮＣＢ１－ＢＬＡＳＴ－２（Ａｌｔｓｃｈｕｌ等，Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．２５：３３８９－３４０２（１９９７））を用いて記載することもできる。ＮＣＢ１－ＢＬＡＳＴ２配列比較プログラムはｈｔｔｐ：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｏｖからダウンロードすることができる。ＮＣＢ１－ＢＬＡＳＴ２は幾つかの検索パラメータを使用し、それら検索パラメータの全ては、例えば、ｕｎｍａｓｋ＝可、鎖＝全て、予測される発生＝１０、最小低複合長＝１５／５、マルチパスｅ－値＝０．０１、マルチパスの定数＝２５、最終ギャップアラインメントのドロップオフ＝２５、及びスコアリングマトリクス＝ＢＬＯＳＵＭ６２を含むデフォルト値に設定される。

配列比較にＮＣＢ１－ＢＬＡＳＴ２が用いれれる状況では、与えられた核酸配列Ｃの、与えられた核酸配列Ｄとの、又はそれに対する％核酸配列同一性（あるいは、与えられた核酸配列Ｄと、又はそれに対して或る程度の％核酸配列同一性を持つ又は含む与えられた核酸配列Ｃと言うこともできる）は次のように計算される：

分率Ｗ／Ｚの１００倍

ここで、Ｗは配列アラインメントプログラムＮＣＢ１－ＢＬＡＳＴ２のＣ及びＤのアラインメントによって同一であると一致したスコアのヌクレオチドの数であり、ＺはＤのヌクレオチドの全数である。核酸配列Ｃの長さが核酸配列Ｄの長さと異なる場合、ＣのＤに対する％核酸配列同一性は、ＤのＣに対する％核酸配列同一性とは異なることは理解されるであろう。

【０１１４】

他の実施態様では、ＰＲＯ変異体ポリヌクレオチドは活性なＰＲＯポリペプチドをコードし、ここに開示された全長ＰＲＯポリペプチドをコード化するヌクレオチド配列に、好ましくは緊縮性ハイブリッド形成及び洗浄条件下でハイブリッド形成可能である各散文詩である。

「陽性（ポジティブ）」という用語は、上記のように実施された配列比較においては、比較された配列において同一ではないが類似の特性を有している残基（例えば、保存的置換の結果として、下記の表１参照）を含む。ここでの目的のために、陽性の％値は、（ａ）天然ＰＲＯポリペプチドから誘導された配列を有する対象とするＰＲＯポリペプチドのアミノ酸配列と、対象とする比較アミノ酸配列（即ち、ＰＲＯポリペプチド配列が比較されるアミノ酸配列）との間の、ＷＵ－ＢＬＡＳＴ－２のＢＬＯＳＵＭ６２マトリクス内でポジティブな値が付けられたアミノ酸残基の数を、（ｂ）対象とするＰＲＯポリペプチドのアミノ酸残基の総数で除した商によって決定される。

特に断らない限り、陽性の％値は、すぐ前の段落に記載したようにして計算される。しかしながら、ＡＬＩＧＮ－２及びＮＣＢＩ－ＢＬＡＳＴ２について上記したように実施され

るアミノ酸配列同一性においては、同一のみならず類似した特性をもつ配列におけるアミノ酸残基を含む。対象とするアミノ酸残基に対して陽性とされたアミノ酸残基は、対象とするアミノ酸残基と同一であるか、又は対象とするアミノ酸残基の好ましい置換（下記の表１に定義）であるものである。
ＡＬＩＧＮ－２又はＮＣＢＩ－ＢＬＡＳＴ２を用いたアミノ酸配列比較では、与えられたアミノ酸配列Ａの、与えられたアミノ酸配列Ｂとの、又はそれに対する陽性の％値（あるいは、与えられたアミノ酸配列Ｂと、又はそれに対して或る程度の％アミノ酸配列同一性を持つ又は含む与えられたアミノ酸配列Ａと言うこともできる）は次のように計算される：
分率Ｘ／Ｙの１００倍
ここで、Ｘは配列アラインメントプログラムＡＬＩＧＮ－２又はＮＣＢＩ－ＢＬＡＳＴ２のＡ及びＢのアラインメントによって陽性であるとされたスコアのアミノ酸残基の数であり、ＹはＢの全アミノ酸残基数である。アミノ酸配列Ａの長さがアミノ酸配列Ｂの長さと等しくない場合、ＡのＢに対する％陽性は、ＢのＡに対する％陽性とは異なることが理解されるであろう。

【０１１５】

「単離された」とは、ここで開示された種々のポリペプチドを記述するために使用するときは、その自然環境の成分から同定され分離され及び／又は回収されたポリペプチドを意味する。その自然環境の汚染成分とは、そのポリペプチドの診断又は治療への使用を典型的には妨害する物質であり、酵素、ホルモン、及び他のタンパク質様又は非タンパク質様溶質が含まれる。好ましい実施態様において、ポリペプチドは、（１）スピニングカップシークエネーターを使用することにより、少なくとも１５残基のＮ末端あるいは内部アミノ酸配列を得るのに充分なほど、あるいは、（２）クーマシーブルーあるいは好ましくは銀染色を用いた非還元あるいは還元条件下でのＳＤＳ－ＰＡＧＥによる均一性まで精製される。単離されたポリペプチドには、ＰＲＯポリペプチドの自然環境の少なくとも１つの成分が存在しないため、組換え細胞内のインサイツのタンパク質が含まれる。しかしながら、通常は、単離されたポリペプチドは少なくとも１つの精製工程により調製される。
「単離された」ＰＲＯポリペプチドコード化核酸は、同定され、ＰＲＯポリペプチドをコードする核酸の天然源に通常付随している少なくとも１つの汚染核酸分子から分離された核酸分子である。単離されたＰＲＯポリペプチドコード化核酸分子は、天然に見出される形態あるいは設定以外のものである。ゆえに、単離されたＰＲＯポリペプチドコード化核酸分子は、天然の細胞中に存在するＰＲＯポリペプチドコード化核酸分子とは区別される。しかし、単離されたＰＲＯポリペプチドコード化核酸分子は、例えば、核酸分子が天然細胞のものとは異なった染色体位置にあるＰＲＯポリペプチドを通常発現する細胞に含まれるＰＲＯポリペプチド核酸分子を含む。
「コントロール配列」という表現は、特定の宿主生物において作用可能に結合したコード配列を発現するために必要なＤＮＡ配列を指す。例えば原核生物に好適なコントロール配列は、プロモーター、場合によってはオペレータ配列、及びリボソーム結合部位を含む。真核生物の細胞は、プロモーター、ポリアデニル化シグナル及びエンハンサーを利用することが知られている。
核酸は、他の核酸配列と機能的な関係にあるときに「作用可能に結合し」ている。例えば、プレ配列あるいは分泌リーダーのＤＮＡは、ポリペプチドの分泌に参画するプレタンパク質として発現されているなら、そのポリペプチドのＤＮＡに作用可能に結合している；プロモーター又はエンハンサーは、配列の転写に影響を及ぼすならば、コード配列に作用可能に結合している；又はリボソーム結合部位は、もしそれが翻訳を容易にするような位置にあるなら、コード配列と作用可能に結合している。一般的に、「作用可能に結合している」とは、結合したＤＮＡ配列が近接しており、分泌リーダーの場合には近接していて読みフェーズにあることを意味する。しかし、エンハンサーは必ずしも近接している必要はない。結合は簡便な制限部位でのライゲーションにより達成される。そのような部位が存在しない場合は、従来の手法に従って、合成オリゴヌクレオチドアダプターあるいはリ

ンカーが使用される。

【０１１６】

「抗体」という用語は最も広い意味において使用され、例えば、単一の抗－ＰＲＯポリペプチドモノクローナル抗体（アゴニスト、アンタゴニスト、及び中和抗体を含む）、多エピトープ特異性を持つ抗－ＰＲＯ抗体組成物、一本鎖抗－ＰＲＯ抗体、及び抗－ＰＲＯ抗体の断片を包含している（下記参照）。ここで使用される「モノクローナル抗体」という用語は、実質的に均一な抗体の集団、すなわち、構成する個々の抗体が、少量存在しうる自然に生じる可能性のある突然変異を除いて同一である集団から得られる抗体を称する。ハイブリッド形成反応の「緊縮性」は、当業者によって容易に決定され、一般的にプローブ長、洗浄温度、及び塩濃度に依存する経験的な計算である。一般に、プローブが長くなると適切なアニーリングのための温度が高くなり、プローブが短くなると温度は低くなる。ハイブリッド形成は、一般的に、相補的鎖がその融点に近いがそれより低い環境に存在する場合における変性ＤＮＡの再アニールする能力に依存する。プローブとハイブリッド形成可能な配列との間の所望の相同性の程度が高くなると、使用できる相対温度が高くなる。その結果、より高い相対温度は、反応条件をより緊縮性にするが、低い温度は緊縮性を低下させる。さらに、緊縮性は塩濃度に逆比例する。ハイブリッド形成反応の緊縮性の更なる詳細及び説明は、Ａｕｓｕｂｅｌ等，Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙ，Ｗｉｌｅｙ Ｉｎｔｅｒｓｃｉｅｎｃｅ Ｐｕｂｌｉｓｈｅｒｓ，（１９９５）を参照のこと。

ここで定義される「緊縮性条件」又は「高度の緊縮性条件」は、（１）洗浄のために低イオン強度及び高温度、例えば、５０℃において０．０１５Ｍの塩化ナトリウム／０．００１５Ｍのクエン酸ナトリウム／０．１％のドデシル硫酸ナトリウムを用いるもの；（２）ハイブリッド形成中にホルムアミド等の変性剤、例えば、４２℃において５０％（ｖ／ｖ）ホルムアミドと０．１％ウシ血清アルブミン／０．１％フィコール／０．１％のポリビニルピロリドン／５０ｍＭのｐＨ６．５のリン酸ナトリウムバッファー、及び７５０ｍＭの塩化ナトリウム、７５ｍＭクエン酸ナトリウムを用いるもの；（３）４２℃における５０％ホルムアミド、５ｘＳＳＣ（０．７５ＭのＮａＣｌ、０．０７５Ｍのクエン酸ナトリウム）、５０ｍＭのリン酸ナトリウム（ｐＨ６．８）、０．１％のピロリン酸ナトリウム、５ｘデンハード液、超音波処理サケ精子ＤＮＡ（５０μｇ／ｍｌ）、０．１％ＳＤＳ、及び１０％のデキストラン硫酸と、４２℃における０．２ｘＳＳＣ（塩化ナトリウム／クエン酸ナトリウム）中の洗浄及び５５℃での５０％ホルムアミド、次いで５５℃におけるＥＤＴＡを含む０．１ｘＳＳＣからなる高緊縮性洗浄を用いるものによって同定される。

「中程度の緊縮性条件」は、Ｓａｍｂｒｏｏｋ等，Ｍｏｌｅｃｕｌａｒ Ｃｌｏｎｉｎｇ：Ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ（Ｎｅｗ Ｙｏｒｋ：Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｌａｂｏｒａｔｏｒｙ Ｐｒｅｓｓ，１９８９に記載されているように同定され、上記の緊縮性より低い洗浄溶液及びハイブリッド形成条件（例えば、温度、イオン強度及び％ＳＤＳ）の使用を含む。中程度の緊縮性条件は、２０％ホルムアミド、５ｘＳＳＣ（１５０ｍＭのＮａＣｌ、１５ｍＭのクエン酸三ナトリウム）、５０ｍＭリン酸ナトリウム（ｐＨ７．６）、５ｘデンハード液、１０％デキストラン硫酸、及び２０ｍｇ／ｍＬの変性剪断サケ精子ＤＮＡを含む溶液中の３７℃での終夜インキュベーション、次いで１ｘＳＳＣ中３７－５０℃でのフィルターの洗浄といった条件である。当業者であれば、プローブ長などの因子に適合させる必要に応じて、どのようにして温度、イオン強度等を調節するかを認識するであろう。

【０１１７】

「エピトープタグ」なる用語は、ここで用いられるときは、「タグポリペプチド」に融合したＰＲＯポリペプチド、又はそれらのドメイン配列を含んでなるキメラポリペプチドを指す。タグポリペプチドは、その抗体が産生され得るエピトープ、又は幾つかの他の試薬によって同定できるエピトープを提供するに十分な数の残基を有しているが、その長さは対象とするＰＲＯポリペプチドの活性を阻害しないよう充分に短い。また、タグポリペプチドは、好ましくは、抗体が他のエピトープと実質的に交差反応をしないようにかなり独

特である。適切なタグポリペプチドは、一般に、少なくとも６のアミノ酸残基、通常は約８～約５０のアミノ酸残基（好ましくは約１０～約２０の残基）を有する。
ここで用いられる「イムノアドヘシン」なる用語は、異種タンパク質（「アドヘシン」）の結合特異性と免疫グロブリン定常ドメインとを結合した抗体様分子を指す。構造的には、イムノアドヘシンは、所望の結合特異性を持ち、抗体の抗原認識及び結合部位以外である（即ち「異種の」）アミノ酸配列と、免疫グロブリン定常ドメイン配列との融合物を含む。イムノアドヘシン分子のアドヘシン部分は、典型的には少なくともレセプター又はリガンドの結合部位を含む隣接アミノ酸配列である。イムノアドヘシンの免疫グロブリン定常ドメイン配列は、ＩｇＧ－１、ＩｇＧ－２、ＩｇＧ－３又はＩｇＧ－４サブタイプ、ＩｇＡ（ＩｇＡ－１及びＩｇＡ－２を含む）、ＩｇＥ、ＩｇＤ又はＩｇＭなどの任意の免疫グロブリンから得ることができる。
ここで意図している「活性な」及び「活性」とは、天然又は天然発生ＰＲＯポリペプチドの生物学的及び／又は免疫学的活性を保持するＰＲＯの形態を意味し、「生物学的」活性とは、天然又は天然発生ＰＲＯによって生ずる（阻害性又は刺激性の）生物学的機能であって、天然又は天然発生ＰＲＯが有する抗原性エピトープに対して抗体を生成する能力を除くものを意味し、「免疫学的」活性とは、天然又は天然発生ＰＲＯが有する抗原性エピトープに対して抗体を生成する能力を意味する。
「アンタゴニスト」なる用語は最も広い意味で用いられ、ここに開示した天然ＰＲＯポリペプチドの生物学的活性を阻止、阻害、又は中和する任意の分子を指す。同様に「アゴニスト」なる用語は最も広い意味で用いられ、ここに開示した天然ＰＲＯポリペプチドの生物学的活性を模倣する任意の分子を指す。好適なアゴニスト又はアンタゴニスト分子は特に、アゴニスト又はアンタゴニスト抗体又は抗体断片、天然ＰＲＯポリペプチドの断片又はアミノ酸配列変異体、ペプチド、有機小分子、などを含む。ＰＲＯポリペプチドのアゴニスト又はアンタゴニストの同定方法は、ＰＲＯポリペプチドを候補アンタゴニスト又はアゴニストと接触させ、ＰＲＯポリペプチドに通常付随する一又は複数の生物学的活性の変化を測定することを含みうる。
【０１１８】
ここで使用される「治療」とは、治癒的処置、予防的療法及び防止的療法の両方を意味し、患者は標的とする病理学的状態又は疾患を防止又は低下（減少）させられる。治療が必要なものとは、既に疾患に罹っているもの、並びに疾患に罹りやすいもの又は疾患が防止されているものを含む。
「慢性」投与とは、急性様式とは異なり連続的な様式での薬剤を投与し、初期の治療効果（活性）を長時間に渡って維持することを意味する。「間欠」投与とは、中断無く連続的になされるのではなく、むしろ本質的に周期的になされる処理である。
治療の対象のための「哺乳動物」は、ヒト、家庭及び農業用動物、動物園、スポーツ、又はペット動物、例えばイヌ、ネコ、ウシ、ウマ、ヒツジ、ブタ、ウサギなどを含む哺乳類に分類される任意の動物を意味する。好ましくは、哺乳動物はヒトである。
一又は複数の治療薬と「組み合わせた」投与とは、同時（同時期）及び任意の順序での連続した投与を含む。
ここで用いられる「担体」は、製薬的に許容されうる担体、賦形剤、又は安定化剤を含み、用いられる用量及び濃度でそれらに暴露される細胞又は哺乳動物に対して非毒性である。生理学的に許容されうる担体は、水性ｐＨ緩衝溶液であることが多い。生理学的に許容されうる担体の例は、リン酸塩、クエン酸塩、及び他の有機酸塩のバッファー；アスコルビン酸を含む酸化防止剤；低分子量（約１０残基未満）ポリペプチド；タンパク質、例えば血清アルブミン、ゼラチン、又は免疫グロブリン；疎水性ポリマー、例えばポリビニルピロリドン；アミノ酸、例えばグリシン、グルタミン、アスパラギン、アルギニン又はリシン；グルコース、マンノース又はデキストランを含む単糖類、二糖類、及び他の炭水化物；ＥＤＴＡ等のキレート剤；マンニトール又は祖ルビトール等の糖アルコール；ナトリウム等の塩形成対イオン；及び／又は非イオン性界面活性剤、例えばＴＷＥＥＮ^{TM}、ポリエチレングリコール（ＰＥＧ）、及びＰＬＵＲＯＮＩＣＳ（商品名）を含む。

【０１１９】
「抗体断片」は、無傷の抗体の一部、好ましくは無傷の抗体の抗原結合又は可変領域を含む。抗体断片の例は、Ｆａｂ、Ｆａｂ’、Ｆ（ａｂ’）$_2$、及びＦｖ断片；ダイアボディ（ｄｉｓｂｏｄｉｅｓ）；直鎖状抗体（Ｚａｐａｔａ等，Ｐｒｏｔｅｉｎ Ｅｎｇ．８（１０）：１０５７－１０６２［１９９５］）；一本鎖抗体分子；及び抗体断片から形成された多重特異性抗体を含む。
抗体のパパイン消化は、「Ｆａｂ」断片と呼ばれる２つの同一の抗体結合断片を生成し、その各々は単一の抗原結合部位を持ち、残りは容易に結晶化する能力を反映して「Ｆｃ」断片と命名される。ペプシン処理はＦ（ａｂ’）$_2$断片を生じ、それは２つの抗原結合部位を持ち、抗原を交差結合することができる。
「Ｆｖ」は、完全な抗原認識及び結合部位を含む最小の抗体断片である。この領域は、密接に非共有結合した１本の重鎖と１本の軽鎖の可変領域の二量体からなる。この配置において各ドメインの３つのＣＤＲが相互作用してＶＨ－ＶＬに量体の表面に抗原結合部位を決定する。集合的には、６つのＣＤＲｓが抗体に対する抗原結合特異性を持つ。しかしながら、単一の可変ドメイン（又は抗原に特異的な３つのＣＤＲのみを含んでなるＦｖの半分）でさえ、結合部位全体よりは低い親和性であるが、抗原を認識し結合する能力を持つ。
またＦａｂ断片は、軽鎖の定常ドメイン及び重鎖の第１の定常ドメイン（ＣＨ１）も含む。Ｆａｂ断片は、抗体ヒンジ領域からの一又は複数のシステインを含む重鎖ＣＨ１ドメインのカルボキシ末端に幾つかの残基が付加されていることによりＦａｂ断片と相違する。ここで、Ｆａｂ’－ＳＨは、定常ドメインのシステイン残基が遊離のチオール基を持つＦａｂ’を表す。Ｆ（ａｂ’）$_2$抗体断片は、最初はＦａｂ’断片の対として生成され、それらの間にヒンジシステインを有する。抗体断片の他の化学的結合も知られている。
任意の脊椎動物種からの抗体（免疫グロブリン）の「軽鎖」は、それらの定常ドメインのアミノ酸配列に基づいて、カッパ及びラムダと呼ばれる二つの明らかに異なる型の一方に分類される。
それらの重鎖の定常ドメインのアミノ酸配列に依存して、免疫グロブリンは異なるクラスに分類できる。免疫グロブリンの五つの主要なクラス：ＩｇＡ、ＩｇＤ、ＩｇＥ、ＩｇＧ及びＩｇＭがあり、それらの幾つかは更にサブクラス（アイソタイプ）、例えばＩｇＧ１、ＩｇＧ２、ＩｇＧ３，ＩｇＧ４、ＩｇＡ及びＩｇＡ２に分類される。
【０１２０】
「一本鎖Ｆｖ」又は「ｓＦｖ」抗体断片は、抗体のＶＨ及びＶＬドメインを含む抗体断片を含み、これらのドメインは単一のポリペプチド鎖に存在する。好ましくは、Ｆｖポリペプチドは ＶＨ及びＶＬドメイン間にポリペプチドリンカーを更に含み、それはｓＦＶが抗原結合に望まれる構造を形成するのを可能にする。ｓＦｖの概説については、Ｔｈｅ Ｐｈａｒｍａｃｏｌｏｇｙ ｏｆ Ｍｏｎｏｃｌｏｎａｌ Ａｎｔｉｂｏｄｉｅｓ，ｖｏｌ．１１３，Ｒｏｓｅｎｂｕｒｇ及びＭｏｏｒｅ編，Ｓｐｒｉｎｇｅｒ－Ｖｅｒｌａｇ，Ｎｅｗ Ｙｏｒｋ，ｐｐ．２６９－３１５（１９９４）のＰｌｕｃｋｔｈｕｎを参照のこと。
用語「ダイアボディ（ｄｉａｂｏｄｉｅｓ）」は、二つの抗原結合部位を持つ小型の抗体断片を指し、その断片は同じポリペプチド鎖（ＶＨ－ＶＬ）内で軽鎖可変ドメイン（ＶＬ）に結合した重鎖可変ドメイン（ＶＨ）を含む。同じ鎖の二つのドメイン間に対形成するには短すぎるリンカーを用いることにより、ドメインは強制的に他の鎖の相補的ドメインと対形成して二つの抗原結合部位を生成する。ダイアボディは、例えば、ＥＰ ４０４，０９７；ＷＯ ９３／１１１６１；及びＨｏｌｌｉｎｇｅｒ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，９０：６４４４－６４４８（１９９３）により十分に記載されている。
「単離された」抗体は、その自然環境の成分から同定され分離及び／又は回収されたものである。その自然環境の汚染成分とは、その抗体の診断又は治療への使用を妨害する物質であり、酵素、ホルモン、及び他のタンパク質様又は非タンパク質様溶質が含まれる。好

ましい実施態様において、抗体は、（１）ローリ法（Ｌｏｗｒｙ ｍｅｔｈｏｄ）で測定した場合９５％を越える抗体、最も好ましくは９９重量％を越えるまで、（２）スピニングカップシークエネーターを使用することにより、少なくとも１５残基のＮ末端あるいは内部アミノ酸配列を得るのに充分なほど、あるいは、（３）クーマシーブルーあるいは好ましくは銀染色を用いた非還元あるいは還元条件下でのＳＤＳ－ＰＡＧＥによる均一性まで精製される。単離された抗体には、抗体の自然環境の少なくとも１つの成分が存在しないため、組換え細胞内のインサイツの抗体が含まれる。しかしながら、通常は、単離された抗体は少なくとも１つの精製工程により調製される。

「標識」なる語は、ここで用いられる場合、抗体に直接又は間接的に抱合して「標識」抗体を生成する検出可能な化合物又は組成物を意味する。標識は、それ自身検出可能でもよく（例えば、放射性標識又は蛍光標識）、又は酵素標識の場合、検出可能な基質化合物又は組成物の化学変換を触媒してもよい。

「固相」とは、本発明の抗体がそれに付着することのできる非水性マトリクスを意味する。ここに意図する固相の例は、部分的又は全体的に、ガラス（例えば、孔制御ガラス）、多糖類（例えばアガロース）、ポリアクリルアミド、ポリスチレン、ポリビニルアルコール及びシリコーンから形成されたものを含む。或る種の実施態様では、内容に応じて、固相はアッセイプレートのウェルを構成することができ；その他では精製カラム（例えばアフィニティクロマトグラフィーカラム）とすることもできる。また、この用語は、米国特許第４，２７５，１４９号に記載されたような、別個の粒子の不連続な固相も包含する。

「リポソーム」は、種々の型の脂質、リン脂質及び／又は界面活性剤からなる小型の小胞であり、哺乳動物への薬物（ＰＲＯポリペプチド又はその抗体など）の輸送に有用である。リポソームの成分は、通常は生体膜の脂質配列に類似する二層形式に配列させる。

「小分子」とは、ここで、約５００ダルトン未満の分子量を持つと定義される。

「ＰＲＯ３１７関連疾患」とはＰＲＯ３１７が過剰又は過小発現している病理状態又は疾患を意味する。そのような障害は、肥厚、子宮内膜炎、子宮内膜症を含む、女性生殖器路又は哺乳動物の子宮内膜の疾患を含み、そこでは患者は子宮内膜因子による不妊、子宮内膜腫、及び子宮体ガン、特に婦人科疾患のような異常出血を含む疾患の危険がある。それらはまた血管新生を含む疾患を含み、そこでは血管新生は充実性腫瘍を含むガンのような病理状態になる（該疾患に対する治療法は低減した血管新生と様々な腫瘍の成長と転位の減少となる）。あるいは、血管新生は例えば虚血、特に冠動脈虚血に対して有益であり得る。よって、これらの障害は、血管形成術及び冠動脈バイパス術に対して最適な候補者ではない冠動脈疾患患者を含む、心臓が機能しているがアテローム硬化型冠動脈疾患のために血液供給が阻害された患者、及び機能しているが灌流が低下した心臓を持つ患者に見いだされるものを含む。障害はまた腎臓に関する又は腎臓組織由来の疾患、例えば多嚢胞性腎疾患及び慢性及び急性腎不全を含む。

【０１２１】

ＩＩ．　本発明の組成物と方法

Ａ．全長ＰＲＯポリペプチド

本発明は、本出願でＰＲＯポリペプチドと呼ばれるポリペプチドをコードする新規に同定され単離された核酸配列を提供する。特に下記の実施例でさらに詳細に説明するように、種々のＰＲＯポリペプチドをコードするｃＤＮＡが同定され単離された。別々の発現ラウンドで生成されたタンパク質には異なるＰＲＯ番号が与えられるが、ＵＮＱ番号は全ての与えられたＤＮＡ及びコード化タンパク質に独特であり、変わることはないことを記しておく。しかしながら、単純化のために、本明細書において、ここに開示した全長天然核酸分子にコードされるタンパク質並びに上記のＰＲＯの定義に含まれるさらなる天然相同体及び変異体は、それらの起源又は調製形式に関わらず、「ＰＲＯ／番号」で呼称する。

下記の実施例に開示するように、種々のｃＤＮＡクローンがＡＴＣＣに寄託されている。これらのクローンの正確なヌクレオチド配列は、この分野で日常的な方法を用いて寄託されたクローンを配列決定することにより容易に決定することができる。予測されるアミノ酸配列は、ヌクレオチド配列から常套的技量を用いて決定できる。ここに記載したＰＲＯ

ポリペプチド及びコード化核酸について、本出願人は、現時点で入手可能な配列情報と最も良く一致するリーディングフレームであると考えられるものを同定した。

【０１２２】

１．全長ＰＲＯ２１１及びＰＲＯ２１７ポリペプチド

本発明は、本出願においてＰＲＯ２１１及びＰＲＯ２１７と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２１１及びＰＲＯ２１７ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ（ＦａｓｔＡフォーマット）配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ２１１及びＰＲＯ２１７をコードしているｃＤＮＡ配列がＥＧＦ様ドメインを有する既知のタンパク質と相同性を有していることを見出した。特に、ｃＤＮＡ配列ＤＮＡ３２２９２－１１３１（Ｆｉｇ１、配列番号：１）はＰＡＣ６＿ＲＡＴとある同一性及び２０９のＢｌａｓｔスコアを、アイソフォームｃ前駆体、フィブリン－１とある同一性と２０６のＢｌａｓｔスコアを有している。ｃＤＮＡ配列ＤＮＡ３３０９４－１１３１（Ｆｉｇ３、配列番号：３）はイースタンｎｅｗｔテネイシンと３６％の同一性及び３３６のＢｌａｓｔスコアを、ヒトテネイシン－Ｘ前駆体と３７％の同一性と３３１のＢｌａｓｔスコアを有している。従って、現在では、本出願で開示されるＰＲＯ２１１及びＰＲＯ２１７ポリペプチドはＥＧＦ様タンパク質ファミリーの新たに同定されたメンバーであり、ＥＧＦ様タンパク質ファミリーに典型的な性質を有すると考えられる。

【０１２３】

２．全長ＰＲＯ２３０ポリペプチド

本発明は、本出願においてＰＲＯ２３０と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２３０ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ２３０をコードしているｃＤＮＡ配列がウサギ尿細管間質性腎炎抗原レセプターと４８％のアミノ酸同一性を有していることを見出した。従って、現在では、本出願で開示されるＰＲＯ２３０ポリペプチドは尿細管間質性腎炎抗原ファミリーの新たに同定されたメンバーであり、尿細管間質性腎炎のある種の型のヒト自己抗体により認識される能力を有すると考えられる。

【０１２４】

３．全長ＰＲＯ２３２ポリペプチド

本発明は、本出願においてＰＲＯ２３２と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２３２ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ２３２の一部分（Ｆｉｇ９と配列番号：１８に示される）がニワトリからの幹細胞表面抗原と３５％の配列同一性を有していることを見出した。従って、現在では、本出願で開示されるＰＲＯ２３２ポリペプチドは新たに同定された幹細胞抗原であると考えられる。

【０１２５】

４．全長ＰＲＯ１８７ポリペプチド

本発明は、本出願においてＰＲＯ１８７と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ１８７ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ１８７（Ｆｉｇ１５に示す）が様々なアンドロゲン誘導成長因子及びＦＧＦ－８と７４％のアミノ酸同一性及び３１０のＢＬＡＳＴスコアを有していることを見出した。従って、現在では、本出願で開示されるＰＲＯ１８７ポリペプチドはＦＧＦ－８タンパク質ファミリーの新たに同定されたメンバーであり、ＦＧＦ－８様タ

ンパク質ファミリーに典型的な活性又は性質を有すると考えられる。

【０１２６】

５．全長ＰＲＯ２６５ポリペプチド

本発明は、本出願においてＰＲＯ２６５と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２６５ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、フィブロモジュリンタンパク質及びフィブロモジュリン前駆体タンパク質と有意な相同性を有していることを見いだした。本出願人はまたＰＲＯ２６５ポリペプチドをコードしているＤＮＡが皮膚と創傷修復に関与するロイシンリッチ反復タンパク質ファミリーのメンバーである血小板糖タンパク質Ｖと有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２６５ポリペプチドはロイシンリッチ反復タンパク質ファミリーの新たに同定されたメンバーであり、このファミリーに典型的なタンパク質間結合能を有し並びに皮膚と創傷修復への関与していると考えられる。

【０１２７】

６．全長ＰＲＯ２１９ポリペプチド

本発明は、本出願においてＰＲＯ２１９と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２１９ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２１９ポリペプチドの様々な部分がマウス及びヒトのマトリリン－２前駆体ポリペプチドと有意な相同性を有していることを見出した。従って、現在では、本出願で開示されるＰＲＯ２１９ポリペプチドはマトリリン－２前駆体ポリペプチドに関連していると考えられる。

【０１２８】

７．全長ＰＲＯ２４６ポリペプチド

本発明は、本出願においてＰＲＯ２４６と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２４６ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ヒト細胞表面タンパク質ＨＣＡＲと有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２４６ポリペプチドは新たに同定された膜結合ウィルスレセプター又は腫瘍細胞特異的抗原であると考えられる。

【０１２９】

８．全長ＰＲＯ２２８ポリペプチド

本発明は、本出願においてＰＲＯ２２８と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２２８ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２２８ポリペプチドの様々な部分がＥＭＲ１タンパク質と有意な相同性を有していることを見いだした。本出願人は、またＰＲＯ２２８ポリペプチドをコードしているＤＮＡがラトロフィリン、マクロファージ制限細胞表面糖タンパク質、Ｂ０４５７．１及び白血球抗原ＣＤ９７前駆体と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２２８ポリペプチドは７つの膜貫通スーパーファミリーの新たに同定されたメンバーであり、このファミリーに典型的な特性及び機能性を有していると考えられる。特に、ＰＲＯ２２８はＣＤ９７及びＥＭＲ１が属するこのファミリー内の亜群の新しいメンバーであると考えられる。

【０１３０】

９．全長ＰＲＯ５３３ポリペプチド

本発明は、本出願においてＰＲＯ５３３と称されるポリペプチドをコードする、新規に同

定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ５３３ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ５３３（Ｆｉｇ２２及び配列番号：５９に示す）が線維芽細胞成長因子（ＦＧＦ）と５０９のＢｌａｓｔスコア及び５３％のアミノ酸配列同一性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ５３３ポリペプチドは線維芽細胞成長因子ファミリーの新たに同定されたメンバーであり、そのようなポリペプチドに典型的な活性を有していると考えられる。

【０１３１】

１０．全長ＰＲＯ２４５ポリペプチド

本発明は、本出願においてＰＲＯ２４５と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２４５ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２４５ポリペプチドのアミノ酸配列の一部がヒトｃ－ｍｙｂタンパク質と６０％のアミノ酸同一性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２４５ポリペプチドは膜貫通プロテインチロシンキナーゼファミリーの新たに同定されたメンバーであると考えられる。

【０１３２】

１１．全長ＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７ポリペプチド

本発明は、本出願においてＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、ＰＲＯ２２０はロイシンリッチタンパク質のアミノ酸配列とアミノ酸同一性を有しており、同一性は８７％である。ＰＲＯ２２０はまたニューロンのロイシンリッチタンパク質のアミノ酸配列とアミノ酸同一性を有しており、同一性は５５％である。ニューロンのロイシンリッチタンパク質は更にＴａｇｕｃｈｉ等，Ｍｏｌ．Ｂｒａｉｎ Ｒｅｓ．，３５：３１－４０（１９９６）に記載されている。

ＰＲＯ２２１はＳＬＩＴタンパク質前駆体とアミノ酸同一性を有しており、これら２種のタンパク質の異なった部分は３９％、３８％、３４％、３１％、及び３０％の各パーセント同一性を有している。

ＰＲＯ２２７は血小板糖タンパク質Ｖ前駆体のアミノ酸配列とアミノ酸同一性を有している。同じ結果がヒト糖タンパク質Ｖに対して得られた。これらの２種のタンパク質の異なった部分は３０％、２８％、２８％、３１％、３５％、３９％及び２７％のパーセント同一性を示す。

従って、現在では、本出願で開示されるＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７ポリペプチドはロイシンリッチ反復タンパク質スーパーファミリーの新たに同定されたメンバーであり、それぞれロイシンリッチ反復タンパク質スーパーファミリーに典型的なタンパク－タンパク結合能を有していると考えられる。また、それらはＳＬＩＴ、ロイシンリッチ反復タンパク質及びヒト糖タンパク質Ｖのものと同様な能力を有していると考えられる。

【０１３３】

１２．全長ＰＲＯ２５８ポリペプチド

本発明は、本出願においてＰＲＯ２５８と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２５８ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２５８ポリペプチドの様々な部分がＣＲＴＡＭ及びポリオウィルスレセ

プターと有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２５８ポリペプチドはＩｇスーパーファミリーの新たに同定されたメンバーであり、このファミリーに典型的なものとしてウィルスレセプター能力を有し又は免疫機能を調節すると考えられる。

【０１３４】

１３．全長ＰＲＯ２６６ポリペプチド

本発明は、本出願においてＰＲＯ２６６と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２６６ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２６６ポリペプチドの様々な部分がショウジョウバエからのＳＬＩＴタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２６６ポリペプチドはロイシンリッチ反復ファミリーの新たに同定されたメンバーであり、このファミリーに典型的なリガンド－リガンド結合活性とニューロン発達を有していると考えられる。ＳＬＩＴはアルツハイマー病、上掲の研究と治療に有用であることが示されており、よってＰＲＯ２６６はこの疾患の研究と治療に関与しうる。

【０１３５】

１４．全長ＰＲＯ２６９ポリペプチド

本発明は、本出願においてＰＲＯ２６９と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２６９ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ２６９（Ｆｉｇ３５と配列番号：９５に示される）のヌクレオチド３１４から１７８３によりコードされるアミノ酸配列が、それが整列化されたヒト尿トロンボモジュリンと様々なトロンボモジュリン類似体に対して有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２６９ポリペプチドはトロンボモジュリンファミリーの新たに同定されたメンバーであると考えられる。

【０１３６】

１５．全長ＰＲＯ２８７ポリペプチド

本発明は、本出願においてＰＲＯ２８７と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２８７ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２８７ポリペプチドの様々な部分が１型プロコラーゲンＣ－プロテイナーゼエンハンサータンパク質前駆体及び１型プロコラーゲンＣ－プロテイナーゼエンハンサータンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２８７ポリペプチドはＣ－プロテイナーゼエンハンサータンパク質ファミリーの新たに同定されたメンバーであると考えられる。

【０１３７】

１６．全長ＰＲＯ２１４ポリペプチド

本発明は、本出願においてＰＲＯ２１４と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２１４ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ２１４ポリペプチド（Ｆｉｇ４０と配列番号：１０９に示される）がＥＧＦ－ファミリーの既知のメンバーであるＨＴタンパク質と４９％のアミノ酸配列同一性を有していることを見いだした。比較は９２０のＢＬＡＳＴスコアで、１５０のヌクレオチドの一致という結果になった。従って、現在では、本出願で開示されるＰＲＯ２１４ポリペプチドはＥＧＦドメインを含んでなるファミリーの新たに同定されたメンバーであり、ＥＧＦ－ドメイン含有ファミリーに典型的な活性又は性質を有していると

考えられる。
【０１３８】
１７．全長ＰＲＯ３１７ポリペプチド
本発明は、本出願においてＰＲＯ３１７と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、以下の実施例で更に詳細に開示するように、ＰＲＯ３１７ポリペプチドをコードするｃＤＮＡが同定され単離された。ＢＬＡＳＴ[TM]及びＦａｓｔＡ[TM]配列アラインメントコンピュータプログラムを用いて、全長天然配列ＰＲＯ３１７（Ｆｉｇ４２と配列番号：１１４に示される）がＥＢＡＦ－１と９２％のアミノ酸配列同一性を有していることが見いだされた。更に、それはＴＧＦ－スーパーファミリーの多くの他のメンバーと密に整列される。
従って、現在では、本出願で開示されるＰＲＯ３１７はＴＧＦ－スーパーファミリーの新たに同定されたメンバーであり、子宮出血等の症状に治療的に有用な性質を有しうると考えられる。よって、ＰＲＯ３１７は、例えば子宮摘出術の必要性を避けたり軽減するために、婦人科疾患に関与する異常出血を診断又は治療するのに有用でありうる。ＰＲＯ３１７はまた一般に血管新生に影響を及ぼす薬剤として有用であり、よってＰＲＯ３１７抗腫瘍効能において、あるいは冠動脈虚血症状を治療するのに有用でありうる。
ライブラリソースは、ＰＲＯ３１７プライマー及びプローブを産生するためのコンセンサスＤＮＡを得るのに使用されるＥＳＴが正常な組織（子宮、前立腺、大腸、及び膵臓）、幾つかの腫瘍（大腸、脳（２回）、膵臓、及びミュラー細胞）及び虚血の心臓に見いだされた。ＰＲＯ３１７はまた幾つかの組織中に見られるが、子宮でより濃度が高いようである。よって、ＰＲＯ３１７はＥＢＡＦ－１よりも体により広い用途を有している。少なくとも幾つかの効能に対して、ＰＲＯ３１７はＥＢＡＦ－１とは反対の効果を有しうると考えられる。
【０１３９】
１８．全長ＰＲＯ３０１ポリペプチド
本発明は、本出願においてＰＲＯ３０１と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ３０１ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ３０１（Ｆｉｇ４４と配列番号：１１９に示される）がヒトＡ３３抗原前駆体と３０％のアミノ酸配列同一性に対応する２４６のＢｌａｓｔスコアを有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ３０１はＡ３３抗原タンパク質ファミリーの新たに同定されたメンバーであり、結腸直腸ガンのようなヒト腫瘍性疾患に発現されうると考えられる。
【０１４０】
１９．全長ＰＲＯ２２４ポリペプチド
本発明は、本出願においてＰＲＯ２２４と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２２４ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ２２４（Ｆｉｇ４６、配列番号：１２７）がホモサピエンスからのアポリポタンパク質Ｅレセプター２９０６とアミノ酸同一性を有していることを見いだした。これら２種のポリペプチドの異なった部分のアラインメントはそれぞれ３７％、３６％、３０％、４４％、４４％及び２８％のアミノ酸同一性を示す。全長天然ＰＲＯ２２４（Ｆｉｇ４６、配列番号：１２７）はまた虫こぶ由来の超低密度リポタンパク質レセプター前駆体とアミノ酸同一性を有している。これら２種のポリペプチドの異なった部分のアラインメントはそれぞれ３８％、３７％、４２％、３３％、及び３７％のアミノ酸同一性を示す。また、全長天然ＰＲＯ２２４（Ｆｉｇ４６、配列番号：１２７）はニワトリ由来のニワトリ卵母細胞レセプターＰ９５とアミノ酸同一性を有している。これら２種のポリペプチドの異なった部分のアラインメントはそれぞれ３８％、３７％、４２％、

３３％、及び３７％のアミノ酸同一性を示す。更に、全長天然ＰＲＯ２２４（Ｆｉｇ４６、配列番号：１２７）はヒト由来の超低密度リポタンパク質レセプター短型前駆体とアミノ酸同一性を有している。これら２種のポリペプチドの異なった部分のアラインメントはそれぞれ３２％、３８％、３４％、４５％、及び３１％のアミノ酸同一性を示す。従って、現在では、本出願で開示されるＰＲＯ２２４は低密度リポタンパク質レセプターファミリーの新たに同定されたメンバーであり、低密度リポタンパク質レセプターファミリーに典型的な低密度リポタンパク質を認識し形質膜陥入する機能性能力を有することを必要とする構造的な特徴を有していると考えられる。（上述のアラインメントに対して次のスコアリングパラメータを使用した：Ｔ＝７，　Ｓ＋６５，　Ｓ２＝３６，　Ｍａｔｒｉｘ：　ＢＬＯＳＵＭ６２．）

【０１４１】

２０．全長ＰＲＯ２２２ポリペプチド

本発明は、本出願においてＰＲＯ２２２と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２２２ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ２２２（Ｆｉｇ４８と配列番号：１３２に示される）をコードしている配列がマウス補体因子ｈ前駆体と２５－２６％のアミノ酸同一性を有し、マウス補体Ｃ３ｂレセプター２型長型前駆体と２５－４７％のアミノ酸同一性を有し、ヒト仮想タンパク質ｋｉａａ０２４７と４０％のアミノ酸同一性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２２２は補体レセプターファミリーの新たに同定されたメンバーであり、補体レセプターファミリーに典型的な活性を有すると考えられる。

【０１４２】

２１．全長ＰＲＯ２３４ポリペプチド

本発明は、本出願においてＰＲＯ２３４と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２３４ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ（ＦａｓｔＡ－フォーマット）配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ２３４をコードしているｃＤＮＡ配列がＥ－セレクチン前駆体と３１％の同一性と１３４のＢｌａｓｔスコアを有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２３４はレクチン／セレクチンファミリーの新たに同定されたメンバーであり、レクチン／セレクチンファミリーに典型的な活性を有すると考えられる。

【０１４３】

２２．全長ＰＲＯ２３１ポリペプチド

本発明は、本出願においてＰＲＯ２３１と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２３１ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ２３１ポリペプチド（Ｆｉｇ５２と配列番号：１４２に示される）がヒト及びラット前立腺酸性ホスファターゼ前駆体タンパク質とそれぞれ３０％及び３１％のアミノ酸同一性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２３１ポリペプチドは酸性フォスファターゼタンパク質ファミリーの新たに同定されたメンバーでありうると考えられる。

【０１４４】

２３．全長ＰＲＯ２２９ポリペプチド

本発明は、本出願においてＰＲＯ２２９と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２２９ポリペプチドをコードするｃＤＮＡを同定し単離し

た。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２２９ポリペプチドの様々な部分が抗原ｗｃ１．１、Ｍ１３０抗原、Ｔ細胞表面糖タンパク質ＣＤ６及びＣＤ６と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２２９ポリペプチドは、免疫機能に関与する多数のタンパク質に見いだされる配列モチーフであるスカベンジャーレセプター相同性を含むファミリーの新たに同定されたメンバーであり、よってこのファミリーに典型的な免疫機能及び／又は分解に抗するセグメントを有すると考えられる。

【０１４５】

２４．全長ＰＲＯ２３８ポリペプチド

本発明は、本出願においてＰＲＯ２３８と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２３８ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２３８ポリペプチドの様々な部分がオキシドレダクターゼ及び脂肪アシル－ＣｏＡレダクターゼを含むレダクターゼと有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２３８ポリペプチドは、レダクターゼファミリーの新たに同定されたメンバーであり、レダクターゼファミリーに典型的な還元活性を有すると考えられる。

【０１４６】

２５．全長ＰＲＯ２３３ポリペプチド

本発明は、本出願においてＰＲＯ２３３と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２３３ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２３３ポリペプチドの様々な部分がレダクターゼタンパク質と有意な相同性を有していることを見いだした。出願人はまたＰＲＯ２３３ポリペプチドをコードしているＤＮＡが線虫由来のタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２３３ポリペプチドは、レダクターゼファミリーの新たに同定されたメンバーであり、レダクターゼファミリーに典型的な細胞の酸化還元状態をなす能力を有すると考えられる。

【０１４７】

２６．全長ＰＲＯ２２３ポリペプチド

本発明は、本出願においてＰＲＯ２２３と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２２３ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２２３ポリペプチドが様々なセリンカルボキシペプチダーゼポリペプチドと有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２２３ポリペプチドは、新たに同定されたセリンカルボキシペプチダーゼであると考えられる。

【０１４８】

２７．全長ＰＲＯ２３５ポリペプチド

本発明は、本出願においてＰＲＯ２３５と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２３５ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２３５ポリペプチドの様々な部分が様々なプレキシンタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２３５ポリペプチドは、プレキシンファミリーの新たに同定されたメンバーであり、プレキシンファミリーに典型的な細胞接着性を有していると考えられる。

【０１４９】
２８．全長ＰＲＯ２３６及びＰＲＯ２６２ポリペプチド
本発明は、本出願においてＰＲＯ２３６及びＰＲＯ２６２と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２３６及びＰＲＯ２６２ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２３６及びＰＲＯ２６２ポリペプチドの様々な部分が様々なβ－ガラクトシダーゼ及びβ－ガラクトシダーゼ前駆体ポリペプチドと有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２３６及びＰＲＯ２６２ポリペプチドは、新たに同定されたβ－ガラクトシダーゼ相同体であると考えられる。
【０１５０】
２９．全長ＰＲＯ２３９ポリペプチド
本発明は、本出願においてＰＲＯ２３９と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２３９ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２３９ポリペプチドの様々な部分がデンシンタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２３９ポリペプチドは、デンシンファミリーの新たに同定されたメンバーであり、デンシンファミリーに典型的なように細胞接着性及びシナプス突起をなす能力を有していると考えられる。
【０１５１】
３０．全長ＰＲＯ２５７ポリペプチド
本発明は、本出願においてＰＲＯ２５７と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２５７ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２５７ポリペプチドの様々な部分がエブネリン前駆体及びエブネリンタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２５７ポリペプチドは、エブネリンタンパク質に関連する新たに同定されたタンパク質メンバーであると考えられる。
【０１５２】
３１．全長ＰＲＯ２６０ポリペプチド
本発明は、本出願においてＰＲＯ２６０と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２６０ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２６０ポリペプチドの様々な部分がα－１－フコシダーゼ前駆体と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２６０ポリペプチドは、フコシダーゼファミリーの新たに同定されたメンバーであり、フコシダーゼファミリーに典型的なフコース残基に関連した酵素活性を有していると考えられる。
【０１５３】
３２．全長ＰＲＯ２６３ポリペプチド
本発明は、本出願においてＰＲＯ２６３と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２６３ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２６３ポリペプチドの様々な部分がＣＤ４４抗原及び関連タンパク質と

有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２６３ポリペプチドは、ＣＤ４４抗原ファミリーの新たに同定されたメンバーであり、これらの抗原に関連する性質の少なくとも一つ、すなわち、ガン及びＨＩＶマーカー、細胞－細胞又は細胞－マトリックス相互作用、調節細胞輸送、リンパ節ホーミング、成長シグナルの伝達、及び移動細胞に対するケモカイン及び成長因子の提示を有していると考えられる。

【０１５４】

３３．全長ＰＲＯ２７０ポリペプチド

本発明は、本出願においてＰＲＯ２７０と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２７０ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２７０ポリペプチドの様々な部分が様々なチオレドキシンタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２７０ポリペプチドは、チオレドキシンファミリーの新たに同定されたメンバーであり、チオレドキシンファミリーに典型的な還元－酸化（レドックス）状態をなす能力を有していると考えられる。

【０１５５】

３４．全長ＰＲＯ２７１ポリペプチド

本発明は、本出願においてＰＲＯ２７１と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２７１ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２７１ポリペプチドが様々な連結タンパク質及びその前駆体と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２７１ポリペプチドは、新たに同定された連結タンパク質相同体であると考えられる。

【０１５６】

３５．全長ＰＲＯ２７２ポリペプチド

本発明は、本出願においてＰＲＯ２７２と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２７２ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２７２ポリペプチドの様々な部分がヒトレチキュロカルビンタンパク質及びその前駆体と有意な相同性を有していることを見いだした。出願人はまたＰＲＯ２７２ポリペプチドをコードしているＤＮＡがマウスレチキュロカルビン前駆体タンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２７２ポリペプチドは、レチキュロカルビンファミリーの新たに同定されたメンバーであり、レチキュロカルビンファミリーに典型的なカルシウムに結合する能力を有していると考えられる。

【０１５７】

３６．全長ＰＲＯ２９４ポリペプチド

本発明は、本出願においてＰＲＯ２９４と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２９４ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２９４ポリペプチドの様々な部分が多くのコラーゲンタンパク質の様々な部分と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２９４ポリペプチドは、コラーゲンファミリーの新たに同定されたメンバーであると考えられる。

【０１５８】

３７．全長ＰＲＯ２９５ポリペプチド
本発明は、本出願においてＰＲＯ２９５と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２９５ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２９５ポリペプチドの様々な部分がインテグリンタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２９５ポリペプチドは、インテグリンファミリーの新たに同定されたメンバーであり、インテグリンファミリーに典型的な細胞接着性を有していると考えられる。

【０１５９】

３８．全長ＰＲＯ２９３ポリペプチド
本発明は、本出願においてＰＲＯ２９３と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２９３ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２９３ポリペプチドの部分がニューロンのロイシンリッチ反復タンパク質１及び２（ＮＬＲＲ－１及びＮＬＲＲ－２）、特にＮＬＲＲ－２と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２９３ポリペプチドは、ニューロンのロイシンリッチ反復タンパク質ファミリーの新たに同定されたメンバーであり、ＮＲＬＬタンパク質ファミリーに典型的なリガンド－リガンド結合活性を有していると考えられる。

【０１６０】

３９．全長ＰＲＯ２４７ポリペプチド
本発明は、本出願においてＰＲＯ２４７と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２４７ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２４７ポリペプチドの様々な部分がデンシンと有意な相同性を有していることを見いだした。出願人はまたＰＲＯ２４７ポリペプチドをコードしているＤＮＡは、ＫＩＡＡ０２３１を含む多数の他のタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２４７ポリペプチドは、ロイシンリッチ反復ファミリーの新たに同定されたメンバーであり、このファミリーに典型的なリガンド結合能を有していると考えられる。

【０１６１】

４０．全長ＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３ポリペプチド
本発明は、本出願においてＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３ポリペプチドの様々な部分が様々なプロテアーゼタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３ポリペプチドは、新たに同定されたプロテアーゼタンパク質であると考えられる。

【０１６２】

４１．全長ＰＲＯ３２８ポリペプチド
本発明は、本出願においてＰＲＯ３２８と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に

詳細に開示するような、ＰＲＯ３２８ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ３２８ポリペプチドの様々な部分がヒトグリア芽細胞腫タンパク質（「ＧＬＩＰ」）と有意な相同性を有していることを見いだした。更に、出願人は、ＰＲＯ３２８ポリペプチドの様々な部分が、ＢＬＡＳＴ相同性［ＥＣＣＲＩＳＰ３＿１，　Ｓ６８６８３，　及びＣＲＳ３＿ＨＵＭＡＮ］により同定されるように、システインリッチ分泌タンパク質（「ＣＲＩＳＰ」）と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ３２８ポリペプチドは、ＧＬＩＰ又はＣＲＩＳＰファミリーの新たに同定されたメンバーであり、ＧＬＩＰ又はＣＲＩＳＰファミリーに典型的な転写制御活性を有していると考えられる。

【０１６３】

４２．全長ＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６ポリペプチド

本発明は、本出願においてＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６ポリペプチドの様々な部分がＬＩＧ－１、ＡＬＳ及びＲＯ３３１の場合には更にデコリンと有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６ポリペプチドは、ロイシンリッチ反復スーパーファミリーの新たに同定されたメンバーで、特にＬＩＧ－１に関連しており、ここで検討し参照したようにこのファミリーの生物活性を有していると考えられる。

【０１６４】

４３．全長ＰＲＯ３３２ポリペプチド

本発明は、本出願においてＰＲＯ３３２と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ３３２ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ３３２（Ｆｉｇ１０８及び配列番号：３１０に示す）が、例えば様々な種のフィブロモジュリン及びフィブロモジュリン（ＦＭＯＤ＿ＢＯＶＩＮ，　ＦＭＯＤ＿ＣＨＩＣＫ，　ＦＭＯＤ＿ＲＡＴ，　ＦＭＯＤ＿ＭＯＵＳＥ，　ＦＭＯＤ＿ＨＵＭＡＮ，　ＰＲ３６７７３）、オステオモジュリン配列（ＡＢ０００１１４＿１，　ＡＢ００７８４８＿１）、デコリン配列（ＣＦＵ８３１４１＿１，　ＯＣＵ０３３９４＿１，　Ｐ＿Ｒ４２２６６，　Ｐ＿Ｒ４２２６７，　Ｐ＿Ｒ４２２６０，　Ｐ＿Ｒ８９４３９）、ケラタン硫酸プロテオグリカン（ＢＴＵ４８３６０１，　ＡＦ０２２８９０＿１）、角膜プロテオグリカン（ＡＦ０２２２５６＿１）、及び骨／軟骨プロテオグリカ及びプロテオグリカン前駆体（ＰＧＳ１＿ＢＯＶＩＮ，　ＰＧＳ２＿，ＭＯＵＳＥ，　ＰＳＧ２＿ＨＵＭＡＮ）を含む一連の既知のプロテオグリカン配列と約３０－４０％のアミノ酸配列同一性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ３３２は新規なプロテオグリカン型の分子であり、細胞外マトリックス、軟骨、及び／又は骨機能を調節する役割を果たしうると考えられる。

【０１６５】

４４．全長ＰＲＯ３３４ポリペプチド

本発明は、本出願においてＰＲＯ３３４と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ３３４ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ３３４ポリペプチドの様々な部分がフィブリン及びフィブリリンと有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ

３３４ポリペプチドは、上皮成長因子ファミリーの新たに同定されたメンバーであり、このファミリーに典型的な性質と活性を有していると考えられる。

【０１６６】

４５．全長ＰＲＯ３４６ポリペプチド

本発明は、本出願においてＰＲＯ３４６と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ３４６ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ３４６（Ｆｉｇ１１２及び配列番号：３２０に示される）がガン胎児性抗原と２８％のアミノ酸配列同一性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ３４６は、ガン胎児性タンパク質ファミリーの新たに同定されたメンバーであり、腫瘍性組織疾患に伴って発現されうると考えられる。

【０１６７】

４６．全長ＰＲＯ２６８ポリペプチド

本発明は、本出願においてＰＲＯ２６８と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２６８ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ２６８ポリペプチドの部分が様々なプロテインジスルフィドイソメラーゼタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２６８ポリペプチドは、プロテインジスルフィドイソメラーゼｐ５タンパク質の相同体であると考えられる。

【０１６８】

４７．全長ＰＲＯ３３０ポリペプチド

本発明は、本出願においてＰＲＯ３３０と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ３３０ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ３３０ポリペプチドの様々な部分がマウスプロリル４－ヒドロキシラーゼα－ＩＩサブユニットタンパク質と有意な相同性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ３３０ポリペプチドは、新規なプロリル４－ヒドロキシラーゼサブユニットポリペプチドであると考えられる。

【０１６９】

４８．全長ＰＲＯ３３９及びＰＲＯ３１０ポリペプチド

本発明は、本出願においてＰＲＯ３３９及びＰＲＯ３１０と称されるポリペプチドをコードする、新規に同定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ３３９ポリペプチドをコードするｃＤＮＡを同定し単離した。本出願人は、また以下の実施例で更に詳細に開示するような、ＰＲＯ３１０ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、ＰＲＯ３３９及びＰＲＯ３１０ポリペプチドの様々な部分が線虫由来の小さな分泌タンパク質と有意な相同性を有し、フリンジと遠い関連を有していることを見いだした。ＰＲＯ３３９はまたコラーゲン様ポリマーと相同性を示す。ここでＤＮＡ４０５３３及びＤＮＡ４２２６７と命名されるＰＲＯ３１０を同定するために使用された配列はまた線虫由来のタンパク質と相同性を示している。従って、現在では、本出願で開示されるＰＲＯ３３９及びＰＲＯ３１０ポリペプチドは、発育に関与するタンパク質ファミリーの新たに同定されたメンバーであり、鋸歯を調節するフリンジの能力に類似した調節能力を有しうると考えられる。

【０１７０】

４９．全長ＰＲＯ２４４ポリペプチド

本発明は、本出願においてＰＲＯ２４４と称されるポリペプチドをコードする、新規に同

定され単離されたヌクレオチド配列を提供する。特に、本出願人は、以下の実施例で更に詳細に開示するような、ＰＲＯ２４４ポリペプチドをコードするｃＤＮＡを同定し単離した。ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメントコンピュータプログラムを用いて、本出願人は、全長天然配列ＰＲＯ２４４（Ｆｉｇ１２２及び配列番号：３７７に示される）が肝臓レクチンニワトリ（ＬＥＣＨ－ＣＨＩＣＫ）と４３％のアミノ酸配列同一性を有し、ＨＩＶｇｐ１２０結合Ｃ型レクチン（Ａ４６２７４）と４２％のアミノ酸配列同一性を有していることを見いだした。従って、現在では、本出願で開示されるＰＲＯ２４４は、Ｃ－レクチンスーパーファミリーの新たに同定されたメンバーであり、免疫機能、アポトーシス、又はアテローム性動脈硬化症の病原においてある役割を果たしうると考えられる。また、ＰＲＯ２４４は腫瘍関連エピトープを同定するのに有用である。

【０１７１】

Ｂ．ＰＲＯポリペプチド変異体

ここに記載した全長天然配列ＰＲＯポリペプチドに加えて、ＰＲＯ変異体も調製できると考えられる。ＰＲＯ変異体は、ＰＲＯポリペプチドＤＮＡに適当なヌクレオチド変化を導入することにより、あるいは所望のＰＲＯポリペプチドを合成することにより調製できる。当業者は、グリコシル化部位の数又は位置の変化あるいは膜固着特性の変化などのアミノ酸変化がＰＲＯポリペプチドの翻訳後プロセスを変えうることを理解するであろう。

天然全長配列ＰＲＯ又はここに記載したＰＲＯポリペプチドの種々のドメインにおける変異は、例えば、米国特許第５，３６４，９３４号に記載されている保存的及び非保存的変異についての任意の技術及び指針を用いてなすことができる。変異は、結果として天然配列ＰＲＯと比較してＰＲＯポリペプチドのアミノ酸配列が変化するＰＲＯポリペプチドをコードする一又は複数のコドンの置換、欠失又は挿入であってよい。場合によっては、変異は少なくとも１つのアミノ酸のＰＲＯポリペプチドの一又は複数のドメインの任意の他のアミノ酸による置換である。いずれのアミノ酸残基が所望の活性に悪影響を与えることなく挿入、置換又は欠失されるかの指針は、ＰＲＯポリペプチドの配列を相同性の知られたタンパク質分子の配列と比較し、相同性の高い領域内でなされるアミノ酸配列変化を最小にすることによって見出される。アミノ酸置換は、一のアミノ酸の類似した構造及び／又は化学特性を持つ他のアミノ酸での置換、例えばロイシンのセリンでの置換、即ち保存的アミノ酸置換の結果とすることができる。挿入及び欠失は、場合によっては１から５のアミノ酸の範囲内とすることができる。許容される変異は、配列においてアミノ酸の挿入、欠失又は置換を系統的に作成し、得られた変異体を下記の実施例に記載するインビトロアッセイの任意のもので活性について試験することにより決定される。

【０１７２】

ＰＲＯポリペプチド断片がここに提供される。このような断片は、例えば、全長天然タンパク質と比較した際に、Ｎ－末端又はＣ－末端で切断されてもよく、又は内部残基を欠いていてもよい。或る種の断片は、ＰＲＯポリペプチドの所望の生物学的活性に必須ではないアミノ酸残基を欠いている。

ＰＲＯ断片は、多くの従来技術の任意のものによって調製してよい。所望のペプチド断片は化学合成してもよい。代替的方法は、酵素的消化、例えば特定のアミノ酸残基によって決定される部位のタンパク質を切断することが知られた酵素でタンパク質を処理することにより、あるいは適当な制限酵素でＤＮＡを消化して所望の断片を単離することによるＰＲＯ断片の生成を含む。さらに他の好適な技術は、ポリメラーゼ連鎖反応（ＰＣＲ）により、所望のポリペプチド断片をコードするＤＮＡ断片を単離し増幅することを含む。ＤＮＡ断片の所望の末端を決定するオリゴヌクレオチドは、ＰＣＲの５’及び３’プライマーで用いられる。好ましくは、ＰＲＯポリペプチド断片は、ここに開示した天然ＰＲＯポリペプチドと少なくとも１つの生物学的及び／又は免疫学的活性を共有する。

特別の実施態様では、対象とする保存的置換を、好ましい置換を先頭にして表１に示す。このような置換が生物学的活性の変化をもたらす場合、表１に例示的置換と名前を付けた又は以下にアミノ酸分類でさらに記載するように、より置換的な変化が導入され生成物がスクリーニングされる。

【０１７３】

表１

元の残基	例示的置換	好ましい置換
Ala(A)	val; Leu; ile	val
Arg(R)	lys; gln; asn	lys
Asn(N)	gln; his; lys; arg	gln
Asp(D)	glu	glu
Cys(C)	ser	ser
Gln(Q)	asn	ans
Gln(E)	asp	asp
Gly(G)	pro; ala	ala
His(H)	asn; gln; lys; arg	arg
Ile(I)	leu; val; met; ala; phe;	
	ﾉﾙﾛｲｼﾝ	leu
Leu(L)	ﾉﾙﾛｲｼﾝ; ile; val;	
	met; ala; phe	ile
Lys(K)	arg; gln; asn	arg
Met(M)	leu; phe; ile	leu
Phe(F)	leu; val; ile; ala; tyr	leu
Pro(P)	ala	ala
Ser(S)	thr	thr
Thr(T)	ser	ser
Trp(W)	tyr; phe	tyr
Tyr(Y)	trp; phe; thr; ser	phe
Val(V)	ile; leu; met; phe;	
	ala; ﾉﾙﾛｲｼﾝ	leu

【０１７４】

ＰＲＯポリペプチドの機能及び免疫学的同一性の置換的修飾は、（ａ）置換領域のポリペプチド骨格の構造、例えばシート又は螺旋配置、（ｂ）標的部位の電荷又は疎水性、又は（ｃ）側鎖の嵩を維持しながら、それらの効果において実質的に異なる置換基を選択することにより達成される。天然発生残基は共通の側鎖特性に基づいてグループに分けることができる：

（１）疎水性：ノルロイシン，ｍｅｔ，ａｌａ，ｖａｌ，ｌｅｕ，ｉｌｅ；

（２）中性の親水性：ｃｙｓ，ｓｅｒ，ｔｈｒ；

（３）酸性：ａｓｐ，ｇｌｕ；

（４）塩基性：ａｓｎ，ｇｌｎ，ｈｉｓ，ｌｙｓ，ａｒｇ；

（５）鎖配向に影響する残基：ｇｌｙ，ｐｒｏ；及び

（６）芳香族：ｔｒｐ，ｔｙｒ，ｐｈｅ。

非保存的置換は、これらの分類の一つのメンバーを他の分類に交換することを必要とするであろう。また、そのように置換された残基は、保存的置換部位、好ましくは残された（非保存）部位に導入されうる。

変異は、オリゴヌクレオチド媒介（部位特異的）突然変異誘発、アラニンスキャンニング、及びＰＣＲ突然変異誘発［Ｃａｒｔｅｒ等，Ｎｕｃｌ．Ａｃｉｄｓ Ｒｅｓ．，１３：４３３１（１９８６）；Ｚｏｌｌｅｒ等，Ｎｕｃｌ．Ａｃｉｄｓ Ｒｅｓ．，１０：６４８７（１９８７）］、カセット突然変異誘発［Ｗｅｌｌｓ等，Ｇｅｎｅ，３４：３１５（１９８５）］、制限的選択突然変異誘発［Ｗｅｌｌｓ等，Ｐｈｉｌｏｓ．Ｔｒａｎｓ．Ｒ．Ｓｏｃ．Ｌｏｎｄｏｎ ＳｅｒＡ，３１７：４１５（１９８６）］等のこの分野で知られた方法を用いてなすことができ、又は他の知られた技術をクローニングしたＤＮＡに実施してＰＲＯ変異体ＤＮＡを作成することもできる。

また、隣接配列に沿って一又は複数のアミノ酸を同定するのにスキャンニングアミノ酸分

析を用いることができる。好ましいスキャンニングアミノ酸は比較的小さく、中性のアミノ酸である。そのようなアミノ酸は、アラニン、グリシン、セリン、及びシステインを含む。アラニンは、ベータ炭素を越える側鎖を排除し変異体の主鎖構造を変化させにくいので、この群の中で典型的に好ましいスキャンニングアミノ酸である［Ｃｕｎｉｎｇｈａｍ及びＷｅｌｌｓ，Ｓｃｉｅｎｃｅ，２４４：１０８１－１０８５（１９８９）］。また、アラニンは最もありふれたアミノ酸であるため典型的には好ましい。さらに、それは埋もれた及び露出した位置の両方に見られることが多い［Ｃｒｅｉｇｈｔｏｎ，Ｔｈｅ Ｐｒｏｔｅｉｎｓ，（Ｗ．Ｈ．Ｆｒｅｅｍａｎ &amp；Ｃｏ．，Ｎ．Ｙ．）；Ｃｈｏｔｈｉａ，Ｊ．Ｍｏｌ．Ｂｉｏｌ．，１５０：１（１９７６）］。アラニン置換が十分な量の変異体を生じない場合は、アイソテリック（ｉｓｏｔｅｒｉｃ）アミノ酸を用いることができる。

【０１７５】

Ｃ．ＰＲＯの修飾

ＰＲＯポリペプチドの共有結合的修飾は本発明の範囲内に含まれる。共有結合的修飾の一型は、ＰＲＯポリペプチドの標的とするアミノ酸残基を、ＰＲＯポリペプチドの選択された側鎖又はＮ又はＣ末端残基と反応できる有機誘導体化試薬と反応させることである。二官能性試薬での誘導体化が、例えばＰＲＯを水不溶性支持体マトリクスあるいは抗－ＰＲＯ抗体の精製方法又はその逆で用いるための表面に架橋させるのに有用である。通常用いられる架橋剤は、例えば、１，１－ビス（ジアゾアセチル）－２－フェニルエタン、グルタルアルデヒド、Ｎ－ヒドロキシスクシンイミドエステル、例えば４－アジドサリチル酸、３，３’－ジチオビス（スクシンイミジルプロピオネート）等のジスクシンイミジルエステルを含むホモ二官能性イミドエステル、ビス－Ｎ－マレイミド－１，８－オクタン等の二官能性マレイミド、及びメチル－３－［（ｐ－アジドフェニル）－ジチオ］プロピオイミダート等の試薬を含む。

他の修飾は、グルタミニル及びアスパラギニル残基の各々対応するグルタミル及びアスパルチルへの脱アミノ化、プロリン及びリシンのヒドロキシル化、セリル又はトレオニル残基のヒドロキシル基のリン酸化、リシン、アルギニン、及びヒスチジン側鎖のα－アミノ基のメチル化［Ｔ．Ｅ．Ｃｒｅｉｇｈｔｏｎ，Ｐｒｏｔｅｉｎｓ：Ｓｔｒｕｃｔｕｒｅ ａｎｄ Ｍｏｌｅｃｕｌａｒ Ｐｒｏｐｅｒｔｉｅｓ，Ｗ．Ｈ．Ｆｒｅｅｍａｎ &amp；amp；Ｃｏ．，Ｓａｎ Ｆｒａｎｃｉｓｃｏ，ｐｐ．７９－８６（１９８３）］、Ｎ末端アミンのアセチル化、及び任意のＣ末端カルボキシル基のアミド化を含む。

【０１７６】

本発明の範囲内に含まれるＰＲＯポリペプチドの共有結合的修飾の他の型は、ポリペプチドの天然グリコシル化パターンの変更を含む。「天然グリコシル化パターンの変更」とは、ここで意図されるのは、天然配列ＰＲＯに見られる１又は複数の炭水化物部分の欠失（存在するグリコシル化部位の除去又は化学的及び／又は酵素的手段によるグリコシル化の削除のいずれかによる）、及び／又は天然配列ＰＲＯに存在しない１又は複数のグリコシル化部位の付加を意味する。さらに、この文節は、存在する種々の炭水化物部分の性質及び特性の変化を含む、天然タンパク質のグリコシル化における定性的変化を含む。

ＰＲＯポリペプチドへのグリコシル化部位の付加はアミノ酸配列の変更を伴ってもよい。この変更は、例えば、１又は複数のセリン又はトレオニン残基の天然配列ＰＲＯ（Ｏ－結合グリコシル化部位）への付加、又は置換によってなされてもよい。ＰＲＯアミノ酸配列は、場合によっては、ＤＮＡレベルでの変化、特に、ＰＲＯポリペプチドをコードするＤＮＡを予め選択された塩基において変異させ、所望のアミノ酸に翻訳されるコドンを生成させることを通して変更されてもよい。

【０１７７】

ＰＲＯポリペプチド上に炭水化物部分の数を増加させる他の手段は、グリコシドのポリペプチドへの化学的又は酵素的結合による。このような方法は、この技術分野において、例えば、１９８７年９月１１日に発行されたＷＯ ８７／０５３３０、及びＡｐｌｉｎ及びＷｒｉｓｔｏｎ，ＣＲＣ Ｃｒｉｔ．Ｒｅｖ．Ｂｉｏｃｈｅｍ．，ｐｐ．２５９－３

０６（１９８１）に記載されている。
ＰＲＯポリペプチド上に存在する炭水化物部分の除去は、化学的又は酵素的に、あるいはグルコシル化の標的として提示されたアミノ酸残基をコードするコドンの変異的置換によってなすことができる。化学的脱グリコシル化技術は、この分野で知られており、例えば、Ｈａｋｉｍｕｄｄｉｎ等，Ａｒｃｈ．Ｂｉｏｃｈｅｍ．Ｂｉｏｐｈｙｓ．，２５９：５２（１９８７）により、及びＥｄｇｅ等，Ａｎａｌ．Ｂｉｏｃｈｅｍ．，１１８：１３１（１９８１）により記載されている。ポリペプチド上の炭水化物部分の酵素的切断は、Ｔｈｏｔａｋｕｒａ等，Ｍｅｔｈ．Ｅｎｚｙｍｏｌ．１３８：３５０（１９８７）に記載されているように、種々のエンド及びエキソグリコシダーゼを用いることにより達成される。
本発明のＰＲＯの共有結合的修飾の他の型は、ＰＲＯポリペプチドの、種々の非タンパク質様ポリマー、例えばポリエチレングリコール、ポリプロピレングリコール、又はポリオキシアルキレンの一つへの、米国特許第４，６４０，８３５号；第４，４９６，６８９号；第４，３０１，１４４号；第４，６７０，４１７号；第４，７９１，１９２号又は第４，１７９，３３７号に記載された方法での結合を含む。
【０１７８】
また、本発明のＰＲＯポリペプチドは、他の異種ポリペプチド又はアミノ酸配列に融合したＰＲＯポリペプチドを含むキメラ分子を形成する方法で修飾してもよい。
一実施態様では、このようなキメラ分子は、抗タグ抗体が選択的に結合できるエピトープを提供するタグポリペプチドとＰＲＯポリペプチドとの融合を含む。エピトープタグは、一般的にはＰＲＯポリペプチドのアミノ又はカルボキシル末端に位置する。このようなＰＲＯポリペプチドのエピトープタグ形態の存在は、タグポリペプチドに対する抗体を用いて検出することができる。また、エピトープタグの提供は、抗タグ抗体又はエピトープタグに結合する他の型の親和性マトリクスを用いたアフィニティ精製によってＰＲＯポリペプチドを容易に精製できるようにする。種々のタグポリペプチド及びそれら各々の抗体はこの分野で良く知られている。例としては、ポリ－ヒスチジン（ｐｏｌｙ－ｈｉｓ）又はポリ－ヒスチジン－グリシン（ｐｏｌｙ－ｈｉｓ－ｇｌｙ）タグ；ｆｌｕ ＨＡタグポリペプチド及びその抗体１２ＣＡ５［Ｆｉｅｌｄ等，Ｍｏｌ．Ｃｅｌｌ．Ｂｉｏｌ．，８：２１５９－２１６５（１９８８）］；ｃ－ｍｙｃタグ及びそれに対する８Ｆ９、３Ｃ７、６Ｅ１０、Ｇ４、Ｂ７及び９Ｅ１０抗体［Ｅｖａｎ等，Ｍｏｌｅｃｕｌａｒ ａｎｄ Ｃｅｌｌｕｌａｒ Ｂｉｏｌｏｇｙ，５：３６１０－３６１６（１９８５）］；及び単純ヘルペスウイルス糖タンパク質Ｄ（ｇＤ）タグ及びその抗体［Ｐａｂｏｒｓｋｙ等，Ｐｒｏｔｅｉｎ Ｅｎｇｉｎｅｅｒｉｎｇ，３（６）：５４７－５５３（１９９０）］を含む。他のタグポリペプチドは、フラッグペプチド［Ｈｏｐｐ等，ＢｉｏＴｅｃｈｎｏｌｏｇｙ，６：１２０４－１２１０（１９８８）］；ＫＴ３エピトープペプチド［Ｍａｒｔｉｎ等，Ｓｃｉｅｎｃｅ，２５５：１９２－１９４（１９９２）］；α－チューブリンエピトープペプチド［Ｓｋｉｎｎｅｒ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２６６：１５１６３－１５１６６（１９９１）］；及びＴ７遺伝子１０タンパク質ペプチドタグ［Ｌｕｔｚ－Ｆｒｅｙｅｒｍｕｔｈ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，８７：６３９３－６３９７（１９９０）］を含む。
それに換わる実施態様では、キメラ分子はＰＲＯの免疫グロブリン又は免疫グロブリンの特定領域との融合体を含んでもよい。キメラ分子の二価形態（「イムノアドヘシン」とも呼ばれる）については、そのような融合体はＩｇＧ分子のＦｃ領域であり得る。Ｉｇ融合体は、好ましくはＩｇ分子内の少なくとも１つの可変領域に換えてＰＲＯポリペプチドの可溶化（膜貫通ドメイン欠失又は不活性化）形態を含む。特に好ましい実施態様では、免疫グロブリン融合体は、ＩｇＧ分子のヒンジ、ＣＨ２及びＣＨ３、又はヒンジ、ＣＨ１、ＣＨ２及びＣＨ３領域を含む。免疫グロブリン融合体の製造については、１９９５年６月２７日発行の米国特許第５，４２８，１３０号を参照のこと。
【０１７９】
Ｄ．ＰＲＯの調製

以下の説明は、主として、ＰＲＯ核酸を含むベクターで形質転換又は形質移入された細胞を培養することによりＰＲＯを生産する方法に関する。もちろん、当該分野においてよく知られている他の方法を用いてＰＲＯを調製することができると考えられる。例えば、ＰＲＯ配列、又はその一部は、固相技術を用いた直接ペプチド合成によって生産してもよい［例えば、Ｓｔｅｗａｒｔ等，Ｓｏｌｉｄ－Ｐｈａｓｅ Ｐｅｐｔｉｄｅ Ｓｙｎｔｈｅｓｉｓ，Ｗ．Ｈ．Ｆｒｅｅｍａｎ Ｃｏ．，Ｓａｎ Ｆｒａｎｃｉｓｃｏ，ＣＡ（１９６９）；Ｍｅｒｒｉｆｉｅｌｄ，Ｊ．Ａｍ．Ｃｈｅｍ．Ｓｏｃ．，８５：２１４９－２１５４（１９６３）参照］。手動技術又は自動によるインビトロタンパク質合成を行ってもよい。自動合成は、例えば、アプライド・バイオシステムズ・ペプチド合成機（Ｆｏｓｔｅｒ Ｃｉｔｙ，ＣＡ）を用いて、製造者の指示により実施してもよい。ＰＲＯの種々の部分は、別々に化学的に合成され、化学的又は酵素的方法を用いて結合させて全長ＰＲＯを生産してもよい。

【０１８０】

１．ＰＲＯをコードするＤＮＡの単離

ＰＲＯをコードするＤＮＡは、ＰＲＯｍＲＮＡを保有していてそれを検出可能なレベルで発現すると考えられる組織から調製されたｃＤＮＡライブラリから得ることができる。従って、ヒトＰＲＯＤＮＡは、実施例に記載されるように、ヒトの組織から調製されたｃＤＮＡライブラリから簡便に得ることができる。またＰＲＯ－コード化遺伝子は、ゲノムライブラリから又は公知の合成方法（例えば、自動化核酸合成）により得ることもできる。ライブラリは、対象となる遺伝子あるいはその遺伝子によりコードされるタンパク質を同定するために設計されたプローブ（ＰＲＯに対する抗体又は少なくとも約２０－８０塩基のオリゴヌクレオチド等）によってスクリーニングできる。選択されたプローブによるｃＤＮＡ又はゲノムライブラリのスクリーニングは、例えばＳａｍｂｒｏｏｋ等，Ｍｏｌｅｃｕｌａｒ Ｃｌｏｎｉｎｇ：Ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ（Ｎｅｗ Ｙｏｒｋ：Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｌａｂｏｒａｔｏｒｙ Ｐｒｅｓｓ，１９８９）に記載されている標準的な手順を使用して実施することができる。所望のＰＲＯポリペプチドをコードする遺伝子を単離する他の方法はＰＣＲ法を使用するものである［Ｓａｍｂｒｏｏｋ等，上掲；Ｄｉｅｆｆｅｎｂａｃｈ等，ＰＣＲ Ｐｒｉｍｅｒ：Ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ（Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｌａｂｏｒａｔｏｒｙ Ｐｒｅｓｓ，１９９５）］。

【０１８１】

下記の実施例には、ｃＤＮＡライブラリのスクリーニング技術を記載している。プローブとして選択されたオリゴヌクレオチド配列は、充分な長さで、疑陽性が最小化されるよう充分に明瞭でなければならない。オリゴヌクレオチドは、スクリーニングされるライブラリ内のＤＮＡとのハイブリッド形成時に検出可能であるように標識されていることが好ましい。標識化の方法は当該分野において良く知られており、^{32}Ｐ標識されたＡＴＰのような放射線標識、ビオチン化あるいは酵素標識の使用が含まれる。中程度の厳密性及び高度の厳密性を含むハイブリッド形成条件は、上掲のＳａｍｂｒｏｏｋ等に与えられている。

このようなライブラリースクリーニング法において同定された配列は、Ｇｅｎｂａｎｋ等の公共データベース又は個人の配列データベースに寄託され公衆に利用可能とされている周知の配列と比較及びアラインメントすることができる。分子の決定された領域内又は全長に渡っての（アミノ酸又は核酸レベルのいずれかでの）配列同一性は、この分野で知られた、そしてここに記載した方法を用いて決定することができる。

タンパク質コード化配列を有する核酸は、初めてここで開示された推定アミノ酸配列を使用し、また必要ならば、ｃＤＮＡに逆転写されなかったｍＲＮＡの生成中間体及び先駆物質を検出する上掲のＳａｍｂｒｏｏｋ等に記述されているような従来のプライマー伸展法を使用し、選択されたｃＤＮＡ又はゲノムライブラリをスクリーニングすることにより得られる。

【０１８２】

２．宿主細胞の選択及び形質転換
宿主細胞を、ここに記載したＰＲＯ生産のための発現又はクローニングベクターで形質移入又は形質転換し、プロモーターを誘導し、形質転換体を選択し、又は所望の配列をコードする遺伝子を増幅するために適当に変性された常套的栄養培地で培養する。培養条件、例えば培地、温度、ｐＨ等々は、過度の実験をすることなく当業者が選ぶことができる。一般に、細胞培養の生産性を最大にするための原理、プロトコール、及び実用技術は、Ｍａｍｍａｌｉａｎ Ｃｅｌｌ Ｂｉｏｔｅｃｈｎｏｌｏｇｙ：ａ Ｐｒａｃｔｉｃａｌ Ａｐｐｒｏａｃｈ，Ｍ．Ｂｕｔｌｅｒ編（ＩＲＬ Ｐｒｅｓｓ，１９９１）及びＳａｍｂｒｏｏｋ等，上掲に見出すことができる。
原核生物細胞形質移入及び真核生物細胞形質移入の方法、例えば、$CaCl_2$、$CaPO_4$、リポソーム媒介及びエレクトロポレーションは当業者に知られている。用いられる宿主細胞に応じて、その細胞に対して適した標準的な方法を用いて形質転換はなされる。前掲のＳａｍｂｒｏｏｋ等に記載された塩化カルシウムを用いるカルシウム処理又はエレクトロポレーションが、一般的に原核生物に対して用いられる。アグロバクテリウム・トゥメファシエンスによる感染が、Ｓｈａｗ等，Ｇｅｎｅ，２３：３１５（１９８３）及び１９８９年６月２９日公開のＷＯ ８９／０５８５９に記載されているように、或る種の植物細胞の形質転換に用いられる。このような細胞壁のない哺乳動物の細胞に対しては、Ｇｒａｈａｍ及びｖａｎ ｄｅｒ Ｅｂ，Ｖｉｒｏｌｏｇｙ，５２：４５６－４５７（１９７８）のリン酸カルシウム沈降法が好ましい。哺乳動物細胞の宿主系形質転換の一般的な態様は米国特許第４，３９９，２１６号に記載されている。酵母菌中への形質転換は、典型的には、Ｖａｎ ｓｏｌｉｎｇｅｎ等，Ｊ．Ｂａｃｔ．，１３０：９４６（１９７７）及びＨｓｉａｏ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，７６：３８２９（１９７９）の方法に従って実施される。しかしながら、ＤＮＡを細胞中に導入する他の方法、例えば、核マイクロインジェクション、エレクトロポレーション、無傷の細胞、又はポリカチオン、例えばポリブレン、ポリオルニチン等を用いる細菌プロトプラスト融合もまた用いることもできる。哺乳動物細胞を形質転換するための種々の技術については、Ｋｅｏｗｎ等，Ｍｅｔｈｏｄｓ ｉｎ Ｅｎｚｙｍｏｌｏｇｙ，１８５：５２７－５３７（１９９０）及びＭａｎｓｏｕｒ等，Ｎａｔｕｒｅ，３３６：３４８－３５２（１９８８）を参照のこと。
【０１８３】
ここに記載のベクターにＤＮＡをクローニングあるいは発現するために適切な宿主細胞は、原核生物、酵母菌、又は高等真核生物細胞である。適切な原核生物は、限定するものではないが、真正細菌、例えばグラム陰性又はグラム陽性生物体、例えば大腸菌のような腸内細菌科を含む。種々の大腸菌株が公衆に利用可能であり、例えば、大腸菌Ｋ１２株ＭＭ２９４（ＡＴＣＣ３１，４４６）；大腸菌Ｘ１７７６（ＡＴＣＣ３１，５３７）；大腸菌株Ｗ３１１０（ＡＴＣＣ２７，３２５）及びＫ５７７２（ＡＴＣＣ５３，６３５）である。他の好ましい原核動物宿主細胞は、大腸菌、例えば、Ｅ．ｃｏｌｉ、エンテロバクター、エルビニア（Ｅｒｗｉｎｉａ）、クレブシエラ（Ｋｌｅｂｓｉｅｌｌａ）、プロテウス（Ｐｒｏｔｅｕｓ）、サルモネラ、例えば、ネズミチフス菌、セラチア、例えば、セラチアマルセサンス（Ｓｅｒｒａｔｉａ ｍａｒｃｅｓｃａｎｓ）、及び赤痢菌、並びに桿菌、例えばバシリスブチリス（Ｂ．ｓｕｂｔｉｌｉｓ）及びバシリリチェニフォルミス（Ｂ．ｌｉｃｈｅｎｉｆｏｒｍｉｓ）（例えば、１９８９年４月１２日発行のＤＤ ２６６，７１０に記載されたバシリリチェニフォルミス４１Ｐ）、シュードモナス、例えば緑膿筋及びストレプトマイセスなどの腸内細菌科を含む。これらの例は限定ではなく例示である。株Ｗ３１１０は、組換えＤＮＡ生産発行のための共通の宿主株であるので一つの特に好ましい宿主又は親宿主である。好ましくは、宿主細胞は最小量のタンパク質分解酵素を分泌する。例えば、株Ｗ３１１０は、細胞に外来のタンパク質をコードする遺伝子における遺伝子変異をするように修飾してもよく、そのような宿主の例としては、完全な遺伝子型ｔｏｎＡを有する大腸菌Ｗ３１１０株１Ａ２；完全な遺伝子型ｔｏｎＡ ｐｔｒ３を有する大腸菌Ｗ３１１０株９Ｅ４；完全な遺伝子型ｔｏｎＡ ｐｒｔ３ ｐｈｏＡ Ｅ１

５（ａｒｇＦ－ｌａｃ）１６９ ｄｅｇＰ ｏｍｐＴｋａｎを有する大腸菌Ｗ３１１０株２７Ｃ７（ＡＴＣＣ ５５，２４４）；完全な遺伝子型ｔｏｎＡ ｐｔｒ３ ｐｈｏＡ Ｅ１５ （ａｌｇＦ－ｌａｃ）１６９ ｄｅｇＰ ｏｍｐＴ ｒｂｓ７ｉｌｖＧｋａｎを有する大腸菌Ｗ３１１０株３７Ｄ６；非カナマイシン耐性ｄｅｇＰ欠失変異を持つ３７Ｄ６株である大腸菌Ｗ３１１０株４０Ｂ４；及び１９９０年８月７日発行の米国特許第４，９４６，７８３号に開示された変異周辺質プロテアーゼを有する大腸菌株を含む。あるいは、クローニングのインビトロ法、例えばＰＣＲ又は他の核酸ポリメラーゼポリメラーゼ反応が好ましい。

【０１８４】

原核生物に加えて、糸状菌又は酵母菌のような真核微生物は、ＰＲＯポリペプチドコード化ベクターのための適切なクローニング又は発現宿主である。サッカロミセス・セレヴィシアは、通常用いられる下等真核生物宿主微生物である。他に、シゾサッカロミセスプロンブ（Ｓｃｈｉｚｏｓａｃｃｈａｒｏｍｙｃｅｓ ｐｒｏｍｂｅ）（Ｂｅａｃｈ及びＮｕｒｓｅ，Ｎａｔｕｒｅ，２９０：１４０［１９８１］；１９８５年5月2日発行のＥＰ １３９，３８３）；クルベロミセスホスツ（Ｋｌｕｖｅｒｏｍｙｃｅｓ ｈｏｓｔｓ）（米国特許第４，９４３，５２９号；Ｆｌｅｅｒ等，Ｂｉｏ／Ｔｅｃｈｎｏｌｏｇｙ，９：９６８－９７５（１９９１））、例えばケーラクチス（Ｋ．ｌａｃｔｉｓ）（ＭＷ９８－８Ｃ，ＣＢＳ６８３，ＣＢＳ４５７４；Ｌｏｕｖｅｎｃｏｕｒｔ等，Ｊ．Ｂａｃｔｅｒｉｏｌ．７３７［１９８３］）、ケーフラギリス（Ｋ．ｆｒａｇｉｌｉｓ）（ＡＴＣＣ １２，４２４）、ケーブルガリクス（Ｋ．ｂｕｌｇａｒｉｃｕｓ）（ＡＴＣＣ １６，０４５）、ケーウィケラミイ（Ｋ．ｗｉｃｋｅｒａｍｉｉ）（ＡＴＣＣ ２４，１７８）、ケーワルチイ（Ｋ．ｗａｌｔｉｉ）（ＡＴＣＣ ５６，５００）、ケードロソフィラルム（Ｋ．ｄｒｏｓｏｐｈｉｌａｒｕｍ）（ＡＴＣＣ ３６，９０６；Ｖａｎ ｄｅｎ Ｂｅｒｇ等，Ｂｉｏ／Ｔｅｃｈｎｏｌｏｇｙ，８：１３５（１９９０））、ケーテモトレランス（Ｋ．ｔｈｅｍｏｔｏｌｅｒａｎｓ）及びケーマルキシアナス（Ｋ．ｍａｒｘｉａｎｕｓ）；ヤロウィア（ｙａｒｒｏｗｉａ）（ＥＰ ４０２，２２６）；ピッチャパストリス（Ｐｉｃｈｉａ ｐａｓｔｏｒｉｓ）（ＥＰ １８３，０７０；Ｓｈｅｅｋｒｉｓｈｎａ等，Ｊ．Ｂａｓｉｃ Ｍｉｃｒｏｂｉｏｌ，２８：２６５－２７８［１９８８］）；カンジダ；トリコデルマレーシア（ｒｅｅｓｉａ）（ＥＰ ２４４，２３４）；アカパンカビ（Ｃａｓｅ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，７６：５２５９－５２６３［１９７９］）；シュワニオマイセス（ｓｃｈｗａｎｎｉｏｍｙｃｅｓ）、例えばシュワニオマイセスオクシデンタリス（ｏｃｃｉｄｅｎｔａｌｉｓ）（１９９０年１０月３１日発行のＥＰ ３９４，５３８）；及び糸状真菌、例えば、ニューロスポラ、ペニシリウム、トリポクラジウム（Ｔｏｌｙｐｏｃｌａｄｉｕｍ）（１９９１年１月１０日発行のＷＯ ９１／００３５７）；及びコウジ菌、例えば偽巣性コウジ菌（Ｂａｌｌａｎｃｅ等，Ｂｉｏｃｈｅｍ．Ｂｉｏｐｈｙｓ．Ｒｅｓ．Ｃｏｍｍｕｎ．，１１２：２８４－２８９［１９８３］；Ｔｉｌｂｕｒｎ等，Ｇｅｎｅ，２６：２０５－２２１［１９８３］；Ｙｅｌｔｏｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，８１：１４７０－１４７４［１９８４］）及びクロカビ（Ｋｅｌｌｙ及びＨｙｎｅｓ，ＥＭＢＯ Ｊ．，４：４７５－４７９［１９８５］）が含まれる。ここで好ましいメチロトロピック（ｍｅｔｈｙｌｏｔｒｏｐｉｃ）酵母は、これらに限られないが、ハンセヌラ（Ｈａｎｓｅｎｕｌａ）、カンジダ、クロエケラ（Ｋｌｏｅｃｋｅｒａ）、ピチア（Ｐｉｃｈｉａ）、サッカロミセス、トルロプシス（Ｔｏｒｕｌｏｐｓｉｓ）、及びロドトルラ（Ｒｈｏｄｏｔｏｒｕｌａ）からなる属から選択されるメタノールで成長可能な酵母を含む。この酵母の分類の例示である特定の種のリストは、Ｃ．Ａｎｔｈｏｎｙ，Ｔｈｅ Ｂｉｏｃｈｅｍｉｓｔｒｙ ｏｆ Ｍｅｔｈｙｌｏｔｒｏｐｈｓ，２６９（１９８２）に記載されている。

【０１８５】

グリコシル化ＰＲＯの発現に適切な宿主細胞は、多細胞生物から誘導される。無脊椎動物細胞の例としては、ショウジョウバエＳ２及びスポドスペラＳｆ９等の昆虫細胞並びに植

物細胞が含まれる。有用な哺乳動物宿主株化細胞の例は、チャイニーズハムスター卵巣（ＣＨＯ）及びＣＯＳ細胞を含む。より詳細な例は、ＳＶ４０によって形質転換されたサル腎臓ＣＶ１株（ＣＯＳ－７，ＡＴＣＣ ＣＲＬ １６５１）；ヒト胚腎臓株（２９３又は懸濁培養での増殖のためにサブクローン化された２９３細胞、Ｇｒａｈａｍ等，Ｊ．Ｇｅｎ Ｖｉｒｏｌ．，３６：５９（１９７７））；チャイニーズハムスター卵巣細胞／－ＤＨＦＲ（ＣＨＯ，Ｕｒｌａｕｂ及びＣｈａｓｉｎ，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，７７：４２１６（１９８０））；マウスのセルトリ細胞（ＴＭ４，Ｍａｔｈｅｒ，Ｂｉｏｌ．Ｒｅｐｒｏｄ．，２３：２４３－２５１（１９８０））ヒト肺細胞（Ｗ１３８，ＡＴＣＣ ＣＣＬ ７５）；ヒト肝細胞（Ｈｅｐ Ｇ２，ＨＢ ８０６５）；及びマウス乳房腫瘍細胞（ＭＭＴ ０６０５６２，ＡＴＴＣ ＣＣＬ ５１）を含む。適切な宿主細胞の選択は、この分野の技術常識内にある。

【０１８６】

３．複製可能なベクターの選択及び使用

ＰＲＯをコードする核酸（例えば、ｃＤＮＡ又はゲノムＤＮＡ）は、クローニング（ＤＮＡの増幅）又は発現のために複製可能なベクター内に挿入される。様々なベクターが公的に入手可能である。ベクターは、例えば、プラスミド、コスミド、ウイルス粒子、又はファージの形態とすることができる。適切な核酸配列が、種々の手法によってベクターに挿入される。一般に、ＤＮＡはこの分野で周知の技術を用いて適当な制限エンドヌクレアーゼ部位に挿入される。ベクター成分としては、一般に、これらに制限されるものではないが、一又は複数のシグナル配列、複製開始点、一又は複数のマーカー遺伝子、エンハンサーエレメント、プロモーター、及び転写終結配列を含む。これらの成分の一又は複数を含む適当なベクターの作成には、当業者に知られた標準的なライゲーション技術を用いる。

ＰＲＯは直接的に組換え手法によって生産されるだけではなく、シグナル配列あるいは成熟タンパク質あるいはポリペプチドのＮ－末端に特異的切断部位を有する他のポリペプチドである異種性ポリペプチドとの融合ペプチドとしても生産される。一般に、シグナル配列はベクターの成分であるか、ベクターに挿入されるＰＲＯ－コード化ＤＮＡの一部である。シグナル配列は、例えばアルカリフォスファターゼ、ペニシリナーゼ、ｌｐｐあるいは熱安定性エンテロトキシンＩＩリーダーの群から選択される原核生物シグナル配列であってよい。酵母の分泌に関しては、シグナル配列は、酵母インベルターゼリーダー、アルファ因子リーダー（酵母菌属（Ｓａｃｃｈａｒｏｍｙｃｅｓ）及びクルイベロマイシス（Ｋｌｕｙｖｅｒｏｍｙｃｅｓ）α因子リーダーを含み、後者は米国特許第５，０１０，１８２号に記載されている）、又は酸ホスフォターゼリーダー、白体（Ｃ．ａｌｂｉｃａｎｓ）グルコアミラーゼリーダー（１９９０年４月４日発行のＥＰ３６２１７９）、又は１９９０年１１月１５日に公開されたＷＯ ９０／１３６４６に記載されているシグナルであり得る。哺乳動物細胞の発現においては、哺乳動物シグナル配列は、同一あるいは関連ある種の分泌ポリペプチド由来のシグナル配列並びにウイルス分泌リーダーのようなタンパク質の直接分泌に使用してもよい。

【０１８７】

発現及びクローニングベクターは共に一又は複数の選択された宿主細胞においてベクターの複製を可能にする核酸配列を含む。そのような配列は多くの細菌、酵母及びウイルスに対してよく知られている。プラスミドｐＢＲ３２２に由来する複製開始点は大部分のグラム陰性細菌に好適であり、２μプラスミド開始点は酵母に適しており、様々なウイルス開始点（ＳＶ４０、ポリオーマ、アデノウイルス、ＶＳＶ又はＢＰＶ）は哺乳動物細胞におけるクローニングベクターに有用である。

発現及びクローニングベクターは、典型的には、選べるマーカーとも称される選択遺伝子を含む。典型的な選択遺伝子は、（ａ）アンピシリン、ネオマイシン、メトトレキセートあるいはテトラサイクリンのような抗生物質あるいは他の毒素に耐性を与え、（ｂ）栄養要求性欠陥を補い、又は（ｃ）例えばバシリに対する遺伝子コードＤ－アラニンラセマーゼのような、複合培地から得られない重要な栄養素を供給するタンパク質をコードする。

【０１８８】

哺乳動物細胞に適切な選べるマーカーの例は、ＤＨＦＲあるいはチミジンキナーゼのように、ＰＲＯ－コード化核酸を取り込むことのできる細胞成分を同定することのできるものである。野生型ＤＨＦＲを用いた場合の好適な宿主細胞は、Ｕｒｌａｕｂ 等により，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，７７：４２１６（１９８０）に記載されているようにして調製され増殖されたＤＨＦＲ活性に欠陥のあるＣＨＯ株化細胞である。酵母菌中での使用に好適な選択遺伝子は酵母プラスミドＹＲｐ７に存在するｔｒｐ１遺伝子である［Ｓｔｉｎｃｈｃｏｍｂ等，Ｎａｔｕｒｅ，２８２：３９（１９７９）；Ｋｉｎｇｍａｎ等，Ｇｅｎｅ，７：１４１（１９７９）；Ｔｓｃｈｅｍｐｅｒ等，Ｇｅｎｅ，１０：１５７（１９８０）］。ｔｒｐ１遺伝子は、例えば、ＡＴＣＣ番号４４０７６あるいはＰＥＰ４－１のようなトリプトファン内で成長する能力を欠く酵母菌の突然変異株に対する選択マーカーを提供する［Ｊｏｎｅｓ，Ｇｅｎｅｔｉｃｓ，８５：１２（１９７７）］。

発現及びクローニングベクターは、通常、ＰＲＯ－コード化核酸配列に作用可能に結合し、ｍＲＮＡ合成を制御するプロモーターを含む。種々の可能な宿主細胞により認識される好適なプロモーターが知られている。原核生物宿主での使用に好適なプロモーターはβ－ラクタマーゼ及びラクトースプロモーター系［Ｃａｈｎｇ等，Ｎａｔｕｒｅ，２７５：６１５（１９７８）；Ｇｏｅｄｄｅｌ等，Ｎａｔｕｒｅ，２８１：５４４（１９７９）］、アルカリフォスファターゼ、トリプトファン（ｔｒｐ）プロモーター系［Ｇｏｅｄｄｅｌ，Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．，８：４０５７（１９８０）；ＥＰ ３６，７７６］、及びハイブリッドプロモーター、例えばｔａｃプロモーター［ｄｅＢｏｅｒ 等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，８０：２１－２５（１９８３）］を含む。細菌系で使用するプロモータもまたＰＲＯをコードするＤＮＡと作用可能に結合したシャイン・ダルガーノ（Ｓ．Ｄ．）配列を有する。

【０１８９】

酵母宿主と共に用いて好適なプロモーター配列の例としては、３－ホスホグリセラートキナーゼ［Ｈｉｔｚｅｍａｎ 等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２５５：２０７３（１９８０）］又は他の糖分解酵素［Ｈｅｓｓ 等，Ｊ．Ａｄｖ．Ｅｎｚｙｍｅ Ｒｅｇ．，７：１４９（１９６８）；Ｈｏｌｌａｎｄ，Ｂｉｏｃｈｅｍｉｓｔｒｙ，１７：４９００（１９８７）］、例えばエノラーゼ、グリセルアルデヒド－３－リン酸デヒドロゲナーゼ、ヘキソキナーゼ、ピルビン酸デカルボキシラーゼ、ホスホフルクトキナーゼ、グルコース－６－リン酸イソメラーゼ、３－ホスホグリセレートムターゼ、ピルビン酸キナーゼ、トリオセリン酸イソメラーゼ、ホスホグルコースイソメラーゼ、及びグルコキナーゼが含まれる。

他の酵母プロモーターとしては、成長条件によって転写が制御される付加的効果を有する誘発的プロモーターであり、アルコールデヒドロゲナーゼ２、イソチトクロムＣ、酸フォスファターゼ、窒素代謝と関連する分解性酵素、メタロチオネイン、グリセルアルデヒド－３－リン酸デヒドロゲナーゼ、及びマルトース及びガラクトースの利用を支配する酵素のプロモーター領域がある。酵母菌での発現に好適に用いられるベクターとプロモータはＥＰ ７３，６５７に更に記載されている。

【０１９０】

哺乳動物の宿主細胞におけるベクターからのＰＲＯ転写は、例えば、ポリオーマウィルス、伝染性上皮腫ウィルス（１９８９年７月５日公開のＵＫ ２，２１１，５０４）、アデノウィルス（例えばアデノウィルス２）、ウシ乳頭腫ウィルス、トリ肉腫ウィルス、サイトメガロウィルス、レトロウィルス、Ｂ型肝炎ウィルス及びサルウィルス４０（ＳＶ４０）のようなウィルスのゲノムから得られるプロモーター、異種性哺乳動物プロモーター、例えばアクチンプロモーター又は免疫グロブリンプロモーター、及び熱衝撃プロモーターから得られるプロモーターによって、このようなプロモーターが宿主細胞系に適合し得る限り制御される。

【０１９１】

より高等の真核生物による所望のＰＲＯをコードするＤＮＡの転写は、ベクター中にエン

ハンサー配列を挿入することによって増強され得る。エンハンサーは、通常は約１０から３００塩基対で、プロモーターに作用してその転写を増強するＤＮＡのシス作動要素である。哺乳動物遺伝子由来の多くのエンハンサー配列が現在知られている（グロビン、エラスターゼ、アルブミン、α－フェトプロテイン及びインスリン）。しかしながら、典型的には、真核細胞ウィルス由来のエンハンサーが用いられるであろう。例としては、複製起点の後期側のＳＶ４０エンハンサー（１００－２７０塩基対）、サイトメガロウィルス初期プロモーターエンハンサー、複製起点の後期側のポリオーマエンハンサー及びアデノウィルスエンハンサーが含まれる。エンハンサーは、ＰＲＯコード化配列の５’又は３’位でベクター中にスプライシングされ得るが、好ましくはプロモーターから５’位に位置している。

また真核生物宿主細胞（酵母、真菌、昆虫、植物、動物、ヒト、又は他の多細胞生物由来の有核細胞）に用いられる発現ベクターは、転写の終結及びｍＲＮＡの安定化に必要な配列も含む。このような配列は、真核生物又はウィルスのＤＮＡ又はｃＤＮＡの通常は５’、時には３’の非翻訳領域から取得できる。これらの領域は、ＰＲＯをコードするｍＲＮＡの非翻訳部分にポリアデニル化断片として転写されるヌクレオチドセグメントを含む。組換え脊椎動物細胞培養でのＰＲＯの合成に適応化するのに適切な他の方法、ベクター及び宿主細胞は、Ｇｅｔｈｉｎｇ等，Ｎａｔｕｒｅ，２９３：６２０－６２５（１９８１）；Ｍａｎｔｅｉ等，Ｎａｔｕｒｅ，２８１：４０－４６（１９７９）；ＥＰ １１７，０６０；及びＥＰ １１７，０５８に記載されている。

【０１９２】

４．遺伝子増幅／発現の検出

遺伝子の増幅及び／又は発現は、ここで提供された配列に基づき、適切に標識されたプローブを用い、例えば、従来よりのサザンブロット法、ｍＲＮＡの転写を定量化するノーザンブロット法［Ｔｈｏｍａｓ，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，７７：５２０１－５２０５（１９８０）］、ドットブロット法（ＤＮＡ分析）、又はインサイツハイブリッド形成法によって、直接的に試料中で測定することができる。あるいは、ＤＮＡ二本鎖、ＲＮＡ二本鎖及びＤＮＡ－ＲＮＡハイブリッド二本鎖又はＤＮＡ－タンパク二本鎖を含む、特異的二本鎖を認識することができる抗体を用いることもできる。次いで、抗体を標識し、アッセイを実施することができ、ここで二本鎖は表面に結合しており、その結果二本鎖の表面での形成の時点でその二本鎖に結合した抗体の存在を検出することができる。

あるいは、遺伝子の発現は、遺伝子産物の発現を直接的に定量する免疫学的な方法、例えば細胞又は組織切片の免疫組織化学的染色及び細胞培養又は体液のアッセイによって、測定することもできる。試料液の免疫組織化学的染色及び／又はアッセイに有用な抗体は、モノクローナルでもポリクローナルでもよく、任意の哺乳動物で調製することができる。簡便には、抗体は、天然配列ＰＲＯポリペプチドに対して、又はここで提供されるＤＮＡ配列をベースとした合成ペプチドに対して、又はＰＲＯＤＮＡに融合し特異的抗体エピトープをコードする外因性配列に対して調製され得る。

【０１９３】

５．ポリペプチドの精製

ＰＲＯの形態は、培地又は宿主細胞の溶菌液から回収することができる。膜結合性であるならば、適切な洗浄液（例えばトリトン－Ｘ１００）又は酵素的切断を用いて膜から引き離すことができる。ＰＲＯの発現に用いられる細胞は、凍結融解サイクル、超音波処理、機械的破壊、又は細胞溶解剤などの種々の化学的又は物理的手段によって破壊することができる。

ＰＲＯを、組換え細胞タンパク又はポリペプチドから精製することが望ましい。適切な精製手順の例である次の手順により精製される：すなわち、イオン交換カラムでの分画；エタノール沈殿；逆相ＨＰＬＣ；シリカ又はカチオン交換樹脂、例えばＤＥＡＥによるクロマトグラフィー；クロマトフォーカシング；ＳＤＳ－ＰＡＧＥ；硫酸アンモニウム沈殿；例えばセファデックスＧ－７５を用いるゲル濾過；ＩｇＧのような汚染物を除くプロテイ

ンＡセファロースカラム；及びＰＲＯポリペプチドのエピトープタグ形態を結合させる金属キレート化カラムである。この分野で知られ、例えば、Ｄｅｕｔｃｈｅｒ，Ｍｅｔｈｏｄｅｓ ｉｎ Ｅｎｚｙｍｏｌｏｇｙ，１８２（１９９０）；Ｓｃｏｐｅｓ，Ｐｒｏｔｅｉｎ Ｐｕｒｉｆｉｃａｔｉｏｎ：Ｐｒｉｎｃｉｐｌｅｓ ａｎｄ Ｐｒａｃｔｉｃｅ，Ｓｐｒｉｎｇｅｒ－Ｖｅｒｌａｇ，Ｎｅｗ Ｙｏｒｋ（１９８２）に記載された多くのタンパク質精製方法を用いることができる。選ばれる精製過程は、例えば、用いられる生産方法及び特に生産される特定のＰＲＯの性質に依存する。

【０１９４】

Ｅ．ＰＲＯの用途

ＰＲＯをコードする核酸配列（又はそれらの補体）は、ハイブリッド形成プローブとしての使用を含む分子生物学の分野において、染色体及び遺伝子マッピングにおいて、及びアンチセンスＲＮＡ及びＤＮＡの生成において種々の用途を有している。また、ＰＲＯ核酸も、ここに記載される組換え技術によるＰＲＯポリペプチドの調製に有用であろう。

全長天然配列ＰＲＯ遺伝子又はその一部は、全長ＰＲＯｃＤＮＡの単離又はここに開示したＰＲＯ配列に対して所望の配列同一性を持つ更に他の遺伝子（例えば、ＰＲＯの天然発生変異体又は他の種からのＰＲＯをコードするもの）の単離のためのｃＤＮＡライブラリ用のハイブリッド形成プローブとして使用できる。場合によっては、プローブの長さは約２０～約５０塩基である。ハイブリッド形成プローブは、少なくとも部分的に全長天然ヌクレオチド配列の新規な領域から誘導してもよく、それらの領域は、過度の実験をすることなく、天然配列ＰＲＯのプロモーター、エンハンサー成分及びイントロンを含むゲノム配列から誘導され得る。例えば、スクリーニング法は、ＰＲＯポリペプチド遺伝子のコード化領域を周知のＤＮＡ配列を用いて単離して約４０塩基の選択されたプローブを合成することを含む。ハイブリッド形成プローブは、^{32}Ｐ又は^{35}Ｓ等の放射性ヌクレオチド、又はアビディン／ビオチン結合系を介してプローブに結合したアルカリフォスファターゼ等の酵素標識を含む種々の標識で標識されうる。本発明のＰＲＯポリペプチド遺伝子に相補的な配列を有する標識されたプローブは、ヒトｃＤＮＡ、ゲノムＤＮＡ又はｍＲＮＡのライブラリーをスクリーニングし、そのライブラリーの何れのメンバーがプローブにハイブッド形成するかを決定するのに使用できる。ハイブリッド形成技術は、以下の実施例において更に詳細に記載する。

【０１９５】

本出願で開示する任意のＥＳＴはプローブと同様に、ここに記載した方法で用いることができる。

ＰＲＯ核酸の他の有用な断片は、標的ＰＲＯｍＲＮＡ（センス）又はＰＲＯＤＮＡ（アンチセンス）配列に結合できる一本鎖核酸配列（ＲＮＡ又はＤＮＡのいずれか）を含むアンチセンス又はセンスオリゴヌクレオチドを含む、アンチセンス又はセンスオリゴヌクレオチドは、本発明によると、ＰＲＯＤＮＡのコード化領域の断片を含む。このような断片は、一般的には少なくとも約１４ヌクレオチド、好ましくは約１４から３０ヌクレオチドを含む。与えられたタンパク質をコードするｃＤＮＡ配列に基づく、アンチセンス又はセンスオリゴヌクレオチドを制御する可能性は、例えば、Ｓｔｅｉｎ及びＣｏｈｅｎ（ＣａｎｃｅｒＲｅｓ．４８：２６５９：１９８８）及びｖａｎ ｄｅｒ Ｋｒｏｌ等（ＢｉｏＴｅｃｈｎｉｑｕｅｓ ６：９５８，１９８８）に記載されている。

アンチセンス又はセンスオリゴヌクレオチドの標的核酸配列への結合は二重鎖の形成をもたらし、それは、二重鎖の分解の促進、転写又は翻訳の期外停止を含む幾つかの方法の一つ、又は他の方法により、標的配列の転写又は翻訳を阻止する。よって、アンチセンスオリゴヌクレオチドは、ＰＲＯタンパク質の発現を阻止するのに用いられる。アンチセンス又はセンスオリゴヌクレオチドは、修飾糖－ホスホジエステル骨格（又は他の糖結合、ＷＯ９１／０６６２９に記載のもの等）を有するオリゴヌクレオチドをさらに含み、そのような糖結合は内因性ヌクレアーゼ耐性である。そのような耐性糖結合を持つオリゴヌクレオチドは、インビボで安定であるが（即ち、酵素分解に耐えうるが）、標的ヌクレオチド配列に結合できる配列特異性は保持している。

【０１９６】
センス又はアンチセンスオリゴヌクレオチドの他の例は、ＷＯ ９０／１００４８に記載されているもののような、有機部分、及びオリゴヌクレオチドの標的核酸配列への親和性を向上させる他の部分、例えばポリ－（Ｌ－リジン）に共有結合したオリゴヌクレオチドを含む。さらにまた、エリプチシン等の挿入剤、アルキル化剤又は金属作体をセンス又はアンチセンスオリゴヌクレオチドに結合させ、アンチセンス又はセンスオリゴヌクレオチドの標的ヌクレオチド配列への結合特異性を改変してもよい。
アンチセンス又はセンスオリゴヌクレオチドは、例えば、$CaPO_4$－媒介ＤＮＡ形質移入、エレクトロポレーションを含む任意の遺伝子転換方法により、又はエプスタイン－バーウイルスなどの遺伝子転換ベクターを用いることにより、標的核酸配列を含む細胞に導入される。好ましい方法では、アンチセンス又はセンスオリゴヌクレオチドは、適切なレトロウイルスベクターに挿入される。標的核酸配列を含む細胞は、インビボ又はエキソビボで組換えレトロウイルスベクターに接触させる。好適なレトロウイルスベクターは、これらに限られないが、マウスレトロウイルスＭ－ＭｕＬＶから誘導されるもの、Ｎ２（Ｍ－ＭｕＬＶから誘導されたレトロウイルス）、又はＤＣＴ５Ａ、ＤＣＴ５Ｂ及びＤＣＴ５Ｃと命名されたダブルコピーベクター（ＷＯ ９０／１３６４１参照）を含む。
【０１９７】
また、センス又はアンチセンスオリゴヌクレオチドは、ＷＯ ９１／０４７５３に記載されているように、リガンド結合分子との複合体の形成により標的配列を含む細胞に導入してもよい。適切なリガンド結合分子は、これらに限られないが、細胞表面レセプター、成長因子、他のサイトカイン、又は細胞表面レセプターに結合する他のリガンドを含む。好ましくは、リガンド結合分子の複合体形成は、リガンド結合分子がその対応する分子又はレセプターに結合する、あるいはセンス又はアンチセンスオリゴヌクレオチド又はその複合体の細胞への侵入を阻止する能力を実質的に阻害しない。
あるいは、センス又はアンチセンスオリゴヌクレオチドは、ＷＯ ９０／１０４４８に記載されたように、オリゴヌクレオチド－脂質複合体の形成により標的核酸配列を含む細胞に導入してもよい。センス又はアンチセンスオリゴヌクレオチド－脂質複合体は、好ましくは内因性リパーゼにより細胞内で分解される。
アンチセンスＲＮＡ又はＤＮＡ分子は一般に少なくとも約５塩基長、約１０塩基長、約１５塩基長、約２０塩基長、約２５塩基長、約３０塩基長、約３５塩基長、約４０塩基長、約４５塩基長、約５０塩基長、約５５塩基長、約６０塩基長、約６５塩基長、約７０塩基長、約７５塩基長、約８０塩基長、約８５塩基長、約９０塩基長、約９５塩基長、約１００塩基長、あるいはそれ以上である。
【０１９８】
また、プローブは、ＰＣＲ技術に用いて、密接に関連したＰＲＯコード化配列の同定のための配列のプールを作成することができる。
また、ＰＲＯをコードするヌクレオチド配列は、そのＰＲＯをコードする遺伝子のマッピングのため、及び遺伝子疾患を持つ個体の遺伝子分析のためのハイブリッド形成プローブの作成にも用いることができる。ここに提供されるヌクレオチド配列は、インサイツハイブリッド形成、既知の染色体マーカーに対する結合分析、及びライブラリーでのハイブリッド形成スクリーニング等の周知の技術を用いて、染色体及び染色体の特定領域にマッピングすることができる。
ＰＲＯのコード化配列が他のタンパク質に結合するタンパク質をコードする場合（例えば、ＰＲＯがレセプターである場合）、ＰＲＯは、そのリガンドを同定するアッセイに用いることができる。このような方法により、レセプター／リガンド結合性相互作用の阻害剤を同定することができる。このような結合性相互作用に含まれるタンパク質も、ペプチド又は小分子阻害剤又は結合性相互作用のアゴニストのスクリーニングに用いることができる。また、レセプターＰＲＯは関連するリガンドの単離にも使用できる。スクリーニングアッセイは、天然ＰＲＯ又はＰＲＯのレセプターの生物学的活性に似たリード化合物の発見のために設計される。このようなスクリーニングアッセイは、化学的ライブラリーの高

スループットスクリーニングにも用いられ、小分子候補薬剤の同定に特に適したものとする。考慮される小分子は、合成有機又は無機化合物を含む。アッセイは、この分野で良く知られ特徴付けられているタンパク質－タンパク質結合アッセイ、生物学的スクリーニングアッセイ、免疫検定及び細胞ベースのアッセイを含む種々の型式で実施される。

【０１９９】

また、ＰＲＯ又はその任意の修飾型をコードする核酸は、トランスジェニック動物又は「ノックアウト」動物を産生するのにも使用でき、これらは治療的に有用な試薬の開発やスクリーニングに有用である。トランスジェニック動物（例えばマウス又はラット）とは、出生前、例えば胚段階で、その動物又はその動物の祖先に導入された導入遺伝子を含む細胞を有する動物である。導入遺伝子とは、トランスジェニック動物が発生する細胞のゲノムに組み込まれたＤＮＡである。一実施形態では、ＰＲＯをコードするｃＤＮＡは、確立された技術によりＰＲＯをコードするゲノムＤＮＡをクローン化するために使用することができ、ゲノム配列を、ＰＲＯをコードするＤＮＡを発現する細胞を有するトランスジェニック動物を産生するために使用することができる。トランスジェニック動物、特にマウス又はラット等の特定の動物を産生する方法は当該分野において常套的になっており、例えば米国特許第４，７３６，８６６号や第４，８７０，００９号に記述されている。典型的には、特定の細胞を組織特異的エンハンサーでのＰＲＯ導入遺伝子の導入の標的にする。胚段階で動物の生殖系列に導入されたＰＲＯコード化導入遺伝子のコピーを含むトランスジェニック動物はＰＲＯをコードするＤＮＡの増大した発現の影響を調べるために使用できる。このような動物は、例えばその過剰発現を伴う病理学的状態に対して保護をもたらすと思われる試薬のテスター動物として使用できる。本発明のこの態様においては、動物を試薬で治療し、導入遺伝子を有する未治療の動物に比べ病理学的状態の発症率が低ければ、病理学的状態に対する治療的処置の可能性が示される。

【０２００】

あるいは、ＰＲＯの非ヒト相同体は、動物の胚性細胞に導入されたＰＲＯをコードする変更ゲノムＤＮＡと、ＰＲＯをコードする内在性遺伝子との間の相同的組換えによって、ＰＲＯをコードする欠陥又は変更遺伝子を有するＰＲＯ「ノックアウト」動物を作成するために使用できる。例えば、ＰＲＯをコードするｃＤＮＡは、確立された技術に従い、ＰＲＯをコードするゲノムＤＮＡのクローニングに使用できる。ＰＲＯをコードするゲノムＤＮＡの一部を欠失したり、組み込みを監視するために使用する選択可能なマーカーをコードする遺伝子等の他の遺伝子で置換することができる。典型的には、ベクターは無変化のフランキングＤＮＡ（５’と３’末端の両方）を数キロベース含む［例えば、相同的組換えベクターについてはＴｈｏｍａｓ及びＣａｐｅｃｃｈｉ，Ｃｅｌｌ，５１：５０３（１９８７）を参照のこと］。ベクターは胚性幹細胞に（例えばエレクトロポレーションによって）導入し、導入されたＤＮＡが内在性ＤＮＡと相同的に組換えられた細胞が選択される［例えば、Ｌｉ等，Ｃｅｌｌ，６９：９１５（１９９２）参照］。選択された細胞は次に動物（例えばマウス又はラット）の胚盤胞内に注入されて集合キメラを形成する［例えば、Ｂｒａｄｌｅｙ，Ｔｅｒａｔｏｃａｒｃｉｎｏｍａｓ ａｎｄ Ｅｍｂｒｙｏｎｉｃ Ｓｔｅｍ Ｃｅｌｌｓ：Ａ Ｐｒａｃｔｉｃａｌ Ａｐｐｒｏａｃｈ，Ｅ．Ｊ．Ｒｏｂｅｒｔｓｏｎ，ｅｄ．（ＩＲＬ，Ｏｘｆｏｒｄ，１９８７），ｐｐ．１１３－１５２参照］。その後、キメラ性胚を適切な偽妊娠の雌性乳母に移植し、期間をおいて「ノックアウト」動物をつくり出す。胚細胞に相同的に組換えられたＤＮＡを有する子孫は標準的な技術により同定され、それらを利用して動物の全細胞が相同的に組換えられたＤＮＡを含む動物を繁殖させることができる。ノックアウト動物は、ＰＲＯポリペプチドが不在であることによるある種の病理的状態及びその病理的状態の進行に対する防御能力によって特徴付けられる。

【０２０１】

また、ＰＲＯポリペプチドをコードする核酸は遺伝子治療にも使用できる。遺伝子治療用途においては、例えば欠陥遺伝子を置換するため、治療的有効量の遺伝子産物のインビボ合成を達成するために遺伝子が導入される。「遺伝子治療」とは、１回の処理により継続

的効果が達成される従来の遺伝子治療と、治療的に有効なＤＮＡ又はｍＲＮＡの１回又は繰り返し投与を含む遺伝子治療薬の投与の両方を含む。アンチセンスＲＮＡ及びＤＮＡは、ある種の遺伝子のインビボ発現を阻止する治療薬として用いることができる。短いアンチセンスオリゴヌクレオチドを、細胞膜による制限された取り込みに起因する低い細胞内濃度にもかかわらず、それが阻害剤として作用する細胞中に移入できることは既に示されている（Ｚａｍｅｃｎｉｋ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ８３：４１４３－４１４６［１９８６］）。オリゴヌクレオチドは、それらの負に荷電したリン酸ジエステル基を非荷電基で置換することによって取り込みを促進するように修飾してもよい。

生存可能な細胞に核酸を導入するための種々の技術が存在する。これらの技術は、核酸が培養細胞にインビトロで、あるいは意図する宿主の細胞においてインビボで移入されるかに応じて変わる。核酸を哺乳動物細胞にインビトロで移入するのに適した方法は、リポソーム、エレクトロポレーション、マイクロインジェクション、細胞融合、ＤＥＡＥ－デキストラン、リン酸カルシウム沈殿法などを含む。現在好ましいインビボ遺伝子移入技術は、ウイルス（典型的にはレトロウイルス）ベクターでの形質移入及びウイルス被覆タンパク質－リポソーム媒介形質移入である（Ｄｚａｕ等，Ｔｒｅｎｄｓ，ｉｎ Ｂｉｏｔｅｃｈｎｏｌｏｇｙ１１，２０５－２１０（１９９３））。幾つかの状況では、核酸供給源を、細胞表面膜タンパク質又は標的細胞に特異的な抗体、標的細胞上のレセプターに対するリガンド等の標的細胞を標的化する薬剤とともに提供するのが望ましい。リポソームを用いる場合、エンドサイトーシスを伴って細胞表面膜タンパク質に結合するタンパク質、例えば、特定の細胞型向性のカプシドタンパク質又はその断片、サイクルにおいて内部移行を受けるタンパク質に対する抗体、細胞内局在化を標的とし細胞内半減期を向上させるタンパク質が、標的化及び／又は取り込みの促進のために用いられる。レセプター媒介エンドサイトーシスは、例えば、Ｗｕ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２６２，４４２９－４４３２（１９８７）；及びＷａｇｎｅｒ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ８７，３４１０－３４１４（１９９０）によって記述されている。遺伝子作成及び遺伝子治療のプロトコールの概説については、Ａｎｄｅｒｓｏｎ等，Ｓｃｉｅｎｃｅ２５６，８０８－８１３（１９９２）を参照のこと。

【０２０２】

ここに記載したＰＲＯポリペプチドは、タンパク質電気泳動目的の分子量マーカーとして用いてもよい。

ここに記載したＰＲＯポリペプチド又はその断片をコードする核酸分子は染色体同定に有用である。この点において、実際の配列に基づく染色体マーキング試薬は殆ど利用可能ではないため、新規な染色体マーカーの同定の必要性が存在している。本発明の各ＰＲＯ核酸分子は染色体マーカーとして使用できる。

また本発明のＰＲＯポリペプチド及び核酸分子は組織タイピングに使用でき、本発明のＰＲＯポリペプチドは一方の組織で他方に比較して異なる発現をする。ＰＲＯ核酸分子は、ＰＣＲ、ノーザン分析、サザン分析及びウェスタン分析のプローブ生成のための用途が見出されるであろう。

ここに記載したＰＲＯポリペプチドは治療薬として用いてもよい。本発明のＰＲＯポリペプチドは、製薬的に有用な組成物を調製するのに知られた方法に従って製剤され、それにより、このＰＲＯ生成物は製薬的に許容される担体媒体と混合される。治療用製剤は、凍結乾燥された製剤又は水性溶液の形態で、任意的な製薬上許容可能なキャリア、賦形剤又は安定剤と、所望の精製度を有する活性成分とを混合することにより（Ｒｅｍｉｎｇｔｏｎ’ｓ Ｐｈａｒｍａｃｅｕｔｉｃａｌ Ｓｃｉｅｎｃｅｓ，１６ｔｈ ｅｄｉｔｉｏｎ，Ａ．Ｏｓｏｌ，Ｅｄ．，［１９８０］）、調製され保管される。許容される担体、賦形剤又は安定剤は、用いる投与量及び濃度ではレシピエントに対して無毒性であり、リン酸、クエン酸及び他の有機酸等の緩衝液；アスコルビン酸を含む抗酸化剤；低分子量（残基数１０個未満）ポリペプチド；血清アルブミン、ゼラチン又は免疫グロブリン等のタンパク質；ポリビニルピロリドン等の親水性重合体；グリシン、グルタミン、アスパラギン

、アルギニン又はリシン等のアミノ酸；グルコース、マンノース又はデキストリン等の単糖類、二糖類又は他の炭水化物、ＥＤＴＡ等のキレート剤、マンニトール又はソルビトール等の糖類、ナトリウム等の塩形成対イオン；及び／又はＴＷＥＥＮ（商品名）、ＰＬＵＲＯＮＩＣＳ（商品名）又はポリエチレングリコール（ＰＥＧ）等の非イオン性界面活性剤を含む。

【０２０３】

インビボ投与に使用される製剤は滅菌されていなくてはならない。これは、凍結乾燥及び再構成の前又は後に、滅菌フィルター膜を通す濾過により容易に達成される。

ここで、本発明の製薬組成物は一般に、無菌のアクセスポートを具備する容器、例えば、皮下注射針で貫通可能なストッパーを持つ静脈内バッグ又はバイアル内に配される。

投与経路は周知の方法、例えば、静脈内、腹膜内、脳内、筋肉内、眼内、動脈内又は病巣内経路での注射又は注入、局所投与、又は徐放系による。

本発明の製薬組成物の用量及び望ましい薬物濃度は、意図する特定の用途に応じて変化する。適切な用量又は投与経路の決定は、通常の内科医の技量の範囲内である。動物実験は、ヒト治療のための有効量の決定についての信頼できるガイダンスを提供する。有効量の種間スケーリングは、Ｔｏｘｉｃｏｋｉｎｅｔｉｃｓ ａｎｄ Ｎｅｗ Ｄｒｕｇ Ｄｅｖｅｌｏｐｍｅｎｔ，Ｙａｃｏｂｉ等，編，Ｐｅｒｇａｍｏｎ Ｐｒｅｓｓ，Ｎｅｗ Ｙｏｒｋ １９８９，ｐｐ．４２－９６のＭｏｒｄｅｎｔｉ，Ｊ．及びＣｈａｐｐｅｌｌ，Ｗ．「Ｔｈｅ ｕｓｅ ｏｆ ｉｎｔｅｒｓｐｅｃｉｅｓ ｓｃａｌｉｎｇ ｉｎ ｔｏｘｉｃｏｋｉｎｅｔｉｃｓ」に記載された原理に従って実施できる。

【０２０４】

ＰＲＯポリペプチド又はそのアゴニスト又はアンタゴニストのインビボ投与が用いられる場合、正常な投与量は、投与経路に応じて、哺乳動物の体重当たり１日に約１０ｎｇ／ｋｇから１００ｍｇ／ｋｇまで、好ましくは約１μｇ／ｋｇ／日から１０ｍｇ／ｋｇ／日である。特定の用量及び輸送方法の指針は文献に与えられている；例えば、米国特許第４，６５７，７６０号、第５，２０６，３４４号、又は第５，２２５，２１２号参照。異なる製剤が異なる治療用化合物及び異なる疾患に有効であること、例えば一つの器官又な組織を標的とする投与には他の器官又は組織とは異なる方式で輸送することは必要であることが予想される。

ＰＲＯポリペプチドの投与を必要とする任意の疾患又は疾病の治療に適した放出特性を持つ製剤でＰＲＯポリペプチドの持続放出が望まれる場合、ＰＲＯポリペプチドのマイクロカプセル化が考えられる。持続放出のための組換えタンパク質のマイクロカプセル化は、ヒト成長ホルモン（ｒｈＧＨ）、インターフェロン－（ｒｈＩＦＮ－）、インターロイキン－２、及びＭＮｒｇｐ１２０で成功裏に実施されている。Ｊｏｈｎｓｏｎ等，Ｎａｔ．Ｍｅｄ．，２：７９５－７９９（１９９６）；Ｙａｓｕｄａ，Ｂｉｏｍｅｄ．Ｔｈｅｒ．，２７：１２２１－１２２３（１９９３）；Ｈｏｒａ等，Ｂｉｏ／Ｔｅｃｈｎｏｌｏｇｙ，８：７５５－７５８（１９９０）；Ｃｌｅｌａｎｄ，「Ｄｅｓｉｇｎ ａｎｄ Ｐｒｏｄｕｃｔｉｏｎ ｏｆ Ｓｉｎｇｌｅ Ｉｍｍｕｎｉｚａｔｉｏｎ Ｖａｃｃｉｎｅｓ Ｕｓｉｎｇ Ｐｏｌｙａｃｔｉｄｅ Ｐｏｌｙｇｌｙｃｏｌｉｄｅ Ｍｉｃｒｏｓｐｈｅｒｅ Ｓｙｓｔｅｍｓ」Ｖａｃｃｉｎｅ Ｄｅｓｉｇｎ：Ｔｈｅ Ｓｕｂｕｎｉｔ ａｎｄ Ａｄｊｕｖａｎｔ Ａｐｐｒｏａｃｈ，Ｐｏｗｅｌｌ 及び Ｎｅｗｍａｎ編，（Ｐｌｅｎｕｍ Ｐｒｅｓｓ：Ｎｅｗ Ｙｏｒｋ，１９９５），ｐ．４３９－４６２；ＷＯ ９７／０３６９２，ＷＯ ９６／４００７２，ＷＯ ９６／０７３９９；及び米国特許第５，５６４，０１０号。

【０２０５】

これらのタンパク質の持続放出製剤は、ポリ－乳酸－コグリコール酸（ＰＬＧＡ）ポリマーを用い、その生体適合性及び広範囲の生分解特性に基づいて開発された。ＰＬＧＡの分解生成物である乳酸及びグリコール酸は、ヒト身体内で即座にクリアされる。さらに、このポリマーの分解性は、分子量及び組成に依存して数ヶ月から数年まで調節できる。Ｌｅｗｉｓ，「Ｃｏｎｔｒｏｌｌｅｄ ｒｅｌｅａｓｅ ｏｆ ｂｉｏａｃｔｉｖｅ ａｇｅｎ

ｔｓ ｆｒｏｍ ｌａｃｔｉｄｅ／ｇｌｙｃｏｌｉｄｅ ｐｏｌｙｍｅｒ」：Ｍ．Ｃｈａｓｉｎ及びＲ．Ｌａｎｇｅｒ（編），Ｂｉｏｄｅｇｒａｄａｂｌｅ Ｐｏｌｙｍｅｒｓ ａｓ Ｄｒｕｇ Ｄｅｌｉｖｅｒｙ Ｓｙｓｔｅｍｓ（Ｍａｒｃｅｌ Ｄｅｋｋｅｒ：Ｎｅｗ Ｙｏｒｋ，１９９０），ｐｐ．１－４１。

本発明は、ＰＲＯポリペプチドに類似する（アゴニスト）又はＰＲＯポリペプチドの効果を阻害する（アンタゴニスト）ものを同定するための化合物のスクリーニング方法も包含する。アンタゴニスト候補薬のスクリーニングアッセイは、ここに同定した遺伝子にコードされるＰＲＯポリペプチドと結合又は複合体形成する化合物、又は他にコード化ポリペプチドの他の細胞性タンパク質との相互作用を阻害する化合物を同定するために設計される。このようなスクリーニングアッセイは、それを特に小分子候補薬の同定に適したものにする、化学的ライブラリの高スループットスクリーニングに適用可能なアッセイを含む。

該アッセイは、タンパク質－タンパク質結合アッセイ、生化学的スクリーニングアッセイ、イムノアッセイ、及び細胞ベースのアッセイで、この分野で知られたものを含む種々の方式で実施される。

アンタゴニストについての全てのアッセイは、それらが候補薬をここで同定された核酸にコードされるＰＲＯポリペプチドと、これら２つの成分が相互作用するのに十分な条件下及び時間で接触させることを必要とすることにおいて共通する。

【０２０６】

結合アッセイにおいて、相互作用は結合であり、形成された複合体は単離されるか、又は反応混合物中で検出される。特別な実施態様では、ここに同定された遺伝子にコードされるＰＲＯポリペプチド又は候補薬が、共有又は非共有結合により固相、例えばミクロタイタープレートに固定化される。非共有結合は、一般的に固体表面をポリペプチドの溶液で被覆し乾燥させることにより達成される。あるいは、固定化されるＰＲＯポリペプチドに特異的な固定化抗体、例えばモノクローナル抗体を、それを固体表面に固着させるために用いることができる。アッセイは、固定化成分、例えば固着成分を含む被覆表面に、検出可能な標識で標識されていてもよい非固定化成分を添加することにより実施される。反応が完了したとき、未反応成分を例えば洗浄により除去し、固体表面に固着した複合体を検出する。最初の非固定化成分が検出可能な標識を有している場合、表面に固定化された標識の検出は複合体形成が起こったことを示す。最初の非固定化成分が標識を持たない場合は、複合体形成は、例えば、固定化された複合体に特異的に結合する標識抗体によって検出できる。

候補化合物が相互作用するがここに同定した遺伝子にコードされる特定のＰＲＯポリペプチに結合しない場合、そのポリペプチドとの相互作用は、タンパク質－タンパク質相互作用を検出するために良く知られた方法によってアッセイすることができる。そのようなアッセイは、架橋、同時免疫沈降、及び勾配又はクロマトグラフィカラムを通す同時精製などの伝統的な手法を含む。さらに、タンパク質－タンパク質相互作用は、Ｆｉｅｌｄｓ及び共同研究者等［Ｆｉｅｌｓ及びＳｏｎｇ，Ｎａｔｕｒｅ（Ｌｏｎｄｏｎ）３４０，２４５－２４６（１９８９）；Ｃｈｉｅｎ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ８８，９５７８－９５８２（１９９１）］に記載された酵母菌ベースの遺伝子系を用いることにより、Ｃｈｅｖｒａｙ及びＮａｔｈａｎｓ［Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ８９，５７８９－５７９３（１９９１）］に開示されている酵母菌ベースの系を用いて監視することができる。酵母菌ＧＡＬ４などの多くの転写活性化剤は、２つの物理的に別個のモジュラードメインからなり、一方はＤＮＡ結合ドメインとして作用し、他方は転写活性化ドメインとして機能する。以前の文献に記載された酵母菌発現系（一般に「２－ハイブリッド系」と呼ばれる）は、この特性の長所を利用して、２つのハイブリッドタンパク質を用い、一方では標的タンパク質がＧＡＬ４のＤＮＡ結合ドメインに融合し、他方では、候補となる活性化タンパク質が活性化ドメインに融合している。ＧＡＬ１－ｌａｃＺリポーター遺伝子のＧＡＬ４活性化プロモーターの制御下での発現は、タンパク質－タンパク質相互作用を介したＧＡＬ４活性の再構成に依存す

る。相互作用するポリペプチドを含むコロニーは、β－ガラクトシダーゼに対する色素生産性物質で検出される。２－ハイブリッド技術を用いた２つの特定なタンパク質間のタンパク質－タンパク質相互作用を同定するための完全なキット（ＭＡＴＣＨＭＡＫＥＲ（商品名））は、Ｃｌｏｎｔｅｃｈから商業的に入手可能である。この系は、特定のタンパク質相互作用に含まれるタンパク質ドメインのマッピング、並びにこの相互作用にとって重要なアミノ酸残基の特定にも拡張することができる。

【０２０７】

ここで同定されたＰＲＯポリペプチドをコードする遺伝子と細胞内又は細胞外成分との相互作用を阻害する化合物は、次のように試験できる：通常は反応混合物は、遺伝子産物と細胞外又は細胞内成分を、それら２つの生成物が相互作用及び結合する条件下及び時間で含むように調製される。候補化合物が結合を阻害する能力を試験するために、反応は試験化合物の不存在及び存在下で実施される。さらに、プラシーボを第３の反応混合物に添加してポジティブ対照を提供してもよい。混合物中に存在する試験化合物と細胞内又は細胞外成分との結合（複合体形成）は上記のように監視される。試験化合物を含有する反応混合物ではなく対照反応における複合体の形成は、試験化合物が試験化合物とその結合パートナーとの相互作用を阻害することを示す。

アンタゴニストを検定するために、ＰＲＯポリペプチドを特定の活性についてスクリーニングする化合物とともに添加してよく、ＰＲＯポリペプチド存在下での対象とする活性を阻害する化合物の能力が化合物がＰＲＯポリペプチドのアンタゴニストであることを示す。あるいは、アンタゴニストは、ＰＲＯポリペプチド及び膜結合ＰＲＯポリペプチドレセプター又は組換えレセプターを持つ潜在的アンタゴニストを、競合的阻害アッセイに適した条件下で結合させることにより検出してもよい。ＰＲＯポリペプチドは、放射活性等で標識でき、レセプターに結合したＰＲＯポリペプチドの数を潜在的アンタゴニストの有効性を決定するのに使用できる。レセプターをコードする遺伝子は、当業者に知られた多くの方法、例えばリガンドパンニング及びＦＡＣＳソートにより同定できる。Ｃｏｌｉｇａｎ等，Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｉｍｍｕｎ．，１（２）：Ｃｈａｐｔｅｒ ５（１９９１）。好ましくは、発現クローニングが用いられ、ポリアデニル化ＲＮＡがＰＲＯポリペプチドに反応性の細胞から調製され、このＲＮＡから生成されたｃＤＮＡライブラリがプールに分配され、ＣＯＳ細胞又はＰＲＯポリペプチド反応性でない他の細胞の形質移入に使用される。スライドガラスで成長させた形質移入細胞を標識したＰＲＯポリペプチドに暴露する。ＰＲＯポリペプチドは、ヨウ素化又は部位特異的タンパク質キナーゼの認識部位の包含を含む種々の手段で標識できる。固定及びインキュベーションの後、スライドにオートラジオグラフィ分析を施す。ポジティブプールを同定し、相互作用サブプール化及び再スクリーニング工程を用いてサブプールを調製して再形質移入し、結果的に推定レセプターをコードする単一のクローンを生成する。

【０２０８】

レセプター同定の代替的方法として、標識下ＰＲＯポリペプチドをレセプター分子を発現する細胞膜又は抽出調製物に光親和性結合させることができる。架橋材料はＰＡＧＥに溶解させ、Ｘ線フィルムに暴露する。レセプターを含む標識複合体を励起し、ペプチド断片に分離し、タンパク質マイクロ配列決定を施すことができる。マイクロ配列決定から得たアミノ酸配列は、推定レセプターをコードする遺伝子を同定するｃＤＮＡライブラリをスクリーニングする分解性オリゴヌクレオチドプローブの組の設計に用いられる。

アンタゴニストの他の検定では、レセプターを発現する哺乳動物細胞又は膜調製物を、候補化合物の存在下で標識ＰＲＯポリペプチドとともにインキュベートする。次いで、この相互作用を促進又は阻止する化合物の能力を測定する。

潜在的なアンタゴニストのより特別な例は、免疫グロブリンとＰＲＯポリペプチドとの融合体に結合するオリゴヌクレオチド、特に、限られないが、ポリペプチド－及びモノクローナル抗体及び抗体断片、一本鎖抗体、抗－イディオタイプ抗体、及びこれらの抗体又は断片のキメラ又はヒト化形態、並びにヒト抗体及び抗体断片を含む抗体を含んでいる。あるいは、潜在的アンタゴニストは、密接に関連したタンパク質、例えば、レセプターを認

識するが効果を与えず、従ってＰＲＯポリペプチドの作用を競合的に阻害するＰＲＯポリペプチドの変異形態であってもよい。
他の潜在的なＰＲＯポリペプチドアンタゴニストは、アンチセンス技術を用いて調製されたアンチセンスＲＮＡ又はＤＮＡ作成物であり、例えば、アンチセンスＲＮＡ又はＤＮＡは、標的ｍＲＮＡにハイブリッド形成してタンパク質翻訳を妨害することによりｍＲＮＡの翻訳を直接阻止するように作用する。アンチセンス技術は、トリプルヘリックス形成又はアンチセンスＤＮＡ又はＲＮＡを通して遺伝子発現を制御するのに使用でき、それらの方法はともに、ポリペプチドヌクレオチドのＤＮＡ又はＲＮＡへの結合に基づく。例えば、ここでの成熟ＰＲＯポリペプチドをコードするポリペプチドヌクレオチド配列の５’コード化部分は、約１０から４０塩基対長のアンチセンスＲＮＡオリゴヌクレオチドの設計に使用される。ＤＮＡオリゴヌクレオチドは、転写に含まれる遺伝子の領域に相補的であるように設計され（トリプルヘリックス－Ｌｅｅ等，Ｎｕｃｌ，Ａｃｉｄ Ｒｅｓ．，６：３０７３（１９７９）；Ｃｏｏｎｅｙ等，Ｓｃｉｅｎｃｅ，２４１：４５６（１９８８）；Ｄｅｒｖａｎ等，Ｓｃｉｅｎｃｅ，２５１：１３６０（１９９１）参照）、それによりＰＲＯポリペプチドの転写及び生成を防止する。アンチセンスＲＮＡオリゴヌクレオチドはインビボでｍＲＮＡにハイブリッド形成してｍＲＮＡ分子のＰＲＯポリペプチドへの翻訳を阻止する（アンチセンス－Ｏｋａｎｏ，Ｎｅｕｒｏｃｈｅｍ．，５６：５６０（１９９１）；Ｏｌｉｇｏｄｅｏｘｙｎｕｃｌｅｏｔｉｄｅｓ ａｓ Ａｎｔｉｓｅｎｓｅ Ｉｎｈｉｂｉｔｏｒｓ ｏｆ Ｇｅｎｅ Ｅｘｐｒｅｓｓｉｏｎ（ＳＲＳ Ｐｒｅｓｓ：Ｂｏｃａ Ｒａｔｏｎ，ＦＬ，１９８８））。上記のオリゴヌクレオチドは、細胞に輸送され、アンチセンスＲＮＡ又はＤＮＡをインビボで発現させて、ＰＲＯポリペプチドの生産を阻害することもできる。アンチセンスＤＮＡが用いられる場合、翻訳開始部位、例えば標的遺伝子ヌクレオチド配列の－１０から＋１０位置の間から誘導されるオリゴデオキシリボヌクレオチドが好ましい。
【０２０９】
潜在的アンタゴニストは、ＰＲＯポリペプチドの活性部位、レセプター結合部位、又は成長因子又は他の関連結合部位に結合し、それによりＰＲＯポリペプチドの正常な生物学的活性を阻止する小分子を含む。小分子の例は、これらに限られないが、小型ペプチド又はペプチド様分子、好ましくは可溶性ペプチド、及び合成非ペプチド有機又は無機化合物を含む。
リボザイムは、ＲＮＡの特異的切断を触媒できる酵素的ＲＮＡ分子である。リボザイムは、相補的標的ＲＮＡへの配列特異的ハイブリッド形成、次いでヌクレオチド鎖切断的切断により作用する。潜在的ＲＮＡ標的内の特異的リボザイム切断部位は、既知の技術で同定できる。更なる詳細は、例えば、Ｒｏｓｓｉ，Ｃｕｒｒｅｎｔ Ｂｉｏｌｏｇｙ ４：４６９－４７１（１９９４）及びＰＣＴ公報、番号ＷＯ ９７／３３５５１（１９９７年９月１８日公開）を参照。
転写阻害に用いられるトリプルヘリックス形成における核酸分子は一本鎖でデオキシヌクレオチドからなる。これらのオリゴヌクレオチドの基本組成は、フーグスチン塩基対則を介するトリプルヘリックス形成を促進するように設計され、それは一般に二重鎖の一方の鎖上のプリン又はピリミジンのサイズ変更可能な伸展を必要とする。さらなる詳細は、例えば、ＰＣＴ公報、番号ＷＯ ９７／３３５５１，上掲を参照。
これらの小分子は、上記で検討したスクリーニングアッセイの一又は複数の任意のものにより及び／又は当業者に良く知られた他の任意のスクリーニング技術により同定できる。
【０２１０】
ＰＲＯ２１１及びＰＲＯ２１７ポリペプチドに関して、治療適応症は胃腸粘膜の保存と維持及び急性及び慢性病変（例えば、全腸炎、ゾリンジャー・エリソン症候群、胃腸潰瘍形成及び先天性微絨毛萎縮症）、異常なケラチノサイト分化に関連した皮膚疾患（例えば、乾癬、上皮ガン、例えば肺扁平上皮ガン、外陰部と神経膠腫の類表皮ガン）に関連した障害を含む。
ＰＲＯ２３２ポリペプチドとそれをコードしている核酸は細胞表面幹細胞抗原及びそのコ

ード核酸と配列相同性を有しているので、ＰＲＯ２３２ヌクレオチド配列に基づくプローブを、他の新規な幹細胞表面抗原タンパク質を同定するために使用することができる。ＰＲＯ２３２ポリペプチドの可溶型はインビトロとインビボの双方において膜結合ＰＲＯ２３２活性のアンタゴニストとして使用することができる。ＰＲＯ２３２ポリペプチドは天然ＰＲＯ２３２ポリペプチドのアゴニスト又はアンタゴニストを同定するように設計されたスクリーニングアッセイ法において使用することができ、そのようなアッセイは任意の一般的な細胞型又は生化学結合アッセイの形態を取ることができる。更に、ＰＲＯ２３２ポリペプチドは該ポリペプチドが特異的に発現されている組織に対する分子マーカーとなりうる。

ここに開示されたＰＲＯ１８７ポリペプチドに関し、ＦＧＦ－８は、肢形成中に現れるパターン形成を含む、細胞性分化及び胚形成に関与していた。従って、本発明のＦＧＦ－８及びＰＲＯ１８７分子は細胞成長と発達に強力な効果を有しているように思われる。従って、細胞の成長と分化に関する疾患はＦＧＦ－８に類似の機能性に基づく治療法の好適な標的である。例えば、パーキンソン病、アルツハイマー病、ＡＬＳ、神経障害を含む神経細胞の成長又は生存に関与する疾患である。また、制御されない細胞増殖に関連する疾患、例えばガンもまた予想される治療標的である。

ここに開示したＰＲＯ２６５ポリペプチドに関しては、ＰＲＯ２６５の他の使用方法はＲｕｏｓｌａｈｔｉ等の米国特許５，６５４，２７０に記載されている。特に、ＰＲＯ２６５はそこに開示されたフィブロモジュリンと比較して使用して、皮膚の瘢痕化を低減するその効果及びどこに位置しているかと何に結合し何に結合しないかを含むそこに記載されたフィブロモジュリンの他の性質を比較する。

血液凝固カスケードにおいて調節の役割を担う本発明のＰＲＯ２１９ポリペプチドはインビボでの治療目的並びにインビトロでの目的に使用されてもよい。当業者であればそのような目的に如何にしてＰＲＯ２１９ポリペプチドを使用するかは分かるであろう。

一又は複数のウィルスに対する細胞表面レセプターとなる本発明のＰＲＯ２４６ポリペプチドには他の用途が見いだされる。例えば、これらのＰＲＯ２４６ポリペプチドから取り出された細胞外ドメインはウィルス感染を低減するためのインビボでの治療に使用することができる。腫瘍特異的抗原になるＰＲＯ２４６ポリペプチドは抗腫瘍薬の治療標的として利用できる等々である。当業者であればそのような用途へのＰＲＯ２４６ポリペプチドの使用方法はよく理解できるであろう。

結合性成長因子と他の成長因子が通常使用されるアッセイはＰＲＯ２６１で実施されるべきである。ＴＧＦβがＰＲＯ２６１を誘導してガンにおける役割を示すかどうかを決定するアッセイは当該分野において知られているようにして実施される。創傷治癒と組織成長アッセイがまたＰＲＯ２６１で実施される。結果は相応に適用される。

ＰＲＯ２２８ポリペプチドは、ＥＭＲ１、ＣＤ９７及びラトロフィリンがその相対活性を決定するために使用されるアッセイにおいて使用されるべきである。結果は相応に利用できる。例えば、ＰＲＯ２２８とＣＤ９７の競合結合アッセイがＣＤ９７、ＣＤ５５のリガンドと共に実施することができる。

【０２１１】

天然のＰＲＯ５３３は残基１－２２がシグナル配列である２１６のアミノ酸ポリペプチドである。残基３から２１６は、線維芽細胞成長因子と５３％の相同性に対応する５０９のＢｌａｓｔスコアを有する。ヌクレオチドレベルで、全長ＤＮＡ９４３５－１２１９を単離するためのＰＣＲオリゴが産生されるＥＳＴであるＤＮＡ４７４１２は１１ｐ１５にマッピングされることが観察されている。１１ｐ１５座位と相同な配列は、ＰＲＯ５３３がアッシャー症候群又は萎縮限域症（Ａｔｒｏｐｈｉａ ａｒｅａｔａ）の治療に有用性を有していることを示している。

これまでに述べられているように、線維芽細胞成長因子は分裂促進的及び非分裂促進的の両方の形で細胞に作用しうる。これらの因子は、顆粒膜細胞、副腎皮質細胞、軟骨細胞、筋芽細胞、角膜及び血管内皮細胞（ウシ又はヒト）、血管平滑筋細胞、水晶体、網膜及び前立腺上皮細胞、オリゴデンドロサイト、アストロサイト、クロンドサイト（ｃｈｒｏｎ

ｄｏｃｙｔｅ）、筋芽細胞及び骨芽細胞を含む広範囲の正常な二倍体中胚葉由来及び神経冠由来細胞に対して分裂促進的である。
線維芽細胞成長因子の非分裂促進的作用は、創傷領域への細胞遊走の促進（走化性）、新しい血管の形成の開始（血管新生）、神経再生と生存の変調（神経栄養性）、内分泌腺機能の変調、及び特異的細胞性タンパク質発現、細胞外生産及び細胞生存の刺激又は抑制を含む。Ｂａｉｒｄ，Ａ．及びＢｏｈｌｅｎ，Ｐ．，Ｈａｎｄｂｏｏｋ ｏｆ Ｅｘｐ．Ｐｈｒｍａｃｏｌ．９５（１）：３６９－４１８（１９９０）。これらの性質は創傷、神経修復、側枝血管形成等々を加速する治療的アプローチにおいて線維芽細胞成長因子を使用するための基礎を提供する。例えば、線維芽細胞成長因子は心臓疾患と手術における心筋損傷を最小化することが示唆されている（米国特許４，３７８，４３７）。
ＰＲＯ２４５ポリペプチドとそれをコードしている核酸は膜貫通プロテインチロシンキナーゼタンパク質とそのコード核酸と相同な配列を有しているので、ＰＲＯ２４５に基づくプローブを用いて他の新規な膜貫通チロシンキナーゼタンパク質を同定するために使用することができる。ＰＲＯ２４５ポリペプチドの可溶型はインビトロとインビボの双方において膜結合ＰＲＯ２４５活性のアンタゴニストとして使用することができる。ＰＲＯ２４５ポリペプチドは天然ＰＲＯ２４５ポリペプチドのアゴニスト又はアンタゴニストを同定するように設計されたスクリーニングアッセイ法において使用することができ、そのようなアッセイは任意の一般的な細胞型又は生化学結合アッセイの形態を取ることができる。更に、ＰＲＯ２４５ポリペプチドは該ポリペプチドが特異的に発現されている組織に対する分子マーカーとなりうる。

【０２１２】

ＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７は全てロイシンリッチ反復を有している。また、ＰＲＯ２２０及びＰＲＯ２２１はＳＬＩＴとロイシンリッチ反復タンパク質と相同性を有している。従って、これらのタンパク質はＳＬＩＴとロイシンリッチ反復タンパク質が使用される上掲の文献に記載されたアッセイにおいて有用である。ＳＬＩＴタンパク質に関しては、ＰＲＯ２２７は神経変性疾患に対するＰＲＯ２２７の影響を決定するアッセイに使用することができる。また、ＰＲＯ２２７はヒト糖タンパク質Ｖと相同性を有している。ＰＲＯ２２７の場合は、このポリペプチドは出血、凝固、組織修復及び瘢痕化に対するその影響を決定するアッセイにおいて使用される。
ＰＲＯ２６６ポリペプチドは神経変性疾患における役割を有しているか又はその逆かを決定するアッセイにおいて使用することができる。
ＰＲＯ２６９ポリペプチド及びトロンビンの活性をもたらすその部分はまたインビボでの治療目的並びに様々なインビトロでの応用に対して有用である。また、ＰＲＯ２６９ポリペプチドとその部分は、ヘパリンと比較して出血の危険が低減した抗血栓剤としての治療用途を有している。トロンボモジュリンと相同性を有しているペプチドは特に望ましい。
ＰＲＯ２８７ポリペプチドと骨形態形成タンパク質「ＢＭＰ１」／プロコラーゲンＣ－プロテイナーゼ（ＰＣＰ）の活性をもたらすその部分はまたインビボでの治療目的並びに様々なインビトロでの応用に対して有用である。また、ＰＲＯ２８７ポリペプチドとその部分は、創傷治癒と組織修復における治療用途を有している。プロコラーゲンＣ－プロテイナーゼエンハンサータンパク質及びその前駆体と相同性を有しているペプチドは骨及び／又は軟骨形成を誘導するために使用でき、従って科学界と医学界にとって特に興味深い。
ＰＲＯ２１４ポリペプチドの治療適応症には、胃腸粘膜の保存と維持及び急性及び慢性病変（例えば、全腸炎、ゾリンジャー・エリソン症候群、胃腸潰瘍形成及び先天性微絨毛萎縮症）、異常なケラチノサイト分化に関連した皮膚疾患（例えば、乾癬、上皮ガン、例えば肺扁平上皮ガン、外陰部と神経膠腫の類表皮ガン）に関連した障害を含む。
ＴＧＦ－スーパーファミリーのメンバーを欠損したマウスの産生と分析に関する研究がＭａｔｚｕｋ，Ｔｒｅｎｄｓ ｉｎ Ｅｎｄｃｒｉｎｏｌ．ａｎｄ Ｍｅｔａｂｏｌ．，６：１２０－１２７（１９９５）に報告されている。
ＰＲＯ３１７ポリペプチド並びにＰＲＯ３１７特異的抗体、阻害剤、アゴニスト、レセプター又はその類似体はここにおいてＰＲＯ３１７関連疾患を治療するのに有用である。よ

って、それらは例えば子宮内膜出血血管新生を変調するのに使用でき、また腎臓組織に効果を持ちうる。子宮内膜出血は異常出血の子宮内膜ガンのような婦人科疾患において生じうる。よって、ここでの組成物は子宮摘出の必要性を減じ又はなくすることのような、子宮内膜における異常出欠状態を診断し治療する用途がまた見いだされる。ここでの分子はまた血管新生用途、例えば血管子内皮成長因子に対する抗体が使用されるか抗腫瘍適応症、又は逆に、血管内皮成長因子が使用される虚血適応症に用途が見いだされうる。

【０２１３】

ＰＲＯ３１７又はそのアゴニスト又はアンタゴニストを含有する生物活性組成物は、哺乳動物種で最大の耐容量を決定し正常なヒト被験者で安全な用量を決定するための臨床実験を含む幾つかの方法の任意のものによって決定される好適な治療用量で投与することができる。また、生物活性剤は様々なよく確立された化合物又は安定性又は半減期のような薬理学的性質を亢進する組成物で複合化してもよい。治療用の生物活性組成物は血流内への静脈内注入又は腎臓、子宮、子宮内膜、血管、又は関連組織、例えば心臓又は生殖器官の治療に使用できる任意の他の効果的な手段により送達させることができると考えられる。製薬組成物でのＰＲＯ３１７、ＰＲＯ３１７アゴニスト、又はＰＲＯ３１７アンタゴニストの用量及び投与は臨床薬理学又は薬物動態学の当業者により決定されうる。例えば、Ｍｏｒｄｅｎｔｉ及びＲｅｓｃｉｇｎｏ，Ｐｈａｒｍａｃｅｕｔｉｃａｌ Ｒｅｓｅａｒｃｈ，９：１７－２５（１９９２）；Ｍｏｒｅｎｔｉ等，Ｐｈａｒｍａｃｅｕｔｉｃａｌ Ｒｅｓｅａｒｃｈ，８：１３５１－１３５９（１９９１）；及びＭｏｒｄｅｎｔｉ及びＣｈａｐｐｅｌｌ，「ｅ ｕｓｅ ｏｆ ｉｎｔｅｒｓｐｅｃｉｅｓ ｓｃａｌｉｎｇ ｉｎ ｔｏｘｉｃｏｋｉｎｅｔｉｃｓ」ｉｎ Ｔｏｘｉｃｏｋｉｎｅｔｉｃｓ ａｎｄ Ｎｅｗ Ｄｒｕｇ Ｄｅｖｅｌｏｐｅｍｅｎｔ，Ｙａｃｏｂｉ等（ｅｄｓ）（Ｐｅｒｇａｍｏｎ Ｐｒｅｓｓ：ＮＹ，１９８９），ｐｐ．４２－９６を参照されたい。治療的に使用されるＰＲＯ３１７、ＰＲＯ３１７アゴニスト、又はＰＲＯ３１７アンタゴニストの有効量は、例えば治療目的、投与経路、及び哺乳動物の症状に依存する。従って、最適の治療効果を得るために必要に応じて療法士が用量力価を決め投与経路を修正することが必要であろう。典型的な毎日の用量は約１０ｎｇ／ｋｇから１００ｍｇ／ｋｇ哺乳動物体重／日あるいはそれ以上、好ましくは約１μｇ／ｋｇ／日から１０ｍｇ／ｋｇ／日の範囲であろう。典型的には臨床医は、上述の疾患の治療に対して所望の効果を達成する用量に到達するまでＰＲＯ３１７、ＰＲＯ３１７アゴニスト、又はＰＲＯ３１７アンタゴニストを投与する。

ＰＲＯ３１７、ＰＲＯ３１７アゴニスト、又はＰＲＯ３１７アンタゴニストは単独で又は所望の薬理効果を達成するために他のものと組み合わせて投与できる。ＰＲＯ３１７自体、又はＰＲＯ３１７のアゴニスト又はアンタゴニストは治療的に投与されたときに異なった効果をもたらしうる。治療のためのそのような化合物は非毒性、不活性で製薬的に許容可能な水性の担体媒体中に好ましくは約５から８、より好ましくは６から８のｐＨで処方されるが、ｐＨは処方されるＰＲＯ３１７、アゴニスト又はアンタゴニスト及び治療される症状に応じて変わりうる。治療化合物の特性は、分子の溶解度、半減期、及び抗原性／免疫原生を含む；これらと他の特性が効果的な担体を定めるのに役立ちうる。

ＰＲＯ３１７、ＰＲＯ３１７アゴニスト、又はＰＲＯ３１７アンタゴニストは、局所用クリーム及びゲル；経粘膜スプレー及びエアゾール、経皮性パッチ及び包帯；注射、静脈内、及び洗浄製剤；及び特に胃酸及び酵素に抗するように処方された経口投与液体及び丸剤を含む既知の投与経路によって送達されうる。特定の処方、正確な投薬量、及び投与経路は主治医師により決定され、それぞれの特定の状況に応じて変動するであろう。

そのような投与の決定は、治療される病状、治療される哺乳動物の種類、投与される化合物、及び特定の治療化合物の薬物動態学特性のような複数の変動要因を考慮してなされる。考慮されうる更なる要因は患者の病気の状態（例えば重症度）、年齢、体重、性別、食餌、投与時間、薬物組み合わせ、反応感受性、及び治療に対する耐性／応答を含む。長く作用する治療化合物製剤（例えばリポソームにカプセル化したＰＲＯ３１７又はＰＥＧ化ＰＲＯ３１７又はＰＲＯ３１７ポリマーミクロスフィア、例えばポリ乳酸ベースミクロス

フィア）は特定の治療化合物の半減期及びクリアランス速度に応じて３から４日毎、毎週、あるいは２週間毎に投与することができるかもしれない。
通常の投薬量は、投与経路に応じて、約１０ｎｇ／ｋｇから１００ｍｇ／ｋｇ哺乳動物体重／日あるいはそれ以上、好ましくは約１μｇ／ｋｇ／日から１０ｍｇ／ｋｇ／日の範囲で変化しうる。特定の投薬量及び送達方法に関する指導は文献において提供されている；例えば米国特許４，６５７，７６０；５，２０６，３４４；又は５，２２５，２１２を参照されたい。異なった製剤は異なった治療化合物に対して効果的であり、例えば子宮を標的とする投与には他の器官又は組織、例えば心臓組織へのものとは異なった形で送達することを必要とすると予想される。
【０２１４】
ＰＲＯ３１７の投与を必要とする任意の疾患又は障害の治療に好適な放出特性を持つ製剤によるＯＲＯ３１７の徐放投与が望ましい場合、ＰＲＯ３１７のマイクロカプセル化が考えられる。徐放のための組換えタンパク質のマイクロカプセル化はヒト成長ホルモン（ｒｈＧＨ）、インターフェロン－（ｒｈＩＦＮ－）、インターロイキン－２、及びＭＮｒｇｐ１２０成功裏に実施されている。Ｊｏｈｎｓｏｎ等，Ｎａｔ．Ｍｅｄ．，２：７９５－７９９（１９９６）；Ｙａｓｕｄａ，Ｂｉｏｍｅｄ．Ｔｈｅｒ．，２７：１２２１－１２２３（１９９３）；Ｈｏｒａ等，Ｂｉｏ／Ｔｅｃｈｎｏｌｏｇｙ，８：７５５－７５８（１９９０）；Ｃｌｅｌａｎｄ，「Ｄｅｓｉｇｎ ａｎｄ Ｐｒｏｄｕｃｔｉｏｎ ｏｆ Ｓｉｎｇｌｅ Ｉｍｍｕｎｉｚａｔｉｏｎ Ｖａｃｃｉｎｅｓ Ｕｓｉｎｇ Ｐｏｌｙｌａｃｔｉｄｅ Ｐｏｌｙｇｌｙｃｏｌｉｄｅ Ｍｉｃｒｏｓｐｈｅｒｅ Ｓｙｓｔｅｍｓ」ｉｎ Ｖａｃｃｉｎｅ Ｄｅｓｉｇｎ：Ｔｈｅ Ｓｕｂｕｎｉｔ ａｎｄ Ａｄｊｕｖａｎｔ Ａｐｐｒｏａｃｈ，Ｐｏｗｅｌｌ及びＮｅｗｍａｎ，ｅｄｓ．（Ｐｌｅｎｕｍ Ｐｒｅｓｓ：Ｎｅｗ Ｙｏｒｋ，１９９５），ｐｐ．４３９－４６２；ＷＯ９７／０３６９２，ＷＯ９６／４００７２，ＷＯ９６／０７３９９；及び米国特許５，６５４，０１０。
子宮、子宮内膜組織、又は他の生殖組織又は心臓組織の病状又は疾患は、ＰＲＯ３１７の発現が病理状態では減少する場合はＰＲＯ３１７又はＰＲＯアゴニストで；あるいはＰＲＯ３１７の発現が病理状態では増加する場合はＰＲＯ３１７の抗体又は他のＰＲＯ３１７アンタゴニストで治療可能である損傷を引き起こしうる。これらの症状又は疾患は器官の機能に影響を及ぼす生理的及び病理的問題に対して上述した探索試験により特異的に診断することができる。
ＰＲＯ３１７、ＰＲＯ３１７アゴニスト、又はＰＲＯ３１７アンタゴニストは、他の生物活性剤と共に、別個にあるいは同じ製剤で哺乳動物に投与して、それらが好適な共通の適応症を治療してもよい。例えば、ＰＲＯ３１７はＥＮＢＡＦ－１と共にそれらが同じ定性的生物学的効果を示す適応症に対して投与することができると考えられる。あるいは、それらが反対の効果を有する場合は、ＥＢＡＦ－１はＰＲＯ３１７に対するアンタゴニスト、例えば抗ＰＲＯ３１７抗体と共に投与することができる。更に、ＰＲＯ３１７は適応が保証されている冠動脈虚血症に対してＶＥＧＦと共に、あるいは保証されているガンに対して抗ＶＥＧＦと共に投与することができ、あるいは逆に、ＰＲＯ３１７のアンタゴニストが、冠動脈虚血症に対してＶＥＧＦと共に又は保証されているガンの治療のために抗ＶＥＧＦと共に投与されうる。これらの投与はかかる障害を治療するのに効果的な量でなされる。
天然ＰＲＯ３０１（配列番号：１１９）はＡ３３抗原前駆体のＡ３３＿ＨＵＭＡＮとＦｉｇ４４の残基２４から２８２で３０％の相同性と２４６のＢｌａｓｔスコアを有している。背景技術において説明したように、Ａ３３抗原前駆体は腫瘍特異的抗原であり、よって大腸癌の診断と治療の認められているマーカー及び治療標的である。腫瘍特異的抗原の発現はしばしば新生物組織疾患の進行に関連している。天然ＰＲＯ３０１（配列番号：１１９）とＡ３３＿ＨＵＭＡＮはまた如何にして間隙が比較配列に挿入されるかに依存した変異であるＡ３３＿ＨＵＭＡＮとＦｉｇ４４の残基２１から２８２で３０％の相同性と２４５のＢｌａｓｔスコアを示す。天然ＰＲＯ３０１（配列番号：１１９）はまたヒトコクサ

ッキー及びアデノウィルスレセプタータンパク質で細胞表面タンパク質ＨＣＡＲとしても知られているＨＳ４６ＫＤＡ＿１とＦｉｇ４４の残基６０から２５５で２９％の相同性と１６５のＢｌａｓｔスコアを有している。ＰＲＯ３０１のこの領域はまたＨＳＵ９０７１６＿１と類似のＢｌａｓｔスコアと相同性を示す。このようなタンパク質の発現は通常ウィルス感染が付随し、従ってそのような感染を防止する治療法が考えられうる。背景技術において述べたように、ウィルスレセプターの発現はしばしば新生物腫瘍に付随する。

【０２１５】

本発明のＰＲＯ２３４ポリペプチドの治療的用途は、白血球ホーミング又は炎症性応答の間の白血球と内皮間の相互作用に関連した治療を含む。例としては、喘息、関節リウマチ、乾癬及び多発性硬化症が含まれる。

ＰＲＯ２３１ポリペプチドとそれをコードしている核酸は推定酸性ホスファターゼとそのコード核酸に対して相同性を有しているので、ＰＲＯ２３１ヌクレオチド配列に基づくプローブを、他の新規なホスファターゼタンパク質を同定するために使用することができる。ＰＲＯ２３１ポリペプチドの可溶型はインビトロとインビボの双方において膜結合ＰＲＯ２３１活性のアンタゴニストとして使用することができる。ＰＲＯ２３１ポリペプチドは天然ＰＲＯ２３１ポリペプチドのアゴニスト又はアンタゴニストを同定するように設計されたスクリーニングアッセイ法において使用することができ、そのようなアッセイは任意の一般的な細胞型又は生化学結合アッセイの形態を取ることができる。更に、ＰＲＯ２３１ポリペプチドは該ポリペプチドが特異的に発現されている組織に対する分子マーカーとなりうる。

ＰＲＯ２２９ポリペプチドは、融合ペプチドが対象のペプチドに対して増加した半減期を有しているかどうかを決定するために対象のペプチドと融合させることができる。従って、ＰＲＯ２２９ポリペプチドは対象のポリペプチドの半減期を増加させるために使用することができる。半減期の増加を引き起こすＰＲＯ２２９の部分はここで本発明の実施態様である。

ＰＲＯ２３８は酸素とアセチル－ＣｏＡを含む基質を還元するその能力を測定するアッセイ、特に酸素フリーラジカルをつくるＰＲＯ２３８の能力を測定するアッセイにおいて使用することができる。

ＰＲＯ２３８は更にその還元能を阻止、低減又は逆転させる候補を検定するために使用することができる。これは、ＰＲＯ２３８と還元する基質を有する一つのアッセイに候補が添加され、他のアッセイに添加されず、候補の存在がなかったら同じである並行アッセイを実施することによりなされる。

ＰＲＯ２３３ポリペプチドとレダクターゼと相同性を有するその部分がまたインビボ治療目的並びに様々な他の用途に対して有用である。新規なレダクターゼタンパク質と関連分子の同定は、多数のヒト障害、例えば炎症性疾患、器官不全、アテローム性動脈硬化症、心臓傷害、不妊症、奇形児、早老、エイズ、ガン、糖尿病合併症及び突然変異一般に関連している。酸素フリーラジカルと抗酸化剤が多数の疾患プロセスにおいて重要な役割を担っていることが分かっているので、新規なレダクターゼタンパク質とレダクターゼ様分子を同定することは、そのようなタンパク質が様々な異なったヒトの障害に対する強力な治療法となる点で特に重要である。そのようなポリペプチドはまたバイオテクノロジーと医療研究、並びに様々な産業上の利用において重要な役割を果たしうる。その結果、新規な分子、例えばＰＲＯ２３３には特定の科学的及び医学的興味がある。

セリンカルボキシペプチダーゼ活性を示す本発明のＰＲＯ２２３ポリペプチドはインビボでの治療目的に対して並びにインビトロでの目的に対して使用することができる。当業者であればそのような用途に対してＰＲＯ２２３ポリペプチドを如何にして使用するかは分かるであろう。

【０２１６】

ＰＲＯ２３５ポリペプチドと細胞接着に関与しうるその部分がまたインビボ治療目的並びに様々なインビトロでの用途に対して有用である。また、ＰＲＯ２３５ポリペプチドとその部分は細胞接着に関与する疾患状態における治療用途を有している。インビボでの細胞

接着機構の生理的な重要性が分かっているので、細胞接着に関与する新規な天然タンパク質を同定する努力が現在なされている。従って、プレキシンと相同性を有するペプチドが科学界と医学界にとって特に興味深い。
ここに開示されたＰＲＯ２３６及びＰＲＯ２６２ポリペプチドは様々な既知のβ－ガラクトシダーゼタンパク質と相同であるので、ここに開示されたＰＲＯ２３６及びＰＲＯ２６２は、ガラクトシル化プロドラッグからの活性薬物の発生（例えばプロドラッグβ－ガラクトシル－５－フルオロウリジンからの５－フルオロウリジンの発生）による腫瘍細胞の特異的死滅のためのポリペプチドとモノクローナル抗体のコンジュゲートに用途が見いだされる。ここに開示されたＰＲＯ２３６及びＰＲＯ２６２ポリペプチドは、インビボとインビトロの両方での様々な用途が見いだされ、その用途はβ－ガラクトシダーゼタンパク質が現在使用されている用途と同様又は同一である。当業者であればそのような用途に対してＰＲＯ２２３ポリペプチドを如何にして使用するかは分かるであろう。
ＰＲＯ２３９ポリペプチドとデンシンに対して相同性を有しているその部分はまたインビボでの治療目的、並びに様々なインビトロでの用途に対して有用である。また、ＰＲＯ２３９ポリペプチドとその部分はシナプス機構、再生又は細胞接着に関与する疾患状態における治療用途を有している。インビボでのシナプス突起、再生及び細胞接着機構の生理学的な重要性が分かっているので、シナプス機構と細胞接着に関与する新規な天然のタンパク質を同定するための努力が現在なされている。従って、デンシンと相同性を有するペプチドが科学界と医学界にとって特に興味深い。
ここに記載されたＰＲＯ２６０はフコシダーゼに対するその関係を決定するアッセイにおいて使用することができる。特に、ＰＲＯ２６０ポリペプチドはプロテオグリカンからフコース又は他の糖残基を取り除く能力を決定するアッセイに使用することができる。ＰＲＯ２６０ポリペプチドは、それらがフコシダーゼと機能的又は位置的類似性を有しているかどうかを決定するために検定することができる。ついで、ＰＲＯ２６０はそれらが完全である系を制御するために使用することができる。
ＰＲＯ２６３はＣＤ４４抗原がＣＤ４４のものに対するＰＲＯ２６３の活性を決定するために一般に使用されるアッセイにおいて使用することができる。結果は相応に使用することができる。
ＰＲＯ２７０と還元－酸化（レドックス）をもたらすその部分はまたインビボでの治療目的、並びに様々なインビトロでの用途に対して有用でありうる。より詳細には、ＰＲＯ２７０ポリペプチドは、エイズ、ガン、アテローム性動脈硬化症、糖尿病合併症及び炎症及び脳卒中のような酸化ストレスに関与する病理状態に関与していると思われる非常に多様な遺伝子の発現に影響を及ぼしうる。また、ＰＲＯ２７０ポリペプチドとその部分はアポトーシスに役割を担う遺伝子の発現に影響を及ぼしうる。従って、ヒオレドキシンと相同性を有するペプチドは科学界と医学界にとって特に望ましい。
ＰＲＯ２７２とカルシウムに結合する能力を有するその部分はまた数多くのインビボでの治療用途、並びに様々なインビトロでの用途を有しうる。従って、レティキュロカルビンと相同性を有するペプチドが特に望ましい。当業者であればそのような目的に対してＰＲＯ２７２ポリペプチドとその部分を如何にして使用するかを理解できるであろう。
【０２１７】
ＰＲＯ２９４ポリペプチドとコラーゲンに対して相同性を有しているその部分はまたインビボでの治療目的、並びに様々なインビトロでの用途に対して有用である。新規なコラーゲン及びコラーゲン様分子の同定は多数のヒト障害と関連性を有している。よって、新規なコラーゲンとコラーゲン様分子の同定は、そのようなタンパク質が様々な異なったヒト障害に対して強力な治療法となりうる点で特に重要である。そのようなポリペプチドはまたバイオテクノロジー及び医学研究並びに様々な産業上の用途において重要な役割を担っている。コラーゲンの非常に多くの用途が分かっているので、コラーゲン分子と相同性を持つポリペプチドには大きな興味がある。
ＰＲＯ２９５ポリペプチドとインテグリンに対して相同性を有しているその部分はまたインビボでの治療目的、並びに様々なインビトロでの用途に対して有用である。新規なイン

テグリン及びインテグリン様分子の同定は多数のヒト障害、例えば免疫系の細胞の結合又は活性の変調と関連性を有している。よって、新規なインテグリンとインテグリン様分子の同定は、そのようなタンパク質が様々な異なったヒト障害に対して強力な治療法となりうる点で特に重要である。そのようなポリペプチドはまたバイオテクノロジー及び医学研究並びに様々な産業上の用途において重要な役割を担っている。その結果、例えばＰＲＯ２９５のような新規の分子に特に科学的かつ医学的興味がある。

ＰＲＯ２９３ポリペプチドは明らかにロイシンリッチ反復ポリペプチド相同体であるので、該ペプチドは、既知のＮＬＲＲ－１及びＮＬＲＲ－２ポリペプチドが使用される全ての用途に使用することができる。活性はこれらのペプチド間で比較することができ、よって適宜適用される。

ここに記載されたＰＲＯ２４７ポリペプチドは、デンシンがデンシン又はこれら他のタンパク質に対するＰＲＯ２４７の活性を決定するために使用されるアッセイにおいて使用することができる。結果はＰＲＯ２４７での診断及び／又は治療用とにおいて適宜使用することができる。

プロテアーゼ活性を有する本発明のＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３ポリペプチドは、インビボでの治療目的となインビトロの双方で使用されうる。当業者であれば、このような目的に対して本発明のＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３ポリペプチドを如何にして使用するかは分かっているであろう。

ＰＲＯ３２８ポリペプチドとＧＬＩＰ及びＣＲＩＳＰに対して相同性を有しているその部分はまたインビボでの治療目的、並びに様々な他の用途に対して有用である。新規なＧＬＩＰ及びＣＲＩＳＰ様分子の同定は、転写制御に関与するかヒト腫瘍に過剰発現される多数のヒト障害と関連性を有している。よって、新規なＧＬＩＰ及びＣＲＩＳＰ様分子の同定は、そのようなタンパク質が様々な異なったヒト障害に対して強力な治療法となりうる点で特に重要である。そのようなポリペプチドはまたバイオテクノロジー及び医学研究並びに様々な産業上の用途において重要な役割を担っている。その結果、例えばＰＲＯ３２８のような新規の分子に特に科学的かつ医学的興味がある。

【０２１８】

ＰＲＯ３３５、ＰＲＯ３３１又はＰＲＯ３２６の用途はＬＩＧ－１、ＡＬＳ及びデコリンでの競合アッセイにおいてその相対活性を決定するための用途を含む。結果は適宜使用することができる。ＰＲＯ３３５、ＰＲＯ３３１又はＰＲＯ３２６はＬＩＧ－１が同じ効果を受けるかどうかを決定するために使用されるアッセイにおいて使用することもまたできる。

ＰＲＯ３３２はＦｉｇ１０８（配列番号：３１０）のアミノ酸位置６２５－６２８にＧＡＧ反復（ＧＫＥＫ）を含む。そのような反復におけるずれはヒト疾患と関連しうる。従って、ＰＲＯ３３２は遺伝子治療による、つまり正しいＧＫＥＫ配列モチーフを含む遺伝子を導入することによるそのような疾患状態を治療するのに有用でありうる。

ＰＲＯ３３４の他の用途はフィブリリン又はフィブリンがフィブリリン又はフィブリンに対するＰＲＯ３３４の相対活性を決定するために使用されるアッセイにおける使用を含む。特に、ＰＲＯ３３４は、上皮成長因子反復により付与される機構を必要とするアッセイにおいて使用することができる。

天然ＰＲＯ３４６（配列番号：３２０）は、ガン胎児抗原ｃｇｍ６前駆体であるＣＧＭ６＿ＨＵＭＡＮの残基３５から１０４０とアミノ酸残基２１から３４３の間に２７％の相同性に対応する２３０のＢｌａｓｔスコアを有している。この相同領域は、幾つかの膜貫通残基（３４０－３４３）に加えて、ＰＲＯ３４６の残基１４８から３３９で免疫グロブリンスーパーファミリー相同性を含む殆ど全てであるが２つのＮ末端細胞外ドメイン残基を含む。ガン胎児抗原前駆体は背景技術において説明されているように腫瘍特異的抗原であり、よって大腸ガンの診断と治療に対する認められたマーカー及び治療標的である。腫瘍特異的抗原の発現はしばしば新生物組織疾患の進行に関連している。天然ＰＲＯ３４６（配列番号：３２０）とＰ＿Ｗ０６８７４、ヒトガン胎児抗原ＣＥＡ－ｄはそれぞれ残基２

から３４３と６７から３４２の間に２８％の相同性を２２４のＢｌａｓｔスコアを有している。この相同性は天然のＰＲＯ３４６の全細胞外ドメイン残基、マイナス開始メチオニン（残基２から１８）並びに膜貫通残基（３４０－３４３）を含む。
プロテインジスルフィドイソメラーゼ活性を有しているＰＲＯ２６８ポリペプチドは、例えば組換え生産されたタンパク質において正しく折り畳まれたタンパク質の収量を増大させるように正しいジスルフィド結合形成を促進する際の用途を含む、プロテインジスルフィドイソメラーゼ活性が望まれる多くの応用において有用である。当業者であれはそのような目的に対してかかるＰＲＯ２６８ポリペプチドを如何にして使用するかは即座に理解できるであろう。
プロピル４－ヒドロキシラーゼαサブユニットタンパク質のものと関連した生物活性を有する本発明のＰＲＯ３３０ポリペプチドは治療目的のインビボとインビトロの両方において使用することができる。当業者であれはそのような目的に対して本発明のＰＲＯ３３０ポリペプチドを如何にして使用するかはよく知っているであろう。

【０２１９】

Ｆ．抗－ＰＲＯ抗体

本発明は、さらに抗－ＰＲＯ抗体を提供するものである。抗体の例としては、ポリクローナル、モノクローナル、ヒト化、二重特異性及びヘテロ抱合体抗体が含まれる。

１．ポリクローナル抗体

抗－ＰＲＯ抗体はポリクローナル抗体を含む。ポリクローナル抗体の調製方法は当業者に知られている。哺乳動物においてポリクローナル抗体は、例えば免疫化剤、及び所望するのであればアジュバントを、一又は複数回注射することで発生させることができる。典型的には、免疫化剤及び／又はアジュバントを複数回皮下又は腹腔内注射により、哺乳動物に注射する。免疫化剤は、ＰＲＯポリペプチド又はその融合タンパク質を含みうる。免疫化剤を免疫化された哺乳動物において免疫原性が知られているタンパク質に抱合させるのが有用である。このような免疫原タンパク質の例は、これらに限られないが、キーホールリンペットヘモシアニン、血清アルブミン、ウシサイログロブリン及び大豆トリプシンインヒビターが含まれる。使用され得るアジュバントの例には、フロイント完全アジュバント及びＭＰＬ－ＴＤＭアジュバント（モノホスホリル脂質Ａ、合成トレハロースジコリノミコラート）が含まれる。免疫化プロトコールは、過度の実験なく当業者により選択されるであろう。

【０２２０】

２．モノクローナル抗体

あるいは、抗－ＰＲＯ抗体はモノクローナル抗体であってもよい。モノクローナル抗体は、Ｋｏｈｌｅｒ及びＭｉｌｓｔｅｉｎ，Ｎａｔｕｒｅ，２５６：４９５（１９７５）に記載されているようなハイブリドーマ法を使用することで調製することができる。ハイブリドーマ法では、マウス、ハムスター又は他の適切な宿主動物を典型的には免疫化剤により免疫化することで、免疫化剤に特異的に結合する抗体を生成するかあるいは生成可能なリンパ球を誘発する。また、リンパ球をインビトロで免疫化することもできる。
免疫化剤は、典型的には対象とするＰＲＯポリペプチド又はその融合タンパク質を含む。一般にヒト由来の細胞が望まれる場合には末梢血リンパ球（「ＰＢＬ」）が使用され、あるいは非ヒト哺乳動物源が望まれている場合は、脾臓細胞又はリンパ節細胞が使用される。次いで、ポリエチレングリコール等の適当な融合剤を用いてリンパ球を不死化株化細胞と融合させ、ハイブリドーマ細胞を形成する［Ｇｏｄｉｎｇ，Ｍｏｎｏｃｌｏｎａｌ Ａｎｔｉｂｏｄｉｅｓ：Ｐｒｉｎｃｉｐｌｅｓ ａｎｄ Ｐｒａｃｔｉｃｅ，Ａｃａｄｅｍｉｃ Ｐｒｅｓｓ，（１９８６）ｐｐ．５９－１０３］。不死化株化細胞は、通常は、形質転換した哺乳動物細胞、特に齧歯動物、ウシ、及びヒト由来の骨髄腫細胞である。通常、ラット又はマウスの骨髄腫株化細胞が使用される。ハイブリドーマ細胞は、好ましくは、未融合の不死化細胞の生存又は成長を阻害する一又は複数の物質を含有する適切な培地で培養される。例えば、親細胞が、酵素のヒポキサンチングアニンホスホリボシルトランスフェラーゼ（ＨＧＰＲＴ又はＨＰＲＴ）を欠いていると、ハイブリドーマの培地は、

典型的には、ヒポキサチン、アミノプチリン及びチミジンを含み（「ＨＡＴ培地」）、この物質がＨＧＰＲＴ欠乏性細胞の増殖を阻止する。
好ましい不死化株化細胞は、効率的に融合し、選択された抗体生成細胞による安定した高レベルの抗体発現を支援し、ＨＡＴ培地のような培地に対して感受性である。より好ましい不死化株化細胞はマウス骨髄腫株であり、これは例えばカリフォルニア州サンディエゴのＳａｌｋ Ｉｎｓｔｉｔｕｔｅ Ｃｅｌｌ Ｄｉｓｔｒｉｂｕｔｉｏｎ Ｃｅｎｔｅｒやメリーランド州ロックビルのアメリカン・タイプ・カルチャー・コレクションより入手可能である。ヒトモノクローナル抗体を生成するためのヒト骨髄腫及びマウス－ヒト異種骨髄腫株化細胞も開示されている［Ｋｏｚｂｏｒ，Ｊ．Ｉｍｍｕｎｏｌ．，１３３：３００１（１９８４）、Ｂｒｏｄｅｕｒ等，Ｍｏｎｏｃｌｏｎａｌ Ａｎｔｉｂｏｄｙ Ｐｒｏｄｕｃｔｉｏｎ Ｔｅｃｈｎｉｑｕｅｓ ａｎｄ Ａｐｐｌｉｃａｔｉｏｎｓ，Ｍａｒｃｅｌ Ｄｅｋｋｅｒ，Ｉｎｃ．，Ｎｅｗ Ｙｏｒｋ，（１９８７）ｐｐ．５１－６３］。
次いでハイブリドーマ細胞が培養される培養培地を、ＰＲＯに対するモノクローナル抗体の存在について検定する。好ましくは、ハイブリドーマ細胞によって生成されたモノクローナル抗体の結合特異性は免疫沈降又はラジオイムノアッセイ（ＲＩＡ）や酵素結合免疫測定法（ＥＬＩＳＡ）等のインビトロ結合検定法によって測定する。このような技術及びアッセイは、当該分野において公知である。モノクローナル抗体の結合親和性は、例えばＭｕｎｓｏｎ及びＰｏｌｌａｒｄ，Ａｎａｌ．Ｂｉｏｃｈｅｍ．，１０７：２２０（１９８０）によるスキャッチャード分析法によって測定することができる。
所望のハイブリドーマ細胞が同定された後、クローンを制限希釈工程によりサブクローニングし、標準的な方法で成長させることができる［Ｇｏｄｉｎｇ，上掲］。この目的のための適当な培地には、例えば、ダルベッコの改変イーグル培地及びＲＰＭＩ－１６４０倍地が含まれる。あるいは、ハイブリドーマ細胞は哺乳動物においてインビボで腹水として成長させることもできる。
サブクローンによって分泌されたモノクローナル抗体は、例えばプロテインＡ－セファロース法、ヒドロキシルアパタイトクロマトグラフィー法、ゲル電気泳動法、透析法又はアフィニティークロマトグラフィー等の従来の免疫グロブリン精製方法によって培養培地又は腹水液から単離又は精製される。
また、モノクローナル抗体は、組換えＤＮＡ法、例えば米国特許第４，８１６，５６７号に記載された方法により作成することができる。本発明のモノクローナル抗体をコードするＤＮＡは、常套的な方法を用いて（例えば、マウス抗体の重鎖及び軽鎖をコードする遺伝子に特異的に結合可能なオリゴヌクレオチドプローブを使用して）、容易に単離し配列決定することができる。本発明のハイブリドーマ細胞はそのようなＤＮＡの好ましい供給源となる。ひとたび単離されたら、ＤＮＡは発現ベクター内に配することができ、これが宿主細胞、例えばサルＣＯＳ細胞、チャイニーズハムスター卵巣（ＣＨＯ）細胞、あるいは免疫グロブリンタンパク質を生成等しない骨髄腫細胞内に形質移入され、組換え宿主細胞内でモノクローナル抗体の合成をすることができる。また、ＤＮＡは、例えば相同マウス配列に換えてヒト重鎖及び軽鎖定常ドメインのコード配列を置換することにより［米国特許第４，８１６，５６７号；Ｍｏｒｒｉｓｏｎ等，上掲］、又は免疫グロブリンコード配列に非免疫グロブリンポリペプチドのコード配列の一部又は全部を共有結合することにより修飾することができる。このような非免疫グロブリンポリペプチドは、本発明の抗体の定常ドメインに置換でき、あるいは本発明の抗体の１つの抗原結合部位の可変ドメインに置換でき、キメラ性二価抗体を生成する。
抗体は一価抗体であってもよい。一価抗体の調製方法は当該分野においてよく知られてる。例えば、一つの方法は免疫グロブリン軽鎖と修飾重鎖の組換え発現を含む。重鎖は一般的に、重鎖の架橋を防止するようにＦｃ領域の任意の点で切断される。あるいは、関連するシステイン残基を他のアミノ酸残基で置換するか欠失させて架橋を防止する。
一価抗体の調製にはインビトロ法がまた適している。抗体の消化による、その断片、特にＦａｂ断片の生成は、当該分野において知られている慣用的技術を使用して達成できる。

【０２２１】
３．ヒト及びヒト化抗体
本発明の抗－ＰＲＯ抗体は、さらにヒト化抗体又はヒト抗体を含む。非ヒト（例えばマウス）抗体のヒト化形とは、キメラ免疫グロブリン、免疫グロブリン鎖あるいはその断片（例えばＦｖ、Ｆａｂ、Ｆａｂ’、Ｆ（ａｂ’）$_2$あるいは抗体の他の抗原結合サブ配列）であって、非ヒト免疫グロブリンに由来する最小配列を含むものである。ヒト化抗体はレシピエントの相補性決定領域（ＣＤＲ）の残基が、マウス、ラット又はウサギのような所望の特異性、親和性及び能力を有する非ヒト種（ドナー抗体）のＣＤＲの残基によって置換されたヒト免疫グロブリン（レシピエント抗体）を含む。幾つかの例では、ヒト免疫グロブリンのＦｖフレームワーク残基は、対応する非ヒト残基によって置換されている。また、ヒト化抗体は、レシピエント抗体にも、移入されたＣＤＲもしくはフレームワーク配列にも見出されない残基を含んでいてもよい。一般に、ヒト化抗体は、全てあるいはほとんど全てのＣＤＲ領域が非ヒト免疫グロブリンのものに対応し、全てあるいはほとんど全てのＦＲ領域がヒト免疫グロブリンコンセンサス配列のものである、少なくとも１つ、典型的には２つの可変ドメインの実質的に全てを含む。ヒト化抗体は、最適には免疫グロブリン定常領域（Ｆｃ）、典型的にはヒトの免疫グロブリンの定常領域の少なくとも一部を含んでなる［Ｊｏｎｅｓ等，Ｎａｔｕｒｅ，３２１：５２２－５２５（１９８６）；Ｒｉｅｃｈｍａｎｎ等，Ｎａｔｕｒｅ，３３２：３２３－３２９（１９８８）；及びＰｒｅｓｔａ，Ｃｕｒｒ．Ｏｐ Ｓｔｒｕｃｔ．Ｂｉｏｌ．，２：５９３－５９６（１９９２）］。
非ヒト抗体をヒト化する方法はこの分野でよく知られている。一般的に、ヒト化抗体には非ヒト由来の一又は複数のアミノ酸残基が導入される。これら非ヒトアミノ酸残基は、しばしば、典型的には「移入」可変ドメインから得られる「移入」残基と称される。ヒト化は基本的に齧歯動物のＣＤＲ又はＣＤＲ配列でヒト抗体の該当する配列を置換することによりウィンター（ｗｉｎｔｅｒ）及び共同研究者［Ｊｏｎｅｓ等，Ｎａｔｕｒｅ，３２１：５２２－５２５（１９８６）；Ｒｉｅｃｈｍａｎｎ等，Ｎａｔｕｒｅ，３３２：３２３－３２７（１９８８）；Ｖｅｒｈｏｅｙｅｎ等，Ｓｃｉｅｎｃｅ，２３９：１５３４－１５３６（１９８８）］の方法に従って、齧歯類ＣＤＲ又はＣＤＲ配列をヒト抗体の対応する配列に置換することにより実施される。よって、このような「ヒト化」抗体は、無傷のヒト可変ドメインより実質的に少ない分が非ヒト種由来の対応する配列で置換されたキメラ抗体（米国特許第４，８１６，５６７号）である。実際には、ヒト化抗体は典型的には幾つかのＣＤＲ残基及び場合によっては幾つかのＦＲ残基が齧歯類抗体の類似する部位からの残基によって置換されたヒト抗体である。
また、ヒト抗体は、ファージ表示ライブラリ［Ｈｏｏｇｅｎｂｏｏｍ及びＷｉｎｔｅｒ，Ｊ．Ｍｏｌ．Ｂｉｏｌ．，２２７：３８１（１９９２）；Ｍａｒｋｓ等，Ｊ．Ｍｏｌ．Ｂｉｏｌ．，２２２：５８１（１９９１）］を含むこの分野で知られた種々の方法を用いて作成することもできる。また、Ｃｏｌｅ等及びＢｏｅｒｎｅｒ等の方法も、ヒトモノクローナル抗体の調製に利用することができる［Ｃｏｌｅ等，Ｍｏｎｏｃｌｏｎａｌ Ａｎｔｉｂｏｄｉｅｓ ａｎｄ Ｃａｎｃｅｒ Ｔｈｅｒａｐｙ，Ａｌａｎ Ｒ．Ｌｉｓｓ．ｐ．７７（１９８５）及びＢｏｅｒｎｅｒ等，Ｊ．Ｉｍｍｕｎｏｌ．，１４７（１）：８６－９５（１９９１）］。同様に、ヒト抗体はヒト免疫グロブリン座位をトランスジェニック動物、例えば内在性免疫グロブリン遺伝子は部分的又は完全に不活性化されたマウスに導入することにより産生することができる。投与の際に、遺伝子再配列、組立、及び抗体レパートリーを含むあらゆる観点においてヒトに見られるものに非常に類似しているヒト抗体の生産が観察される。このアプローチは、例えば米国特許第５，５４５，８０７号；同第５，５４５，８０６号；同第５，５６９，８２５号；同第５，６２５，１２６号；同第５，６３３，４２５号；同第５，６６１，０１６号、及び次の科学文献：Ｍａｒｋｓ等，Ｂｉｏ／Ｔｅｃｈｎｏｌｏｇｙ １０，７７９－７８３（１９９２）；Ｌｏｎｂｅｒｇ等，Ｎａｔｕｒｅ ３６８ ８５６－８５９（１９９４）；Ｍｏｒｒｉｓｏｎ，Ｎａｔｕｒｅ ３６８，８１２－１３（１９９４）；Ｆｉｓｈｗｉｌ

ｄ等，Ｎａｔｕｒｅ Ｂｉｏｔｅｃｈｎｏｌｏｇｙ １４，８４５－５１（１９９６）；Ｎｅｕｂｅｒｇｅｒ，Ｎａｔｕｒｅ Ｂｉｏｔｅｃｈｎｏｌｏｇｙ １４，８２６（１９９６）；Ｌｏｎｂｅｒｇ及びＨｕｓｚａｒ，Ｉｎｔｅｒｎ．Ｒｅｖ．Ｉｍｍｕｎｏｌ．１３ ６５－９３（１９９５）に記載されている。

【０２２２】

４．二重特異性抗体

二重特異性抗体は、少なくとも２つの異なる抗原に対して結合特異性を有するモノクローナル抗体、好ましくはヒトもしくはヒト化抗体である。本発明の場合において、結合特異性の一方はＰＲＯに対してであり、他方は任意の他の抗原、好ましくは細胞表面タンパク質又はレセプター又はレセプターサブユニットに対してである。

二重特異性抗体を作成する方法は当該技術分野において周知である。伝統的には、二重特異性抗体の組換え生産は、二つの重鎖が異なる特異性を持つ二つの免疫グロブリン重鎖／軽鎖対の同時発現に基づく［Ｍｉｌｓｔｅｉｎ及びＣｕｅｌｌｏ，Ｎａｔｕｒｅ，３０５：５３７－５３９（１９８３）］。免疫グロブリンの重鎖と軽鎖を無作為に取り揃えるため、これらハイブリドーマ（クアドローマ）は１０種の異なる抗体分子の潜在的混合物を生成し、その内一種のみが正しい二重特異性構造を有する。正しい分子の精製は、アフィニティークロマトグラフィー工程によって通常達成される。同様の手順が１９９３年５月１３日公開のＷＯ ９３／０８８２９、及びＴｒａｕｎｅｃｋｅｒ等，ＥＭＢＯ Ｊ．，１０：３６５５－３６５６（１９９１）に開示されている。

所望の結合特異性（抗体－抗原結合部位）を有する抗体可変ドメインを免疫グロブリン定常ドメイン配列に融合できる。融合は、好ましくは少なくともヒンジ部、ＣＨ２及びＣＨ３領域の一部を含む免疫グロブリン重鎖定常ドメインとのものである。少なくとも一つの融合には軽鎖結合に必要な部位を含む第一の重鎖定常領域（ＣＨ１）が存在することが望ましい。免疫グロブリン重鎖融合をコードするＤＮＡ、及び望むのであれば免疫グロブリン軽鎖を、別々の発現ベクターに挿入し、適当な宿主生物に同時形質移入する。二重特異性抗体を作成するための更なる詳細については、例えばＳｕｒｅｓｈ等，Ｍｅｔｈｏｄｓ ｉｎ Ｅｎｚｙｍｏｌｏｇｙ，１２１：２１０（１９８６）を参照されたい。

【０２２３】

ＷＯ ９６／２７０１１に記載された他の方法によれば、一対の抗体分子間の界面を操作して組換え細胞培養から回収される異種二量体の割合を最大にすることができる。好適な界面は抗体定常ドメインのＣＨ３領域の少なくとも一部を含む。この方法では、第１抗体分子の界面からの一又は複数の小さいアミノ酸側鎖がより大きな側鎖（例えばチロシン又はトリプトファン）と置き換えられる。大きな側鎖と同じ又は類似のサイズの相補的「キャビティ」を、大きなアミノ酸側鎖を小さいもの（例えばアラニン又はスレオニン）と置き換えることにより第２の抗体分子の界面に作り出す。これにより、ホモダイマーのような不要の他の最終産物に対してヘテロダイマーの収量を増大させるメカニズムが提供される。

二重特異性抗体は、全長抗体又は抗体断片（例えば、Ｆ（ａｂ’）$_2$二重特異性抗体）として調製できる。抗体断片から二重特異性抗体を産生する技術もまた文献に記載されている。例えば、化学結合を使用して二重特異性抗体を調製することができる。Ｂｒｅｎｎａｎ等，Ｓｃｉｅｎｃｅ，２２９：８１（１９８５）は無傷の抗体をタンパク分解性に切断してＦ（ａｂ’）$_2$断片を産生する手順を記述している。これらの断片は、ジチオール錯体形成剤亜砒酸ナトリウムの存在下で還元して近接ジチオールを安定化させ、分子間ジスルフィド形成を防止する。産生されたＦａｂ’断片はついでチオニトロベンゾアート（ＴＮＢ）誘導体に転換される。Ｆａｂ’－ＴＮＢ誘導体の一つをついでメルカプトエチルアミンでの還元によりＦａｂ’－チオールに再転換し、他のＦａｂ’－ＴＮＢ誘導体の等モル量と混合して二重特異性抗体を形成する。作られた二重特異性抗体は酵素の選択的固定化用の薬剤として使用することができる。

大腸菌からＦａｂ’フラグメントを直接回収でき、これは化学的に結合して二重特異性抗体を形成することができる。Ｓｈａｌａｂｙ等，Ｊ．Ｅｘｐ．Ｍｅｄ．，１７５：２

１７－２２５（１９９２）は完全にヒト化された二重特異性抗体Ｆ（ａｂ’）$_2$分子の製造を記述している。各Ｆａｂ’フラグメントは大腸菌から別個に分泌され、インビトロで定方向化学共役を受けて二重特異性抗体を形成する。このようにして形成された二重特異性抗体は、正常なヒトＴ細胞及びＥｒｂＢ２レセプターを過剰発現する細胞に結合可能で、ヒト乳房腫瘍標的に対するヒト細胞障害性リンパ球の細胞溶解活性の誘因となる。

【０２２４】

組換え細胞培養から直接的に二重特異性抗体断片を作成し分離する様々な方法もまた記述されている。例えば、二重特異性抗体はロイシンジッパーを使用して生産されている。Ｋｏｓｔｅｌｎｙら，Ｊ．Ｉｍｍｕｎｏｌ．１４８（５）：１５４７－１５５３（１９９２）。Ｆｏｓ及びＪｕｎタンパク質からのロイシンジッパーペプチドを遺伝子融合により二つの異なった抗体のＦａｂ’部分に結合させる。抗体ホモダイマーをヒンジ領域で還元してモノマーを形成し、ついで再酸化して抗体ヘテロダイマーを形成する。この方法はまた抗体ホモダイマーの生産に対して使用することができる。Ｈｏｌｌｉｎｇｅｒら，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，９０：６４４４－６４４８（１９９３）により記述された「ダイアボディ」技術は二重特異性抗体断片を作成する別のメカニズムを提供した。断片は、同一鎖上の２つのドメイン間の対形成を可能にするには十分に短いリンカーにより軽鎖可変ドメイン（V_L）に重鎖可変ドメイン（V_H）を結合してなる。従って、一つの断片のV_H及びV_Lドメインは他の断片の相補的V_L及びV_Hドメインと強制的に対形成させられ、２つの抗原結合部位を形成する。単鎖Ｆｖ（ｓＦｖ）ダイマーの使用により二重特異性抗体断片を製造する他の方策もまた報告されている。Ｇｒｕｂｅｒら，Ｊ．Ｉｍｍｕｎｏｌ．１５２：５３６８（１９９４）を参照されたい。二価より多い抗体も考えられる。例えば、三重特異性抗体を調製することができる。Ｔｕｔｔら Ｊ．Ｉｍｍｕｎｏｌ．１４７：６０（１９９１）。

例示的な二重特異性抗体は、ここに与えられたＰＲＯポリペプチドの２つの異なるエピトープに結合しうる。あるいは、抗－ＰＲＯポリペプチドのアームは、特定のＰＲＯポリペプチド発現細胞に細胞防御メカニズムを集中させるように、Ｔ細胞レセプター分子（例えばＣＤ２、ＣＤ３、ＣＤ２８、又はＢ７）等の白血球上のトリガー分子又はＦｃγＲＩ（ＣＤ６４）、ＦｃγＲＩＩ（ＣＤ３２）及びＦｃγＲＩＩＩ（ＣＤ１６）等のＩｇＧ（ＦｃγＲ）に対するＦｃレセプターに結合するアームと結合しうる。また、二重特異性抗体は特定のＰＲＯポリペプチドを発現する細胞に細胞毒性薬を局在化させるためにも使用されうる。これらの抗体はＰＲＯ結合アーム及び細胞毒性薬又は放射性キレート化剤、例えばＥＯＴＵＢＥ、ＤＰＴＡ、ＤＯＴＡ、又はＴＥＴＡと結合するアームを有する。興味の対象となる他の二重特異性抗体はＰＲＯポリペプチドに結合し、そしてさらに組織因子（ＴＦ）に結合する。

【０２２５】

５．ヘテロ抱合体抗体

ヘテロ抱合抗体もまた本発明の範囲に入る。ヘテロ抱合抗体は、２つの共有結合した抗体からなる。このような抗体は、例えば、免疫系細胞を不要な細胞に対してターゲティングさせるため［米国特許第４，６７６，９８０号］及びＨＩＶ感染の治療のために［ＷＯ９１／００３６０；ＷＯ ９２／２００３７３；ＥＰ ０３０８９］提案されている。この抗体は、架橋剤に関連したものを含む合成タンパク化学における既知の方法を使用して、インビトロで調製することができると考えられる。例えば、ジスルフィド交換反応を使用するか又はチオエーテル結合を形成することにより、免疫毒素を作成することができる。この目的に対して好適な試薬の例には、イミノチオレート及びメチル－４－メルカプトブチリミデート、及び例えば米国特許第４，６７０，９８０号に開示されているものが含まれる。

６．エフェクター機能の加工

本発明の抗体をエフェクター機能について改変し、例えば癌治療における抗体の有効性を向上させるのが望ましい。例えば、システイン残基をＦｃ領域に導入し、それにより、この領域に鎖間ジスルフィド結合を形成させるようにしてもよい。そのようにして生成され

た同種二量体抗体は、向上した内部移行能力及び／又は増加した補体媒介細胞殺傷及び抗体－依存性細胞性細胞毒性（ＡＤＣＣ）を有しうる。Ｃａｒｏｎ等，Ｊ．Ｅｘｐ．Ｍｅｄ．１７６：１１９１－１１９５（１９９２）及びＳｈｏｐｅｓ，Ｂ．Ｊ．Ｉｍｍｕｎｏｌ．１４８：２９１８－２９２２（１９９２）参照。向上した抗腫瘍活性を持つ同種二量体抗体はまた、Ｗｏｌｆｆ等，Ｃａｎｃｅｒ ｒｅｓｅａｒｃｈ ５３：２５６０－２５６５（１９９３）に記載されたような異種二官能性架橋を用いても調製しうる。あるいは、抗体は、２つのＦｃ領域を有するように加工して、それにより補体溶解及びＡＤＣＣ能力を向上させることもできる。Ｓｔｅｖｅｎｓｏｎ等，Ａｎｔｉ－Ｃａｎｃｅｒ Ｄｒｕｇ Ｄｅｓｉｇｎ ３：２１９－２３０（１９８９）参照。

７．免疫複合体

本発明はまた、化学治療薬、毒素（例えば、細菌、真菌、植物又は動物由来の酵素活性毒素、又はその断片）などの細胞毒性薬、あるいは放射性同位体（即ち、放射性抱合）に抱合された抗体を含む免疫複合体にも関する。

【０２２６】

このような免疫複合体の生成に有用な化学治療薬は上記した。用いることのできる酵素活性毒素及びその断片は、ジフテリアＡ鎖、ジフテリア毒素の非結合活性断片、（緑膿菌からの）外毒素Ａ鎖、リシンＡ鎖、アブリンＡ鎖、モデクシン（ｍｏｄｅｃｃｉｎ）Ａ鎖、アルファ－サルシン、アレウリテス・フォーディ（Ａｌｅｕｒｉｔｅｓ ｆｏｒｄｉｉ）タンパク質、ジアンチン（ｄｉａｎｔｈｉｎ）タンパク質、フィトラカ・アメリカーナ（Ｐｈｙｔｏｌａｃａ ａｍｅｒｉｃａｎａ）タンパク質（ＰＡＰＩ、ＰＡＰＩＩ、及びＰＡＰ－Ｓ）、モモルディカ・チャランチア（ｍｏｍｏｒｄｉｃａ ｃｈａｒａｎｔｉａ）インヒビター、クルシン（ｃｕｒｃｉｎ）、クロチン（ｃｒｏｔｉｎ）、サパオナリア・オフィシナリス（ｓａｐａｏｎａｒｉａ ｏｆｉｃｉｎａｌｉｓ）インヒビター、ゲロニン（ｇｅｌｏｎｉｎ）、ミトゲリン（ｍｉｔｏｇｅｌｌｉｎ）、レストリクトシン（ｒｅｓｔｒｉｃｔｏｃｉｎ）、フェノマイシン（ｐｈｅｎｏｍｙｃｉｎ）、エノマイシン（ｅｎｏｍｙｃｉｎ）及びトリコテセン（ｔｒｉｃｏｔｈｅｃｅｎｅ）を含む。様々な放射性ヌクレオチドが放射性抱合抗体の生成に利用可能である。例として、^{212}Ｂｉ、^{131}Ｉ、^{131}Ｉｎ、^{90}Ｙ及び^{186}Ｒｅを含む。抗体及び細胞毒性薬の複合体は、種々の二官能性タンパク質カップリング剤、例えば、Ｎ－スクシンイミジル－３－（２－ピリジルジチオール）プロピオネート（ＳＰＤＰ）、イミノチオラン（ＩＴ）、イミドエステルの二官能性誘導体（ジメチルアジピミデートＨＣＬ等）、活性エステル（ジスクシンイミジルスベレート等）、アルデヒド（グルタルアルデヒド等）、ビス－アジド化合物（ビス（ｐ－アジドベンゾイル）ヘキサンジアミン等）、ビス－ジアゾニウム誘導体（ビス－（ｐ－ジアゾニウムベンゾイル）－エチレンジアミン等）、ジイソシアネート（トリエン２，６－ジイソシアネート等）、及びビス－活性フッ素化合物（１，５－ジフルオロ－２，４－ジニトロベンゼン等）を用いて作成できる。例えば、リシン免疫毒素は、Ｖｉｔｅｔｔａ等，Ｓｃｉｅｎｃｅ ２３８：１０９８（１９８７）に記載されたように調製することができる。カーボン－１４－標識１－－イソチオシアナトベンジル－３－メチルジエチレントリアミン五酢酸（ＭＸ－ＤＴＰＡ）は、放射性ヌクレオチドの抗体への抱合のためのキレート剤の例である。ＷＯ ９４／１１０２６参照。

他の実施態様では、腫瘍の予備標的化で使用するために、抗体は「レセプター」（ストレプトアビジン等）に抱合されてもよく、抗体－レセプター複合体は患者に投与され、次いで清澄化剤を用いて未結合複合体を循環から除去し、次に細胞毒性薬（例えば、放射性ヌクレオチド等）に抱合された「リガンド」（アビジン等）を投与する。

８．免疫リポソーム

また、ここに開示する抗体は、免疫リポソームとして調製してもよい。抗体を含むリポソームは、Ｅｐｓｔｅｉｎ等，Ｐｒｏｃ．Ｎａｔｌ．ａｃａｄ．Ｓｃｉ．ＵＳＡ，８２：３６８８（１９８５）；Ｈｗａｎｇ等，Ｐｒｏｃ．ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ，７７：４０３０（１９８０）；及び米国特許第４，４８５，０４５号及び第４，５４４，５４５号に記載されたような、この分野で知られた方法で調製され

る。向上した循環時間を持つリポソームは、米国特許第５，０１３，５５６号に開示されている。
特に有用なリポソームは、ホスファチジルコリン、コレステロール及びＰＥＧ－誘導ホスファチジルエタノールアミン（ＰＥＧ－ＰＥ）を含む脂質組成物での逆相蒸発法によって生成される。リポソームは、所定サイズのフィルターを通して押し出され、所望の径を有するリポソームが生成される。本発明の抗体のＦａｂ’断片は、Ｍａｒｔｉｎ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２５７：２８６－２８８（１９８２）に記載されているように、ジスルフィド交換反応を介してリポソームに抱合され得る。化学治療薬（ドキソルビシン等）は、場合によってはリポソーム内に包含される。Ｇａｂｉｚｏｎ等，Ｊ．Ｎａｔｉｏｎａｌ Ｃａｎｃｅｒ Ｉｎｓｔ．８１（１９）１４８４（１９８９）参照。

【０２２７】
９．抗体の製薬組成物
ここで同定されるＰＲＯポリペプチドに特異的に結合する抗体、並びに上記に開示したスクリーニングアッセイで同定された他の分子は、種々の疾患の治療のために、製薬組成物の形態で投与することができる。
ＰＲＯポリペプチドが細胞内であり、全抗体が阻害剤として用いられる場合、内在化抗体が好ましい。しかし、リポフェクション又はリポソームも抗体、又は抗体断片を細胞に導入するのに使用できる。抗体断片が用いられる場合、標的タンパク質の結合ドメインに特異的に結合する最小阻害断片が好ましい。例えば、抗体の可変領域配列に基づいて、標的タンパク質配列に結合する能力を保持したペプチド分子が設計できる。このようなペプチドは、化学的に合成でき、又は組換えＤＮＡ技術によって生成できる。例えば、Ｍａｒａｓｃｏ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ９０，７８８９－７８９３（１９９３）参照。
ここでの製剤は、治療すべき特定の徴候に必要な場合に１以上の活性化合物、好ましくは互いに悪影響を及ぼさない相補的活性を持つものも含んでよい。あるいは、又はそれに加えて、組成物は、細胞毒性薬、サイトカイン又は成長阻害剤を含んでもよい。これらの分子は、適切には、意図する目的に有効な量の組み合わせで存在する。
また、活性成分は、例えばコアセルベーション技術により又は界面重合により調製されたマイクロカプセル、例えば、各々ヒドロキシメチルセルロース又はゼラチン－マイクロカプセル及びポリ（メタクリル酸メチル）マイクロカプセル中、コロイド状薬物送達系（例えば、リポソーム、アルブミン小球、マイクロエマルション、ナノ粒子及びナノカプセル）中、又はマイクロエマルション中に包括されていてもよい。これらの技術は、上掲のＲｅｍｉｎｇｔｏｎ’ｓ Ｐｈａｒｍａｃｅｕｔｉｃａｌ Ｓｃｉｅｎｃｅに開示されている。
インビボ投与に使用される製剤は無菌でなけらばならない。これは、滅菌濾過膜を通した濾過により容易に達成される。
徐放性製剤を調製してもよい。徐放性製剤の好適な例は、抗体を含有する固体疎水性ポリマーの半透性マトリクスを含み、このマトリクスは成形された物品、例えばフィルム、又はマイクロカプセルの形状である。除放性マトリクスの例は、ポリエステルヒドロゲル（例えば、ポリ（２－ヒドロキシエチル－メタクリレート）又はポリ（ビニルアルコール））、ポリアクチド（米国特許第３，７７３，９１９号）、Ｌ－グルタミン酸及びγ－エチル－Ｌ－グルタメート、非分解性エチレン－酢酸ビニル、ＬＵＰＲＯＮ ＤＥＰＯＴ（商品名）（乳酸－グリコール酸コポリマーと酢酸リュープロリドの注射可能な小球）などの分解性乳酸－グリコール酸コポリマー、ポリ－（Ｄ）－３－ヒドロキシブチル酸を含む。エチレン－酢酸ビニル及び乳酸－グリコール酸などのポリマーは分子を１００日に渡って放出することができるが、ある種のヒドロゲルはより短時間でタンパク質を放出してしまう。カプセル化された抗体が身体内に長時間残ると、それらは３７℃の水分に露出されることにより変性又は凝集し、その結果、生物学的活性の低下及び起こりうる免疫原性の変化をもたらす。合理的な方法は、含まれる機構に依存する安定化について工夫することができる。例えば、凝集機構がチオ－ジスルフィド交換を通した分子間Ｓ－Ｓ結合形成であ

ると発見された場合、安定化はスルフヒドリル残基の修飾、酸性溶液からの凍結乾燥、水分含有量の制御、適切な添加剤の付加、及び特異的ポリマーマトリクス組成物の開発によって達成されうる。

【０２２８】

Ｇ．抗－ＰＲＯ抗体の用途

本発明の抗－ＰＲＯ抗体は様々な有用性を有している。例えば、抗－ＰＲＯ抗体は、ＰＲＯの診断アッセイ、例えばその特定細胞、組織、又は血清での発現の検出に用いられる。競合的結合アッセイ、直接又は間接サンドウィッチアッセイ及び不均一又は均一相で行われる免疫沈降アッセイ［Ｚｏｌａ，Ｍｏｎｏｃｌｏｎａｌ Ａｎｔｉｂｏｄｉｅｓ：Ａ Ｍａｎｕａｌ ｏｆ Ｔｅｃｈｎｉｑｕｅｓ，ＣＲＣ Ｐｒｅｓｓ，Ｉｎｃ．（１９８７）ｐｐ．１４７－１５８］等のこの分野で知られた種々の診断アッセイ技術が使用される。診断アッセイで用いられる抗体は、検出可能な部位で標識される。検出可能な部位は、直接又は間接に、検出可能なシグナルを発生しなければならない。例えば、検出可能な部位は、^{3}Ｈ、^{14}Ｃ、^{32}Ｐ、^{35}Ｓ又は^{125}Ｉ等の放射性同位体、フルオレセインイソチオシアネート、ローダミン又はルシフェリン等の蛍光又は化学発光化合物、あるいはアルカリホスファターゼ、β－ガラクトシダーゼ又はセイヨウワサビペルオキシダーゼ等の酵素であってよい。抗体に検出可能な部位を抱合させるためにこの分野で知られた任意の方法が用いられ、それにはＨｕｎｔｅｒ等 Ｎａｔｕｒｅ，１４４：９４５（１９６２）；Ｄａｖｉｄ等，Ｂｉｏｃｈｅｍｉｓｔｒｙ，１３：１０１４（１９７４）；Ｐａｉｎ等，Ｊ．Ｉｍｍｕｎｏｌ．Ｍｅｔｈ．，４０：２１９（１９８１）；及びＮｙｇｒｅｎ，Ｊ．Ｈｉｓｔｏｃｈｅｍ．ａｎｄ Ｃｙｔｏｃｈｅｍ．，３０：４０７（１９８２）に記載された方法が含まれる。

また、抗－ＰＲＯ抗体は組換え細胞培養又は天然供給源からのＰＲＯのアフィニティー精製にも有用である。この方法においては、ＰＲＯに対する抗体を、当該分野でよく知られている方法を使用して、セファデックス樹脂や濾紙のような適当な支持体に固定化する。次に、固定化された抗体を、精製するＰＲＯを含む試料と接触させた後、固定化された抗体に結合したＰＲＯ以外の試料中の物質を実質的に全て除去する適当な溶媒で支持体を洗浄する。最後に、ＰＲＯを抗体から脱離させる他の適当な溶媒で支持体を洗浄する。

以下の実施例は例示するためにのみ提供されるものであって、本発明の範囲を決して限定することを意図するものではない。

本明細書で引用した全ての特許及び参考文献の全体を、出典明示によりここに取り込む。

【０２２９】

（実施例）

実施例で言及されている全ての市販試薬は、特に示さない限りは製造者の使用説明に従い使用した。ＡＴＣＣ登録番号により以下の実施例及び明細書全体を通して特定されている細胞の供給源はアメリカン・タイプ・カルチャー・コレクション、ロックビル、メリーランドである。

実施例１：新規なポリペプチド及びそれをコードするｃＤＮＡを同定するための細胞外ドメイン相同性スクリーニング

Ｓｗｉｓｓ－Ｐｒｏｔ公的データベースからの約９５０の既知の分泌タンパク質からの細胞外ドメイン（ＥＣＤ）配列（あれば、分泌シグナル配列を含む）を、ＥＳＴデータベースの検索に使用した。ＥＳＴデータベースは、公的データベース（例えば、Ｄａｙｈｏｆｆ、ＧｅｎＢａｎｋ）及び企業のデータベース（例えば、ＬＩＦＥＳＥＱ（商品名）、Ｉｎｃｙｔｅ Ｐｈａｒｍａｃｅｕｔｉｃａｌｓ、Ｐａｌｏ Ａｌｔｏ，ＣＡ）を含む。検索は、コンピュータプログラムＢＬＡＳＴ又はＢＬＡＳＴ－２（Ａｌｔｓｃｈｕｌ等，Ｍｅｔｈｏｄｓ ｉｎ Ｅｎｚｙｍｏｌｏｇｙ ２６６：４６０－４８０（１９９６）；ｈｔｔｐ：／／ｂｌａｓｔ．ｗｕｓｔｌ／ｅｄｕ／ｂｌａｓｔ／ＲＥＡＤＭＥ．ｈｔｍｌ）を用いて、ＥＣＤタンパク質配列のＥＳＴ配列の６フレーム翻訳との比較として実施した。既知のタンパク質をコードせず、ＢＬＡＳＴスコア７０（９０の場合もある）又はそれ以上を持つ比較物は、プログラム「ｐｈｒａｐ」（Ｐｈｉｌ Ｇｒｅｅｎ，Ｕｎｉｖｅ

ｒｓｉｔｙ ｏｆ Ｗａｓｈｉｎｇｔｏｎ, Ｓｅａｔｔｌｅ, ＷＡ）で集団化してコンセンサスＤＮＡ配列に構築した。
この細胞外ドメイン相同性スクリーニングを用いて、ｐｈｒａｐを用いて他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を構築した。さらに、得られたコンセンサスＤＮＡ配列を、しばしば（常にではない）ＢＬＡＳＴ又はＢＬＡＳＴ－２及びｐｈｒａｐの繰り返しサイクルを用いて伸長し、コンセンサス配列を上で議論したＥＳＴ配列の供給源を用いて可能な限り伸長させた。
上記のように得られたコンセンサス配列に基づいて、次いでオリゴヌクレオチドを合成し、ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及びＰＲＯポリペプチドの全長コード化配列のクローンを単離するプローブとして用いるために使用した。正方向（．ｆ）及び逆方向（．ｒ）ＰＣＲプライマーは一般的に２０から３０ヌクレオチドの範囲であり、しばしば約１００－１０００ｂｐ長のＰＣＲ産物を与えるために設計される。プローブ（．ｐ）配列は、典型的に４０－５５ｂｐ長である。或る場合には、コンセンサス配列が約１－１．５ｋｂｐより大きいときに付加的なオリゴヌクレオチドが合成される。全長クローンについて幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを、Ａｕｓｕｂｅｌ等, Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙ, のように、ＰＣＲプライマー対でのＰＣＲによりスクリーニングした。ポジティブライブラリを、次いで、プローブオリゴヌクレオチド及びプライマー対の一方を用いて対象とする遺伝子をコードするクローンの単離するのに使用した。
ｃＤＮＡクローンの単離に用いたｃＤＮＡライブラリは、Ｉｎｖｉｔｒｏｇｅｎ, Ｓａｎ Ｄｉｅｇｏ, ＣＡからのもの等の市販試薬を用いて標準的な方法によって作成した。ｃＤＮＡは、ＮｏｔＩ部位を含むオリゴｄＴでプライムし、平滑末端でＳａｌＩヘミキナーゼアダプターに結合させ、ＮｏｔＩで切断し、ゲル電気泳動でおよそのサイズ分類し、そして適切なクローニングベクター（ｐＲＫ５Ｂ又はｐＲＫ５Ｄ等；ｐＲＫ５ＢはＳｆｉＩ部位を含まないｐＲＫ５Ｄの前駆体である；Ｈｏｌｍｅｓ等, Ｓｃｉｅｎｃｅ, ２５３：１２７８－１２８０（１９９１）参照）に、独特のＸｈｏＩ及びＮｏｔＩ部位において、所定の方向でクローニングした。
【０２３０】
実施例２：ヒトＰＲＯ２１１及びＰＲＯ２１７をコードするｃＤＮＡクローンの単離
上記実施例１に記載したようにしてコンセンサスＤＮＡ配列を組み立て、それぞれＤＮＡ２８７３０とＤＮＡ２８７６０と命名した。これらのコンセンサス配列に基づいて、オリゴヌクレオチドを合成し、ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及びＰＲＯ２１１及びＰＲＯ２１７ポリペプチドの全長コード化配列のクローンを単離するプローブとして使用した。ＤＮＡ３２２９２－１１３１及びＤＮＡ３３０９４－１１３１を単離するために使用したライブラリは胎児肺ライブラリであった。
ｃＤＮＡクローンはその全体を配列決定した。ＰＲＯ２１１（ＤＮＡ３２２９２－１１３１）及びＰＲＯ２１７（ＵＮＱ１９１）の全ヌクレオチド配列をＦｉｇ１（配列番号：１）とＦｉｇ３（配列番号：３）にそれぞれ示す。予測されるポリペプチドはそれぞれ３５３及び３７９アミノ酸長であり、それぞれ約３８，１９０及び４１，５２０ダルトンの分子量を有する。
上記の手順において使用されたオリゴヌクレオチド配列は次の通りであった：
２８７３０．ｐ（ＯＬＩ ５１６）（配列番号：５）
５’－ＡＧＧＧＡＧＣＡＣＧＧＡＣＡＧＴＧＴＧＣＡＧＡＴＧＴＧＧＡＣＧＡＧＴＧＣＴＣＡＣＴＡＧＣＡ－３’
２８７３０．ｆ（ＯＬＩ ５１７）（配列番号：６）
５’－ＡＧＡＧＴＧＴＡＴＣＴＣＴＧＧＣＴＡＣＧＣ－３’
２８７３０．ｒ（ＯＬＩ ５１８）（配列番号：７）
５’－ＴＡＡＧＴＣＣＧＧＣＡＣＡＴＴＡＣＡＧＧＴＣ－３’
２８７６０．ｐ（ＯＬＩ ６１７）（配列番号：８）

５’－ＣＣＣＡＣＧＡＴＧＴＡＴＧＡＡＴＧＧＴＧＧＡＣＴＴＴＧＴＧＴＧＡＣＴＣＣＴＧＧＴＴＴＣＴＧＣＡＴＣ－３’
２８７６０．ｆ（ＯＬＩ　６１８）（配列番号：９）
５’－ＡＡＡＧＡＣＧＣＡＴＣＴＧＣＧＡＧＴＧＴＣＣ－３’
２８７６０．ｒ（ＯＬＩ　６１９）（配列番号：１０）
５’－ＴＧＣＴＧＡＴＴＴＣＡＣＡＣＴＧＣＴＣＴＣＣＣ－３’

【０２３１】

実施例３：ヒトＰＲＯ２３０をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列をここでＤＮＡ３０８５７と命名した。ジェネンテクの私的なＥＳＴをコンセンサスアセンブリに用いた。ＥＳＴはＤＮＡ２００８８と命名され、Ｆｉｇ７に示されるようなヌクレオチド配列を有している（配列番号：１３）。

ＤＮＡ３０８５７コンセンサス配列に基づいて、ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及びＰＲＯ２３０の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー　５’－ＴＴＣＧＡＧＧＣＣＴＣＴＧＡＧＡＡＧＴＧＧＣＣＣ－３’（配列番号：１４）

逆方向ＰＣＲプライマー　５’－ＧＧＣＧＧＴＡＴＣＴＣＴＣＴＧＧＣＣＴＣＣＣ－３’（配列番号：１５）

さらに、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０８５７配列から作成した。

ハイブリッド形成プローブ

５’－ＴＴＣＴＣＣＡＣＡＧＣＡＧＣＴＧＴＧＧＣＡＴＣＣＧＡＴＣＧＴＧＴＣＴＣＡＡＴＣＣＡＴＴＣＴＣＴＧＧＧ－３’（配列番号：１６）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２３０遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリー作成のためのＲＮＡは、ヒト胎児肺組織から単離した。上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２３０の全長ＤＮＡ配列（ここではＤＮＡ３３２２３－１１３６と命名）及びＰＲＯ２３０の誘導タンパク質配列が得られた。

ＤＮＡ３３２２３－１１３６の全ヌクレオチド配列をＦｉｇ５（配列番号：１１）に示す。クローンＤＮＡ３３２２３－１１３６は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１００－１０３に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１５０１－１５０４の停止コドンで終端する（Ｆｉｇ５；配列番号：１１）。予測されるポリペプチド前駆体は４６７アミノ酸長である（Ｆｉｇ６）。

【０２３２】

実施例４：ヒトＰＲＯ２３２をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列をここでＤＮＡ３０９３５と命名した。ＤＮＡ３０９３５コンセンサス配列に基づいて、ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及びＰＲＯ２３２の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー　５’－ＴＧＣＴＧＴＧＣＴＡＣＴＣＣＴＧＣＡＡＡＧＣＣＣ－３’（配列番号：１９）

逆方向ＰＣＲプライマー　５’－ＴＧＣＡＣＡＡＧＴＣＧＧＴＧＴＣＡＣＡＧＣＡＣＧ－

３’（配列番号：２０）
さらに、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０９３５配列から作成した。
ハイブリッド形成プローブ
５’－ＡＧＣＡＡＣＧＡＧＧＡＣＴＧＣＣＴＧＣＡＧＧＴＧＧＡＧＡＡＣＴＧＣＡＣＣＣＡＧＣＴＧＧＧ－３’（配列番号：２１）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２３２遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリー作成のためのＲＮＡは、ヒト胎児腎臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２３２の全長ＤＮＡ配列［ここではＤＮＡ３４４３５－１１４０と命名］及びＰＲＯ２３０の誘導タンパク質配列が得られた。
ＤＮＡ３４４３５－１１４０の全ヌクレオチド配列をＦｉｇ８（配列番号：１７）に示す。クローンＤＮＡ３４４３５－１１４０は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１７－１９に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置３５９－３６１の停止コドンで終端する（Ｆｉｇ８；配列番号：１７）。予測されるポリペプチド前駆体は１１４アミノ酸長である（Ｆｉｇ９）。クローンＤＮＡ３４４３５－１１４０は１９９７年９月１６日にＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９２５０が付与されている。
全長ＰＲＯ２３２のアミノ酸配列の分析により、ニワトリ由来の幹細胞表面抗原と３５％の配列同一性を有していることが示唆される。
【０２３３】
実施例５：ヒトＰＲＯ１８７をコードするｃＤＮＡクローンの単離
企業の発現配列タグ（ＥＳＴ）ＤＮＡデータベース（ＬＩＦＥＳＥＱ^{TM}、Ｉｎｃｙｔｅ Ｐｈａｒｍａｃｅｕｔｉｃａｌｓ、Ｐａｌｏ Ａｌｔｏ，ＣＡ）を検索して、アンドロゲン誘発成長因子としても知られている線維芽細胞成長因子（ＦＧＦ－８）と相同性を示したＥＳＴ（＃８４３１９３）を同定した。ｍＲＮＡはＩｎｖｉｔｒｏｇｅｎ，Ｓａｎ Ｄｉｅｇｏ，ＣＡからの試薬とプロトコールを使用してヒト胎児肺組織から単離した（ファーストトラック２）。ｃＤＮＡクローンを単離するために使用したｃＤＮＡライブラリは、商業的に入手できる試薬（例えば、Ｉｎｖｉｔｒｏｇｅｎ，Ｓａｎ Ｄｉｅｇｏ，ＣＡ，Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ，Ｇａｉｔｈｅｒｓｂｕｒｇ，ＭＤ）を使用する標準的な方法により作成した。ｃＤＮＡは、ＮｏｔＩ部位を含むオリゴｄＴでプライムし、平滑末端でＳａｌＩヘミキナーゼアダプターに結合させ、ＮｏｔＩで切断し、ゲル電気泳動で適当にサイズ分類し、そしてＬｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ，Ｇａｉｔｈｅｒｓｂｕｒｇ，ＭＤからの試薬とプロトコール（Ｓｕｐｅｒ Ｓｃｒｉｐｔ Ｐｌａｓｍｉｄ Ｓｙｓｔｅｍ）を使用してクローニングベクターｐＲＫ５Ｄに定まった方向でクローニングした。二本鎖ｃＤＮＡを１０００ｂｐより大きくなるようにサイズ調整し、ＳａｌＩ／ＮｏｔＩ結合ｃＤＮＡをＸｈｏＩ／ＮｏｔＩ切断ベクター中にクローニングした。ｐＲＫ５Ｄはｓｐ６転写開始部位に続いてＳｆｉＩ制限酵素部位と続いてＸｈｏＩ／ＮｏｔＩｃＤＮＡクローニング部位があるクローニングベクターである。
様々な組織供給源からの幾つかのライブラリを次のオリゴヌクレオチドプローブでのＰＣＲ増幅によりスクリーニングした：
ＩＮ８４３１９３．ｆ（ＯＬＩ ３１５）（配列番号：２４）
５’－ＣＡＧＴＡＣＧＴＧＡＧＧＧＡＣＣＡＧＧＧＣＧＣＣＡＴＧＡ－３’
ＩＮ８４３１９３．ｒ（ＯＬＩ ３１７）（配列番号：２５）
５’－ＣＣＧＧＴＧＡＣＣＴＧＣＡＣＧＴＧＣＴＴＧＣＣＡ－３’
次いで、ポジティブライブラリを、上記のオリゴヌクレオチドの一方と次のオリゴヌクレ

オチドプローブを用いてＰＲＯ１８７遺伝子をコードするクローンを単離するのに使用した。
ＩＮ８４３１９３．ｐ（ＯＬＩ　３１６）（配列番号：２６）
５’－ＧＣＧＧＡＴＣＴＧＣＣＧＣＣＴＧＣＴＣＡＮＣＴＧＧＴＣＧＧＴＣＡＴＧＧＣＧＣＣＣＴ－３’
ｃＤＮＡクローンはその全体を配列決定した。ＰＲＯ１８７（ＤＮＡ２７８６４－１１５５）の全ヌクレオチド配列をＦｉｇ１０（配列番号：２２）に示す。クローンＤＮＡ２７８６４－１１５５は単一のオープンリーディングフレームを含み、ヌクレオチド位置１に見かけの翻訳開始部位を持つ（Ｆｉｇ１０；配列番号：２２）。予測されるポリペプチドは２０５アミノ酸長である。クローンＤＮＡ２７８６４－１１５５はＡＴＣＣに寄託され（命名：ＤＮＡ２７８６４－１１５５）、ＡＴＣＣ寄託番号２０９３７５が付与された。全長配列の（ＡＬＩＧＮコンピュータプログラムを使用する）ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメント分析法に基づいて、ＰＲＯ１８７は、ヒト線維芽細胞成長因子－８（アンドロゲン誘発成長因子）と７４％のアミノ酸配列同一性（Ｂｌａｓｔスコア３１０）を示した。
【０２３４】
実施例６：ヒトＰＲＯ２６５をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列を、ここでＤＮＡ３３６７９と命名する。ＤＮＡ３３６７９コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２６５の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
ＰＣＲプライマー（２つの正及び１つの逆）を合成した：
正方向ＰＣＲプライマーＡ：５’－ＣＧＧＴＣＴＡＣＣＴＧＴＡＴＧＧＣＡＡＣＣ－３’（配列番号：２９）
正方向ＰＣＲプライマーＢ：５’－ＧＣＡＧＧＡＣＡＡＣＣＡＧＡＴＡＡＡＣＣＡＣ－３’（配列番号：３０）
逆方向ＰＣＲプライマー　５’－ＡＣＧＣＡＧＡＴＴＴＧＡＧＡＡＧＧＣＴＧＴＣ－３’（配列番号：３１）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３３６７９配列から作成した。
ハイブリッド形成プローブ
５’－ＴＴＣＡＣＧＧＧＣＴＧＣＴＣＴＴＧＣＣＣＡＧＣＴＣＴＴＧＡＡＧＣＴＴＧＡＡＧＡＧＣＴＧＣＡＣ－３’（配列番号：３２）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２６５遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児脳ライブラリから単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２６５に対する全長ＤＮＡ配列［ここでＤＮＡ３６３５０－１１５８と命名］（配列番号：２７）及びＰＲＯ２６５の誘導タンパク質配列が得られた。
ＤＮＡ３６３５０－１１５８の全コード配列をＦｉｇ１２（配列番号：２７）に示す。クローンＤＮＡ３６３５０－１１５８は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３５２－３５４に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置２３３２－２３３４の停止コドンで終端する（Ｆｉｇ１２）。予測されるポリペプチド前駆体は６６０アミノ酸長である（Ｆｉｇ１３）。クローンＤＮＡ３６３５０－１１５８はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３７８が付与された。
全長ＰＲＯ２６５ポリペプチドのアミノ酸配列の分析により、その部分がフィブロモジュ

リン及びフィブロモジュリン前駆体に対して有意な相同性を有しており、よってＰＲＯ２６５を、特にフィブロモジュリンに関するロイシンリッチ反復ファミリーの新規なメンバーでありうることを示していることが示唆される。

【０２３５】

実施例７：ヒトＰＲＯ２１９をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列を、ここでＤＮＡ２８７２９と命名する。ＤＮＡ２８７２９コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２１９の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー　５’－ＧＴＧＡＣＣＣＴＧＧＴＴＧＴＧＡＡＴＡＣＴＣＣ－３’（配列番号：３５）

逆方向ＰＣＲプライマー　５’－ＡＣＡＧＣＣＡＴＧＧＴＣＴＡＴＡＧＣＴＴＧＧ－３’（配列番号：３６）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ２８７２９配列から作成した。

ハイブリッド形成プローブ

５’－ＧＣＣＴＧＴＣＡＧＴＧＴＣＣＴＧＡＧＧＧＡＣＡＣＧＴＧＣＴＣＣＧＣＡＧＣＧＡＴＧＧＧＡＡＧ－３’（配列番号：３７）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２１９遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２１９に対する全長ＤＮＡ配列［ここでＤＮＡ３２２９０－１１６４と命名］（配列番号：３３）及びＰＲＯ２１９の誘導タンパク質配列が得られた。

ＤＮＡ３２２９０－１１６４の全コード配列をＦｉｇ１４（配列番号：３３）に示す。クローンＤＮＡ３２２９０－１１６４は、単一のオープンリーディングフレームを含み、ヌクレオチド位置２０４－２０６に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置２９４９－２９５１の停止コドンで終端する（Ｆｉｇ１４）。予測されるポリペプチド前駆体は９１５アミノ酸長である（Ｆｉｇ１５）。クローンＤＮＡ３２２９０－１１６４はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３８４が付与された。

全長ＰＲＯ２１９ポリペプチドのアミノ酸配列の分析により、その部分がマウス及びヒトのマトリリン－２前駆体と有意な相同性を有していることが示唆される。

【０２３６】

実施例８：ヒトＰＲＯ２４６をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列を、ここでＤＮＡ３０９５５と命名する。ＤＮＡ３０９５５コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２４６の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー　５’－ＡＧＧＧＴＣＴＣＣＡＧＧＡＧＡＡＡＧＡＣＴＣ－３’（配列番号：４０）

逆方向ＰＣＲプライマー　５’－ＡＴＴＧＴＧＧＧＣＣＴＴＧＣＡＧＡＣＡＴＡＧＡＣ－３’（配列番号：４１）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持

つコンセンサスＤＮＡ３０９５５配列から作成した。
ハイブリッド形成プローブ
５’－ＧＧＣＣＡＣＡＧＣＡＴＣＡＡＡＡＣＣＴＴＡＧＡＡＣＴＣＡＡＴＧＴＡＣＴＧＧＴＴＣＣＴＣＣＡＧＣＴＣＣ－３’（配列番号：４２）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２４６遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２４６に対する全長ＤＮＡ配列［ここでＤＮＡ３５６３９－１１７２と命名］（配列番号：３８）及びＰＲＯ２４６の誘導タンパク質配列が得られた。
ＤＮＡ３５６３９－１１７２の全コード配列をＦｉｇ１６（配列番号：３８）に示す。クローンＤＮＡ３５６３９－１１７２は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１２６－１２８に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１２９６－１２９８の停止コドンで終端する（Ｆｉｇ１６）。予測されるポリペプチド前駆体は３９０アミノ酸長である（Ｆｉｇ１７）。クローンＤＮＡ３５６３９－１１７２はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３９６が付与された。
全長ＰＲＯ２４６ポリペプチドのアミノ酸配列の分析により、それがヒトの細胞表面タンパク質ＨＣＡＲと有意な相同性を有しており、よってＰＲＯ２４６は新規な細胞表面ウィルスレセプターでありうることが示される。
【０２３７】
実施例９：ヒトＰＲＯ２２８をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列を、ここでＤＮＡ２８７５８と命名する。コンセンサス構築にジェネンテクの企業のＥＳＴを用いた。このＥＳＴはＦｉｇ２０（配列番号：５０）に示され、ここでＤＮＡ２１９５１と命名される。
ＤＮＡ２８７５８コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２２８の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
ＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＧＧＴＡＡＴＧＡＧＣＴＣＣＡＴＴＡＣＡＧ－３’（配列番号：５１）
正方向ＰＣＲプライマー　５’－ＧＧＡＧＴＡＧＡＡＡＧＣＧＣＡＴＧＧ－３’（配列番号：５２）
正方向ＰＣＲプライマー　５’－ＣＡＣＣＴＧＡＴＡＣＣＡＴＧＡＡＴＧＧＣＡＧ－３’（配列番号：５３）
逆方向ＰＣＲプライマー　５’－ＣＧＡＧＣＴＣＧＡＡＴＴＡＡＴＴＣＧ－３’（配列番号：５４）
逆方向ＰＣＲプライマー　５’－ＧＧＡＴＣＴＣＣＴＧＡＧＣＴＣＡＧＧ－３’（配列番号：５５）
逆方向ＰＣＲプライマー　５’－ＣＣＴＡＧＴＴＧＡＧＴＧＡＴＣＣＴＴＧＴＡＡＧ－３’（配列番号：５６）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ２８７５８配列から作成した。
ハイブリッド形成プローブ
５’－ＡＴＧＡＧＡＣＣＣＡＣＡＣＣＴＣＡＴＧＣＣＧＣＴＧＴＡＡＴＣＡＣＣＴＧＡＣＡＣＡＴＴＴＴＧＣＡＡＴＴ－３’（配列番号：５７）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブ

ラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２２８遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２２８に対する全長ＤＮＡ配列［ここでＤＮＡ３３０９２－１２０２と命名］（配列番号：４８）及びＰＲＯ２２８の誘導タンパク質配列が得られた。

ＤＮＡ３３０９２－１２０２の全コード配列をＦｉｇ１８（配列番号：４８）に示す。クローンＤＮＡ３３０９２－１２０２は、単一のオープンリーディングフレームを含み、配列番号：４８のヌクレオチド位置２４－２６に見かけの翻訳開始部位を持ち、そして配列番号：４８のヌクレオチド位置２０９３の後の停止コドンで終端する。予測されるポリペプチド前駆体は６９０アミノ酸長である（Ｆｉｇ１９）。クローンＤＮＡ３３０９２－１２０２はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４２０が付与された。

全長ＰＲＯ２２８ポリペプチドのアミノ酸配列の分析により、その部分がセクレチン関連タンパク質ＣＤ９７及びＥＭＲ１並びにセクレチンメンバー、ラトロフィリンと有意な相同性を有しており、よってＰＲＯ２２８はセクレチン関連タンパク質の新規なメンバーでありうることが示される。

【０２３８】

実施例１０：ヒトＰＲＯ５３３をコードするｃＤＮＡクローンの単離

ＥＳＴ配列寄託番号ＡＦ００７２６８であるマウス線維芽細胞成長因子（ＦＧＦ－１５）を様々な公的ＥＳＴデータベース（例えば、ＧｅｎＢａｎｋ，　Ｄａｙｈｏｆｆ等）を検索するのに使用した。検索は、コンピュータプログラムＢＬＡＳＴ又はＢＬＡＳＴ－２［Ａｌｔｓｃｈｕｌ等，Ｍｅｔｈｏｄｓ　ｉｎ　Ｅｎｚｙｍｏｌｏｇｙ　２６６：４６０－４８０（１９９６）；ｈｔｔｐ：／／ｂｌａｓｔ．ｗｕｓｔｌ／ｅｄｕ／ｂｌａｓｔ／ＲＥＡＤＭＥ．ｈｔｍｌ）を用いて、ＥＣＤタンパク質配列のＥＳＴ配列の６フレーム翻訳との比較として実施した。検索の結果ＧｅｎＢａｎｋＥＳＴ　ＡＡ２２０９９４がヒットし、ストラタジーンＮＴ２ニューロン前駆体９３７２３０として同定された。

ＧｅｎＢａｎｋＥＳＴ　ＡＡ２２０９９４配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。正及び逆ＰＣＲプライマーは２０～３０のヌクレオチド（典型的には約２４）の範囲とでき、１００－１０００ｂｐの長さのＰＣＲ産物が得られるように設計される。プローブ配列は典型的には４０－５５ｂｐ（典型的には約５０）の長さである。全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡをＰＣＲプライマー対で、Ａｕｓｕｂｅｌ等，Ｃｕｒｒｅｎｔ　Ｐｒｏｔｏｃｏｌｓ　ｉｎ　Ｍｏｌｅｃｕｌａｒ　Ｂｉｏｌｏｇｙに従って、ＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いて対象の遺伝子をコードするクローンを単離するのに使用した。

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ５３３遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児網膜から単離した。ｃＤＮＡクローンの単離に用いたｃＤＮＡライブラリは、市販試薬（例えばＩｎｖｉｔｒｏｇｅｎ，　Ｓａｎ　Ｄｉｅｇｏ，　ＣＡ；Ｃｌｏｎｔｅｃｈ等々）を用いて標準的な方法によって作成した。ｃＤＮＡは、ＮｏｔＩ部位を含むオリゴｄＴでプライムし、平滑末端でＳａｌＩヘミキナーゼアダプターに結合させ、ＮｏｔＩで切断し、ゲル電気泳動でおよそのサイズ分類し、そして適切なクローニングベクター（例えばｐＲＫＢ又はｐＲＫＤ等；ｐＲＫ５Ｂは

ＳｆｉＩ部位を含まないｐＲＫ５Ｄの前駆体である；Ｈｏｌｍｅｓ等，Ｓｃｉｅｎｃｅ，２５３：１２７８－１２８０（１９９１）参照）に、独特のＸｈｏＩ及びＮｏｔＩ部位において、所定の方向でクローニングした。
ｃＤＮＡクローンはその全体を配列決定した。ＰＲＯ５３３の全長ヌクレオチド列はＦｉｇ２１（配列番号：５８）に示す。クローンＤＮＡ４９４３５－１２１９は単一のオープンリーディングフレームを含み、ヌクレオチド位置４５９－４６１に見かけの翻訳開始部位を持つ（Ｆｉｇ２１；配列番号：５８）。予測されるポリペプチド前駆体は２１６アミノ酸長である。クローンＤＮＡ４７４１２－１２１９はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４８０が付与された。
全長配列のＢＬＡＳＴ－２及びＦａｓｔＡ配列アラインメント分析に基づくと、ＰＲＯ５３３は線維芽細胞成長因子とアミノ酸配列同一性を示す（５３％）。上記の手順で使用したオリゴヌクレオチド配列は次のものであった：
ＦＧＦ１５．正：５’－ＡＴＣＣＧＣＣＣＡＧＡＴＧＧＣＴＡＣＡＡＴＧＴＧＴＡ－３’（配列番号：６０）；
ＦＧＦ１５．プローブ：５’－ＧＣＣＴＣＣＣＧＧＴＣＴＣＣＣＴＧＡＧＣＡＧＴＧＣＣＡＡＡＣＡＧＣＧＧＣＡＧＴＧＴＡ－３’（配列番号：６１）：
ＦＧＦ１５．逆：５’－ＣＣＡＧＴＣＣＧＧＴＧＡＣＡＡＧＣＣＣＡＡＡ－３’（配列番号：６２）。
【０２３９】
実施例１１：ヒトＰＲＯ２４５をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列を、ここでＤＮＡ３０９５４と命名した。
ＤＮＡ３０９５４コンセンサス配列に基づいて、ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及びＰＲＯ２４５の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＡＴＣＧＴＴＧＴＧＡＡＧＴＴＡＧＴＧＣＣＣＣ－３’（配列番号：６５）
逆方向ＰＣＲプライマー　５’－ＡＣＣＴＧＣＧＡＴＡＴＣＣＡＡＣＡＧＡＡＴＴＧ－３’（配列番号：６６）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０９５４配列から作成した。
ハイブリッド形成プローブ
５’－ＧＧＡＡＧＡＧＧＡＴＡＣＡＧＴＣＡＣＴＣＴＧＧＡＡＧＴＡＴＴＡＧＴＧＧＣＴＣＣＡＧＣＡＧＴＴＣＣ－３’（配列番号：６７）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２４５遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２４５に対する全長ＤＮＡ配列［ここでＤＮＡ３５６３８－１１４１と命名］及びＰＲＯ２４５の誘導タンパク質配列が得られた。
ＤＮＡ３５６３８－１１４１の全コード配列をＦｉｇ２３（配列番号：６３）に示す。クローンＤＮＡ３５６３８－１１４１は、単一のオープンリーディングフレームを含み、ヌクレオチド位置８９－９１に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１０２５－１０２７の停止コドンで終端する（Ｆｉｇ２３；配列番号：６３）。予測されるポリペプチド前駆体は３１２アミノ酸長である（Ｆｉｇ２４）。クローンＤＮＡ３５６３８－１１４１は１９９７年９月１６日ＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９２６５

が付与された。
全長ＰＲＯ２４５のアミノ酸配列の分析により、その部分がヒトのｃ－ｍｙｂタンパク質と６０％のアミノ酸同一性を有しており、よって膜貫通タンパク質レセプターチロシンキナーゼファミリーの新しいメンバーであることが示唆される。
【０２４０】
実施例１２：ヒトＰＲＯ２２０、ＰＲＯ２２１及びＰＲＯ２２７をコードするｃＤＮＡクローンの単離
（ａ）ＰＲＯ２２０
上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列を、ここでＤＮＡ２８７４９と命名した。
ＤＮＡ２８７４９コンセンサス配列に基づいて、ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及びＰＲＯ２２０の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＴＣＡＣＣＴＧＧＡＧＣＣＴＴＴＡＴＴＧＧＣＣ－３’（配列番号：７４）
逆方向ＰＣＲプライマー　５’－ＡＴＡＣＣＡＧＣＴＡＴＡＡＣＣＡＧＧＣＴＧＣＧ－３’（配列番号：７５）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ２８７４９配列から作成した。
ハイブリッド形成プローブ
５’－ＣＡＡＣＡＧＴＡＡＧＴＧＧＴＴＴＧＡＴＧＣＴＣＴＴＣＣＡＡＡＴＣＴＡＧＡＧＡＴＴＣＴＧＡＴＧＡＴＴＧ－３’（配列番号：７６）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２２０遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２２０に対する全長ＤＮＡ配列［ここでＤＮＡ３２２９８－１１３２と命名］及びＰＲＯ２２０の誘導タンパク質配列が得られた。
ＤＮＡ３２２９８－１１３２の全コード配列をＦｉｇ２５（配列番号：６８）に示す。クローンＤＮＡ３２２９８－１１３２は、単一のオープンリーディングフレームを含み、ヌクレオチド位置４８０－４８２に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置２６０４－２６０６の停止コドンで終端する（Ｆｉｇ２５）。予測されるポリペプチド前駆体は７０８アミノ酸長である（Ｆｉｇ２６）。クローンＤＮＡ３２２９８－１１３２はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９２５７が付与された。
全長ＰＲＯ２２０のアミノ酸配列の分析により、それが、ロイシンリッチ反復タンパク質及びニューロンのロイシンリッチ反復タンパク質１を含むロイシンリッチ反復タンパク質スーパーファミリーのメンバーと相同性を有していることが示される。
（ｂ）ＰＲＯ２２１
上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列を、ここでＤＮＡ２８７５６と命名した。
ＤＮＡ２８７５６コンセンサス配列に基づいて、ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及びＰＲＯ２２１の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＣＣＡＴＧＴＧＴＣＴＣＣＴＣＣＴＡＣＡＡＡＧ－３’（配列番号：７７）

逆方向ＰＣＲプライマー　５’－ＧＧＧＡＡＴＡＧＡＴＧＴＧＡＴＣＴＧＡＴＴＧＧ－３’（配列番号：７８）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ２８７５６配列から作成した。
ハイブリッド形成プローブ
５’－ＣＡＣＣＴＧＴＡＧＣＡＡＴＧＣＡＡＡＴＣＴＣＡＡＧＧＡＡＡＴＡＣＣＴＡＧＡＧＡＴＣＴＴＣＣＴＣＣＴＧ－３’（配列番号：７９）　全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２２１遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２２１に対する全長ＤＮＡ配列［ここでＤＮＡ３３０８９－１１３２と命名］及びＰＲＯ２２１の誘導タンパク質配列が得られた。
ＤＮＡ３３０８９－１１３２の全コード配列をＦｉｇ２７（配列番号：７０）に示す。クローンＤＮＡ３３０８９－１１３２は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１７９－１８１に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置９５６－９５８の停止コドンで終端する（Ｆｉｇ２７）。予測されるポリペプチド前駆体は２５９アミノ酸長である（Ｆｉｇ２８）。ＰＲＯ２２１はアミノ酸２０６－２２５に膜貫通領域を有すると考えられる。クローンＤＮＡ３３０８９－１１３２はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９２６２が付与された。
全長ＰＲＯ２２１のアミノ酸配列の分析により、それが、ＳＬＩＴタンパク質を含むロイシンリッチ反復タンパク質スーパーファミリーのメンバーと相同性と有することが示される。
（ｃ）ＰＲＯ２２７
上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列を、ここでＤＮＡ２８７４０と命名した。
ＤＮＡ２８７４０コンセンサス配列に基づいて、ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及びＰＲＯ２２７の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＡＧＣＡＡＣＣＧＣＣＴＧＡＡＧＣＴＣＡＴＣＣ－３’（配列番号：８０）
逆方向ＰＣＲプライマー　５’－ＡＡＧＧＣＧＣＧＧＴＧＡＡＡＧＡＴＧＴＡＧＡＣＧ－３’（配列番号：８１）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ２８７４０配列から作成した。
ハイブリッド形成プローブ
５’－ＧＡＣＴＡＣＡＴＧＴＴＴＣＡＧＧＡＣＣＴＧＴＡＣＡＡＣＣＴＣＡＡＧＴＣＡＣＴＧＧＡＧＧＴＴＧＧＣＧＡ－３’（配列番号：８２）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２２７遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２２７に対する全長ＤＮＡ配列［ここでＤＮＡ３３７８６－１１３２と命名］及びＰＲＯ２２７の誘導タンパク質配列が得られた。

ＤＮＡ３３７８６－１１３２の全コード配列をＦｉｇ２９（配列番号：７２）に示す。クローンＤＮＡ３３７８６－１１３２は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３３－３５に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１８９３－１８９５の停止コドンで終端する（Ｆｉｇ２９）。予測されるポリペプチド前駆体は６２０アミノ酸長である（Ｆｉｇ３０）。ＰＲＯ２２７は膜貫通ドメインを有していると考えられる。クローンＤＮＡ３３７８６－１１３２はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９２５３が付与された。

全長ＰＲＯ２２７のアミノ酸配列の分析により、それが、血小板糖タンパク質Ｖ前駆体及びヒト糖タンパク質Ｖを含むロイシンリッチ反復タンパク質スーパーファミリーのメンバーと相同性と有することが示される。

【０２４１】

実施例１３：ヒトＰＲＯ２５８をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列を、ここでＤＮＡ２８７４６と命名する。ＤＮＡ２８７４６コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２５８の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

ＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー　５’－ＧＣＴＡＧＧＡＡＴＴＣＣＡＣＡＧＡＡＧＣＣＣ－３’（配列番号：８５）

逆方向ＰＣＲプライマー　５’－ＡＡＣＣＴＧＧＡＡＴＧＴＣＡＣＣＧＡＧＣＴＧ－３’（配列番号：８６）

逆方向ＰＣＲプライマー　５’－ＣＣＴＡＧＣＡＣＡＧＴＧＡＣＧＡＧＧＧＡＣＴＴＧＧＣ－３’（配列番号：８７）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ２８７４０配列から作成した。

ハイブリッド形成プローブ

５’－ＡＡＧＡＣＡＣＡＧＣＣＡＣＣＣＴＡＡＡＣＴＧＴＣＡＧＴＣＴＴＣＴＧＧＧＡＧＣＡＡＧＣＣＴＧＣＡＧＣＣ－３’（配列番号：８８）

５’－ＧＣＣＣＴＧＧＣＡＧＡＣＧＡＧＧＧＣＧＡＧＴＡＣＡＣＣＴＧＣＴＣＡＡＴＣＴＴＣＡＣＴＡＴＧＣＣＴＧＴ－３’（配列番号：８９）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２５８遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織から単離した。上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２５８に対する全長ＤＮＡ配列［ここでＤＮＡ３５９１８－１１７４と命名］（配列番号：８３）及びＰＲＯ２５８の誘導タンパク質配列が得られた。

ＤＮＡ３５９１８－１１７４の全コード配列をＦｉｇ３１（配列番号：８３）に示す。クローンＤＮＡ３５９１８－１１７４は、単一のオープンリーディングフレームを含み、配列番号：８３のヌクレオチド位置１４７－１４９に見かけの翻訳開始部位を持ち、そして配列番号８３のヌクレオチド位置１３４０の後の停止コドンで終端する（Ｆｉｇ３１）。予測されるポリペプチド前駆体は３９８アミノ酸長である（Ｆｉｇ３２）。クローンＤＮＡ３５９１８－１１７４はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４０２が付与された。

全長ＰＲＯ２５８ポリペプチドのアミノ酸配列の分析により、それがＣＲＴＡＭ及びポリオウィルスレセプターと有意な相同性を有しＩｇドメインを有しており、よってＰＲＯ２５８はＩｇスーパーファミリーの新規なメンバーであることが示される。

【０２４２】
実施例１４：ヒトＰＲＯ２６６をコードするｃＤＮＡクローンの単離
ＳＬＩＴと相同性を有するＥＳＴに対して発現タグデータベースを検索し、ここでＴ７３９９６と命名する単一のＥＳＴ配列を同定した。Ｔ７３９９６ＥＳＴ配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２６６の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＧＴＴＧＧＡＴＣＴＧＧＧＣＡＡＣＡＡＴＡＡＣ－３’（配列番号：９２）
逆方向ＰＣＲプライマー　５’－ＡＴＴＧＴＴＧＴＧＣＡＧＧＣＴＧＡＧＴＴＴＡＡＧ－３’（配列番号：９３）
また、次のヌクレオチド配列を持つ合成オリゴヌクレオチドハイブリッド形成プローブを作成した。
ハイブリッド形成プローブ
５’－ＧＧＴＧＧＣＴＡＴＡＣＡＴＧＧＡＴＡＧＣＡＡＴＴＡＣＣＴＧＧＡＣＡＣＧＣＴＧＴＣＣＣＧＧＧ－３’（配列番号：９４）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２６６遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児脳織から単離した。上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２６６に対する全長ＤＮＡ配列［ここでＤＮＡ３７１５０－１１７８と命名］（配列番号：９０）及びＰＲＯ２６６の誘導タンパク質配列が得られた。
ＤＮＡ３７１５０－１１７８の全コード配列をＦｉｇ３３（配列番号：９０）に示す。クローンＤＮＡ３７１５０－１１７８は、単一のオープンリーディングフレームを含み、配列番号：９０のヌクレオチド位置１６７－１６９に見かけの翻訳開始部位を持ち、そしてそのヌクレオチド位置２２５４の後の停止コドンで終端する。予測されるポリペプチド前駆体は６９６アミノ酸長である（Ｆｉｇ３４）。クローンＤＮＡ３７１５０－１１７８はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４０１が付与された。
全長ＰＲＯ２６６ポリペプチドのアミノ酸配列の分析により、その部分がＳＬＩＴタンパク質と有意な相同性を有しており、よってＰＲＯ２６６は新規なロイシンリッチ反復タンパク質のメンバーであることが示される。
【０２４３】
実施例１５：ヒトＰＲＯ２６９をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列を、ここでＤＮＡ３５７０５と命名する。ＤＮＡ３５７０５コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２６９の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
ＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー（．ｆ１）　５’－ＴＧＧＡＡＧＧＡＧＡＴＧＣＧＡＴＧＣＣＡＣＣＴＧ－３’（配列番号：９７）
正方向ＰＣＲプライマー（．ｆ２）　５’－ＴＧＡＣＣＡＧＴＧＧＧＧＡＡＧＧＡＣＡ－３’（配列番号：９８）
正方向ＰＣＲプライマー（．ｆ３）　５’－ＡＣＡＧＡＧＣＡＧＡＧＧＧＴＧＣＣＴＴＧ－３’（配列番号：９９）
逆方向ＰＣＲプライマー（．ｒ１）　５’－ＴＣＡＧＧＧＡＣＡＡＧＴＧＧＴＧＴＣＴＣ

ＴＣＣＣ－３’（配列番号：１００）
逆方向ＰＣＲプライマー（．ｒ２）　５’－ＴＣＡＧＧＧＡＡＧＧＡＧＴＧＴＧＣＡＧＴＴＣＴＧ－３’（配列番号：１０１）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３５７０５配列から作成した。
ハイブリッド形成プローブ
５’－ＡＣＡＧＣＴＣＣＣＧＡＴＣＴＣＡＧＴＴＡＣＴＴＧＣＡＴＣＧＣＧＧＡＣＧＡＡＡＴＣＧＧＣＧＣＴＣＧＣＴ－３’（配列番号：１０２）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２６９遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織から単離した。上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２６９に対する全長ＤＮＡ配列［ここでＤＮＡ３８２６０－１１８０と命名］（配列番号：９５）及びＰＲＯ２６９の誘導タンパク質配列が得られた。
ＤＮＡ３８２６０－１１８０の全コード配列をＦｉｇ３５（配列番号：９５）に示す。クローンＤＮＡ３８２６０－１１８０は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３１４－３１６に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１７８４－１７８６の停止コドンで終端する（Ｆｉｇ３５；配列番号：９５）。予測されるポリペプチド前駆体は４９０アミノ酸長である（Ｆｉｇ３６）。クローンＤＮＡ３８２６０－１１８０はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３９７が付与された。
全長ＰＲＯ２６９ポリペプチドのアミノ酸配列の分析により、その部分がヒトトロンボモジュリンタンパク質と有意な相同性を有しており、よってＰＲＯ２６９は一又は複数のトロンボモジュリン様ドメインを有しうることが示唆される。
【０２４４】
実施例１６：ヒトＰＲＯ２８７をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列を、ここでＤＮＡ２８７２８と命名した。ＤＮＡ２８７２８コンセンサス配列に基づいて、ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及びＰＲＯ２８７の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＣＣＧＡＴＴＣＡＴＡＧＡＣＣＴＣＧＡＧＡＧＴ－３’（配列番号：１０５）
逆方向ＰＣＲプライマー　５’－ＧＴＣＡＡＧＧＡＧＴＣＣＴＣＣＡＣＡＡＴＡＣ－３’（配列番号：１０６）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ２８７２８配列から作成した。
ハイブリッド形成プローブ
５’－ＧＴＧＴＡＣＡＡＴＧＧＣＣＡＴＧＣＣＡＡＴＧＧＣＣＡＧＣＧＣＡＴＴＧＧＣＣＧＣＴＴＣＴＧＴ－３’（配列番号：１０７）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２８７遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２８７に対する全長ＤＮ

Ａ配列［ここでＤＮＡ３９９６９－１１８５と命名、配列番号：１０３］及びＰＲＯ２８７の誘導タンパク質配列が得られた。
ＤＮＡ３９９６９－１１８５の全コード配列をＦｉｇ３７（配列番号：１０３）に示す。クローンＤＮＡ３９９６９－１１８５は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３０７－３０９に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１５５２－１５５４の停止コドンで終端する（Ｆｉｇ３７；配列番号：１０３）。予測されるポリペプチド前駆体は４１５アミノ酸長である（Ｆｉｇ３８）。クローンＤＮＡ３９９６９－１１８５はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４００が付与された。
全長ＰＲＯ２８７のアミノ酸配列の分析により、それが、一又は複数のプロコラーゲンＣ－プロテイナーゼエンハンサータンパク質前駆体又はプロコラーゲンＣ－プロテイナーゼエンハンサータンパク質様ドメインを有しうることが示唆される。全長配列のＢＬＡＳＴ及びＦａｓｔＡアラインメント分析に基づいて、ＰＲＯ２８７はプロコラーゲンＣ－プロテイナーゼエンハンサータンパク質前駆体及びプロコラーゲンＣ－プロテイナーゼエンハンサータンパク質と核酸配列同一性を示す（それぞれ４７及び５４％）。
【０２４５】
実施例１７：ヒトＰＲＯ２１４をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いてコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列を、ここでＤＮＡ２８７４４と命名する。このコンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２１４遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡクローンはその全体を配列決定した。ＤＮＡ３２２８６－１１９１の全長ヌクレオチド配列はＦｉｇ３９（配列番号：１０８）に示す。ＤＮＡ３２２８６－１１９１は単一のオープンリーディングフレームを含み、ヌクレオチド位置１０３に見かけの翻訳開始部位を持つ（Ｆｉｇ３９；配列番号：１０８）。予測されるポリペプチド前駆体は４２０アミノ酸長である（Ｆｉｇ１０９）。
全長配列のＢＬＡＳＴ及びＦａｓｔＡ配列アラインメント分析に基づいて、ＰＲＯ２１４はＨＴタンパク質及び／又はフィブリンとアミノ酸配列同一性を示す（ぞれぞれ４９％及び３８％）。
上記の手順において使用されたオリゴヌクレオチド配列は次の通りであった：２８７４４．ｐ（ＯＬＩ５５５）
５’－ＣＣＴＧＧＣＴＡＴＣＡＧＣＡＧＧＴＧＧＧＣＴＣＣＡＡＧＴＧＴＣＴＣＧＡＴＧＴＧＧＡＴＧＡＧＴＧＴＧＡ－３’（配列番号：１１０）
２８７４４．ｆ（ＯＬＩ５５６）
５’－ＡＴＴＣＴＧＣＧＴＧＡＡＣＡＣＴＧＡＧＧＧＣ－３’（配列番号：１１１）
２８７４４．ｒ（ＯＬＩ５５７）
５’－ＡＴＣＴＧＣＴＴＧＴＡＧＣＣＣＴＣＧＧＣＡＣ－３’（配列番号：１１２）
【０２４６】
実施例１８：ヒトＰＲＯ３１７をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いてコンセンサスＤＮＡ配列を組み立て、コンセンサス配列をここでＤＮＡ２８７２２と命名した。このコンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。この目的のために合成した正及び逆ＰＣＲプライマーは次のもので

あった。
５’－ＡＧＧＡＣＴＧＣＣＡＴＡＡＣＴＴＧＣＣＴＧ（ＯＬＩ４８９）（配列番号：１１５）及び
５’－ＡＴＡＧＧＡＧＴＴＧＡＡＧＣＡＧＣＧＣＴＧＣ（ＯＬＩ４９０）（配列番号：１１６）。
この目的のために合成したプローブは次のものであった。
５’－ＴＧＴＧＴＧＧＡＣＡＴＡＧＡＣＧＡＧＴＧＣＣＧＣＴＡＣＣＧＣＴＡＣＴＧＣＣＡＧＣＡＣＣＧＣ（ＯＬＩ４８８）（配列番号：１１７）
ｃＤＮＡライブラリの作成のためのｍＲＮＡはヒト胎児腎臓組織から単離した。
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡをＰＣＲプライマー対で、Ａｕｓｕｂｅｌ等，Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙに従って、ＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３１７遺伝子を含むクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織から単離された。
ｃＤＮＡクローンはその全体を配列決定した。（ＰＲＯ３１７をコードしている）ＤＮＡ３３４６１－１１９９の全長ヌクレオチド配列はＦｉｇ４１（配列番号：１１３）に示す。クローンＤＮＡ３３４６１－１１９９は単一のオープンリーディングフレームを含み、ヌクレオチド位置６８－７０に見かけの翻訳開始部位を持つ（Ｆｉｇ４１；配列番号：１１３）。予測されるポリペプチド前駆体は３６６アミノ酸長である。予想されるシグナル配列はＦｉｇ４２（配列番号：１１４）のアミノ酸１－１８である。アミノ酸残基１６０に一つの予想Ｎ結合グリコシル化部位がある。クローンＤＮＡ３３４６１－１１９９はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３６７が付与された。
全長ＰＲＯ３１７配列の（ＡＬＩＧＮ™コンピュータプログラムを用いる）ＢＬＡＳＴ™及びＦａｓｔＡ™配列アラインメント分析に基づくと、ＰＲＯ３１７はＥＢＡＦ－１と最高のアミノ酸配列同一性を有している（９２％）。結果はまたヒトＰＲＯ３１７及びマウスＬＥＦＴＹタンパク質との間の有意な相同性を証明している。ＰＲＯ３１７のＣ末端はＴＧＦ－スーパーファミリーのメンバーであると期待されるパターンと一致した多くの保存配列を含む。
以下に記載されるようにして実施されたヒト組織中でのインサイツ発現分析により、膵臓組織中でのＰＲＯ３１７ポリペプチドの明確な強い発現があることが証明される。
【０２４７】
実施例１９：ヒトＰＲＯ３０１をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３０１遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓から単離した。
ｃＤＮＡクローンはその全体を配列決定した。天然配列ＰＲＯ３０１の全長ヌクレオチド配列をＦｉｇ４３（配列番号：１１８）に示す。クローンＤＮＡ４０６２８－１２１６は単一のオープンリーディングフレームを含み、ヌクレオチド位置５２－５４に見かけの翻訳開始部位を持つ（Ｆｉｇ４３；配列番号：１１８）。予測されるポリペプチド前駆体は２９９アミノ酸長で、３２，５８３ダルトンの予想分子量と８．２９のｐＩを持つ。クローンＤＮＡ４０６２８－１２１６はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４３２

が付与された。
全長配列のＢＬＡＳＴ及びＦａｓｔＡ配列アラインメント分析に基づいて、ＰＲＯ３０１はＡ３３抗原前駆体（３０％）及びコクサッキー及びアデノウィルスレセプタータンパク質（２９％）とアミノ酸配列同一性を示す。
上記の手順において使用されたオリゴヌクレオチド配列は次の通りであった：
ＯＬＩ２１６２（３５９３６．ｆ１）　５’－ＴＣＧＣＧＧＡＧＣＴＧＴＧＴＴＣＴＧＴＴＴＣＣＣ－３’（配列番号：１２０）
ＯＬＩ２１６３（３５９３６．ｐ１）
５’－ＴＧＡＴＣＧＣＧＡＴＧＧＧＧＡＣＡＡＡＧＧＣＧＣＡＡＧＣＴＣＧＡＧＡＧＧＡＡＡＣＴＧＴＴＧＴＧＣＣＴ－３’（配列番号：１２１）
ＯＬＩ２１６４（３５９３６．ｆ２）
５’－ＡＣＡＣＣＴＧＧＴＴＣＡＡＡＧＡＴＧＧＧ－３’（配列番号：１２２）
ＯＬＩ２１６５（３５９３６．ｒ１）
５’－ＴＡＧＧＡＡＧＡＧＴＴＧＣＴＧＡＡＧＧＣＡＣＧＧ－３’（配列番号：１２３）
ＯＬＩ２１６６（３５９３６．ｆ３）
５’－ＴＴＧＣＣＴＴＡＣＴＣＡＧＧＴＧＣＴＡＣ－３’（配列番号：１２４）
ＯＬＩ２１６７（３５９３６．ｒ２）
５’－ＡＣＴＣＡＧＣＡＧＴＧＧＴＡＧＧＡＡＡＧ－３’（配列番号：１２５）
【０２４８】
実施例２０：ヒトＰＲＯ２２４をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列を、ここでＤＮＡ３０８４５と命名した。ＤＮＡ３０８４５コンセンサス配列に基づいて、対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及びＰＲＯ２２４の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＡＡＧＴＴＣＣＡＧＴＧＣＣＧＣＡＣＣＡＧＴＧＧＣ－３’（配列番号：１２８）
逆方向ＰＣＲプライマー　５’－ＴＴＧＧＴＴＣＣＡＣＡＧＣＣＧＡＧＣＴＣＧＴＣＧ－３’（配列番号：１２９）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０８４５配列から作成した。
ハイブリッド形成プローブ
５’－ＧＡＧＧＡＧＧＡＧＴＧＣＡＧＧＡＴＴＧＡＧＣＣＡＴＧＴＡＣＣＣＡＧＡＡＡＧＧＧＣＡＡＴＧＣＣＣＡＣＣ－３’（配列番号：１３０）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２２４遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２２４に対する全長ＤＮＡ配列［ここでＤＮＡ３３２２１－１１３３と命名］及びＰＲＯ２２４の誘導タンパク質配列が得られた。
ＤＮＡ３３２２１－１１３３の全ヌクレオチド配列をＦｉｇ４５（配列番号：１２６）に示す。クローンＤＮＡ３３２２１－１１３３は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３３－３５に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置８７９－８９９の停止コドンで終端する（Ｆｉｇ４５；配列番号：１２６）。膜貫通領域の開始はヌクレオチド位置７７７で始まる。予測されるポリペプチド前駆体は２８２アミノ酸長である（Ｆｉｇ４６）。クローンＤＮＡ３３２２１－１１３３はＡＴＣＣに寄

託され、ＡＴＣＣ寄託番号２０９２６３が付与された。
全長ＰＲＯ２２４のアミノ酸配列の分析により、それが、超低密度リポタンパク質レセプター、アポリポタンパク質Ｅレセプター及びニワトリ卵母細胞レセプターＰ９５と相同性を有することが示唆される。全長配列のＢＬＡＳＴ及びＦａｓｔＡ配列アラインメント分析に基づくと、ＰＲＯ２２４は２８％から４５％の範囲でこれらのタンパク質の部分とアミノ酸同一性を有し、３３％から３９％の範囲でこれらのタンパク質と全体的同一性を有している。
【０２４９】
実施例２１：ヒトＰＲＯ２２２をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列を、ここでＤＮＡ２８７７１と命名した。ＤＮＡ２８７７１コンセンサス配列に基づいて、対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及びＰＲＯ２２２の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＡＴＣＴＣＣＴＡＴＣＧＣＴＧＣＴＴＴＣＣＣＧＧ－３’（配列番号：１３３）
逆方向ＰＣＲプライマー　５’－ＡＧＣＣＡＧＧＡＴＣＧＣＡＧＴＡＡＡＡＣＴＣＣ－３’（配列番号：１３４）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ２８７７１配列から作成した。
ハイブリッド形成プローブ
５’－ＡＴＴＴＡＡＡＣＴＴＧＡＴＧＧＧＴＣＴＧＣＧＴＡＴＣＴＴＧＡＧＴＧＣＴＴＡＣＡＡＡＡＣＣＴＴＡＴＣＴ－３’（配列番号：１３５）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２２２遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２２２に対する全長ＤＮＡ配列［ここでＤＮＡ３３１０７－１１３５と命名］及びＰＲＯ２２２の誘導タンパク質配列が得られた。
ＤＮＡ３３１０７－１１３５の全ヌクレオチド配列をＦｉｇ４７（配列番号：１３１）に示す。クローンＤＮＡ３３１０７－１１３５は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１５９－１６１に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１６２９－１６３１の停止コドンで終端する（Ｆｉｇ４７；配列番号：１３１）。予測されるポリペプチド前駆体は４９０アミノ酸長である（Ｆｉｇ４８）。クローンＤＮＡ３３１０７－１１３５はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９２５１が付与された。
全長配列のＢＬＡＳＴ及びＦａｓｔＡ配列アラインメント分析に基づくと、ＰＲＯ２２２は、マウス補体因子ｈ前駆体（２５－２６％）、補体レセプター（２７－２９％）、マウス補体Ｃ３ｂレセプター２型長型前駆体（２５－４７％）及びヒト仮想タンパク質ｋｉａａ０２４７（４０％）とアミノ酸配列同一性を示している。
【０２５０】
実施例２２：ヒトＰＲＯ２３４をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いてコンセンサスＤＮＡ配列を組み立てた（ＤＮＡ３０９２６）。このコンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

ｃＤＮＡライブラリの作成のためのＲＮＡは標準的な単離プロトコール、例えばＡｕｓｕｂｅｌ等，Ｃｕｒｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙを使用して、組織又は株化細胞から単離されるか、商業的供給源（例えばＣｌｏｎｔｅｃｈ）から購入した。ｃＤＮＡクローンを単離するために使用したｃＤＮＡライブラリは商業的に利用できる試薬（例えばＩｎｖｉｔｒｏｇｅｎ）を使用して標準的な方法（例えばＡｕｓｕｂｅｌ等）により作成された。このライブラリは２２週齢の胎児脳組織から取り出された。

ｃＤＮＡクローンはその全体を配列決定した。ＰＲＯ２３４の全長ヌクレオチド配列をＦｉｇ４９（配列番号：１３６）に示す。予測されるポリペプチド前駆体は３８２アミノ酸長で、４３．１ｋＤａの算定分子量を持つ。

上記の手順において使用されたオリゴヌクレオチド配列は次の通りであった：

３０９２６．ｐ（ＯＬＩ８２６）（配列番号：１３８）：　５’－ＧＴＴＣＡＴＴＧＡＡＡＣＣＴＣＴＴＧＣＣＡＴＣＴＧＡＴＧＧＴＧＡＣＴＴＣＴＧＧＡＴＴＧＧＧＣＴＣＡ－３’

３０９２６．ｆ（ＯＬＩ８２７）（配列番号：１３９）：　５’－ＡＡＧＣＣＡＡＡＧＡＡＧＣＣＴＧＣＡＧＧＡＧＧＧ－３’

３０９２６．ｒ（ＯＬＩ８２８）（配列番号：１４０）：　５’－ＣＡＧＴＣＣＡＡＧＣＡＴＡＡＡＧＧＴＣＣＴＧＧＣ－３’

【０２５１】

実施例２３：ヒトＰＲＯ２３１をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列を、ここでＤＮＡ３０９３３と命名した。ＤＮＡ３０９３３コンセンサス配列に基づいて、対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及びＰＲＯ２３１の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー１　５’－ＣＣＡＡＣＴＡＣＣＡＡＡＧＣＴＧＣＴＧＧＡＧＣＣ－３’（配列番号：１４３）

正方向ＰＣＲプライマー２　５’－ＧＣＡＧＣＴＣＴＡＴＴＡＣＣＡＣＧＧＧＡＡＧＧＡ－３’（配列番号：１４４）

逆方向ＰＣＲプライマー　　５’－ＴＣＣＴＴＣＣＣＧＴＧＧＴＡＡＴＡＧＡＧＣＴＧＣ－３’（配列番号：１４５）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０９３３配列から作成した。

ハイブリッド形成プローブ

５’－ＧＧＣＡＧＡＧＡＡＣＣＡＧＡＧＧＣＣＧＧＡＧＧＡＧＡＣＴＧＣＣＴＣＴＴＴＡＣＡＧＣＣＡＧＧ－３’（配列番号：１４６）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２３１遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２３１に対する全長ＤＮＡ配列［ここでＤＮＡ３４４３４－１１３９と命名］及びＰＲＯ２３１の誘導タンパク質配列が得られた。

ＤＮＡ３４４３４－１１３９の全ヌクレオチド配列をＦｉｇ５１（配列番号：１４１）に示す。クローンＤＮＡ３４４３４－１１３９は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１７３－１７５に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１４５７－１４５９の停止コドンで終端する（Ｆｉｇ５１；配列番号：１４１）

。予測されるポリペプチド前駆体は４２８アミノ酸長である（Ｆｉｇ５２）。クローンＤＮＡ３４４３４－１１３９は１９９７年９月１６日にＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９２５２が付与された。
全長ＰＲＯ２３１のアミノ酸配列の分析によると、それが、ヒト及びラット前立腺酸性ホスファターゼ前駆体タンパク質とそれぞれ３０％及び３１％のアミノ酸同一性を有していることが示唆される。
【０２５２】
実施例２４：ヒトＰＲＯ２２９をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ２８７６２と命名した。ＤＮＡ２８７６２コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２２９の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＴＴＣＡＧＣＴＣＡＴＣＡＣＣＴＴＣＡＣＣＴＧＣＣ－３’（配列番号：１４９）
逆方向ＰＣＲプライマー　５’－ＧＧＣＴＣＡＴＡＣＡＡＡＡＴＡＣＣＡＣＴＡＧＧＧ－３’（配列番号：１５０）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ２８７６２配列から作成した。
ハイブリッド形成プローブ
５’－ＧＧＧＣＣＴＣＣＡＣＣＧＣＴＧＴＧＡＡＧＧＧＣＧＧＧＴＧＧＡＧＧＴＧＧＡＡＣＡＧＡＡＡＧＧＣＣＡＧＴ－３’（配列番号：１５１）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２２９遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２２９に対する全長ＤＮＡ配列［ここでＤＮＡ３３１００－１１５９と命名］（配列番号：１４７）及びＰＲＯ２２９の誘導タンパク質配列が得られた。
ＤＮＡ３３１００－１１５９の全ヌクレオチド配列をＦｉｇ５３（配列番号：１４７）に示す。クローンＤＮＡ３３１００－１１５９は、単一のオープンリーディングフレームを含み、ヌクレオチド位置９８－１００に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１１３９－１１４１の停止コドンで終端する（Ｆｉｇ５３）。予測されるポリペプチド前駆体は３４７アミノ酸長である（Ｆｉｇ５４）。クローンＤＮＡ３３１００－１１５９はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３７７が付与された。
全長ＰＲＯ２２９ポリペプチドのアミノ酸配列の分析によると、その部分が抗原ｗｃ１．１、Ｍ１３０抗原及びＣＤ６と有意な相同性を有していることが示唆される。
【０２５３】
実施例２５：ヒトＰＲＯ２３８をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列を、ここでＤＮＡ３０９０８と命名した。ＤＮＡ３０９０８コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２３８の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
ＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー１　５’－ＧＧＴＧＣＴＡＡＡＣＴＧＧＴＧＣＴＣＴＧＴＧＧＣ－３’（配列番号：１５４）

正方向ＰＣＲプライマー２　５’－ＣＡＧＧＧＣＡＡＧＡＴＧＡＧＣＡＴＴＣＣ－３’（配列番号：１５５）
逆方向ＰＣＲプライマー　５’－ＴＣＡＴＡＣＴＧＴＴＣＣＡＴＣＴＣＧＧＣＡＣＧＣ－３’（配列番号：１５６）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０９０８配列から作成した。
ハイブリッド形成プローブ
５’－ＡＡＴＧＧＴＧＧＧＧＣＣＣＴＡＧＡＡＧＡＧＣＴＣＡＴＣＡＧＡＧＡＡＣＴＣＡＣＣＧＣＴＴＣＴＣＡＴＧＣ－３’（配列番号：１５７）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２３８遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２３８に対する全長ＤＮＡ配列及びＰＲＯ２３８の誘導タンパク質配列が得られた。
ＤＮＡ３５６００－１１６２の全ヌクレオチド配列をＦｉｇ５５（配列番号：１５２）に示す。クローンＤＮＡ３５６００－１１６２は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１３４－１３６に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１０６４－１０６６の停止コドンで終端する（Ｆｉｇ５５）。予測されるポリペプチド前駆体は３１０アミノ酸長である（Ｆｉｇ５６）。クローンＤＮＡ３５６００－１１６２はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３７０が付与された。
全長ＰＲＯ２３８のアミノ酸配列の分析によると、その部分が、レダクターゼ、特にオキシドレダクターゼと有意な相同性を有していることが示唆され、ＰＲＯ２３８が新規なレダクターゼでありうることを示している。
【０２５４】
実施例２６：ヒトＰＲＯ２３３をコードするｃＤＮＡクローンの単離
Ｓｗｉｓｓ－Ｐｒｏｔ公的データベースからの約９５０の既知の分泌タンパク質からの細胞外ドメイン（ＥＣＤ）配列（あれば、分泌シグナルを含む）を、発現配列タグ（ＥＳＴ）データベースの検索に使用した。ＥＳＴデータベースは、公的ＥＳＴデータベース（例えば、ＧｅｎＢａｎｋ）及び企業のＥＳＴＤＮＡデータベース（ＬＩＦＥＳＥＱ（商品名）、Ｉｎｃｙｔｅ Ｐｈａｒｍａｃｅｕｔｉｃａｌｓ、Ｐａｌｏ Ａｌｔｏ，ＣＡ）を含む。検索は、コンピュータプログラムＢＬＡＳＴ又はＢＬＡＳＴ－２（Ａｌｔｓｃｈｕｌ等，Ｍｅｔｈｏｄｓ ｉｎ Ｅｎｚｙｍｏｌｏｇｙ ２６６：４６０－４８０（１９９６））を用いて、ＥＣＤタンパク質配列のＥＳＴ配列の６フレーム翻訳との比較として実施した。既知のタンパク質をコードせず、ＢＬＡＳＴスコア７０（９０の場合もある）又はそれ以上を持つ比較物は、プログラム「ｐｈｒａｐ」（Ｐｈｉｌ Ｇｒｅｅｎ，Ｕｎｉｖｅｒｓｉｔｙ ｏｆ Ｗａｓｈｉｎｇｔｏｎ，Ｓｅａｔｔｌｅ，ＷＡ；ｈｔｔｐ：／／ｂｏｚｅｍａｎ．ｍｂｔ．ｗａｓｈｉｎｇｔｏｎ．ｅｄｕ／ｐｈｒａｐ．ｄｏｃｓ／ｐｈｒａｐ．ｈｔｍｌ）で集団化してコンセンサスＤＮＡ配列に構築した。
発現配列タグ（ＥＳＴ）はＥＳＴデータベース検索により確認され、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列を、ここでＤＮＡ３０９４５と命名した。ＤＮＡ３０９４５コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２３３の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＧＧＴＧＡＡＧＧＣＡＧＡＡＡＴＴＧＧＡＧＡＴＧ－３’（配列番号：１６０）

逆方向ＰＣＲプライマー　５’－ＡＴＣＣＣＡＴＧＣＡＴＣＡＧＣＣＴＧＴＴＴＡＣＣ－３’（配列番号：１６１）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０９４５配列から作成した。
ハイブリッド形成プローブ
５’－ＧＣＴＧＧＴＧＴＡＧＴＣＴＡＴＡＣＡＴＣＡＧＡＴＴＴＧＴＴＴＧＣＴＡＣＡＣＡＡＧＡＴＣＣＴＣＡＧ－３’（配列番号：１６２）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２３３遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児脳組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２３３に対する全長ＤＮＡ配列［ここでＤＮＡ３４４３６－１２３８と命名］（配列番号：１５８）及びＰＲＯ２３３の誘導タンパク質配列が得られた。
ＤＮＡ３４４３６－１２３８の全ヌクレオチド配列をＦｉｇ５７（配列番号：１５８）に示す。クローンＤＮＡ３４４３６－１２３８は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１０１－１０３に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１００１－１００３の停止コドンで終端する（Ｆｉｇ５７）。予測されるポリペプチド前駆体は３００アミノ酸長である（Ｆｉｇ５８）。Ｆｉｇ５８に示す全長ＰＲＯ２３３プロテインは約３２，９６４ダルトン及び約９．５２のｐｌの算定分子量を有する。クローンＤＮＡ３４４３６－１２３８はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９５２３が付与された。
全長ＰＲＯ２３３ポリペプチドのアミノ酸配列の分析によると、その部分がレダクターゼタンパク質と有意な相同性を有していることが示唆され、ＰＲＯ２３３が新規なレダクターゼでありうることを示している。
【０２５５】
実施例２７：ヒトＰＲＯ２２３をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３０８３６と命名した。ＤＮＡ３０８３６コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２２３の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＴＴＣＣＡＴＧＣＣＡＣＣＴＡＡＧＧＧＡＧＡＣＴＣ－３’（配列番号：１６５）
逆方向ＰＣＲプライマー１　５’－ＴＧＧＡＴＧＡＧＧＴＧＴＧＣＡＡＴＧＧＣＴＧＧＣ－３’（配列番号：１６６）
逆方向ＰＣＲプライマー２　５’－ＡＧＣＴＣＴＣＡＧＡＧＧＣＴＧＧＴＣＡＴＡＧＧＧ－３’（配列番号：１６７）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０８３６配列から作成した。
ハイブリッド形成プローブ
５’－ＧＴＣＧＧＣＣＣＴＴＴＣＣＣＡＧＧＡＣＴＧＡＡＣＡＴＧＡＡＧＡＧＴＴＡＴＧＣＣＧＧＣＴＴＣＣＴＣＡＣ－３’（配列番号：１６８）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２２３遺伝子をコードするクローンを単離するのに使用した

。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２２３に対する全長ＤＮＡ配列［ここでＤＮＡ３３２０６－１１６５と命名］（配列番号：１６３）及びＰＲＯ２２３の誘導タンパク質配列が得られた。
ＤＮＡ３３２０６－１１６５の全ヌクレオチド配列をＦｉｇ５９（配列番号：１６３）に示す。クローンＤＮＡ３３２０６－１１６５は、単一のオープンリーディングフレームを含み、ヌクレオチド位置９７－９９に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１５２５－１５２７の停止コドンで終端する（Ｆｉｇ５９）。予測されるポリペプチド前駆体は４７６アミノ酸長である（Ｆｉｇ６０）。クローンＤＮＡ３３２０６－１１６５はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３７２が付与された。
全長ＰＲＯ２２３のアミノ酸配列の分析によると、それが、様々なセリンカルボキシペプチダーゼタンパク質と有意な相同性を有していることが示唆され、ＰＲＯ２２３が新規なセリンカルボキシペプチダーゼでありうることを示している。
【０２５６】
実施例２８：ヒトＰＲＯ２３５をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３０９２７と命名した。ＤＮＡ３０９２７コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２３５の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＴＧＧＡＡＴＡＣＣＧＣＣＴＣＣＴＧＣＡＧ－３’（配列番号：１７１）
逆方向ＰＣＲプライマー　５’－ＣＴＴＣＴＧＣＣＣＴＴＴＧＧＡＧＡＡＧＡＴＧＧＣ－３’（配列番号：１７２）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０９２７配列から作成した。
ハイブリッド形成プローブ
５’－ＧＧＡＣＴＣＡＣＴＧＧＣＣＣＡＧＧＣＣＴＴＣＡＡＴＡＴＣＡＣＣＡＧＣＣＡＧＧＡＣＧＡ－３’（配列番号：１７３）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２３５遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２３５に対する全長ＤＮＡ配列［ここでＤＮＡ３５５５８－１１６７と命名］（配列番号：１６９）及びＰＲＯ２３５の誘導タンパク質配列が得られた。
ＤＮＡ３５５５８－１１６７の全ヌクレオチド配列をＦｉｇ６１（配列番号：１６９）に示す。クローンＤＮＡ３５５５８－１１６７は、単一のオープンリーディングフレームを含み、ヌクレオチド位置６６７－６６９に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置２３２３－２３２５の停止コドンで終端する（Ｆｉｇ６１）。予測されるポリペプチド前駆体は５５２アミノ酸長である（Ｆｉｇ６２）。クローンＤＮＡ３５５５８－１１６７はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３７４が付与された。
全長ＰＲＯ２３５のアミノ酸配列の分析によると、その部分がヒト、マウス及びアフリカツメガエルのプレキシンタンパク質と有意な相同性を有しており、ＰＲＯ２３５が新規なプレキシンタンパク質でありうることを示している。
【０２５７】

実施例２９：ヒトＰＲＯ２３６及びヒトＰＲＯ２６２をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３０９０１及びＤＮＡ３０８４７と命名した。ＤＮＡ３０９０１及びＤＮＡ３０８４７コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２３６及びＰＲＯ２６２の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

ＤＮＡ３０９０１コンセンサス配列に基づいて、一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー　５’－ＴＧＧＣＴＡＣＴＣＣＡＡＧＡＣＣＣＴＧＧＣＡＴＧ－３’（配列番号：１７８）

逆方向ＰＣＲプライマー　５’－ＴＧＧＡＣＡＡＡＴＣＣＣＣＴＴＧＣＴＣＡＧＣＣＣ－３’（配列番号：１７９）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０９０１配列から作成した。

ハイブリッド形成プローブ

５’－ＧＧＧＣＴＴＣＡＣＣＧＡＡＧＣＡＧＴＧＧＡＣＣＴＴＴＡＴＴＴＴＧＡＣＣＡＣＣＴＧＡＴＧＴＣＣＡＧＧＧ－３’（配列番号：１８０）

ＤＮＡ３０８４７コンセンサス配列に基づいて、一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー　５’－ＣＣＡＧＣＴＡＴＧＡＣＴＡＴＧＡＴＧＣＡＣＣ－３’（配列番号：１８１）

逆方向ＰＣＲプライマー　５’－ＴＧＧＣＡＣＣＣＡＧＡＡＴＧＧＴＧＴＴＧＧＣＴＣ－３’（配列番号：１８２）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０８４７配列から作成した。

ハイブリッド形成プローブ

５’－ＣＧＡＧＡＴＧＴＣＡＴＣＡＧＣＡＡＧＴＴＣＣＡＧＧＡＡＧＴＴＣＣＴＴＴＧＧＧＡＣＣＴＴＴＡＣＣＴＣＣ－３’（配列番号：１８３）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２３６及びＰＲＯ２６２遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはＰＲＯ２３６に対してはヒト胎児肺組織及びＰＲＯ２６２に対してはヒト胎児肝臓組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２３６に対する全長ＤＮＡ配列［ここでＤＮＡ３５５９９－１１６８と命名］（配列番号：１７４）、ＰＲＯ２３６の誘導タンパク質配列、ＰＲＯ２６２に対する全長ＤＮＡ配列［ここでＤＮＡ３６９９２－１１６８と命名］（配列番号：１７６）及びＰＲＯ２６２の誘導タンパク質配列が得られた。

ＤＮＡ３５５９９－１１６８の全ヌクレオチド配列をＦｉｇ６３（配列番号：１７４）に示す。クローンＤＮＡ３５５９９－１１６８は、単一のオープンリーディングフレームを含み、ヌクレオチド位置６９－７１に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１９７７－１９７９の停止コドンで終端する（Ｆｉｇ６３）。予測されるポリペプチド前駆体は６３６アミノ酸長である（Ｆｉｇ６４）。クローンＤＮＡ３５５９９－１１６８はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３７３が付与された。

ＤＮＡ３６９９２－１１６８の全ヌクレオチド配列をＦｉｇ６５（配列番号：１７６）に示す。クローンＤＮＡ３６９９２－１１６８は、単一のオープンリーディングフレームを

含み、ヌクレオチド位置２４０－２４２に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置２２０２－２２０４の停止コドンで終端する（Ｆｉｇ６５）。予測されるポリペプチド前駆体は６５４アミノ酸長である（Ｆｉｇ６６）。クローンＤＮＡ３６９９２－１１６８はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３８２が付与された。
全長ＰＲＯ２３６及びＰＲＯ２６２ポリペプチドのアミノ酸配列の分析によると、これらのポリペプチドの部分が、様々なβ－ガラクトシダーゼタンパク質と有意な相同性を有していることが示唆され、ＰＲＯ２３６及びＰＲＯ２６２が新規なβ－ガラクトシダーゼ相同体でありうることを示している。

【０２５８】

実施例３０：ヒトＰＲＯ２３９をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３０９０９と命名した。ＤＮＡ３０９０９コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２３９の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＣＣＴＣＣＣＴＣＴＡＴＴＡＣＣＣＡＴＧＴＣ－３’（配列番号：１８６）
逆方向ＰＣＲプライマー　５’－ＧＡＣＣＡＡＣＴＴＴＣＴＣＴＧＧＧＡＧＴＧＡＧＧＣ－３’（配列番号：１８７）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０９０９配列から作成した。
ハイブリッド形成プローブ
５’－ＧＴＣＡＣＴＴＴＡＴＴＴＣＴＣＴＡＡＣＡＡＣＡＡＧＣＴＣＧＡＡＴＣＣＴＴＡＣＣＡＧＴＧＧＣＡＧ－３’（配列番号：１８８）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２３９遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２３９に対する全長ＤＮＡ配列［ここでＤＮＡ３４４０７－１１６９と命名］（配列番号：１８４）及びＰＲＯ２３９の誘導タンパク質配列が得られた。
ＤＮＡ３４４０７－１１６９の全ヌクレオチド配列をＦｉｇ６７（配列番号：１８４）に示す。クローンＤＮＡ３４４０７－１１６９は、単一のオープンリーディングフレームを含み、ヌクレオチド位置７２－７４に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１５７５－１５７７の停止コドンで終端する（Ｆｉｇ６７）。予測されるポリペプチド前駆体は５０１アミノ酸長である（Ｆｉｇ６８）。クローンＤＮＡ３４４０７－１１６９はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３８３が付与された。
全長ＰＲＯ２３９ポリペプチドのアミノ酸配列の分析によると、その部分がデンシンタンパク質と有意な相同性を有していることが示唆され、ＰＲＯ２３９がデンシンファミリーの新規なタンパク質であり得ることを示している。

【０２５９】

実施例３１：ヒトＰＲＯ２５７をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ２８７３１と命名した。ＤＮＡ２８７３１コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２５７の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＴＣＴＣＴＡＴＴＣＣＡＡＡＣＴＧＴＧＧＣＧ－３’（配列番号：１９１）
逆方向ＰＣＲプライマー　５’－ＴＴＴＧＡＴＧＡＣＧＡＴＴＣＧＡＡＧＧＴＧＧ－３’（配列番号：１９２）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ２８７３１配列から作成した。
ハイブリッド形成プローブ
５’－ＧＧＡＡＧＧＡＴＣＣＴＴＣＡＣＣＡＧＣＣＣＣＡＡＴＴＡＣＣＣＡＡＡＧＣＣＧＣＡＴＣＣＴＧＡＧＣ－３’（配列番号：１９３）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２５７遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２５７に対する全長ＤＮＡ配列［ここでＤＮＡ３５８４１－１１７３と命名］（配列番号：１８９）及びＰＲＯ２５７の誘導タンパク質配列が得られた。
ＤＮＡ３５８４１－１１７３の全ヌクレオチド配列をＦｉｇ６９（配列番号：１８９）に示す。クローンＤＮＡ３５８４１－１１７３は、単一のオープンリーディングフレームを含み、ヌクレオチド位置９６４－９６６に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置２７８５－２７８７の停止コドンで終端する（Ｆｉｇ６９）。予測されるポリペプチド前駆体は６０７アミノ酸長である（Ｆｉｇ７０）。クローンＤＮＡ３５８４１－１１７３はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４０３が付与された。
全長ＰＲＯ２５７のアミノ酸配列の分析によると、その部分がエブネリンタンパク質と有意な相同性を有していることが示唆され、ＰＲＯ２５７がエブネリンタンパク質に関連した新規なタンパク質メンバーでありうることを示している。
【０２６０】
実施例３２：ヒトＰＲＯ２６０をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３０８３４と命名した。ＤＮＡ３０８３４コンセンサス配列に基づいて、１）ＰＣＲにより対象とする配列を含むｃＤＮＡライブラリを同定するため、及び２）ＰＲＯ２６０の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＴＧＧＴＴＴＧＡＣＣＡＧＧＣＣＡＡＧＴＴＣＧＧ－３’（配列番号：１９６）
逆方向ＰＣＲプライマーＡ　５’－ＧＧＡＴＴＣＡＴＣＣＴＣＡＡＧＧＡＡＧＡＧＣＧＧ－３’（配列番号：１９７）
逆方向ＰＣＲプライマーＢ　５’－ＡＡＣＴＴＧＣＡＧＣＡＴＣＡＧＣＣＡＣＴＣＴＧＣ－３’（配列番号：１９８）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０８３４配列から作成した。
ハイブリッド形成プローブ
５’－ＴＴＣＣＧＴＧＣＣＣＡＧＣＴＴＣＧＧＴＡＧＣＧＡＧＴＧＧＴＴＣＴＧＧＴＧＧＴＡＴＴＧＧＣＡ－３’（配列番号：１９９）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプラ

イマーの一方を用いてＰＲＯ２６０遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２６０に対する全長ＤＮＡ配列［ここでＤＮＡ３３４７０－１１７５と命名］（配列番号：１９４）及びＰＲＯ２６０の誘導タンパク質配列が得られた。
ＤＮＡ３３４７０－１１７５の全ヌクレオチド配列をＦｉｇ７１（配列番号：１９４）に示す。クローンＤＮＡ３３４７０－１１７５は、単一のオープンリーディングフレームを含み、ヌクレオチド位置６７－６９に見かけの翻訳開始部位を持ち、そして１４６８－１４７０の停止コドンで終端する（Ｆｉｇ７１）。予測されるポリペプチド前駆体は４６７アミノ酸長である（Ｆｉｇ７２）。クローンＤＮＡ３３４７０－１１７５はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３９８が付与された。
全長ＰＲＯ２６０のアミノ酸配列の分析によると、その部分がα－Ｉ－フコシダーゼ前駆体と有意な相同性を有していることが示唆され、ＰＲＯ２６０が新規なフコシダーゼでありうることを示している。
【０２６１】
実施例３３：ヒトＰＲＯ２６３をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴに対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３０９１４と命名した。ＤＮＡ３０９１４コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ２６３の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー１　５’－ＧＡＧＣＴＴＴＣＣＡＴＣＣＡＧＧＴＧＴＣＡＴＧＣ－３’（配列番号：２０２）
正方向ＰＣＲプライマー２　５’－ＧＴＣＡＧＴＧＡＣＡＧＴＡＣＣＴＡＣＴＣＧＧ－３’（配列番号：２０３）
逆方向ＰＣＲプライマー　５’－ＴＧＧＡＧＣＡＧＧＡＧＧＡＧＴＡＧＴＡＧＴＡＧＧ２４－３’（配列番号：２０４）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３０９１４配列から作成した。
ハイブリッド形成プローブ
５’－ＡＧＧＡＧＧＣＣＴＧＴＡＧＧＣＴＧＣＴＧＧＧＡＣＴＡＡＧＴＴＴＧＧＣＣＧＧＣＡＡＧＧＡＣＣＡＡＧＴＴ－３’（配列番号：２０５）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２６３遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２６３に対する全長ＤＮＡ配列［ここでＤＮＡ３４４３１－１１７７と命名］（配列番号：２００）及びＰＲＯ２６３の誘導タンパク質配列が得られた。
ＤＮＡ３４４３１－１１７７の全ヌクレオチド配列をＦｉｇ７３（配列番号：２００）に示す。クローンＤＮＡ３４４３１－１１７７は、単一のオープンリーディングフレームを含み、配列番号：２００のヌクレオチド位置１６０－１６２に見かけの翻訳開始部位を持ち、そして配列番号：２００の位置１１２６－１１２８の停止コドンで終端する（Ｆｉｇ７３；配列番号：２００）。予測されるポリペプチド前駆体は３２２アミノ酸長である（Ｆｉｇ７４）。クローンＤＮＡ３４４３１－１１７７はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３９９が付与された。

全長ＰＲＯ２６３ポリペプチドのアミノ酸配列の分析によると、その部分がＣＤ４４抗原と有意な相同性を有していることが示唆され、ＰＲＯ２６３が新規な細胞表面接着分子でありうることを示している。

【０２６２】

実施例３４：ヒトＰＲＯ２７０をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、他の同定されたＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立て、コンセンサス配列を、ここでＤＮＡ３５７１２と命名した。このＤＮＡ３５７１２コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ２７０の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

ＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー（．ｆ１）　５’－ＧＣＴＴＧＧＡＴＡＴＴＣＧＣＡＴＧＧＧＣＣＴＡＣ－３’（配列番号：２０８）

正方向ＰＣＲプライマー（．ｆ２）　５’－ＴＧＧＡＧＡＣＡＡＴＡＴＣＣＣＴＧＡＧＧ－３’（配列番号：２０９）

逆方向ＰＣＲプライマー（．ｒ１）　５’－ＡＡＣＡＧＴＴＧＧＣＣＡＣＡＧＣＡＴＧＧＣＡＧＧ－３’（配列番号：２１０）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３５７１２配列から作成した。

ハイブリッド形成プローブ

５’－ＣＣＡＴＴＧＡＴＧＡＧＧＡＡＣＴＡＧＡＡＣＧＧＧＡＣＡＡＧＡＧＧＧＴＣＡＣＴＴＧＧＡＴＴＧＴＧＧＡＧ－３’（配列番号：２１１）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２７０遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２７０に対する全長ＤＮＡ配列［ここでＤＮＡ３９５１０－１１８１と命名］（配列番号：２０６）及びＰＲＯ２７０の誘導タンパク質配列が得られた。

ＤＮＡ３９５１０－１１８１の全ヌクレオチド配列をＦｉｇ７５（配列番号：２０６）に示す。クローンＤＮＡ３９５１０－１１８１は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３－５に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置８９１－８９３の停止コドンで終端する（Ｆｉｇ７５；配列番号：２０６）。予測されるポリペプチド前駆体は２９６アミノ酸長である（Ｆｉｇ７６）。クローンＤＮＡ３９５１０－１１８１はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３９２が付与された。

全長ＰＲＯ２７０のアミノ酸配列の分析によると、その部分がチオレドキシンタンパク質と有意な相同性を有していることがが示唆され、ＰＲＯ２７０がチオレドキシンファミリーの新規なメンバーでありうることを示している。

【０２６３】

実施例３５：ヒトＰＲＯ２７１をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３５７３７と命名した。このＤＮＡ３５７３７コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ２７１の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

ＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー１　５’－ＴＧＣＴＴＣＧＣＴＡＣＴＧＣＣＣＴＣ－３’（配列番号：２１４）

正方向ＰＣＲプライマー２　５’－ＴＴＣＣＣＴＴＧＴＧＧＧＴＴＧＧＡＧ－３’（配列番号：２１５）
正方向ＰＣＲプライマー３　５’－ＡＧＧＧＣＴＧＧＡＡＧＣＣＡＧＴＴＣ－３’（配列番号：２１６）
逆方向ＰＣＲプライマー１　５’－ＡＧＣＣＡＧＴＧＡＧＧＡＡＡＴＧＣＧ－３’（配列番号：２１７）
逆方向ＰＣＲプライマー２　５’－ＴＧＴＣＣＡＡＡＧＴＡＣＡＣＡＣＣＴＧＡＧＧ－３’（配列番号：２１８）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３５７３７配列から作成した。
ハイブリッド形成プローブ
５’－ＧＡＴＧＣＣＡＣＧＡＴＣＧＣＣＡＡＧＧＴＧＧＧＡＣＡＧＣＴＣＴＴＴＧＣＣＧＣＣＴＧＧＡＡＧ－３’（配列番号：２１９）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２７１遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児脳組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２７１に対する全長ＤＮＡ配列［ここでＤＮＡ３９４２３－１１８１と命名］（配列番号：２１２）及びＰＲＯ２７１の誘導タンパク質配列が得られた。
ＤＮＡ３９４２３－１１８１の全ヌクレオチド配列をＦｉｇ７７（配列番号：２１２）に示す。クローンＤＮＡ３９４２３－１１８１は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１０１－１０３に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１１８１－１１８３の停止コドンで終端する（Ｆｉｇ７７）。予測されるポリペプチド前駆体は３６０アミノ酸長である（Ｆｉｇ７８）。クローンＤＮＡ３９４２３－１１８１はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３８７が付与された。
全長ＰＲＯ２７１ポリペプチドのアミノ酸配列の分析によると、それが、プロテオグリカン連結タンパク質と有意な相同性を有していることがが示唆され、ＰＲＯ２７１が連結タンパク質相同体でありうることを示している。
【０２６４】
実施例３６：ヒトＰＲＯ２７２をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３６４６０と命名した。このＤＮＡ３６４６０コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ２７２の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー（．ｆ１）　５’－ＣＧＣＡＧＧＣＣＣＴＣＡＴＧＧＣＣＡＧＧ－３’（配列番号：２２２）
正方向ＰＣＲプライマー（．ｆ２）　５’－ＧＡＡＡＴＣＣＴＧＧＧＴＡＡＴＴＧＧ－３’（配列番号：２２３）
逆方向ＰＣＲプライマー　５’－ＧＴＧＣＧＣＧＧＴＧＣＴＣＡＣＡＧＣＴＣＡＴＣ－３’（配列番号：２２４）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３６４６０配列から作成した。
ハイブリッド形成プローブ
５’－ＣＣＣＣＣＣＴＧＡＧＣＧＡＣＧＣＴＣＣＣＣＣＡＴＧＡＴＧＡＣＧＣＣＣＡＣＧＧＧＡＡＣＴＴ－３’（配列番号：２２５）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２７２遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２７２に対する全長ＤＮＡ配列［ここでＤＮＡ４０６２０－１１８３と命名］（配列番号：２２０）及びＰＲＯ２７２の誘導タンパク質配列が得られた。

ＤＮＡ４０６２０－１１８３の全ヌクレオチド配列をＦｉｇ７９（配列番号：２２０）に示す。クローンＤＮＡ４０６２０－１１８３は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３５－３７に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１０１９－１０２１の停止コドンで終端する（Ｆｉｇ７９）。予測されるポリペプチド前駆体は３２８アミノ酸長である（Ｆｉｇ８０）。クローンＤＮＡ４０６２０－１１８３はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３８８が付与された。

全長ＰＲＯ２７２ポリペプチドのアミノ酸配列の分析によると、その部分が、ヒト及びマウスのレティキュロカルビンタンパク質と有意な相同性を有していることがことが示唆され、ＰＲＯ２７２が新規なレティキュロカルビンタンパク質であり得ることを示している。

【０２６５】

実施例３７：ヒトＰＲＯ２９４をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３５７３１と命名した。このＤＮＡ３５７３１コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ２９４の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

ＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー（．ｆ１）　５’－ＴＧＧＴＣＴＣＧＣＡＣＡＣＣＧＡＴＣ－３’（配列番号：２２８）

正方向ＰＣＲプライマー（．ｆ２）　５’－ＣＴＧＣＴＧＴＣＣＡＣＡＧＧＧＧＡＧ－３’（配列番号：２２９）

正方向ＰＣＲプライマー（．ｆ３）　５’－ＣＣＴＴＧＡＡＧＣＡＴＡＣＴＧＣＴＣ－３’（配列番号：２３０）

正方向ＰＣＲプライマー（．ｆ４）　５’－ＧＡＧＡＴＡＧＣＡＡＴＴＴＣＣＧＣＣ－３’（配列番号：２３１）

逆方向ＰＣＲプライマー（．ｒ１）　５’－ＴＴＣＣＴＣＡＡＧＡＧＧＧＣＡＧＣＣ－３’（配列番号：２３２）

逆方向ＰＣＲプライマー（．ｒ２）　５’－ＣＴＴＧＧＣＡＣＣＡＡＴＧＴＣＣＧＡＧＡＴＴＴＣ－３’（配列番号：２３３）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３５７３１配列から作成した。

ハイブリッド形成プローブ

５’－ＧＣＴＣＴＧＡＧＧＡＡＧＧＴＧＡＣＧＣＧＣＧＧＧＧＣＣＴＣＣＧＡＡＣＣＣＴＴＧＧＣＣＴＴＧ－３’（配列番号：２３４）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２９４遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児脳組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２９４に対する全長ＤＮＡ配列［ここでＤＮＡ４０６０４－１１８７と命名］（配列番号：２２６）及びＰＲＯ２９４の誘導タンパク質配列が得られた。
ＤＮＡ４０６０４－１１８７の全ヌクレオチド配列をＦｉｇ８１（配列番号：２２６）に示す。クローンＤＮＡ４０６０４－１１８７は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３９６－３９８に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置２０４６－２０４８の停止コドンで終端する（Ｆｉｇ８１）。予測されるポリペプチド前駆体は５５０アミノ酸長である（Ｆｉｇ８２）。クローンＤＮＡ４０６０４－１１８７はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３９４が付与された。
全長ＰＲＯ２９４ポリペプチドのアミノ酸配列の分析によると、その部分が様々なコラーゲンタンパク質と有意な相同性を有していることが示唆され、ＰＲＯ２９４がコラーゲン様分子であり得ることを示している。

【０２６６】

実施例３８：ヒトＰＲＯ２９５をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３５８１４と命名した。このＤＮＡ３５８１４コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ２９５の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
ＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー（．ｆ１）　５’－ＧＣＡＧＡＧＣＧＧＡ　ＧＡＴＧＣＡＧＣＧＧ　ＣＴＴＧ－３’（配列番号：２３８）
正方向ＰＣＲプライマー（．ｆ２）　５’－ＣＣＣＡＧＣＡＴＧＴＡＣＴＧＣＣＡＧ－３’（配列番号：２３９）
正方向ＰＣＲプライマー（．ｆ３）　５’－ＴＴＧＧＣＡＧＣＴＴＣＡＴＧＧＡＧＧ－３’（配列番号：２４０）
正方向ＰＣＲプライマー（．ｆ４）　５’－ＣＣＴＧＧＧＣＡＡＡＡＡＴＧＣＡＡＣ－３’（配列番号：２４１）
逆方向ＰＣＲプライマー（．ｒ１）　５’－ＣＴＣＣＡＧＣＴＣＣＴＧＧＣＧＣＡＣＣＴＣＣＴＣ－３’（配列番号：２４２）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３５８１４配列から作成した。
ハイブリッド形成プローブ
５’－ＧＧＣＴＣＴＣＡＧＣＴＡＣＣＧＣＧＣＡＧＧＡＧＣＧＡＧＧＣＣＡＣＣＣＴＣＡＡＴＧＡＧＡＴＧ－３’（配列番号：２４３）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２９５遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２９５に対する全長ＤＮＡ配列［ここでＤＮＡ３８２６８－１１８８と命名］（配列番号：２３５）及びＰＲＯ２９５の誘導タンパク質配列が得られた。
ＤＮＡ３８２６８－１１８８の全ヌクレオチド配列をＦｉｇ８３（配列番号：２３５）に示す。クローンＤＮＡ３８２６８－１１８８は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１５３－１５５に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１２０２－１２０４の停止コドンで終端する（Ｆｉｇ８３）。予測されるポリペプチド前駆体は３５０アミノ酸長である（Ｆｉｇ８４）。クローンＤＮＡ３８２６８－１１８８はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４２１が付与された。

全長ＰＲＯ２９５のアミノ酸配列の分析によると、その部分がインテグリンタンパク質と有意な相同性を有していることがが示唆され、ＰＲＯ２９５が新規なインテグリンでありうることを示している。

【０２６７】

実施例３９：ヒトＰＲＯ２９３をコードするｃＤＮＡクローンの単離

Ｓｗｉｓｓ－Ｐｒｏｔ公的タンパク質データベースからの約９５０の既知の分泌タンパク質からの細胞外ドメイン（ＥＣＤ）配列（あれば、分泌シグナルを含む）を、発現配列タグ（ＥＳＴ）データベースの検索に使用した。ＥＳＴデータベースは、公的ＥＳＴデータベース（例えば、ＧｅｎＢａｎｋ）及び企業のＥＳＴＤＮＡデータベース（例えば、ＬＩＦＥＳＥＱ（商品名）、Ｉｎｃｙｔｅ Ｐｈａｒｍａｃｅｕｔｉｃａｌｓ、Ｐａｌｏ Ａｌｔｏ, ＣＡ）を含む。検索は、コンピュータプログラムＢＬＡＳＴ又はＢＬＡＳＴ－２（Ａｌｔｓｃｈｕｌ等, Ｍｅｔｈｏｄｓ ｉｎ Ｅｎｚｙｍｏｌｏｇｙ ２６６：４６０－４８０（１９９６））を用いて、ＥＣＤタンパク質配列のＥＳＴ配列の６フレーム翻訳との比較として実施した。既知のタンパク質をコードせず、ＢＬＡＳＴスコア７０（９０の場合もある）又はそれ以上を持つ比較物は、プログラム「ｐｈｒａｐ」（Ｐｈｉｌ Ｇｒｅｅｎ, Ｕｎｉｖｅｒｓｉｔｙ ｏｆ Ｗａｓｈｉｎｇｔｏｎ, Ｓｅａｔｔｌｅ, ＷＡ；ｈｔｔｐ：／／ｂｏｚｅｍａｎ．ｍｂｔ．ｗａｓｈｉｎｇｔｏｎ．ｅｄｕ／ｐｈｒａｐ．ｄｏｃｓ／ｐｈｒａｐ．ｈｔｍｌ）で集団化してコンセンサスＤＮＡ配列に構築した。

上記分析で証明された、Ｔ０８２９４とここで命名された発現配列タグに基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ２９３の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー　５’－ＡＡＣＡＡＧＧＴＡＡＧＡＴＧＣＣＡＴＣＣＴＧ－３’（配列番号：２４６）

逆方向ＰＣＲプライマー　５’－ＡＡＡＣＴＴＧＴＣＧＡＴＧＧＡＧＡＣＣＡＧＣＴＣ－３’（配列番号：２４７）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つ発現配列タグから作成した。

ハイブリッド形成プローブ

５’－ＡＧＧＧＧＣＴＧＣＡＡＡＧＣＣＴＧＧＡＧＡＧＣＣＴＣＴＣＣＴＴＣＴＡＴＧＡＣＡＡＣＣＡＧＣ－３’（配列番号：２４８）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２９３遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児脳組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２９３に対する全長ＤＮＡ配列［ここでＤＮＡ３７１５１－１１９３と命名］（配列番号：２４４）及びＰＲＯ２９３の誘導タンパク質配列が得られた。

ＤＮＡ３７１５１－１１９３の全ヌクレオチド配列をＦｉｇ８５（配列番号：２４４）に示す。クローンＤＮＡ３７１５１－１１９３は、単一のオープンリーディングフレームを含み、ヌクレオチド位置８８１－８８３に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置３０１９の停止コドンで終端する（Ｆｉｇ８５；配列番号：２４４）。予測されるポリペプチド前駆体は７１３アミノ酸長である（Ｆｉｇ８６）。クローンＤＮＡ３７１５１－１１９３はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３９３が付与された。

全長ＰＲＯ２９３のアミノ酸配列の分析によると、その部分が、ＮＬＲＲタンパク質と有意な相同性を有していることが示唆され、ＰＲＯ２９３が新規なＮＬＲＲタンパク質であ

り得ることを示している。
【０２６８】
実施例４０：ヒトＰＲＯ２４７をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３３４８０と命名した。このＤＮＡ３３４８０コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ２４７の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＣＡＡＣＡＡＴＧＡＧＧＧＣＡＣＣＡＡＧＣ－３’（配列番号：２５１）
逆方向ＰＣＲプライマー　５’－ＧＡＴＧＧＣＴＡＧＧＴＴＣＴＧＧＡＧＧＴＴＣＴＧ－３’（配列番号：２５２）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つＤＮＡ３３４８０発現配列から作成した。
ハイブリッド形成プローブ
５’－ＣＡＡＣＣＴＧＣＡＧＧＡＧＡＴＴＧＡＣＣＴＣＡＡＧＧＡＣＡＡＣＡＡＣＣＴＣＡＡＧＡＣＣＡＴＣＧ－３’（配列番号：２５３）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２４７遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児脳組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２４７に対する全長ＤＮＡ配列［ここでＤＮＡ３５６７３－１２０１と命名］（配列番号：２４９）及びＰＲＯ２４７の誘導タンパク質配列が得られた。
ＤＮＡ３５６７３－１２０１の全ヌクレオチド配列をＦｉｇ８９（配列番号：２４９）に示す。クローンＤＮＡ３５６７３－１２０１は、単一のオープンリーディングフレームを含み、ヌクレオチド位置８０－８２に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１７１７の停止コドンで終端する（Ｆｉｇ８９；配列番号：２４９）。予測されるポリペプチド前駆体は５４６アミノ酸長である（Ｆｉｇ８８）。クローンＤＮＡ３５６７３－１２０１はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４１８が付与された。
全長ＰＲＯ２４７ポリペプチドのアミノ酸配列の分析によると、その部分がデンシン分子及びＫＩＡＡ０２３１と有意な相同性を有していることが示唆され、ＰＲＯ２４７が新規なロイシンリッチ反復タンパク質でありうることを示している。
【０２６９】
実施例４１：ヒトＰＲＯ３０２、ＰＲＯ３０３、ＰＲＯ３０４、ＰＲＯ３０７及びＰＲＯ３４３をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３５９５３、ＤＮＡ３５９５５、ＤＮＡ３５９５８、ＤＮＡ３７１６０及びＤＮＡ３０８９５と命名した。このＤＮＡ３５９５３コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ３０２の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
ＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー１　５’－ＧＴＣＣＧＣＡＡＧＧＡＴＧＣＣＴＡＣＡＴＧＴＴＣ－３’（配列番号：２６４）
正方向ＰＣＲプライマー２　５’－ＧＣＡＧＡＧＧＴＧＴＣＴＡＡＧＧＴＴＧ－３’（配列番号：２６５）

逆方向ＰＣＲプライマー　５’－ＡＧＣＴＣＴＡＧＡＣＣＡＡＴＧＣＣＡＧＣＴＴＣＣ－３’（配列番号：２６６）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３５９５３配列から作成した。
ハイブリッド形成プローブ
５’－ＧＣＣＡＣＣＡＡＣＴＣＣＴＧＣＡＡＧＡＡＣＴＴＣＴＣＡＧＡＡＣＴＧＣＣＣＣＴＧＧＴＣＡＴＧ－３’（配列番号：２６７）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３０２遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織から単離した。（ＬＩＢ２２８）
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ３０２に対する全長ＤＮＡ配列［ここでＤＮＡ４０３７０－１２１７と命名］（配列番号：２５４）及びＰＲＯ３０２の誘導タンパク質配列が得られた。
ＤＮＡ４０３７０－１２１７の全ヌクレオチド配列をＦｉｇ８９（配列番号：２５４）に示す。クローンＤＮＡ４０３７０－１２１７は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３４－３６に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１３９０－１３９２の停止コドンで終端する（Ｆｉｇ８９）。予測されるポリペプチド前駆体は４５２アミノ酸長である（Ｆｉｇ９０）。ＰＲＯ３０２タンパク質の様々な独特な態様をＦｉｇ９０に示す。クローンＤＮＡ４０３７０－１２１７は１９９７年１１月２１日にＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４８５が付与された。
このＤＮＡ３５９５５コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ３０３の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＧＧＧＧＡＡＴＴＣＡＣＣＣＴＡＴＧＡＣＡＴＴＧＣＣ－３’（配列番号：２６８）
逆方向ＰＣＲプライマー　５’－ＧＡＡＴＧＣＣＣＴＧＣＡＡＧＣＡＴＣＡＡＣＴＧＧ－３’（配列番号：２６９）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３５９５５配列から作成した。
ハイブリッド形成プローブ
５’－ＧＣＡＣＣＴＧＴＣＡＣＣＴＡＣＡＣＴＡＡＡＣＡＣＡＴＣＣＡＧＣＣＣＡＴＣＴＧＴＣＴＣＣＡＧＧＣＣＴＣ－３’（配列番号：２７０）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３０３遺伝子をコードするクローンを単離するのに使用した。

【０２７０】
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織（ＬＩＢ２５）から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ３０３に対する全長ＤＮＡ配列［ここでＤＮＡ４２５５１－１２１７と命名］（配列番号：２５６）及びＰＲＯ３０３の誘導タンパク質配列が得られた。
ＤＮＡ４２５５１－１２１７の全ヌクレオチド配列をＦｉｇ９１（配列番号：２５６）に示す。クローンＤＮＡ４２５５１－１２１７は、単一のオープンリーディングフレームを

含み、ヌクレオチド位置２０－２２に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置９６２－９６４の停止コドンで終端する（Ｆｉｇ９１）。予測されるポリペプチド前駆体は３１４アミノ酸長である（Ｆｉｇ９２）。ＰＲＯ３０３タンパク質の様々な独特の態様をＦｉｇ９２に示す。クローンＤＮＡ４２５５１－１２１７は１９９７年１１月２１日にＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４８３が付与された。

ＤＮＡ３５９５８コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ３０４の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー１　５’－ＧＣＧＧＡＡＧＧＧＣＡＧＡＡＴＧＧＧＡＣＴＣＣＡＡＧ－３’（配列番号：２７１）

正方向ＰＣＲプライマー２　５’－ＣＡＧＣＣＣＴＧＣＣＡＣＡＴＧＴＧＣ－３’（配列番号：２７２）

正方向ＰＣＲプライマー３　５’－ＴＡＣＴＧＧＧＴＧＧＴＣＡＧＣＡＡＣ－３’（配列番号：２７３）

逆方向ＰＣＲプライマー　５’－ＧＧＣＧＡＡＧＡＧＣＡＧＧＧＴＧＡＧＡＣＣＣＣＧ－３’（配列番号：２７４）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３５９５８配列から作成した。

ハイブリッド形成プローブ

５’－ＧＣＣＣＴＣＡＴＣＣＴＣＴＧＧＣＡＡＡＴＧＣＡＧＴＴＡＣＡＧＣＣＣＧＧＡＧＣＣＣＧＡＣ－３’（配列番号：２７５）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３０４遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡは２２週齢ヒト胎児脳組織（ＬＩＢ１５３）から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ３０４に対する全長ＤＮＡ配列［ここでＤＮＡ３９５２０－１２１７と命名］（配列番号：２５８）及びＰＲＯ３０４の誘導タンパク質配列が得られた。

ＤＮＡ３９５２０－１２１７の全ヌクレオチド配列をＦｉｇ９３（配列番号：２５８）に示す。クローンＤＮＡ３９５２０－１２１７は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３４－３６に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１７０２－１７０４の停止コドンで終端する（Ｆｉｇ９３）。予測されるポリペプチド前駆体は５５６アミノ酸長である（Ｆｉｇ９４）。ＰＲＯ３０４タンパク質の様々な独特の態様をＦｉｇ９４に示す。クローンＤＮＡ３９５２０－１２１７は１９９７年１１月２１日にＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４８２が付与された。

このＤＮＡ３７１６０コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ３０７の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー１　５’－ＧＧＧＣＡＧＧＧＡＴＴＣＣＡＧＧＧＣＴＣＣ－３’（配列番号：２７６）

正方向ＰＣＲプライマー２　５’－ＧＧＣＴＡＴＧＡＣＡＧＣＡＧＧＴＴＣ－３’（配列番号：２７７）

正方向ＰＣＲプライマー３　５’－ＴＧＡＣＡＡＴＧＡＣＣＧＡＣＣＡＧＧ－３’（配列番号：２７８）

逆方向ＰＣＲプライマー　５’－ＧＣＡＴＣＧＣＡＴＴＧＣＴＧＧＴＡＧＡＧＣＡＡＧ

－３’（配列番号：２７９）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３７１６０配列から作成した。
ハイブリッド形成プローブ
５’－ＴＴＡＣＡＧＴＧＣＣＣＣＣＴＧＧＡＡＡＣＣＣＡＣＴＴＧＧＣＣＴＧＣＡＴＡＣＣＧＣＣＴＣＣＣ－３’（配列番号：２８０）
【０２７１】
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３０７遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織（ＬＩＢ２２９）から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ３０７に対する全長ＤＮＡ配列［ここでＤＮＡ４１２２５－１２１７と命名］（配列番号：２６０）及びＰＲＯ３０７の誘導タンパク質配列が得られた。
ＤＮＡ４１２２５－１２１７の全ヌクレオチド配列をＦｉｇ９５（配列番号：２６０）に示す。クローンＤＮＡ４１２２５－１２１７は、単一のオープンリーディングフレームを含み、ヌクレオチド位置９２－９４に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１２４１－１２４３の停止コドンで終端する（Ｆｉｇ９５）。予測されるポリペプチド前駆体は３８３アミノ酸長である（Ｆｉｇ９６）。ＰＲＯ３０７タンパク質の様々な独特の様態をＦｉｇ９６に示す。クローンＤＮＡ４１２２５－１２１７は１９９７年１１月２１日にＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４９１が付与された。
このＤＮＡ３０８９５コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＣＲ３４３の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
正方向ＰＣＲプライマー　５’－ＣＧＴＣＴＣＧＡＧＣＧＣＴＣＣＡＴＡＣＡＧＴＴＣＣＣＴＴＧＣＣＣＣＡ－３’（配列番号：２８１）
逆方向ＰＣＲプライマー　５’－ＴＧＧＡＧＧＧＧＧＡＧＣＧＧＧＡＴＧＣＴＴＧＴＣＴＧＧＧＣＧＡＣＴＣＣＧＧＧＧＧＣＣＣＣＣＴＣＡＴＧ－３’（配列番号：２８２）
また、次のヌクレオチド配列を持つ合成オリゴヌクレオチドハイブリッド形成プローブをコンセンサスＤＮＡ３０８９５配列から作成した。
ハイブリッド形成プローブ
５’－ＣＣＣＴＣＡＧＡＣＣＣＴＧＣＡＧＡＡＧＣＴＧＡＡＧＧＴＴＣＣＴＡＴＣＡＴＣＧＡＣＴＣＧＧＡＡＧＴＣＴ
ＧＣＡＧＣＣＡＴＣＴＧＴＡＣＴＧＧＣＧＧＧＧＡＧＣＡＧＧＡＣＡＧＧＧＡＣＣＣＡＴＣＡＣＴＧＡＧＧＡＣＡＴＧＣＴＧＴＧＴＧＣＣＧＧＣＴＡＣＴ－３’（配列番号：２８３）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３４３遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織（ＬＩＢ２６）から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ３４３に対する全長ＤＮＡ配列［ここでＤＮＡ４３３１８－１２１７と命名］（配列番号：２６２）及びＰＲＯ３４３の誘導タンパク質配列が得られた。

ＤＮＡ４３３１８－１２１７の全ヌクレオチド配列をＦｉｇ９７（配列番号：２６２）に示す。クローンＤＮＡ４３３１８－１２１７は、単一のオープンリーディングフレームを含み、ヌクレオチド位置５３－５５に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１００４－１００６の停止コドンで終端する（Ｆｉｇ９７）。予測されるポリペプチド前駆体は３１７アミノ酸長である（Ｆｉｇ９８）。ＰＲＯ３４３タンパク質の様々な独特の態様をＦｉｇ９８に示す。クローンＤＮＡ４３３１８－１２１７は１９９７年１１月２１日にＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４８１が付与された。

【０２７２】

実施例４２：ヒトＰＲＯ３２８をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３５６１５と命名した。このＤＮＡ３５６１５コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ３２８の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

一対のＰＣＲプライマー（正及び逆）を合成した：

正方向ＰＣＲプライマー　５’－ＴＣＣＴＧＣＡＧＴＴＴＣＣＴＧＡＴＧＣ－３’（配列番号：２８６）

逆方向ＰＣＲプライマー　５’－ＣＴＣＡＴＡＴＴＧＣＡＣＡＣＣＡＧＴＡＡＴＴＣＧ－３’（配列番号：２８７）

また、次のヌクレオチド配列を持つ合成オリゴヌクレオチドハイブリッド形成プローブをコンセンサスＤＮＡ３５６１５配列から作成した。

ハイブリッド形成プローブ

５’－ＡＴＧＡＧＧＡＧＡＡＡＣＧＴＴＴＧＡＴＧＧＴＧＧＡＧＣＴＧＣＡＣＡＡＣＣＴＣＴＡＣＣＧＧＧ－３’（配列番号：２８８）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３２８遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ３２８に対する全長ＤＮＡ配列［ここでＤＮＡ４０５８７－１２３１と命名］（配列番号：２８４）及びＰＲＯ３２８の誘導タンパク質配列が得られた。

ＤＮＡ４０５８７－１２３１の全ヌクレオチド配列をＦｉｇ９９（配列番号：２８４）に示す。クローンＤＮＡ４０５８７－１２３１は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１５－１７に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１４０４－１４０６の停止コドンで終端する（Ｆｉｇ９９）。予測されるポリペプチド前駆体は４６３アミノ酸長である（Ｆｉｇ１００）。クローンＤＮＡ４０５８７－１２３１はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４３８が付与された。

全長ＰＲＯ３２８ポリペプチドのアミノ酸配列の分析によると、その部分が、ヒトグリア芽細胞腫タンパク質及びシステインリッチ分泌性タンパク質と有意な相同性を有していることが示唆され、ＰＲＯ３２８が新規なグリア芽細胞腫タンパク質又はシステインリッチ分泌性タンパク質でありうることが示される。

【０２７３】

実施例４３：ヒトＰＲＯ３３５、ＰＲＯ３３１又はＰＲＯ３２６をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３６６８５と命名した。このＤＮＡ３６６８５コンセンサス配列及びＩｎｃｙｔｅＥＳＴ配列番号２２２８９９０に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため

、及び２）ＰＲＯ３３５、ＰＲＯ３３１又はＰＲＯ３２６の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
ＰＲＯ３３５の決定のために正及び逆ＰＣＲプライマーを合成した：
正方向ＰＣＲプライマー　５’－ＧＧＡＡＣＣＧＡＡＴＣＴＣＡＧＣＴＡ－３’（配列番号：２９５）
正方向ＰＣＲプライマー　５’－ＣＣＴＡＡＡＣＴＧＡＡＣＴＧＧＡＣＣＡ－３’（配列番号：２９６）
正方向ＰＣＲプライマー　５’－ＧＧＣＴＧＧＡＧＡＣＡＣＴＧＡＡＣＣＴ－３’（配列番号：２９７）
正方向ＰＣＲプライマー　５’－ＡＣＡＧＣＴＧＣＡＣＡＧＣＴＣＡＧＡＡＣＡＧＴＧ－３’（配列番号：２９８）
逆方向ＰＣＲプライマー　５’－ＣＡＴＴＣＣＣＡＧＴＡＴＡＡＡＡＡＴＴＴＴＣ－３’（配列番号：２９９）
逆方向ＰＣＲプライマー　５’－ＧＧＧＴＣＴＴＧＧＴＧＡＡＴＧＡＧＧ－３’（配列番号：３００）
逆方向ＰＣＲプライマー　５’－ＧＴＧＣＣＴＣＴＣＧＧＴＴＡＣＣＡＣＣＡＡＴＧＧ－３’（配列番号：３０１）
また、次のヌクレオチド配列を持つ合成オリゴヌクレオチドハイブリッド形成プローブをＰＲＯ３３５の決定のために作成した。
ハイブリッド形成プローブ
５’－ＧＣＧＧＣＣＡＣＴＧＴＴＧＧＡＣＣＧＡＡＣＴＧＴＡＡＣＣＡＡＧＧＧＡＧＡＡＡＣＡＧＣＣＧＴＣＣＴＡＣ－３’（配列番号：３０２）
ＰＲＯ３３１の決定のために正及び逆ＰＣＲプライマーを合成した：
正方向ＰＣＲプライマー　５’－ＧＣＣＴＴＴＧＡＣＡＡＣＣＴＴＣＡＧＴＣＡＣＴＡＧＴＧＧ－３’（配列番号：３０３）
逆方向ＰＣＲプライマー　５’－ＣＣＣＣＡＴＧＴＧＴＣＣＡＴＧＡＣＴＧＴＴＣＣＣ－３’（配列番号：３０４）
また、次のヌクレオチド配列を持つ合成オリゴヌクレオチドハイブリッド形成プローブをＰＲＯ３３１の決定のために作成した。
ハイブリッド形成プローブ
５’－ＴＡＣＴＧＣＣＴＣＡＴＧＡＣＣＴＣＴＴＣＡＣＴＣＣＣＴＴＧＣＡＴＣＡＴＣＴＴＡＧＡＧＣＧＧ－３’（配列番号：３０５）
ＰＲＯ３２６の決定のための正及び逆ＰＣＲプライマーを合成した：
正方向ＰＣＲプライマー　５’－ＡＣＴＣＣＡＡＧＧＡＡＡＴＣＧＧＡＴＣＣＧＴＴＣ－３’（配列番号：３０６）
逆方向ＰＣＲプライマー　５’－ＴＴＡＧＣＡＧＣＴＧＡＧＧＡＴＧＧＧＣＡＣＡＡＣ－３’（配列番号：３０７）
また、次のヌクレオチド配列を持つ合成オリゴヌクレオチドハイブリッド形成プローブをＰＲＯ３２６の決定のため作成した。
ハイブリッド形成プローブ
５’－ＧＣＣＴＴＣＡＣＴＧＧＴＴＴＧＧＡＴＧＣＡＴＴＧＧＡＧＣＡＴＣＴＡＧＡＣＣＴＧＡＧＴＧＡＣＡＡＣＧＣ－３’（配列番号：３０８）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３３５、ＰＲＯ３３１又はＰＲＯ３２６遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織（ＰＲＯ３３５及びＰＲＯ３２６）及びヒト胎児脳（ＰＲＯ３３１）から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ３３５、ＰＲＯ３３１又

はＰＲＯ３２６に対する全長ＤＮＡ配列［ここでそれぞれ配列番号：２８９、２９１及び２９３と命名；それぞれＦｉｇ１０３、１０５及び１０７を参照］及びＰＲＯ３３５、ＰＲＯ３３１又はＰＲＯ３２６の誘導タンパク質配列が得られた。（それぞれＦｉｇ１０４、１０６及び１０８；それぞれ配列番号：２９０、２９２及び２９４）
全ヌクレオチド配列をＦｉｇ１０３、１０５及び１０７に示し、１９９８年６月２日、１９９７年１１月７日及び１９９７年１１月２１日にそれぞれＡＴＣＣに寄託された。
全長ＰＲＯ３３５、ＰＲＯ３３１又はＰＲＯ３２６ポリペプチドのアミノ酸配列の分析によると、その部分が、ＬＩＧ－１タンパク質と有意な相同性を有していることが示唆され、ＰＲＯ３３５、ＰＲＯ３３１及びＰＲＯ３２６が新規なＬＩＧ－１関連タンパク質でありうることを示している。

【０２７４】
実施例４４：ヒトＰＲＯ３３２をコードするｃＤＮＡクローンの単離
上記実施例１に記載したようにして実施されたＥＣＤ相同性検索に基づいて、ここでＤＮＡ３６６８８と命名したコンセンサス配列を構築した。ＤＮＡ３６６８８コンセンサス配列に基づいて、対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及びＰＲＯ３３２の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
５’－ＧＣＡＴＴＧＧＣＣＧＣＧＡＧＡＣＴＴＴＧＣＣ－３’（配列番号：３１１）
５’－ＧＣＧＧＣＣＡＣＧＧＴＣＣＴＴＧＧＡＡＡＴＧ－３’（配列番号：３１２）
プローブも合成した：
５’－ＴＧＧＡＧＧＡＧＣＴＣＡＡＣＣＴＣＡＧＣＴＡＣＡＡＣＣＧＣＡＴＣＡＣＣＡＧＣＣＣＡＣＡＧＧ－３’（配列番号：３１３）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３３２遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。（ＬＩＢ２２９）
上記のように単離したクローンのＤＮＡ配列決定により、ＤＮＡ４０９８２－１２３５に対する全長ＤＮＡ配列及びＰＲＯ３３２の誘導タンパク質配列が得られた。
ＤＮＡ４０９８２－１２３５の全ヌクレオチド配列をＦｉｇ１０９（配列番号：３０９）に示す。クローンＤＮＡ４０９８２－１２３５は、単一のオープンリーディングフレームを含む（Ｆｉｇ１０９に示すように、ヌクレオチド位置３４２－３４４に見かけの翻訳開始部位を持つ）。予測されるポリペプチド前駆体は６４２アミノ酸長であり、７２，０６７ダルトンの算定分子量（ｐｌ：６．６０）を有する。クローンＤＮＡ４０９８２－１２３５はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４３３が付与された。
全長配列のＢＬＡＳＴ及びＦａｓｔＡ配列アラインメント分析法に基づいて、ＰＲＯ３３２は、例えば、様々な種のフィブロモジュリン及びフィブロモジュリン前駆体配列（ＦＭＯＤ＿ＢＯＶＩＮ，　ＦＭＯＤ＿ＣＨＩＣＫ，　ＦＭＯＤ＿ＲＡＴ，　ＦＭＯＤ＿ＭＯＵＳＥ，　ＦＭＯＤ＿ＨＵＭＡＮ，　Ｐ＿Ｒ３６７７３）、オステオモジュリン配列（ＡＢ００１１４＿１，　ＡＢ００７８４８＿１）、デコリン配列（ＣＦＵ８３１４１＿１，　ＯＣＵ０３３９４＿１，　Ｐ＿Ｒ４２２６６，　Ｐ＿Ｒ４２２６７，　Ｐ＿Ｒ４２２６０，　Ｐ＿Ｒ８９４３９）、ケラチン硫酸プロテオグリカン（ＢＴＵ４８３６０＿１，　ＡＦ０２２８９０１）、角膜プロテオグリカン（ＡＦ０２２２５６＿１）、及び骨／軟骨プロテオグリカン及びプロテオグリカン前駆体（ＰＧＳ１＿ＢＯＶＩＮ，　ＰＧＳ２＿ＭＯＵＳＥ，　ＰＧＳ２＿ＨＵＭＡＮ）を含む一連の既知のプロテオグリカン配列と約３０－４０％のアミノ酸配列同一性を示している。

【０２７５】

実施例４５：ヒトＰＲＯ３３４をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ３３４の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
ＰＲＯ３３４の決定に対する正及び逆ＰＣＲプライマーを合成した：
正方向ＰＣＲプライマー　５’－ＧＡＴＧＧＴＴＣＣＴＧＣＴＣＡＡＧＴＧＣＣＣＴＧ－３’（配列番号：３１６）
逆方向ＰＣＲプライマー　５’－ＴＴＧＣＡＣＴＴＧＴＡＧＧＡＣＣＣＡＣＧＴＡＣＧ－３’（配列番号：３１７）
また、次のヌクレオチド配列を持つ合成オリゴヌクレオチドハイブリッド形成プローブをＰＲＯ３３４の決定のために作成した。
ハイブリッド形成プローブ
５’－ＣＴＧＡＴＧＧＧＡＧＧＡＣＣＴＧＴＧＴＡＧＡＴＧＴＴＧＡＴＧＡＡＴＧＴＧＣＴＡＣＡＧＧＡＡＧＡＧＣＣ－３’（配列番号：３１８）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３３４遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡクローンを単離するために使用したヒト胎児腎臓ｃＤＮＡライブラリは、Ｉｎｖｉｔｒｏｇｅｎ，　Ｓａｎ　Ｄｉｅｇｏ，　ＣＡからのもののような商業的に利用できる試薬を使用して標準的な方法で作成した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ３３４に対する全長ＤＮＡ配列［ここでＤＮＡ４１３７９－１２３６と命名］（配列番号：３１４）及びＰＲＯ３３４の誘導タンパク質配列が得られた。
ＤＮＡ４１３７９－１２３６（またＵＮＱ２９５と表す）の全ヌクレオチド配列をＦｉｇ１０９（配列番号：３１４）に示す。クローンＤＮＡ４１３７９－１２３６は、単一のオープンリーディングフレームを含み、ヌクレオチド位置２０３－２０５に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１７３０－１７３２の停止コドンで終端する（Ｆｉｇ１０９）。予測されるポリペプチド前駆体は５０９アミノ酸長である（Ｆｉｇ１１０）。クローンＤＮＡ４１３７９－１２３６はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４８８が付与された。　全長ＰＲＯ３３４ポリペプチドのアミノ酸配列の分析によると、それの部分が、フィブリン及びフィブリリンタンパク質と有意な相同性を有していることが示唆され、ＰＲＯ３３４がＥＧＦタンパク質ファミリーの新規なメンバーでありうることを示している。
【０２７６】
実施例４６：ヒトＰＲＯ３４６をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いてコンセンサスＤＮＡ配列を組み立てた。特に、コンセンサス配列を、ここでＤＮＡ３８２４０と命名した。このＤＮＡ３８２４０コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ３４６の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓から単離した。ｃＤＮＡクローンの単離に用いたｃＤＮＡライブラリは、市販試薬（例えば、Ｉｎｖｉｔｒｏｇｅｎ，　Ｓａｎ　Ｄｉｅｇｏ，　ＣＡ；　Ｃｌｏｎｔｅｃｈ等）を用いる標準的な方法によって作成した。ｃＤＮＡは、ＮｏｔＩ部位を含むオリゴｄＴでプライムし、平滑末端でＳａｌＩヘミキナーゼアダプターに結合させ、ＮｏｔＩで切断し、ゲル電気泳動でおよそのサイズ分類し、そして適切なクローニングベクター（ｐＲＫＢ又はｐＲＫＤ等；ｐＲＫ５ＢはＳｆ

ｉＩ部位を含まないｐＲＫ５Ｄの前駆体である；Ｈｏｌｍｅｓ等，Ｓｃｉｅｎｃｅ，２５３：１２７８－１２８０（１９９１）参照）に、独特のＸｈｏＩ及びＮｏｔＩ部位において、所定の方向でクローニングした。
ｃＤＮＡクローンは全体を配列決定した。ＤＮＡ４４１６７－１２４３の全ヌクレオチド配列をＦｉｇ１１１（配列番号：３１９）に示す。クローンＤＮＡ４４１６７－１２４３は、単一のオープンリーディングフレームを含み、ヌクレオチド位置６４－６６に見かけの翻訳開始部位を持つ（Ｆｉｇ１１３；配列番号：３１９）。予測されるポリペプチド前駆体は４５０アミノ酸長である。クローンＤＮＡ４４１６７－１２４３はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４３４が付与された（ＤＮＡ４４１６７－１２４３と命名）。
全長配列の（ＡＬＩＧＮコンピュータプログラムを使用する）ＢＬＡＳＴ、ＢＬＡＳＴ－２及びＦａｓｔＡ配列アラインメント分析法に基づいて、ＰＲＯ３４６は、ガン胎児性抗原とアミノ酸配列同一性（２８％）を示した。
上記の手順において使用されたオリゴヌクレオチド配列は次の通りであった：ＯＬＩ２６９１（３８２４０．ｆ１）
５’－ＧＡＴＣＣＴＧＴＣＡＣＡＡＡＧＣＣＡＧＴＧＧＴＧＣ－３’（配列番号：３２１）
ＯＬＩ２６９３（３８２４０．ｒ１）
５’－ＣＡＣＴＧＡＣＡＧＧＧＴＴＣＣＴＣＡＣＣＣＡＧＧ－３’（配列番号：３２２）
ＯＬＩ２６９２（３８２４０．ｐ１）
５’－ＣＴＣＣＣＴＣＴＧＧＧＣＴＧＴＧＧＡＧＴＡＴＧＴＧＧＧＧＡＡＣＡＴＧＡＣＣＣＴＧＡＣＡＴＧ－３’（配列番号：３２３）
【０２７７】
実施例４７：ヒトＰＲＯ２６８をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３５６９８と命名した。このＤＮＡ３５６９８コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ２６８の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
正及び逆ＰＣＲプライマーを合成した：
正方向ＰＣＲプライマー１　５’－ＴＧＡＧＧＴＧＧＧＣＡＡＧＣＧＧＣＧＡＡＡＴＧ－３’（配列番号：３２６）
正方向ＰＣＲプライマー２　５’－ＴＡＴＧＴＧＧＡＴＣＡＧＧＡＣＧＴＧＣＣ－３’（配列番号：３２７）
正方向ＰＣＲプライマー３　５’－ＴＧＣＡＧＧＧＴＴＣＡＧＴＣＴＡＧＡＴＴＧ－３’（配列番号：３２８）
逆方向ＰＣＲプライマー　５’－ＴＴＧＡＡＧＧＡＣＡＡＡＧＧＣＡＡＴＣＴＧＣＣＡＣ－３’（配列番号：３２９）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３５６９８配列から作成した。
ハイブリッド形成プローブ
５’－ＧＧＡＧＴＣＴＴＧＣＡＧＴＴＣＣＣＣＴＧＧＣＡＧＴＣＣＴＧＧＴＧＣＴＧＴＴＧＣＴＴＴＧＧＧ－３’（配列番号：３３０）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２６８遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肺組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ２６８に対する全長ＤＮ

Ａ配列［ここでＤＮＡ３９４２７－１１７９と命名］（配列番号：３２４）及びＰＲＯ２６８の誘導タンパク質配列が得られた。
ＤＮＡ３９４２７－１１７９の全ヌクレオチド配列をＦｉｇ１１３（配列番号：３２４）に示す。クローンＤＮＡ３９４２７－１１７９は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１３－１５に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置８５３－８５５の停止コドンで終端する（Ｆｉｇ１１３）。予測されるポリペプチド前駆体は２８０アミノ酸長である（Ｆｉｇ１１４）。クローンＤＮＡ３９４２７－１１７９はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９３９５が付与された。
全長ＰＲＯ２６８のアミノ酸配列の分析によると、それが、プロテインジスルフィドイソメラーゼと有意な相同性を有していることが示唆され、ＰＲＯ２６８が新規なプロテインジスルフィドイソメラーゼでありうることが示される。

【０２７８】

実施例４８：ヒトＰＲＯ３３０をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ３５７３０と命名した。このＤＮＡ３５７３０コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ３３０の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
正及び逆ＰＣＲプライマー合成した：
正方向ＰＣＲプライマー１　５’－ＣＣＡＧＧＣＡＣＡＡＴＴＴＣＣＡＧＡ－３’（配列番号：３３３）
正方向ＰＣＲプライマー２　５’－ＧＧＡＣＣＣＴＴＣＴＧＴＧＴＧＣＣＡＧ－３’（配列番号：３３４）
逆方向ＰＣＲプライマー１　５’－ＧＧＴＣＴＣＡＡＧＡＡＣＴＣＣＴＧＴＣ－３’（配列番号：３３５）
逆方向ＰＣＲプライマー２　５’－ＡＣＡＣＴＣＡＧＣＡＴＴＧＣＣＴＧＧＴＡＣＴＴＧ－３’（配列番号：３３６）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ３５７３０配列から作成した。
ハイブリッド形成プローブ
５’－ＧＧＧＣＡＣＡＴＧＡＣＴＧＡＣＣＴＧＡＴＴＴＡＴＧＣＡＧＡＧＡＡＡＧＡＧＣＴＧＧＴＧＣＡＧ－３’（配列番号：３３７）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３３０遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ３３０に対する全長ＤＮＡ配列［ここでＤＮＡ４０６０３－１２３２と命名］（配列番号：３３１）及びＰＲＯ３３０の誘導タンパク質配列が得られた。
ＤＮＡ４０６０３－１２３２の全ヌクレオチド配列をＦｉｇ１１５（配列番号：３３１）に示す。クローンＤＮＡ４０６０３－１２３２は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１６７－１６９に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１７６６－１７６８の停止コドンで終端する（Ｆｉｇ１１５）。予測されるポリペプチド前駆体は５３３アミノ酸長である（Ｆｉｇ１１６）。クローンＤＮＡ４０６０３－１２３２はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４８６が１９９７年１１月２１日に付与された。
全長ＰＲＯ３３０ポリペプチドのアミノ酸配列の分析によると、その部分が、マウスのプロリル４－ヒドロキシラーゼαサブユニットタンパク質と有意な相同性を有していること

が示唆され、ＰＲＯ３３０が新規なプロリル４－ヒドロキシラーゼαサブユニットポリペプチドでありうることが示される。

【０２７９】

実施例４９：ヒトＰＲＯ３１０をコードするｃＤＮＡクローンの単離

上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。コンセンサス配列を、ここでＤＮＡ４０５５３と命名した。このＤＮＡ４０５５３コンセンサス配列に基づいて、１）対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及び２）ＰＲＯ３１０の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。

性及び逆ＰＣＲプライマーを合成した：

正方向ＰＣＲプライマー１　５’－ＴＣＣＣＣＡＡＧＣＣＧＴＴＣＴＡＧＡＣＧＣＧＧ－３’（配列番号：３４２）

正方向ＰＣＲプライマー２　５’－ＣＴＧＧＴＴＣＴＴＣＣＴＴＧＣＡＣＧ－３’（配列番号：３４３）

逆方向ＰＣＲプライマー　５’－ＧＣＣＣＡＡＡＴＧＣＣＣＴＡＡＧＧＣＧＧＴＡＴＡＣＣＣＣ－３’（配列番号：３４４）

また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサスＤＮＡ４０５５３配列から作成した。

ハイブリッド形成プローブ

５’－ＧＧＧＴＧＴＧＡＴＧＣＴＴＧＧＡＡＧＣＡＴＴＴＴＣＴＧＴＧＣＴＴＴＧＡＴＣＡＣＴＡＴＧＣＴＡＧＧＡＣ－３’（配列番号：３４５）

全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３１０遺伝子をコードするクローンを単離するのに使用した。

ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、ＰＲＯ３１０に対する全長ＤＮＡ配列［ここでＤＮＡ４３０４６－１２２５と命名］（配列番号：３４０）及びＰＲＯ３１０の誘導タンパク質配列が得られた（配列番号：３４１）。

ＤＮＡ４３０４６－１２２５の全ヌクレオチド配列をＦｉｇ１１９（配列番号：３４０）に示す。クローンＤＮＡ４３０４６－１２２５は、単一のオープンリーディングフレームを含み、ヌクレオチド位置８１－８３に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置１０３５－１０３７の停止コドンで終端する（Ｆｉｇ１１９）。予測されるポリペプチド前駆体は３１８アミノ酸長であり（Ｆｉｇ１２０）、３６，３８２ダルトンの算定分子量を有する。クローンＤＮＡ４３０４６－１２２５はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４８４が付与された。

全長ＰＲＯ３１０ポリペプチドのアミノ酸配列の分析によると、その部分が線虫タンパク質及びフリンジと相同性を有していることが示唆され、ＰＲＯ３１０が発達に関与しうること示される。

【０２８０】

実施例５０：ヒトＰＲＯ３３９をコードするｃＤＮＡクローンの単離

発現配列タグ（ＥＳＴ）ＤＮＡデータベース（ＬＩＦＥＳＥＱ（商品名）、Ｉｎｃｙｔｅ Ｐｈａｒｍａｃｅｕｔｉｃａｌｓ、Ｐａｌｏ Ａｌｔｏ，ＣＡ）は検索され、ＥＳＴｓが同定された。上記実施例１に記載したように、ｐｈｒａｐを用いてＩｎｃｙｔｅクローン及びコンセンサス配列のアセンブリを形成した。

アセンブリ形成コンセンサス配列に基づいて正及び逆ＰＣＲプライマーを合成した：

正方向ＰＣＲプライマー１　５’－ＧＧＧＡＴＧＣＡＧＧＴＧＧＴＧＴＣＴＣＡＴＧＧＧＧ－３’（配列番号：３４６）

正方向ＰＣＲプライマー２　５’－ＣＣＣＴＣＡＴＧＴＡＣＣＧＧＣＴＣＣ－３’（配列

番号：３４７）
正方向ＰＣＲプライマー３　５’－ＧＴＧＴＧＡＣＡＣＡＧＣＧＴＧＧＧＣ－３’（配列番号：４３）
正方向ＰＣＲプライマー４　５’－ＧＡＣＣＧＧＣＡＧＧＣＴＴＣＴＧＣＧ－３’（配列番号：４４）
逆方向ＰＣＲプライマー１　５’－ＣＡＧＣＡＧＣＴＴＣＡＧＣＣＡＣＣＡＧＧＡＧＴＧＧ－３’（配列番号：４５）
逆方向ＰＣＲプライマー２　５’－ＣＴＧＡＧＣＣＧＴＧＧＧＣＴＧＣＡＧＴＣＴＣＧＣ－３’（配列番号：４６）
また、合成オリゴヌクレオチドハイブリッド形成プローブを、次のヌクレオチド配列を持つコンセンサス配列から作成した。
ハイブリッド形成プローブ
５’－ＣＣＧＡＣＴＡＣＧＡＣＴＧＧＴＴＣＴＴＣＡＴＣＡＴＧＣＡＧＧＡＴＧＡＣＡＣＡＴＡＴＧＴＧＣ－３’（配列番号：４７）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ３３９遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児肝臓組織から単離した。
ｃＤＮＡクローンの全体を配列決定した。ＤＮＡ４３４６６－１２２５の全ヌクレオチド配列をＦｉｇ１１７（配列番号：３３８）に示す。クローンＤＮＡ４３４６６－１２２５は、単一のオープンリーディングフレームを含み、ヌクレオチド位置３３３－３３５に見かけの翻訳開始部位を持ち、そしてヌクレオチド位置２６４９－２６５１の停止コドンで終端する（Ｆｉｇ１１７；配列番号：３３８）。予測されるポリペプチド前駆体は７７２アミノ酸長であり、８６，２２６ダルトンの算定分子量を有する。クローンＤＮＡ４３４６６－１２２５はＡＴＣＣに寄託され、ＡＴＣＣ寄託番号２０９４９０が付与された。
全長配列の（ＡＬＩＧＮコンピュータプログラムを使用する）ＢＬＡＳＴ及びＦａｓｔＡ配列アラインメント分析法に基づいて、ＰＲＯ３３９は、線虫タンパク質及びコラーゲン様ポリマー配列並びにフリンジと相同性を有しており、ＰＲＯ３３９が発達又は組織成長に関与しうることを示した。
【０２８１】
実施例５１：ヒトＰＲＯ２４４をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列に基づいて、対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及びＰＲＯ２４４の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
５’－ＴＴＣＡＧＣＴＴＣＴＧＧＧＡＴＧＴＡＧＧＧ－３’（３０９２３．ｆ１）（配列番号：３７８）
５’－ＴＡＴＴＣＣＴＡＣＣＡＴＴＴＣＡＣＡＡＡＴＣＣＧ－３’（３０９２３．ｒ１）（配列番号：３７９）
また、プローブも合成した。
５’－ＧＧＡＧＧＡＣＴＧＴＧＣＣＡＣＣＡＴＧＡＧＡＧＡＣＴＣＴＴＣＡＡＡＣＣＣＡＡＧＧＣＡＡＡＡＴＴＧＧ－３’（配列番号：３８０）　全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２４４遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織から単離した。

上記のように単離したクローンのＤＮＡ配列決定により、全長ＤＮＡ配列及びＰＲＯ２４４の誘導タンパク質配列が得られた。
ＰＲＯ２４４の全ヌクレオチド配列をＦｉｇ１２１（配列番号：３７６）に示す。クローンＤＮＡ３５６６８－１１７１は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１０６－１０８に見かけの翻訳開始部位を持つ（Ｆｉｇ１２１）。予測されるポリペプチド前駆体は２１９アミノ酸長である。クローンＤＮＡ３５６６８－１１７１はＡＴＣＣに寄託され（ＤＮＡ３５６６３－１１７１と命名）、ＡＴＣＣ寄託番号２０９３７１が付与された。タンパク質は細胞質ドメイン（ａａ１－２０）、膜貫通ドメイン（ａａ２１－４６）、及び細胞外ドメイン（ａａ４７－２１９）を、ａａ５５－２０６のＣレクチンドメインと共に有している。
全長配列のＢＬＡＳＴ及びＦａｓｔＡ配列アラインメント分析法に基づいて、ＰＲＯ２４４は、肝臓レクチンニワトリ（４３％）、ＨＩＣｈｐ１２０結合Ｃ型レクチン（４２％）、マクロファージレクチン２（ＨＵＭＨＭＬ２－１，　４１％）、及び配列ＰＲ３２１８８（４４％）と顕著なアミノ酸配列同一性を示す。
【０２８２】
実施例５２：ハイブリッド形成プローブとしてのＰＲＯポリペプチドコード核酸の使用
以下の方法は、ＰＲＯポリペプチドをコードするヌクレオチド配列のハイブリッド形成プローブとしての使用を記載する。
ここに開示する対象のＰＲＯポリペプチドのコード化配列を含むＤＮＡは、ヒト組織ｃＤＮＡライブラリ又はヒト組織ゲノムライブラリにおける同種ＤＮＡ類（ＰＲＯポリペプチドの天然発生変異体をコードするものなど）のスクリーニングのためのプローブとして又はプローブを調製する基礎として用いることができる。
何れかのライブラリＤＮＡを含むフィルターのハイブリッド形成及び洗浄は、以下の高い緊縮条件で実施した。放射標識ＰＲＯポリペプチドコード核酸誘導プローブのフィルターへのハイブリッド形成は、５０％ホルムアミド、５ｘＳＳＣ、０．１％ＳＤＳ、０．１％ピロリン酸ナトリウム、５０ｍＭリン酸ナトリウム、ｐＨ６．８、２ｘデンハード液、及び１０％デキストラン硫酸の溶液中で、４２℃において２０時間行った。フィルターの洗浄は、０．１ｘＳＳＣ及び０．１％ＳＤＳの水溶液中、４２℃で行った。
次いで、全長天然配列ＰＲＯポリペプチドをコードするＤＮＡと所望の配列同一性を有するＤＮＡは、この分野で知られた標準的な方法を用いて同定できる。
【０２８３】
実施例５３：大腸菌におけるＰＲＯポリペプチドの発現
この実施例は、大腸菌における組み換え発現による所望のＰＲＯポリペプチドの非グリコシル化形態の調製を例示する。
所望のＰＲＯポリペプチドをコードするＤＮＡ配列は、選択されたＰＣＲプライマーを用いて最初に増幅される。プライマーは、選択された発現ベクターの制限酵素部位に対応する制限酵素部位を持たなければならない。種々の発現ベクターを用いることができる。好適なベクターの例は、ｐＢＲ３２２（大腸菌から誘導されたもの；Ｂｏｌｉｖａｒ等，Ｇｅｎｅ，　２：９５　（１９７７）を参照）であり、アンピシリン及びテトラサイクリン耐性のための遺伝子を含む。ベクターは、制限酵素で消化され、脱リン酸化される。ＰＣＲ増幅された配列を、次いで、ベクターに結合させる。ベクターは、好ましくは抗生物質耐性遺伝子、ｔｒｐプロモーター、ｐｏｌｙｈｉｓリーダー（最初の６つのＳＴＩＩコドン、ｐｏｌｙｈｉｓ配列、及びエンテロキナーゼ切断部位を含む）、特異的ＰＲＯポリペプチドコード化領域、ラムダ転写終結区、及びａｒｇＵ遺伝子を含む。
ライゲーション混合物は、次いで、上掲のＳａｍｂｒｏｏｋ等に記載された方法を用いた選択した大腸菌株の形質転換に使用される。形質転換体は、それらのＬＢプレートで成長する能力により同定され、次いで抗生物質耐性コロニーが選択される。プラスミドＤＮＡが単離され、制限分析及びＤＮＡ配列決定で確認される。
選択されたクローンは、抗生物質を添加したＬＢブロスなどの液体培地で終夜成長させることができる。終夜培養は、続いて大規模な培養の播種に用いられる。次に細胞を所望の

光学密度まで成長させ、その間に発現プロモーターが作動する。
更に数時間の細胞の培養後、細胞を遠心分離により採集することができる。遠心分離で得られた細胞ペレットは、この分野で知られた種々の試薬を用いて可溶化され、次いで可溶化ＰＲＯポリペプチドを金属キレート化カラムを用いてタンパク質を緊密に結合させる条件下で精製することができる。
以下の手法を用いて、大腸菌においてポリＨｉｓタグ形態でＰＲＯ１８７、ＰＲＯ３１７、ＰＲＯ３０１、ＰＲＯ２２４及びＰＲＯ２３８が首尾よく発現された。ＰＲＯ１８７、ＰＲＯ３１７、ＰＲＯ３０１、ＰＲＯ２２４又はＰＲＯ２３８をコードするＤＮＡを、選択したＰＣＲプライマーを用いて最初に増幅した。プライマーは、選択された発現ベクターの制限酵素部位に対応する制限酵素部位、及び効率的で信頼性のある翻訳開始、金属キレートカラムでの迅速な精製、及びエンテロキナーゼでのタンパク質分解的除去をもたらす他の有用な配列を含んでいた。次いでＰＣＲ増幅されたポリ－Ｈｉｓタグ配列を発現ベクターに結合させ、それを株５２（Ｗ３１１０ ｆｕｈＡ（ｔｏｎＡ） ｌｏｎ ｇａｌＥ ｒｐｏＨｔｓ（ｈｔｐＲｔｓ） ｃｌｐＰ（ｌａｃＩｑ））に基づく大腸菌宿主の形質転換に使用した。形質転換体は、最初に５０ｍｇ／ｍｌのカルベニシリンを含有するＬＢ中、３０℃で振盪しながら３－５のＯ．Ｄ．６００に達するまで成長させた。ついで培養をＣＲＡＰ培地（３．５７ｇの（NH_4）$_2$ SO_4、０．７１ｇのクエン酸ナトリウム・２Ｈ２Ｏ、１．０７ｇのＫＣｌ、５．３６ｇのＤｉｆｃｏ酵母抽出物、５００ｍＬ水中の５．３６ｇのＳｈｅｆｆｉｅｌｄ ｈｙｃａｓｅ ＳＦ、並びに１１０ｍＭのＭＰＯＳ、ｐＨ７．３、０．５５％（ｗ／ｖ）のグルコース及び７ｍＭのＭｇSO_4の混合で調製）中に５０－１００倍希釈し、３０℃で振盪させながら約２０－３０時間成長させた。試料を取り出してＳＤＳ－ＰＡＧＥ分析により発現を確認し、バルク培養を遠心分離して細胞をペレット化した。細胞ペレットを精製及び再折りたたみまで凍結させた。
【０２８４】
０．５から１Ｌの発酵（６－１０ｇペレット）からの大腸菌ペーストを、７Ｍのグアニジン、２０ｍＭのトリス、ｐＨ８バッファー中で１０容量（ｗ／ｖ）で再懸濁させた。固体硫酸ナトリウム及びテトラチオン酸ナトリウムを添加して最終濃度を各々０．１Ｍ及び０．０２Ｍとし、溶液を４℃で終夜撹拌した。この工程により、亜硫酸によりブロックされた全てのシステイン残基を持つ変性タンパク質がもたらされた。溶液をＢｅｃｋｍａｎ Ｕｌｔｒａｃｅｎｔｒｉｆｕｇｅ中で４０，０００ｒｐｍで３０分間遠心分離した。上清を金属キレートカラムバッファー（６Ｍのグアニジン、２０ｍＭのトリス、ｐＨ７．４）の３－５容量で希釈し、０．２２ミクロンフィルターを通して濾過して透明化した。透明化抽出物を、金属キレートカラムバッファーで平衡化させた５ｍｌのＱｉａｇｅｎ Ｎｉ－ＮＴＡ金属キレートカラムに充填した。カラムを５０ｍＭのイミダゾール（Ｃａｌｂｉｏｃｈｅｍ，Ｕｔｒｏｌ ｇｒａｄｅ）を含む添加バッファー、ｐＨ７．４で洗浄した。タンパク質を２５０ｍＭのイミダゾールを含有するバッファーで溶離した。所望のタンパク質を含有する画分をプールし、４℃で保存した。タンパク質濃度は、そのアミノ酸配列に基づいて計算した吸光係数を用いて２８０ｎｍにおけるその吸収により算定した。
試料を、２０ｍＭのトリス、ｐＨ８．６、０．３ＭのＮａＣｌ、２．５Ｍの尿素、５ｍＭのシステイン、２０ｍＭのグリシン及び１ｍＭのＥＤＴＡからなる新たに調製した再折りたたみバッファー中に徐々に希釈することによりタンパク質を再折りたたみさせた。リフォールディング容量は、最終的なタンパク質濃度が５０～１００マイクログラム／ｍｌとなるように選択した。リフォールディング溶液を４℃で１２－３６時間ゆっくり撹拌した。リフォールディング反応はＴＦＡを最終濃度０．４％（約３のｐＨ）で添加することにより停止させた。タンパク質をさらに精製する前に、溶液を０．２２ミクロンフィルターを通して濾過し、アセトニトリルを最終濃度２－１０％になるまで添加した。再折りたたみされたタンパク質を、Ｐｏｒｏｓ Ｒ１／Ｈ逆相カラムで、０．１％ＴＦＡの移動バッファーと１０～８０％のアセトニトリル勾配での溶離を用いてクロマトグラフにかけた。Ａ２８０吸収を持つ画分のアリコートをＳＤＳポリアクリルアミドゲルで分析し、相同な再折りたたみされたタンパク質を含有する画分をプールした。一般的に、殆どのタンパク

質の正しく再折りたたみされた種は、これらの種が最もコンパクトであり、その疎水性内面が逆相樹脂との相互作用から遮蔽されているので、アセトニトリルの最低濃度で溶離される。凝集した種は通常、より高いアセトニトリル濃度で溶離される。タンパク質の誤って折りたたまれた形態を所望の形態から除くのに加えて、逆相工程は試料からエンドトキシンも除去する。

所望の折りたたまれたＰＲＯ１８７、ＰＲＯ３１７，ＰＲＯ３０１，ＰＲＯ２２４及びＰＲＯ２３８タンパク質を含有する画分をそれぞれプールし、溶液に向けた窒素の弱い気流を用いてアセトニトリルを除去した。タンパク質を、透析及び調製バッファーで平衡化したＧ２５　Ｓｕｐｅｒｆｉｎｅ（Ｐｈａｒｍａｃｉａ）樹脂でのゲル濾過又は滅菌濾過により、０．１４Ｍの塩化ナトリウム及び４％のマンニトールを含む２０ｍＭのＨｅｐｅｓ、ｐＨ６．８に処方した。

【０２８５】

実施例５４：哺乳動物細胞でのＰＲＯポリペプチドの発現

この実施例は、哺乳動物細胞における組み換え発現による所望のＰＲＯポリペプチドのグリコシル化形態の調製を例示する。

発現ベクターとしてベクターｐＲＫ５（１９８９年３月１５日公開のＥＰ　３０７，２４７参照）を用いた。場合によっては、ＰＲＯポリペプチドコードＤＮＡを選択した制限酵素を持つｐＲＫ５に結合させ、上掲のＳａｍｂｒｏｏｋ等に記載されたようなライゲーション方法を用いてＰＲＯポリペプチドＤＮＡを挿入させる。得られたベクターは、ｐＲＫ５－ＰＲＯポリペプチドと呼ばれる。

一実施態様では、選択された宿主細胞は２９３細胞とすることができる。ヒト２９３細胞（ＡＴＣＣ　ＣＣＬ　１５７３）は、ウシ胎児血清及び場合によっては滋養成分及び／又は抗生物質を添加したＤＭＥＭなどの培地中で組織培養プレートにおいて成長させて集密化した。約１０μｇのｐＲＫ５－ＰＲＯポリペプチドＤＮＡを約１μｇのＶＡ　ＲＮＡ遺伝子コード化ＤＮＡ［Ｔｈｉｍｍａｐｐａｙａ等，Ｃｅｌｌ，３１：５４３（１９８２））］と混合し、５００μｌの１ｍＭトリス－ＨＣｌ、０．１ｍＭＥＤＴＡ、０．２２７ＭＣａＣｌ$_2$に溶解させた。この混合物に、滴状の、５００μｌの５０ｍＭＨＥＰＥＳ（ｐＨ７．３５）、２８０ｍＭのＮａＣｌ、１．５ｍＭのＮａＰＯ$_4$を添加し、２５℃で１０分間析出物を形成させた。析出物を懸濁し、２９３細胞に加えて３７℃で約４時間定着させた。培地を吸引し、２ｍｌのＰＢＳ中２０％グリセロールを３０秒間添加した。２９３細胞は、次いで無血清培地で洗浄し、新鮮な培地を添加し、細胞を約５日間インキュベートした。

形質移入の約２４時間後、培地を除去し、培地（のみ）又は２００μＣｉ／ｍｌ^{35}Ｓ－システイン及び２００μＣｉ／ｍｌ^{35}Ｓ－メチオニンを含む培地で置換した。１２時間のインキュベーションの後、条件培地を回収し、スピンフィルターで濃縮し、１５％ＳＤＳゲルに添加した。処理したゲルを乾燥させ、ＰＲＯポリペプチドの存在を現す選択された時間にわたってフィルムにさらした。形質転換した細胞を含む培地に、更なるインキュベーションを施し（無血清培地で）、培地を選択されたバイオアッセイで試験した。

これに代わる技術において、ＰＲＯポリペプチドは、Ｓｏｍｐａｒｙａｃ等，Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．，１２：７５７５（１９８１）に記載されたデキストラン硫酸法を用いて２９３細胞に一過的に導入される。２９３細胞は、スピナーフラスコ内で最大密度まで成長させ、７００μｇのｐＲＫ５－ＰＲＯポリペプチドＤＮＡを添加する。細胞は、まずスピナーフラスコから遠心分離によって濃縮し、ＰＢＳで洗浄した。ＤＮＡ－デキストラン沈殿物を細胞ペレット上で４時間インキュベートした。細胞を２０％グリセロールで９０秒間処理し、組織培地で洗浄し、組織培地、５μｇ／ｍｌウシインシュリン及び０．１μｇ／ｍｌウシトランスフェリンを含むスピナーフラスコに再度導入した。約４日後に、条件培地を遠心分離して濾過し、細胞及び細胞片を除去した。次いで発現されたＰＲＯポリペプチドを含む試料を濃縮し、透析及び／又はカラムクロマトグラフィー等の選択した方法によって精製した。

他の実施態様では、ＰＲＯポリペプチドをＣＨＯ細胞で発現させることができる。ｐＲＫ

５－ＰＲＯポリペプチドは、$CaPO_4$又はＤＥＡＥ－デキストランなどの公知の試薬を用いてＣＨＯ細胞に形質移入することができる。上記したように、細胞培地をインキュベートし、培地を培養培地（単独）又は^{35}Ｓ－メチオニン等の放射性標識を含む培地に置換することができる。ＰＲＯポリペプチドの存在を同定した後、培地を無血清培地に置換してもよい。好ましくは、培地を約６日間インキュベートし、次いで条件培地を収集する。次いで、発現されたＰＲＯポリペプチドを含む培地を濃縮して、任意の選択した方法によって精製することができる。

また、エピトープタグＰＲＯポリペプチドは、宿主ＣＨＯ細胞において発現させてもよい。ＰＲＯポリペプチドはｐＲＫ５ベクターからサブクローニングしてもよい。サブクローン挿入物は、ＰＣＲを施してバキュロウイルス発現ベクター中のポリ－ｈｉｓタグ等の選択されたエピトープタグを持つ枠に融合できる。ポリ－ｈｉｓタグＰＲＯポリペプチド挿入物は、次いで、安定なクローンの選択のためのＤＨＦＲ等の選択マーカーを含むＳＶ４０誘導ベクターにサブクローニングできる。最後に、ＣＨＯ細胞をＳＶ４０誘導ベクターで（上記のように）形質移入した。発現を確認するために、上記のように標識化を行ってもよい。発現されたポリ－ｈｉｓタグＰＲＯポリペプチドを含む培地は、次いで濃縮し、Ni^{2+}－キレートアフィニティクロマトグラフィー等の選択された方法により精製できる。

ＰＲＯ２１１、ＰＲＯ２１７、ＰＲＯ２３０、ＰＲＯ２１９、ＰＲＯ２４５、ＰＲＯ２２１、ＰＲＯ２５８、ＰＲＯ３０１、ＰＲＯ２２４、ＰＲＯ２２２、ＰＲＯ２３４、ＰＲＯ２２９、ＰＲＯ２２３、ＰＲＯ３２８及びＰＲＯ３３２は、一過性及び安定な発現法の双方によりＣＨＯ細胞中に首尾よく発現した。また、ＰＲＯ２３２、ＰＲＯ２６５、ＰＲＯ２４６、ＰＲＯ２２８、ＰＲＯ２２７、ＰＲＯ２２０、ＰＲＯ２６６、ＰＲＯ２６９、ＰＲＯ２８７、ＰＲＯ２１４、ＰＲＯ２３１、ＰＲＯ２３３、ＰＲＯ２３８、ＰＲＯ２４４、ＰＲＯ２３５、ＰＲＯ２３６、ＰＲＯ２６２、ＰＲＯ２３９、ＰＲＯ２５７、ＰＲＯ２６０、ＰＲＯ２６３、ＰＲＯ２７０、ＰＲＯ２７１、ＰＲＯ２７２、ＰＲＯ２９４、ＰＲＯ２９５、ＰＲＯ２９３、ＰＲＯ２４７、ＰＲＯ３０３及びＰＲＯ２６８はＣＨＯ細胞で一過性に首尾よく発現した。

【０２８６】

ＣＨＯ細胞における安定な発現は以下の方法を用いて実施された。タンパク質は、それぞれのタンパク質の可溶化形態のコード化配列（例えば、細胞外ドメイン）がヒンジ、ＣＨ２及びＣＨ２ドメインを含むＩｇＧ１定常領域配列に融合したＩｇＧ作成物（イムノアドヘシン）、又はポリ－Ｈｉｓタグ形態として発現された。

ＰＣＲ増幅に続いて、各ＤＮＡを、Ａｕｓｕｂｅｌ等，Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｏｆ Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙ，Ｕｎｉｔ ３．１６，Ｊｏｈｎ Ｗｉｌｅｙ ａｎｄ Ｓｏｎｓ（１９９７）に記載されたような標準的技術を用いてＣＨＯ発現ベクターにサブクローニングした。ＣＨＯ発現ベクターは、対象とするＤＮＡの５’及び３’に適合する制限部位を有し、ｃＤＮＡの便利なシャトル化ができるように作成される。ベクターは、Ｌｕｃａｓ等，Ｎｕｃｌ．Ａｃｉｄｓ Ｒｅｓ．２４：９，１７７４－１７７９（１９９６）に記載されたようにＣＨＯ細胞での発現を用い、対象とするｃＤＮＡ及びジヒドロフォレートレダクターゼ（ＤＨＦＲ）の発現の制御にＳＶ４０初期プロモーター／エンハンサーを用いる。ＤＨＦＲ発現は、形質移入に続くプラスミドの安定な維持のための選択を可能にする。

所望のプラスミドＤＮＡの１２マイクログラムを、市販の形質移入試薬Ｓｕｐｅｒｆｅｃｔ（登録商標）（Ｑｕｉａｇｅｎ），Ｄｏｓｐｅｒ（登録商標）及びＦｕｇｅｎｅ（登録商標）（Ｂｏｅｈｒｉｎｇｅｒ Ｍａｎｎｈｅｉｍ）約一千万のＣＨＯ細胞に導入する。細胞は、上掲のＬｕｃａｓ等に記載されているように成長させた。約$3x10^{-7}$細胞を、下記のような更なる成長及び生産のためにアンプル中で凍結させた。

プラスミドＤＮＡを含むアンプルを水槽に配して解凍し、ボルテックスにより混合した。内容物を１０ｍＬの媒質を含む遠心管にピペットして、１０００ｒｐｍで５分間遠心分離した。上清を吸引して細胞を１０ｍＬの選択培地（０．２μｍ濾過ＰＳ２０、５％の０．

２μｍ透析濾過ウシ胎児血清を添加）中に懸濁させた。次いで細胞を９０ｍＬの選択培地を含む１００ｍｌスピナーに分ける。１－２日後、細胞を１５０ｍＬの選択培地を満たした２５０ｍＬスピナーに移し、３７℃でインキュベートする。さらに２－３日後、２５０ｍＬ、５００ｍＬ及び２０００ｍＬのスピナーを３ｘ１０^{5}細胞／ｍＬで播種した。細胞培地を遠心分離により新鮮培地に交換し、生産培地に再懸濁させた。任意の適切なＣＨＯ培地を用いてもよいが、実際には１９９２年６月１６日に発行された米国特許第５，１２２，４６９号に記載された生産培地を使用した。３Ｌの生産スピナーを１．２ｘ１０^{6}細胞／ｍＬで播種した。０日目に、細胞数とｐＨを測定した。１日目に、スピナーをサンプルし、濾過空気での散布を実施した。２日目に、スピナーをサンプルし、温度を３３℃に変え、５００ｇ／Ｌのグルコース及び０．６ｍＬの１０％消泡剤（例えば３５％ポリジメチルシロキサンエマルション、Ｄｏｗ Ｃｏｒｎｉｎｇ ３６５ Ｍｅｄｉｃａｌ Ｇｒａｄｅ Ｅｍｕｌｓｉｏｎ）の３０ｍＬとした。生産を通して、ｐＨは７．２近傍に調節し維持した。１０日後、又は生存率が７０％を下回るまで、細胞培地を遠心分離で回収して０．２２μｍフィルターを通して濾過した。濾過物は、４℃で貯蔵するか、即座に精製用カラムに充填した。

【０２８７】

ポリ－Ｈｉｓタグ作成物について、タンパク質はＮｉ－ＮＴＡカラム（Ｑｉａｇｅｎ）を用いて精製した。精製の前に、イミダゾールを条件培地に５ｍＭの濃度まで添加した。条件培地を、０．３ＭのＮａＣｌ及び５ｍＭイミダゾールを含む２０ｍＭのＨｅｐｅｓ，ｐＨ７．４バッファーで平衡化した６ｍｌのＮｉ－ＮＴＡカラムに４－５ｍｌ／分の流速で４℃においてポンプ供給した。充填後、カラムをさらに平衡バッファーで洗浄し、タンパク質を０．２５Ｍイミダゾールを含む平衡バッファーで溶離した。高度に精製されたタンパク質は、続いて１０ｍＭのＨｅｐｅｓ、０．１４ＭのＮａＣｌ及び４％のマンニトール，ｐＨ６．８，を含む貯蔵バッファー中で２５ｍｌのＧ２５ Ｓｕｐｅｒｆｉｎｅ（Ｐｈａｒｍａｃｉａ）カラムを用いて脱塩し、－８０℃で貯蔵した。

イムノアドヘシン（Ｆｃ含有）作成物を以下のようにして条件培地から精製した。条件培地を、２０ｍＭのリン酸ナトリウムバッファー，ｐＨ６．８で平衡化した５ｍｌのプロテインＡカラム（Ｐｈａｒｍａｃｉａ）に汲み上げた。充填後、カラムを平衡バッファーで強く洗浄した後、１００ｍＭのクエン酸，ｐＨ３．５で溶離した。溶離したタンパク質は、１ｍｌの画分を２７５μＬの１Ｍトリスバッファー，ｐＨ９を含む管に回収することにより即座に中性化した。高度に精製されたタンパク質は、続いてポリ－Ｈｉｓタグタンパク質について上記した貯蔵バッファー中で脱塩した。均一性はＳＤＳポリアクリルアミドゲルで試験し、エドマン（Ｅｄｍａｎ）分解によりＮ－末端アミノ酸配列決定した。

ＰＲＯ２１１、ＰＲＯ２１７、ＰＲＯ２３０、ＰＲＯ２３２、ＰＲＯ１８７、ＰＲＯ２６５、ＰＲＯ２１９、ＰＲＯ２４６、ＰＲＯ２２８、ＰＲＯ５３３、ＰＲＯ２４５、ＰＲＯ２２１、ＰＲＯ２２７、ＰＲＯ２２０、ＰＲＯ２５８、ＰＲＯ２６６、ＰＲＯ２６９、ＰＲＯ２８７、ＰＲＯ２１４、ＰＲＯ３１７、ＰＲＯ３０１、ＰＲＯ２２４、ＰＲＯ２２２、ＰＲＯ２３４、ＰＲＯ２３１、ＰＲＯ２２９、ＰＲＯ２３３、ＰＲＯ２３８、ＰＲＯ２２３、ＰＲＯ２３５、ＰＲＯ２３６、ＰＲＯ２６２、ＰＲＯ２３９、ＰＲＯ２５７、ＰＲＯ２６０、ＰＲＯ２６３、ＰＲＯ２７０、ＰＲＯ２７１、ＰＲＯ２７２、ＰＲＯ２９４、ＰＲＯ２９５、ＰＲＯ２９３、ＰＲＯ２４７、ＰＲＯ３０４、ＰＲＯ３０２、ＰＲＯ３０７、ＰＲＯ３０３、ＰＲＯ３４３、ＰＲＯ３２８、ＰＲＯ３２６、ＰＲＯ３３１、ＰＲＯ３３２、ＰＲＯ３３４、ＰＲＯ３４６、ＰＲＯ２６８、ＰＲＯ３３０、ＰＲＯ３３１及びＰＲＯ３３９はＣＯＳ細胞中で一過性に成功裏に発現された。

【０２８８】

実施例５５：酵母菌でのＰＲＯポリペプチドの発現

以下の方法は、酵母菌中での所望のＰＲＯポリペプチドの組換え発現を記載する。

第１に、ＡＤＨ２／ＧＡＰＤＨプロモーターからのＰＲＯポリペプチドの細胞内生産又は分泌のための酵母菌発現ベクターを作成する。所望のＰＲＯポリペプチドをコードするＤ

ＮＡ、選択されたシグナルペプチド及びプロモーターを、選択したプラスミドの適当な制限酵素部位に挿入してＰＲＯポリペプチドの細胞内発現を指示する。分泌のために、ＰＲＯポリペプチドをコードするＤＮＡを、選択したプラスミドに、ＡＤＨ２／ＧＡＰＤＨプロモーターをコードするＤＮＡ、酵母菌アルファ因子分泌シグナル／リーダー配列、及び（必要ならば）ＰＲＯポリペプチドの発現のためのリンカー配列とともにクローニングすることができる。

酵母菌株ＡＢ１１０等の酵母菌は、次いで上記の発現プラスミドで形質転換し、選択された発酵培地中で培養できる。形質転換した酵母菌上清は、１０％トリクロロ酢酸での沈降及びＳＤＳ－ＰＡＧＥによる分離で分析し、次いでクマシーブルー染色でゲルの染色をすることができる。

続いて組換えＰＲＯポリペプチドは、発酵培地から遠心分離により酵母菌細胞を除去し、次いで選択されたカートリッジフィルターを用いて培地を濃縮することによって単離及び精製できる。ＰＲＯポリペプチドを含む濃縮物は、選択されたカラムクロマトグラフィー樹脂を用いてさらに精製してもよい。

【０２８９】

実施例５６：バキュロウイルス感染昆虫細胞でのＰＲＯポリペプチドの発現

以下の方法は、バキュロウイルス感染昆虫細胞中におけるＰＲＯポリペプチドの組換え発現を記載する。

所望のＰＲＯポリペプチドを、バキュロウイルス発現ベクターに含まれるエピトープタグの上流に融合させた。このようなエピトープタグは、ポリ－ｈｉｓタグ及び免疫グロブリンタグ（ＩｇＧのＦｃ領域など）を含む。ｐＶＬ１３９３（Ｎａｖｏｇｅｎ）などの市販されているプラスミドから誘導されるプラスミドを含む種々のプラスミドを用いることができる。簡単には、ＰＲＯポリペプチド又はＰＲＯポリペプチドの所望の部分（例えば膜貫通タンパク質の細胞外ドメインをコードする配列）が５’及び３’領域に相補的なプライマーでのＰＣＲにより増幅される。５’プライマーは、隣接する（選択された）制限酵素部位を包含していてもよい。生産物は、次いで、選択された制限酵素で消化され、発現ベクターにサブクローニングされる。

組換えバキュロウイルスは、上記のプラスミド及びＢａｃｕｌｏＧｏｌｄ（商品名）ウイルスＤＮＡ（Ｐｈａｒｍｉｎｇｅｎ）を、Ｓｐｏｄｏｐｔｅｒａ ｆｒｕｇｉｐｅｒｄａ（「Ｓｆ９」）細胞（ＡＴＣＣ ＣＲＬ １７１１）中にリポフェクチン（ＧＩＢＣＯ－ＢＲＬから市販）を用いて同時形質移入することにより作成される。２８℃で４－５日インキュベートした後、放出されたウイルスを回収し、更なる増幅に用いた。ウイルス感染及びタンパク質発現は、Ｏ’Ｒｅｉｌｌｅｙ等，Ｂａｃｕｌｏｖｉｒｕｓ ｅｘｐｒｅｓｓｉｏｎ ｖｅｃｔｏｒｓ：Ａ ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ，Ｏｘｆｏｒｄ：Ｏｘｆｏｒｄ Ｕｎｉｖｅｒｓｉｔｙ Ｐｒｅｓｓ（１９９４）に記載されているように実施した。

次に、発現されたポリ－ｈｉｓタグＰＲＯポリペプチドは、例えばNi^{2+}－キレートアフィニティクロマトグラフィーにより次のように精製される。抽出は、Ｒｕｐｅｒｔ等，Ｎａｔｕｒｅ，３６２：１７５－１７９（１９９３）に記載されているように、ウイルス感染した組み換えＳｆ９細胞から調製した。簡単には、Ｓｆ９細胞を洗浄し、超音波処理用バッファー（２５ｍＬのＨｅｐｅｓ、ｐＨ７．９；１２．５ｍＭの$MgCl_2$；０．１ｍＭ ＥＤＴＡ；１０％グリセロール；０．１％のＮＰ－４０；０．４ＭのＫＣｌ）中に再懸濁し、氷上で２回２０秒間超音波処理した。超音波処理物を遠心分離で透明化し、上清を負荷バッファー（５０ｍＭリン酸塩、３００ｍＭのＮａＣｌ、１０％グリセロール、ｐＨ７．８）で５０倍希釈し、０．４５μｍフィルターで濾過した。Ni^{2+}－ＮＴＡアガロースカラム（Ｑｉａｇｅｎから市販）を５ｍＬの総容積で調製し、２５ｍＬの水で洗浄し、２５ｍＬの負荷バッファーで平衡させた。濾過した細胞抽出物は、毎分０．５ｍＬでカラムに負荷した。カラムを、分画回収が始まる点であるA_{280}のベースラインまで負荷バッファーで洗浄した。次に、カラムを、結合タンパク質を非特異的に溶離する二次洗浄バッファー（５０ｍＭリン酸塩；３００ｍＭのＮａＣｌ、１０％グリセロール、

ｐＨ６．０）で洗浄した。A_{280}のベースラインに再度到達した後、カラムを二次洗浄バッファー中で０から５００ｍＭイミダゾール勾配で展開した。１ｍＬの分画を回収し、ＳＤＳ－ＰＡＧＥ及び銀染色又はアルカリホスファターゼ（Ｑｉａｇｅｎ）に複合したNi^{2+}－ＮＴＡでのウェスタンブロットで分析した。溶離したHis_{10}－タグＰＲＯポリペプチドを含む画分をプールして負荷バッファーで透析した。

あるいは、ＩｇＧタグ（又はＦｃタグ）ＰＲＯの精製は、例えば、プロテインＡ又はプロテインＧカラムクロマトグラフィーを含む公知のクロマトグラフィー技術を用いて実施できる。

ＰＲＯ２１１、ＰＲＯ２１７、ＰＲＯ２３０、ＰＲＯ１８７、ＰＲＯ２６５、ＰＲＯ２４６、ＰＲＯ２２８、ＰＲＯ５３３、ＰＲＯ２４５、ＰＲＯ２２１、ＰＲＯ２２０、ＰＲＯ２５８、ＰＲＯ２６６、ＰＲＯ２６９、ＰＲＯ２８７、ＰＲＯ２１４、ＰＲＯ３０１、ＰＲＯ２２４、ＰＲＯ２２２、ＰＲＯ２３４、ＰＲＯ２３１、ＰＲＯ２２９、ＰＲＯ２３５、ＰＲＯ２３９、ＰＲＯ２５７、ＰＲＯ２７２、ＰＲＯ２９４、ＰＲＯ２９５、ＰＲＯ３２８、ＰＲＯ３２６、ＰＲＯ３３１、ＰＲＯ３３４、ＰＲＯ３４６及びＰＲＯ３１０がバキュロウィルス感染Ｓｆ９又はｈｉｇｈ５昆虫細胞中に成功裡に発現された。発現は実際には０．５－２Ｌのスケールでなされたが、より大規模な調製（例えば８Ｌ）に容易にスケールアップすることができる。タンパク質はＩｇＧ作成物（イムノアドヘシン）として発現され、そこではタンパク質細胞外領域は、ヒンジ、ＣＨ２及びＣＨ３ドメインを含むＩｇＧ１定常領域配列に、及び／又はｐｏｌｙ－Ｈｉｓタグ形態に融合された。

ＰＣＲ増幅に続いて、各コード化配列をバキュロウィルス発現ベクター（ＩｇＧ融合のためのｐｂ．ＰＨ．ＩｇＧ及びｐｏｌｙ－Ｈｉｓタグタンパク質のためのｐｂ．ＰＨ．Ｈｉｓ．ｃ）中にサブクローニングし、ベクターとＢａｃｕｌｏｇｏｌｄ（商品名）バキュロウィルスＤＮＡ（Ｐｈａｒｍｉｎｇｅｎ）を１０５のＳｐｏｄｏｐｔｅｒａ ｆｒｕｇｉｐｅｒｄａ（「Ｓｆ９」）細胞（ＡＴＣＣ ＣＲＬ １７１１）中にリポフェクチン（Ｇｉｂｃｏ－ＢＲＬ）を用いて同時形質移入した。ｐｂ．ＰＨ．ＩｇＧ及びｐｂ．ＰＨ．Ｈｉｓは商業的に利用できるバキュロウィルス発現ベクターｐＶＬ１３９３（Ｐｈａｒｍｉｎｇｅｎ）の改変物で、Ｈｉｓ又はＦｃタグ配列を含む修飾ポリリンカー領域を持つ。１０％ＦＢＳを補填したＨｉｎｋのＴＮＭ－ＦＨ培地（Ｈｙｃｌｏｎｅ）で細胞を増殖させた。細胞を３８℃で５日間インキュベートした。上清を収集し、続いて１０％ＦＢＳを補填したＨｉｎｋのＴＮＭ－ＦＨ培地中のＳｆ９細胞を１０のおよその感染効率（ＭＯＩ）で感染させることにより最初のウィルス増幅に使用した。細胞を３８℃で３日間インキュベートした。上清を収集し、バキュロウィルス発現ベクター中における作成物の発現を、ヒスチジンタグタンパク質に対して２５ｍＬのＮｉ－ＮＴＡビード（ＱＩＡＧＥＮ）へ、又はＩｇＧタグタンパク質に対してプロテイン－Ａセファロース ＣＬ－４Ｂビード（Ｐｈａｒｍａｃｉａ）へ１ｍｌの上清をバッチ結合させ、続いてＳＤＳ－ＰＡＧＥ分析を行って、クーマシーブルー染色により既知の濃度のタンパク質標準と比較することにより決定した。

最初のウィルス増幅上清は、０．１のおよそのＭＯＩでＥＳＦ－９２１培地（Ｅｘｐｒｅｓｓｉｏｎ Ｓｙｓｔｅｍｓ ＬＬＣ）中で増殖したＳｆ９細胞の撹拌培養（５００ｍｌ）を感染させるために使用した。細胞を２８℃で３日間インキュベートした。上清を収集し、濾過した。撹拌培養の発現が確認されるまでバッチ結合とＳＤＳ－ＰＡＧＥ分析を必要に応じて繰り返した。

形質移入細胞からの条件培地（０．５から３Ｌ）を遠心分離により収集して細胞を除去し、０．２２ミクロンフィルターで濾過した。ｐｏｌｙ－Ｈｉｓタグ作成物に対しては、Ｎｉ－ＮＴＡカラム（Ｑｉａｇｅｎ）を使用してタンパク質作成物を精製した。精製前に、５ｍＭの濃度まで条件培地にイミダゾールを添加した。条件培地を、４℃で４－５ｍｌ／ｍｉｎの流量で２０ｍＭ Ｈｅｐｅｓ、ｐＨ７．４、０．３ＭのＮａＣｌと５ｍＭのイミダゾールを含むバッファーで平衡化した６ｍｌのＮｉ－ＮＴＡカラムに汲み上げた。充填後、カラムを更なる平衡化バッファーで洗浄し、タンパク質を０．２５Ｍのイミダゾールを含む平衡化バッファーで溶出させた。高度に精製されたタンパク質は続いて２５ｍｌの

Ｇ２５ Ｓｕｐｅｒｆｉｎｅ（Ｐｈａｒｍａｃｉａ）カラムで１０ｍＭ Ｈｅｐｅｓ、０．１４ＭのＮａＣｌ及び４％のマンニトール、ｐＨ６．８を含む貯蔵バッファー中で脱塩し、－８０℃で貯蔵した。

タンパク質のイムノアドヘシン（Ｆｃ含有）作成物を次のようにして条件培地から精製した。２０ｍＭのリン酸ナトリウムバッファー、ｐＨ６．８で平衡化された５ｍｌのプロテインＡカラム（Ｐｈａｒｍａｃｉａ）に条件培地を汲み上げた。充填後、１００ｍＭのクエン酸、ｐＨ３．５での溶出前にカラムを平衡化バッファーで十分に平衡化した。溶出したタンパク質を、２７５ｍＬの１Ｍ Ｔｒｉｓバッファー，ｐＨ９を含むチューブ中に１ｍｌのフラクションを収集することにより即座に中和させた。高度に精製したタンパク質を上述のようにｐｏｌｙ－Ｈｉｓタグタンパク質の貯蔵バッファー中に脱塩した。タンパク質の一様性はＳＤＳポリアクリルアミドゲル（ＰＥＧ）電気泳動法とエドマン分解によるＮ末端アミノ酸配列決定により証明した。

【０２９０】

実施例５７：ＰＲＯポリペプチドに結合する抗体の調製

この実施例は、ＰＲＯポリペプチドに特異的に結合できるモノクローナル抗体の調製を例示する。

モノクローナル抗体の生産のための技術は、この分野で知られており、例えば、上掲のＧｏｄｉｎｇに記載されている。用いられ得る免疫原は、精製ＰＲＯポリペプチド、ＰＲＯポリペプチドを含む融合タンパク質、細胞表面に組換えＰＲＯポリペプチドを発現する細胞を含む。免疫原の選択は、当業者が過度の実験をすることなく選択することができる。Ｂａｌｂ／ｃ等のマウスを、完全フロイントアジュバントに乳化して皮下又は腹腔内に１－１００マイクログラムで注入したＰＲＯポリペプチド免疫原で免疫化する。あるいは、免疫原をＭＰＬ－ＴＤＭアジュバント（Ｒｉｂｉ Ｉｍｍｕｎｏｃｈｅｍｉｃａｌ Ｒｅｓｅａｒｈ，Ｈａｍｉｌｔｏｎ，ＭＴ）に乳化し、動物の後足蹠に注入してもよい。免疫化したマウスは、次いで１０から１２日後に、選択したアジュバント中に乳化した付加的免疫源で追加免疫する。その後、数週間、マウスをさらなる免疫化注射で追加免疫する。抗ＰＲＯポリペプチド抗体の検出のためのＥＬＩＳＡアッセイで試験するために、レトロオービタル出血からの血清試料をマウスから周期的に採取してもよい。

適当な抗体力価が検出された後、抗体に「ポジティブ（陽性）」な動物に、ＰＲＯポリペプチドの静脈内注射の最後の注入をすることができる。３から４日後、マウスを屠殺し、脾臓細胞を取り出した。次いで脾臓細胞を（３５％ポリエチレングリコールを用いて）、ＡＣＴＴから番号ＣＲＬ１５９７で入手可能なＰ３Ｘ６３ＡｇＵ．１等の選択されたマウス骨髄腫株化細胞に融合させた。融合によりハイブリドーマ細胞が生成され、次いで、ＨＡＴ（ヒポキサンチン、アミノプテリン、及びチミジン）培地を含む９６ウェル組織培養プレートに蒔き、非融合細胞、骨髄腫ハイブリッド、及び脾臓細胞ハイブリッドの増殖を阻害した。

ハイブリドーマ細胞は、ＰＲＯポリペプチドに対する反応性についてのＥＬＩＳＡでスクリーニングされる。ＰＲＯポリペプチドに対する所望のモノクローナル抗体を分泌する「ポジティブ（陽性）」ハイブリドーマ細胞の決定は、技術常識の範囲内である。

陽性ハイブリドーマ細胞を同系のＢａｌｂ／ｃマウスに腹腔内注入し、抗ＰＲＯポリペプチドモノクローナル抗体を含む腹水を生成させる。あるいは、ハイブリドーマ細胞を、組織培養フラスコ又はローラーボトルで成長させることもできる。腹水中に生成されたモノクローナル抗体の精製は、硫酸アンモニウム沈降、それに続くゲル排除クロマトグラフィーを用いて行うことができる。あるいは、抗体のプロテインＡ又はプロテインＧへの親和性に基づくアフィニティクロマトグラフィーを用いることもできる。

【０２９１】

実施例５８：キメラＰＲＯポリペプチド

ＰＲＯポリペプチドはタンパク質の精製を容易にするために添加される一又は複数の更なるポリペプチドドメインを持つキメラタンパク質として発現されてもよい。このような精製を容易にするドメインには、限定されるものではないが、金属キレートペプチド、例え

ば固定化された金属上での精製を可能にするヒスチジン－トリプトファンモジュール、固定化免疫グロブリン上での精製を可能にするプロテインＡドメイン、及びＦＬＡＧＳ^{TM}伸展／アフィニティー精製システム（Ｉｍｍｕｎｅｘ Ｃｏｒｐ．，Ｓｅａｔｔｌｅ Ｗａｓｈ．）に使用されるドメインが含まれる。精製ドメインとＰＲＯポリペプチド配列の間へＸＡ因子又はエンテロキナーゼ（Ｉｎｖｔｒｏｇｅｎ，Ｓａｎ Ｄｉｅｇｏ Ｃａｌｉｆ．）のような切断性リンカー配列を含めることは有用であり、ＰＲＯポリペプチドをコードするＤＮＡの発現を容易にする。

【０２９２】

実施例５９：特異的抗体を用いたＰＲＯポリペプチドの精製

天然又は組換えＰＲＯポリペプチドは、タンパク質精製分野の種々の標準的な方法によって精製できる。例えば、プロ－ＰＲＯポリペプチド、成熟ＰＲＯポリペプチド、又はプレ－ＰＲＯポリペプチドは、対象とするＰＲＯポリペプチドに特異的な抗体を用いた免疫親和性クロマトグラフィーによって精製される。一般に、免疫親和性カラムは抗ＰＲＯポリペプチド抗体を活性化クロマトグラフィー樹脂に共有結合させて作成される。

ポリクローナル免疫グロブリンは、硫酸アンモニウムでの沈殿又は固定化プロテインＡ（Ｐｈａｒｍａｃｉａ ＬＫＢ Ｂｉｏｔｅｃｈｎｏｌｏｇｙ，Ｐｉｓｃａｔａｗａｙ，Ｎ．Ｊ．）での精製の何れかにより免疫血清から調製される。同様に、モノクローナル抗体は、硫酸アンモニウム沈殿又は固定化プロテインＡでのクロマトグラフィーによりマウス腹水液から調製される。部分的に精製された免疫グロブリンは、ＣｎＢｒ－活性化セファロース（商品名）（Ｐｈａｒｍａｃｉａ ＬＫＢ Ｂｉｏｔｅｃｈｎｏｌｏｇｙ）等のクロマトグラフィー樹脂に共有結合される。抗体が樹脂に結合され、樹脂がブロックされ、誘導体樹脂は製造者の指示に従って洗浄される。

このような免疫親和性カラムは、可溶化形態のＰＲＯポリペプチドを含有する細胞からの画分を調製することによるＰＲＯポリペプチドの精製において利用される。この調製物は、洗浄剤の添加又はこの分野で公知の方法により微分遠心分離を介して得られる全細胞又は細胞成分画分の可溶化により誘導される。あるいは、シグナル配列を含む可溶化ＰＲＯポリペプチドは、細胞が成長する培地中に有用な量で分泌される。

可溶化ＰＲＯポリペプチド含有調製物は、免疫親和性カラムを通され、カラムはＰＲＯポリペプチドの好ましい吸着をさせる条件下（例えば、洗浄剤存在下の高イオン強度バッファー）で洗浄される。次いで、カラムは、抗体／ＰＲＯポリペプチド結合を分解する条件下（例えば、約２－３といった低ｐＨ、高濃度の尿素又はチオシアン酸イオン等のカオトロープ）で溶離され、ＰＲＯポリペプチドが回収される。

【０２９３】

実施例６０：薬剤スクリーニング

本発明は、ＰＲＯポリペプチド又はその結合断片を種々の薬剤スクリーニング技術において使用することによる化合物のスクリーニングに特に有用である。そのような試験に用いられるＰＲＯポリペプチド又は断片は、溶液中の自由状態でも、固体支持体に固定されても、細胞表面に担持されていても、細胞内に位置していてもよい。薬剤スクリーニングの１つの方法は、ＰＲＯポリペプチド又は断片を発現する組換え核酸で安定に形質移入される真核生物又は原核生物宿主細胞を利用する。薬剤は、そのような形質移入細胞に対して、競合的結合アッセイにおいてスクリーニングされる。そのような細胞は、生存可能又は固定化形態のいずれかにおいて、標準的な結合アッセイに使用できる。例えば、ＰＲＯポリペプチド又は断片と試験される試薬の間での複合体の形成を測定してよい。あるいは、試験する試薬によって生ずるＰＲＯポリペプチドとその標的細胞との間の複合体形成における減少を試験することもできる

しかして、本発明は、ＰＲＯポリペプチド関連疾患又は障害に影響を与えうる薬剤又は任意の他の試薬のスクリーニング方法を提供する。これらの方法は、その試薬をＰＲＯポリペプチド又は断片に接触させ、（Ｉ）試薬とＰＲＯポリペプチド又は断片との間の複合体の存在について、又は（ｉｉ）ＰＲＯポリペプチド又は断片と細胞との間の複合体の存在について検定することを含む。これらの競合結合アッセイでは、ＰＲＯポリペプチド又は

断片が典型的には標識される。適切なインキュベーションの後、自由なＰＲＯポリペプチド又は断片を結合形態のものから分離し、自由又は未複合の標識の量が、特定の試薬がＰＲＯポリペプチドに結合する又はＰＲＯポリペプチド／細胞複合体を阻害する能力の尺度となる。

薬剤スクリーニングのための他の技術は、ポリペプチドに対して適当な結合親和性を持つ化合物についての高スループットスクリーニングを提供し、１９８４年９月１３日に公開されたＷＯ ８４／０３５６４に詳細に記載されている。簡単に述べれば、多数の異なる小型ペプチド試験化合物が、プラスチックピン等の固体支持体又は幾つかの他の表面上で合成される。ＰＲＯポリペプチドに適用すると、ペプチド試験化合物はＰＲＯポリペプチドと反応して洗浄される。結合したＰＲＯポリペプチドはこの分野で良く知られた方法により検出される。精製したＰＲＯポリペプチドは、上記の薬剤スクリーニング技術に使用するためにプレート上に直接被覆することもできる。さらに、非中和抗体は、ペプチドを捕捉し、それを固体支持体上に固定化するのに使用できる。

また、本発明は、ＰＲＯポリペプチドに結合可能な中和抗体がＰＲＯポリペプチド又はその断片について試験化合物と特異的に競合する競合薬剤スクリーニングアッセイも考慮する。この方法において、抗体は、ＰＲＯポリペプチドで、一又は複数の抗原決定基を持つ任意のペプチドの存在を検出するのに使用できる。

【０２９４】

実施例６１：合理的薬剤設計

合理的薬剤設計の目的は、対象とする生物学的活性ポリペプチド（例えば、ＰＲＯポリペプチド）又はそれらが相互作用する小分子、例えばアゴニスト、アンタゴニスト、又はインヒビターの構造的類似物を製造することである。これらの例の任意のものが、ＰＲＯポリペプチドのより活性で安定な形態又はインビボでＰＲＯポリペプチドに機能を促進又は阻害する薬剤の創作に使用できる（参考、Ｈｏｄｇｓｏｎ，Ｂｉｏ／Ｔｅｃｈｎｏｌｏｇｙ，９：１９－２１（１９９１））。

一アプローチ法において、ＰＲＯポリペプチド、又はＰＲＯポリペプチド－インヒビター複合体の三次元構造が、ｘ線結晶学により、コンピュータモデル化により、最も典型的には２つの方法の組み合わせにより決定される。分子の構造を解明し活性部位を決定するためには、ＰＲＯポリペプチドの形状及び電荷の両方が確認されなければならない。数は少ないが、ＰＲＯポリペプチドの構造に関する有用な情報が相同タンパク質の構造に基づいたモデル化によって得られることもある。両方の場合において、関連する構造情報は、類似ＰＲＯポリペプチド様分子の設計又は効果的なインヒビターの同定に使用される。合理的な薬剤設計の有用な例は、Ｂｒａｘｔｏｎ及びＷｅｌｌｓ，Ｂｉｏｃｈｅｍｉｓｔｒｙ，３１：７７９６－７８０１（１９９２）に示されているような向上した活性又は安定性を持つ分子、又はＡｔｈａｕｄａ等，Ｊ．Ｂｉｏｃｈｅｍ．，１１３：７４２－７４６（１９９３）に示されているような天然ペプチドのインヒビター、アゴニスト、又はアンタゴニストとして作用する分子を含む。

また、上記のような機能アッセイによって選択された標的特異的な抗体を単離しその結晶構造を解明することもできる。この方法は、原理的には、それに続く薬剤設計が基礎をおくことのできるファーマコア（ｐｈａｒｍａｃｏｒｅ）を生成する。機能的な薬理学的に活性な抗体に対する抗－イディオタイプ抗体（抗－ｉｄｓ）を生成することにより、タンパク質結晶学をバイパスすることができる。鏡像の鏡像として、抗－ｉｄｓの結合部位は最初のレセプターの類似物であると予測できる。抗－ｉｄは、次いで、化学的又は生物学的に製造したペプチドのバンクからペプチドを同定及び単離するのに使用できる。単離されたペプチドは、ファーマコアとして機能するであろう。

本発明により、十分な量のＰＲＯポリペプチドがＸ線結晶学などの分析実験を実施するために入手可能である。さらに、ここに提供したＰＲＯポリペプチドアミノ酸配列の知識は、ｘ線結晶学に換える、又はそれに加えるコンピュータモデル化技術で用いられる知識を提供する。

【０２９５】

実施例６２：ＰＲＯ３１７ポリペプチド特異的抗体を用いる診断試験
特定の抗ＰＲＯ３１７ポリペプチド抗体は、前病理状態、及び慢性もしくは急性疾患、例えばＰＲＯ３１７の量又は分布の差により特徴づけられる婦人科疾患又は虚血症の診断に有用である。ＰＲＯ３１７はヒト腎臓に発現されることが見いだされ、よってこの器官に影響を及ぼす異常又は病理に関連しているようである。更に、それはＥＢＡＦ－１に密接に関連しているので、子宮内膜と他の生殖組織に影響を及ぼしているようである。更に、ある種のＥＳＴｓのライブラリ源のために、ＰＲＯ３１７は血管形成によく関与しており、よって血管新生の修飾因子であるようである。
ＰＲＯ３１７に対する診断試験は、抗体と標識を用いてヒト体液、組織、又はかかる組織の抽出物においてＰＲＯ３１７を検出する方法を含む。本発明のポリペプチドと抗体は、検出可能なシグナルをもたらす物質とそれらを共有結合的か非共有結合的に結合させることにより標識される。広範囲の標識とコンジュゲーション法が知られており、科学文献と特許文献の双方に広く報告されている。好適な標識には放射性核種、酵素、基質、補助因子、阻害剤、蛍光剤、化学発光剤、磁気粒子等々が含まれる。このような標識の使用を教示する特許には、米国特許３，８１７，８３７；３，８５０，７５２；３，９３９，３５０；３，９９６，３４５；４，２７７，４３７；４，２７５，１４９；及び４，３６６，２４１が含まれる。また、組換え免疫グロブリンを米国特許４，８１６，５６７に示されるようにして産生されてもよい。
可溶型又は膜結合ＰＲＯ３１７を、そのＰＲＯ３１７に特異的なモノクローナル又はポリクローナル抗体の何れかを使用して測定する様々なプロトコールが当該分野において知られている。例には、酵素免疫抗体法（ＥＬＩＳＡ）、ラジオイムノアッセイ（ＲＩＡ）、ラジオレセプターアッセイ（ＲＲＡ）及び蛍光発色セルソーティング（ＦＡＣＳ）が含まれる。ＰＲＯ３１７上の二つの非干渉エピトープに反応性のモノクローナル抗体を使用する二部位モノクローナルベースイムノアッセイが好ましいが、競合結合アッセイを使用してもよい。これらのアッセイはとりわけＭａｄｄｏｘ等，Ｊ．Ｅｘｐ．Ｍｅｄ．，１５８：１２１１（１９８３）に記載されている。
【０２９６】
実施例６３：ＰＲＯ３１７レセプターの同定
精製ＰＲＯ３１７は特異的細胞表面レセプター及び他の結合分子の特徴付けと精製に有用である。物質代謝又は他の特異的応答によりＰＲＯ３１７に応答する細胞はＰＲＯ３１７に対するレセプターを発現しそうである。そのようなレセプターには、限定するものではないが、チロシン及びセリン／スレオニンキナーゼと関連し、これらにより活性化されるレセプターが含まれる。ＴＧＦ－スーパーファミリーに対する既知のレセプターについての概説については上掲のＫｏｌｏｄｚｉｅｊｃｚｙｋ及びＨａｌｌを参照されたい。このスーパーファミリーの候補レセプターはＩ型及びＩＩ型レセプターと名付けられた二つの主要グループに入る。両型ともセリン／スレオニンキナーゼである。適切なリガンドにより活性化すると、Ｉ型及びＩＩ型レセプターは物理的に相互作用をしてヘテロオリゴマーを形成し、続いて細胞内シグナル伝達カスケードを活性化させ、最後に遺伝子の転写と発現を調節する。また、ＴＧＦ－は、シグナル伝達分子に典型的なキナーゼ活性を欠いている膜結合プロテオグリカンである第３のレセプタークラスＩＩＩ型に結合する。
ＰＲＯ３１７レセプター又は他のＰＲＯ３１７結合分子は放射標識ＰＲＯ３１７との相互作用によって同定されてもよい。放射性標識は当該分野で知られている様々な方法によってＰＲＯ３１７中に導入することができる。好適な実施態様は^{125}Ｉボルトン・ハンター試薬でのＰＲＯ３１７中の一級アミノ基の標識であり（Ｂｏｌｔｏｎ及びＨｕｎｔｅｒ，Ｂｉｏｃｈｅｍ．Ｊ．，１３３：５２９（１９７３））、これは生物活性の喪失を随伴させないで他のポリペプチドを標識するために使用されている（Ｈｅｂｅｒｔ等，Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．，２６６：１８９８９（１９９１）；ＭｃＣｏｌｌ等，Ｊ．Ｉｍｍｕｎｏｌ．，１５０：４５５０－４５５５（１９９３））。レセプター担持細胞を標識ＰＲＯ３１７で標識する。ついで、細胞を洗浄して未結合ＰＲＯ３１７を除去し、レセプター結合ＰＲＯ３１７を定量する。異なった濃度のＰＲＯ３１７を使用して

得たデータを用いてレセプターの数と親和性の値を計算する。
標識ＰＲＯ３１７はその特異的レセプターの精製のための試薬として有用である。アフィニティー精製の一実施態様では、ＰＲＯ３１７はクロマトグラフィーカラムに共有結合的に結合される。レセプター担持細胞を抽出し、抽出物をカラムに通す。レセプターはＰＲＯ３１７に対するその生物学的親和性のためにカラムに結合する。レセプターはカラムから回収され、Ｎ末端タンパク質配列決定にかける。ついで、このアミノ酸を使用してレセプター遺伝子をクローニングするための変性オリゴヌクレオチドプローブを設計する。
他の方法では、ｍＲＮＡをレセプター担持細胞から得て、ｃＤＮＡライブラリに構築する。ライブラリを細胞集団にトランスフェクトし、レセプターを発現する細胞を蛍光的に標識したＰＲＯ３１７を使用して選択する。レセプターを、高度に標識した細胞から組換えＤＮＡを回収して配列決定することにより同定する。
更に他の方法では、レセプター担持細胞の表面に対して抗体、特にモノクローナル抗体をつくる。モノクローナル抗体をスクリーニングして標識ＰＲＯ３１７の結合を阻害するものを同定する。ついで、これらのモノクローナル抗体をアフィニティー精製又はレセプターの発現クローニングに使用する。
可溶型レセプター又は他の可溶型結合分子は同様の方法で同定される。標識ＰＲＯ３１７は子宮由来の抽出物又は他の適当な材料と共にインキュベートする。インキュベーション後、精製ＰＲＯ３１７のサイズより大きいＰＲＯ３１７複合体を、サイズ排除クロマトグラフィー又は密度勾配遠心分離のようなサイジング法によって同定し、当該分野で知られている方法によって精製される。可溶型レセプター又は結合タンパク質につきＮ末端配列決定をして、可溶型タンパク質が知られている場合はデータベースの同定に対して、あるいは可溶型タンパク質が未知の場合はクローニングするのに十分な情報を得る。
【０２９７】
実施例６４：ＰＲＯ３１７誘導細胞性応答の決定
例えばＰＲＯ３１７レセプターへの本発明のＰＲＯ３１７の結合により、ＰＲＯ３１７の生物活性を測定する。試験化合物を、レセプターに対するＰＲＯ３１７の結合を阻止する能力についてアンタゴニストとしてスクリーニングする。試験化合物を、ＰＲＯ３１７レセプターに結合しＰＲＯ３１７と同じ生理学的事象に影響を与える能力についてＰＲＯ３１７のアゴニストとして、例えばＳａｄｉｃｋ等，Ａｎａｌｙｔｉｃａｌ Ｂｉｏｃｈｅｍｉｓｔｒｙ，２３５：２０７－２１４（１９９６）により記載され、レセプターチロシンキナーゼの活性化が活性化レセプターの免疫獲得とリガンド誘導リン酸化のレベルの定量によりモニターされるＫＩＲＡ－ＥＬＩＳＡアッセイを使用してスクリーニングする。アッセイは、ＰＲＯ３１７レセプター特異的抗体の使用を通じてＰＲＯ３１７誘導レセプター活性化をモニターして活性化レセプターを捕獲するように適合させることができる。これらの方法はまた個々に記載した他のＰＲＯポリペプチドにも適用できる。
【０２９８】
実施例６５：ＰＲＯ２２４のスクリーニング化合物への使用
ＰＲＯ２２４は膜タンパク質が剥ぎ取られＰＲＯ２２４を発現可能な細胞で発現される。検出可能な標識を有する低密度リポタンパク質を細胞に添加し、エンドサイトーシスに対して十分な時間の間、インキュベートする。細胞を洗浄する。ついで、細胞を、細胞溶解後に膜に結合した細胞内の標識について分析する。細胞内の低密度リポタンパク質の検出は、ＰＲＯ２２４が低密度リポタンパク質のレセプタータンパク質のファミリー内にあることを決定する。このファミリー内に見いだされたメンバーはついでこれらのレセプターに影響を及ぼす化合物、特にこれらのレセプターを介してのコレステロールの摂取をスクリーニングするために使用される。
【０２９９】
実施例６６：血管内皮成長因子（ＶＥＧＦ）で刺激された内皮細胞成長の増殖を阻害するＰＲＯポリペプチドの能力
内皮細胞のＶＥＧＦ刺激増殖を阻害する様々なＰＲＯポリペプチドの能力を試験した。特に、ウシ副腎皮質毛細管内皮（ＡＣＥ）細胞（初代培養からのもの、最大１２－１４継代

）を９６ウェルのマイクロタイタープレート（Ａｍｅｒｓｈａｍ Ｌｉｆｅ Ｓｃｉｅｎｃｅ）上に３ｎｇ／ｍＬのＶＥＧＦを補填した低グルコースＤＭＥＭ、１０％小ウシ血清、２ｍＭグルタミン、１ｘペン／ストレプト及びファンギゾン中に１００μＬ当たり５００細胞／ウェルの密度でプレーティングした。
対照を同じようにしてプレーティングしたが、ＶＥＧＦは含んでいなかった。対象のＰＲＯポリペプチドの試験試料を１００μｌ容積中に添加して２００μｌ最終容積にした。細胞を３７℃で６－７日の間インキュベートした。培地を吸引して細胞をＰＢＳで１ｘ洗浄した。酸性ホスファターゼ反応混合物（１００μＬ，０．１Ｍ酢酸ナトリウム，ｐＨ５．５，０．１％トリトン－１００，１０ｍＭのリン酸ｐ－ニトロフェニル）を添加した。３７℃で２時間のインキュベーション後に、１０μｌの１ＮのＮａＯＨの添加により反応を停止させた。ＯＤを４０５ｎｍでマイクロタイタープレートで測定した。対照は細胞なし、細胞だけ、細胞＋ＦＧＦ（５ｎｇ／ｍＬ）、細胞＋ＶＥＧＦ（３ｎｇ／ｍＬ）、細胞＋ＶＥＧＦ（３ｎｇ／ｍＬ）＋ＴＧＦ－β（１ｎｇ／ｍＬ）、及び細胞＋ＶＥＧＦ（３ｎｇ／ｍＬ）＋ＬＩＦ（５ｎｇ／ｍＬ）であった。（１ｎｇ／ｍｌでのＴＧＦ－βはＶＥＧＦ刺激細胞増殖の７０－９０％を阻止することが知られている。）
結果は、（１）刺激なしの細胞に対して、及び（２）ＶＥＧＦ刺激活性の参照ＴＧＦ－β阻害に対して、ＯＤ４０５ｎｍでの酸性ホスファターゼ活性を測定することにより決定される、ＶＥＧＦ（３ｎｇ／ｍＬ）刺激細胞増殖の阻害パーセントを算定することにより評価した。以下の表２に示す結果は、ガン治療、特に主要血管形成を阻害するＰＲＯポリペプチドの有用性を示す。表２に示される数値（相対阻害度）は、刺激なしの細胞に対するＰＲＯポリペプチドによるＶＥＧＦ刺激増殖の阻害パーセントを算定し、ＶＥＧＦ刺激細胞増殖の７０－９０％を阻害することが知られている１ｎｇ／ｍｌのＴＧＦ－βにより得られた阻害パーセントにそのパーセントを割ることにより決定される。
【０３００】

表２

PRO名	PRO濃度	相対阻害度
PRO211	0.01%	99.0
PRO211	0.01%	1.09
PRO211	0.1%	0.95
PRO211	0.1%	67.0
PRO211	1.0%	0.27
PRO211	1.0%	20.0
PRO217	0.01%	1.06
PRO217	0.1%	0.84
PRO217	1.0%	0.39
PRO217	2.5 μM	0.2
PRO217	25 nM	0.88
PRO217	250 nM	0.58
PRO187	0.01%	0.91
PRO187	0.1%	0.82
PRO187	1.0%	0.44
PRO219	5.7 μM	0.61
PRO219	57 nM	1.09
PRO219	570 nM	0.97
PRO246	0.01%	1.04
PRO246	0.1%	1.0
PRO246	1.0%	0.49
PRO228	0.01%	0.99
PRO228	0.1%	0.93
PRO228	1.0%	0.57
PRO228	0.01%	0.95
PRO228	0.01%	0.98
PRO228	0.1%	0.77

PRO228	0.1%	0.88
PRO228	1.0%	0.16
PRO228	1.0%	0.48
PRO245	0.01%	0.76
PRO245	0.1%	0.35
PRO245	1.0%	0.11
PRO245	0.48 nM	1.03
PRO245	4.8 nM	0.95
PRO245	48 nM	0.49
PRO221	0.01%	1.03
PRO221	0.01%	1.06
PRO221	0.1%	0.82
PRO221	0.1%	0.93
PRO221	1.0%	0.31
PRO221	1.0%	0.43
PRO258	0.01%	0.98
PRO258	0.01%	1.06
PRO258	0.1%	0.95
PRO258	0.1%	1.02
PRO258	1.0%	0.6
PRO258	1.0%	0.69
PRO301	7.0 μM	1.02
PRO301	70 μM	0.88
PRO301	700 μM	0.44
PRO301	0.01%	0.92
PRO301	0.1%	0.85
PRO301	1.0%	0.68
PRO224	0.01%	101.0
PRO224	0.1%	65.0
PRO224	1.0%	23.0
PRO272	0.01%	0.95
PRO272	0.1%	0.57
PRO272	1.0%	0.18
PRO328	0.01%	0.98
PRO328	0.1%	0.96
PRO328	1.0%	0.6
PRO331	0.01%	0.88
PRO331	0.1%	0.82
PRO331	1.0%	0.56

【０３０１】

実施例６７：網膜ニューロンの生存

この実施例は網膜ニューロン細胞の生存を亢進する効果をＰＲＯ２２０ポリペプチドが有していることを証明している。

出生後７日のＳＤラットの子（混合集団：グリア及び網膜ニューロン型）をCO_2麻酔に続いて断頭して殺し、無菌条件下で眼を除去した。神経網膜を色素上皮と他の眼の組織から切り裂き、ついでCa^{2+}、Mg^{2+}フリーのＰＢＳ中の０．２５％トリプシンを使用して単一の細胞懸濁液中に解離させた。網膜を７－１０分間３７℃でインキュベートし、その後、１ｍｌの大豆トリプシン阻害剤を添加することによりトリプシンを不活性化する。細胞をＮ２を補填し、特定の試験ＰＲＯポリペプチドを伴うか伴わないＤＭＥＭ／Ｆ１２中の９６ウェルプレートにウェル当り１００，０００細胞でプレーティングする。全て

の実験に対して細胞は５％のCO_2の水飽和大気下で３７℃にて成長させた。培養２－３日後に細胞をカルセインＡＭで染色し、ついで４％パラホルムアルデヒドを使用して固定し、全細胞計数の決定のためにＤＡＰＩで染色した。全細胞（蛍光）はＣＣＤカメラとマッキントッシュのＮＩＨイメージソフトウェアを使用して２０ｘの対物倍率で定量する。ウェル中のフィールドはランダムに選択される。

ＰＲＯ２２０ポリペプチドの様々な濃度の効果を以下の表３に報告するが、ここで、生存パーセントは培養２－３日でのＤＡＰＩ標識細胞の全数により培養２－３日でのカルセインＡＭ陽性細胞の全数を割ることにより計算される。３０％の生存度を超えるものがポジティブと考えられる。

【０３０２】

表３

PRO名	PRO濃度	生存パーセント
PRO220	0.01%	2.4%
PRO220	0.01%	4.1%
PRO220	0.1%	3.0%
PRO220	0.1%	3.1%
PRO220	1.0%	72.4%
PRO220	1.0%	42.1%

【０３０３】

実施例６８：桿体光受容器の生存

この実施例は桿体光受容器細胞の生存を亢進する効果をＰＲＯ２２０ポリペプチドが有していることを証明している。

出生後７日のＳＤラットの子（混合集団：グリア及び網膜ニューロン型）をCO_2麻酔に続いて断頭して殺し、無菌条件下で眼を除去した。神経網膜を色素上皮と他の眼の組織から切り裂き、ついでCa^{2+}、Mg^{2+}フリーのＰＢＳ中の０．２５％トリプシンを使用して単一の細胞懸濁液中に解離させた。網膜を７－１０分間３７℃でインキュベートし、その後、１ｍｌの大豆トリプシン阻害剤を添加することによりトリプシンを不活性化する。細胞をＮ２を補填し、特定の試験ＰＲＯポリペプチドを伴うか伴わないＤＭＥＭ／Ｆ１２中の９６ウェルプレートにウェル当り１００，０００細胞でプレーティングする。全ての実験に対して細胞は５％のCO_2の水飽和大気下で３７℃にて成長させた。培養２－３日後に細胞を４％パラホルムアルデヒドを使用して固定し、ついでＣｅｌｌＴｒａｃｋｅｒ Ｇｒｅｅｎ ＣＭＦＤＡを使用して染色した。視物質ロドプシンに対して産生されたモノクロナール抗体Ｒｈｏ４Ｄ２（腹水又はＩｇＧ１：１００）が、間接的免疫蛍光法により桿体光受容器細胞を検出するために使用される。結果は、％生存度：培養２－３日でのロドプシン陽性細胞の全数により、培養２－３日でのカルセイン／ＣｅｌｌＴｒａｃｋｅｒ－ロドプシン陽性細胞の全数を割ったものとして報告される。全細胞（蛍光）はＣＣＤカメラとマッキントッシュのＮＩＨイメージソフトウェアを使用して２０ｘの対物倍率で定量する。ウェル中のフィールドはランダムに選択される。

ＰＲＯ２２０ポリペプチドの様々な濃度の効果を以下の表４に報告する。１０％を越える生存度がポジティブと考えられる。

【０３０４】

表４

PRO名	PRO濃度	生存パーセント
PRO220	0.01%	0.0%
PRO220	0.1%	0.0%
PRO220	2.0%	0.0%
PRO220	10%	0.0%
PRO220	20%	66.9%
PRO220	1.0%	56.9%

【０３０５】

実施例６９：内皮細胞アポトーシスの誘導
内皮細胞においてアポトーシスを誘導するＰＲＯ２２８ポリペプチドの能力を、１００ｎｇ／ｍｌのＶＥＧＦを補填した０％血清培地中で９６ウェルフォマットを使用して、ヒト静脈の臍静脈内皮細胞（ＨＵＶＥＣ，Ｃｅｌｌ Ｓｙｓｔｅｍｓ）中で試験した。（ＨＵＶＥＣ細胞はプレーティング表面から容易に取り除かれるのでウェル中の全てのピペット操作は可能な限り穏やかになされなければならない。）
培地を吸引し、細胞をＰＢＳで一度洗浄した。５ｍｌの１ｘトリプシンをＴ－１７５フラスコ中の細胞に添加し、細胞を、プレートから遊離するまで放置した（約５－１０分）。トリプシン処理を５ｍｌの成長培地を添加することにより停止した。細胞を４℃にて５分間１０００ｒｐｍで遠心分離した。培地を吸引し、細胞を、１０ｍｌの１０％血清補完培地（Ｃｅｌｌ Ｓｙｓｔｅｍｓ），１ｘｐｅｎｎ／ｓｔｒｅｐ中に再懸濁させた。
細胞を、１００μｌの全容積中、ウェル当り２ｘ１０^{4}細胞の密度で、１０％血清（ＣＳＧ－培地，Ｃｅｌｌ Ｓｙｓｔｅｍｓ）で、９６ウェルマイクロタイタープレート（Ａｍｅｒｓｈａｍ Ｌｉｆｅ Ｓｃｉｅｎｃｅ，ｃｙｔｏｓｔａｒ－Ｔ シンチレーションマイクロプレート，ＲＰＮＱ１６０，無菌、組織培養処理、個々にラッピング）にプレーティングした。ＰＲＯ２２８ポリペプチドを１％、０．３３％及び０．１１％の希釈で三つ組で添加した。細胞のないウェルをブランクとして使用し、細胞だけのウェルを負の対照として使用した。正の対照として、３ｘ原液のスタウロスポリンの５０μｌの１：３の連続希釈物を使用した。アポトーシスを誘導するＰＲＯ２２８ポリペプチドの能力は、アポトーシスを検出するためにカルシウム及びリン脂質結合タンパク質のメンバーであるアネキシンＶを使用して決定した。
０．２ｍｌのアネキシンＶ－ビオチン原液（１００μｇ／ｍｌ）を、４．６ｍｌの２ｘCa^{2+}結合バッファー及び２．５％ＢＳＡ中に希釈した（１：２５希釈）。５０μｌの希釈アネキシンＶ－ビオチン液を、１．０μｇ／ｍｌの最終濃度になるまで各ウェル（対照を除く）に添加した。試料を^{35}Ｓ－ストレプトアビジンの直接添加の前にアネキシン－ビオチンと共に１０－１５分間インキュベートした。^{35}Ｓ－ストレプトアビジンは２ｘCa^{2+}結合バッファー、２．５％ＢＳＡ中に希釈し、最終濃度が３ｘ１０^{4}ｃｐｍ／ウェルになるまで全てのウェルに添加した。プレートをついで密封し、１０００ｒｐｍで１５分間遠心分離し、２時間の間、軌道シェイカー上に配した。分析は１４５０ Ｍｉｃｒｏｂｅｔａ Ｔｒｉｌｕｘ（Ｗａｌｌａｃ）で実施した。バックグラウンドを越えるパーセントが負の対照を越える毎分当たりのカウントのパーセント量を表す以下の表５に結果を示す。バックグラウンドを越えるパーセントが３０％以上のものがポジティブであると考えられる。

【０３０６】

表５

PRO名	PRO濃度	バックグラウンドを越えるパーセント
PRO228	0.11%	0.7%
PRO228	0.11%	47.6%
PRO228	0.33%	92.2%
PRO228	0.33%	123.7%
PRO228	1.0%	51.4%
PRO228	1.0%	95.3%

【０３０７】
実施例７０：ＰＤＢ１２細胞の阻害
この実施例は、様々なＰＲＯポリペプチドがＰＤＢ１２膵管細胞によるタンパク質の生産を阻害する効能を有していることを証明する。
ＰＤＢ１２膵管細胞を、１００μＬ／１８０μＬ成長培地中にウェル当り１．５ｘ１０^{3}細胞で、フィブロネクチンをコートした９６ウェルプレート上にプレーティングした。ＰＲＯポリペプチド試験試料又はＰＲＯポリペプチドを欠く負の対照を有する１００μＬの成長培地をついで最終容積２００μＬになるまでウェルに添加した。対照はこのアッセイ

において不活性であることが知られているタンパク質を含む成長培地を含む。細胞を３７℃で４日間インキュベートした。ついで、２０μＬのアラマーブルー染料（ＡＢ）を各ウェルに添加し、蛍光読みとりを、５３０ｎｍ励起及び５９０ｎｍ発光にてマイクロタイタープレートリーダーでＡＢ添加後４時間で測定した。用いた標準はウシ下垂体抽出物（ＢＰＥ）なしの細胞と様々な濃度のＢＰＥの細胞である。未知のバッファー又はＣＭ対照は各９６ウェルプレートで２回なされる。

これらのアッセイの結果は以下の表６に示されており、そこでは、タンパク生産の減少パーセントが、負の対照細胞により生産されたアラマーブルー染料による算定タンパク質濃度と、ＰＲＯポリペプチド処理細胞により生産されたアラマーブルー染料による算定タンパク質濃度を比較することにより算定されている。負の対照細胞と比較して２５％以上のタンパク質生産の減少パーセントがポジティブと考えられる。

【０３０８】

表６

PRO名	PRO濃度	タンパク生産の減少パーセント
PRO211	0.1%	0.0%
PRO211	0.01%	0.6%
PRO211	1.0%	59.7%
PRO287	2.0%	22.3%
PRO287	10%	18.2%
PRO287	50%	67.5%
PRO287	2.0%	45.53%
PRO287	10%	57.3%
PRO287	50%	52.24%
PRO301	2.0%	0.0%
PRO301	10%	59.8%
PRO301	50%	65.6%
PRO293	2.0%	0.0%
PRO293	10%	40.4%
PRO293	50%	56.7%

【０３０９】

実施例７１：成人心臓肥大の刺激

このアッセイは成人心臓の肥大を刺激する様々なＰＲＯポリペプチドの能力を測定するように設計された。

成体（２５０ｇ）ＳＤラットから新たに単離した心室筋細胞を２０００細胞／１８０μｌウェル容積でプレーティングした。細胞を単離し、１日目にプレーティングし、ＰＲＯポリペプチド含有試験試料又は成長培地のみ（負の対照）（２０μｌ容積）を2日目に添加し、ついで細胞を固定し、５日目に染色した。染色後、細胞サイズを可視化し、対照細胞と比較して成長亢進のない細胞に０．０の値を与え、対照細胞と比較して小から中程度の成長亢進を示す細胞に０．１の値を与え、対照細胞と比較して大きな成長亢進を示す細胞に２．０の値を与える。負の対照と比較した成長増強のあらゆる度合いがアッセイではポジティブと考えられる。結果を以下の表７に示す。

【０３１０】

表７

PRO名	PRO濃度	成長亢進のスコア
PRO287	20%	1.0
PRO287	20%	1.0
PRO301	20%	1.0
PRO301	20%	1.0
PRO293	20%	1.0
PRO293	20%	1.0
PRO303	20%	1.0
PRO303	20%	1.0

【０３１１】

実施例７２：ＰＤＢ１２細胞の増殖

この実施例は、様々なＰＲＯポリペプチドがＰＤＢ１２膵管細胞の増殖を誘導する効能を有していることを証明する。

ＰＤＢ１２膵管細胞を、１００μＬ／１８０μＬ成長培地中にウェル当り１．５ｘ１０^{3}細胞で、フィブロネクチンをコートした９６ウェルプレート上にプレーティングした。ＰＲＯポリペプチド試験試料又はＰＲＯポリペプチドを欠く負の対照を有する１００μＬの成長培地をついで最終容積２００μＬになるまでウェルに添加した。対照はこのアッセイにおいて不活性であることが知られているタンパク質を含む成長培地を含む。細胞を３７℃で４日間インキュベートした。ついで、２０μＬのアラマーブルー染料（ＡＢ）を各ウェルに添加し、蛍光読みとりを、５３０ｎｍ励起及び５９０ｎｍ発光にてマイクロタイタープレートリーダーでＡＢ添加後４時間で測定した。用いた標準はウシ下垂体抽出物（ＢＰＥ）なしの細胞と様々な濃度のＢＰＥの細胞である。未知のバッファー又は成長培地のみの対照は各９６ウェルプレートで２回なされる。

これらのアッセイの結果は以下の表８に示されており、そこでは、タンパク生産の増加パーセントが、負の対照細胞により生産されたアラマーブルー染料による算定タンパク質濃度と、ＰＲＯポリペプチド処理細胞により生産されたアラマーブルー染料による算定タンパク質濃度を比較することにより算定されている。負の対照細胞と比較して２５％以上のタンパク質生産の増加パーセントがポジティブと考えられる。

【０３１２】

表８

PRO名	PRO濃度	成長亢進のスコア
PRO301	2.0%	44.0%
PRO301	10%	67.4%
PRO301	50%	185.8%
PRO303	2.0%	27.9%
PRO303	10%	174.9%
PRO303	50%	193.1%

【０３１３】

実施例７３：ＰＲＯ２２４により誘導された心臓新生児肥大の亢進

このアッセイは新生児心臓の肥大を刺激するＰＲＯ２２４ポリペプチドの能力を測定するように設計されている。

１日齢のＨａｒｌａｎ Ｓｐｒａｇｕｅ Ｄａｗｌｅｙラットから筋細胞を得た。細胞（７．５ｘ１０^{4}／ｍｌで１８０μｌ、血清＜０．１％、新たに単離）をＤＭＥＭ／Ｆ１２＋４％ＦＣＳで前もってコートした９６ウェルプレートに１日目に添加した。試験ＰＲＯ２２４ポリペプチド又は成長培地のみ（負の対照）（２０μｌ／ウェル）を含む試験試料を１日目にウェルに直接添加した。ついで、ＰＧＦ（２０μｌ／ウェル）を１０$^{-6}$Ｍの最終濃度まで２日目に添加した。ついで、細胞を４日目に染色し、５日目に視覚的にスコアをつけ、負の対照と比較してサイズの増加がない細胞をスコアが０．０とし、負の対照と比較してサイズが小から中程度増加している細胞を１．０とし、負の対照と比較してサイ

ズが大きく増加している細胞を２．０のスコアとした。結果を以下の表９に示す。

【０３１４】

表９

PRO名	PRO濃度	成長亢進のスコア
PRO224	0.01%	0.0
PRO224	0.1%	0.0
PRO224	1.0%	1.0

【０３１５】

実施例７４：インサイツハイブリッド形成

インサイツハイブリッド形成は、細胞又は組織調製物内での核酸配列の検出及び局在化のための強力で多用途の技術である。それは、例えば、遺伝子発現部位の同定、転写物の組織分布の分析、ウイルス感染の同定と局在化、特定のｍＲＮＡ合成における変化の追跡及び染色体マッピングにおける補助に有用である。

インサイツハイブリッド形成は、Ｌｕ及びＧｉｌｌｅｔｔ，Ｃｅｌｌ Ｖｉｓｉｏｎ １：１６９－１７６（１９９４）のプロトコールの最適化バージョンに従って、ＰＣＲ生成^{33}Ｐ－標識リボプローブを用いて実施される。簡単には、ホルマリン固定、パラフィン包埋ヒト組織を切片化し、脱パラフィンし、プロテイナーゼＫ（２０ｇ／ｍｌ）で１５分間３７℃で脱タンパクし、さらに上掲のＬｕ及びＧｉｌｌｅｔｔに記載されたようにインサイツハイブリッド形成する。（^{33}－Ｐ）ＵＴＰ－標識アンチセンスリボプローブをＰＣＲ産物から生成し、５５℃で終夜ハイブリッド形成する。スライドをＫｏｄａｋ ＮＴＢ２核トラックエマルションに浸漬して４週間露出する。

^{33}Ｐ－リボプローブ合成

６．０μｌ（１２５ｍＣｉ）の^{33}Ｐ－ＵＴＰ（Ａｍｅｒｓｈａｍ ＢＦ １００２，ＳＡ＜２０００ Ｃｉ／ｍｍｏｌ）をスピード真空乾燥させた。乾燥^{33}Ｐ－ＵＴＰを含む各管に以下の成分を添加した：

２．０μｌの５ｘ転写バッファー

１．０μｌのＤＴＴ（１００ｍＭ）

２．０μｌのＮＴＰ混合物（２．５ｍＭ：各１０μｌの１０ｍＭ ＧＴＰ，ＣＴＰ及びＡＴＰ＋１０μｌのH_2O）

１．０μｌのＵＴＰ（５０μＭ）

１．０μｌのＲＮＡｓｉｎ

１．０μｌのＤＮＡテンプレート（１μｇ）

１．０μｌのH_2O

１．０μｌのＲＮＡポメラーゼ（ＰＣＲ産物についてＴ３＝ＡＳ，Ｔ７＝Ｓ，通常）

管を３７℃で１時間インキュベートし、１．０μｌのＲＱ１ ＤＮａｓｅを添加し、次いで３７℃で１５分間インキュベートした。９０μｌのＴＥ（１０ｍＭトリスｐＨ７．６／１ｍＭのＥＤＴＡｐＨ８．０）を添加し、混合物をＤＥ８１紙にピペットした。残りの溶液をＭｉｃｒｏｃｏｎ－５０限外濾過ユニットに負荷し、プログラム１０を用いてスピンさせた（６分間）。濾過ユニットを第２の管に変換し、プログラム２を用いてスピンさせた（３分間）。最終回収スピンの後、１００μｌのＴＥを添加した。１μｌの最終生成物をＤＥ８１紙にピペットし６ｍｌのＢＩＯＦＬＵＯＲ ＩＩ^{TM}で数えた。

プローブをＴＢＥ／尿素ゲル上で走らせた。１－３μｌのプローブ又は５μｌのＲＮＡＭｒｋＩＩＩを３μｌの負荷バッファーに添加した。加熱ブロック上で９５℃に３分間加熱した後、ゲルを即座に氷上に置いた。ゲルのウェルをフラッシングし、試料を負荷し、１８０－２５０ボルトで４５分間走らせた。ゲルをサランラップでラップし、－７０℃冷凍機内で補強スクリーンを持つＸＡＲフィルムに１時間から終夜露出した。

【０３１６】

^{33}Ｐ－ハイブリッド形成

Ａ．凍結切片の前処理

スライドを冷凍機から取り出し、アルミニウムトレイに配置して室温で５分間解凍した。

トレイを５５℃のインキュベータに５分間配置して凝結を減らした。スライドを蒸気フード内において４％パラホルムアルデヒド中で５分間固定し、０．５ｘＳＳＣで５分間室温で洗浄した（２５ｍｌ　２０ｘＳＳＣ　＋　９７５ｍｌ　ＳＱ　H_2O）。０．５μｇ／ｍｌのプロテイナーゼ中、３７℃で１０分間の脱タンパクの後（２５０ｍｌの予備加熱ＲＮａｓｅ無しＲＮａｓｅバッファー中の１０ｍｇ／ｍｌストック１２．５μｌ）、切片を０．５ｘＳＳＣで１０分間室温で洗浄した。切片を、７０％、９５％、１００％エタノール中、各２分間脱水した。

Ｂ．パラフィン包埋切片の前処理：

スライドを脱パラフィンし、ＳＱ　H_2O中に配置し、２ｘＳＳＣで室温において各々５分間２回リンスした。切片を２０μｇ／ｍｌのプロテイナーゼＫ（２５０ｍｌのＲＮａｓｅ無しＲＮａｓｅバッファー中１０ｍｇ／ｍｌを５００μｌ；３７℃、１５分間）－ヒト胚又は８ｘプロテイナーゼＫ（２５０ｍｌのＲＮａｓｅバッファー中１００μｌ、３７℃、３０分間）－ホルマリン組織で脱タンパクした。続く０．５ｘＳＳＣでのリンス及び脱水は上記のように実施した。

Ｃ．プレハイブリッド化：

スライドをＢｏｘバッファー（４ｘＳＳＣ、５０％ホルムアミド）－飽和濾紙で列を作ったプラスチックボックスに並べた。組織を５０μｌのハイブリッド形成バッファー（３．７５ｇデキストラン硫酸＋６ｍｌＳＱ　H_2O）で被覆し、ボルテックスし、キャップを外して２分間マイクロ波で加熱した。氷上で冷却した後、１８．７５ｍｌのホルムアミド、３．７５ｍｌの２０ｘＳＳＣ及び９ｍｌのＳＱ　H_2Oを添加し、組織を良くボルテックスし、４２℃で１－４時間インキュベートした。

Ｄ．ハイブリッド形成：

スライド当たり１．０ｘ10^6ｃｐｍのプローブ及び１．０μｌのｔＲＮＡ（５０ｍｇ／ｍｌストック）を９５℃で３分間加熱した。スライドを氷上で冷却し、スライド当たり４８μｌのハイブリッド形成バッファーを添加した。ボルテックスの後、５０μｌの^{33}P混合物をスライド上のプレハイブリッド５０μｌに添加した。スライドを５５℃で終夜インキュベートした。

Ｅ．洗浄：

洗浄は、２ｘ１０分間、２ｘＳＳＣ、ＥＤＴＡで室温で実施し（４００ｍｌの２０ｘＳＳＣ＋１６ｍｌの０．２５Ｍ　ＥＤＴＡ、Ｖｆ＝４Ｌ）、次いでＲＮａｓｅＡ処理を３７℃で３０分間行った（２５０ｍｌＲＮａｓｅバッファー中１０ｍｇ／ｍｌを５００μｌ＝２０μｇ／ｍｌ）。スライドを２ｘ１０分間、ＥＤＴＡで室温において洗浄した。緊縮性洗浄条件は次の通り：５５℃で２時間、０．１ｘＳＳＣ、ＥＤＴＡ（２０ｍｌの２０ｘＳＳＣ＋１６ｍｌのＥＤＴＡ、V_f＝４Ｌ）。

【０３１７】

Ｆ．オリゴヌクレオチド

ここに開示したＤＮＡ配列の二つについてインサイツ分析を実施した。これらの分析に対して用いたオリゴヌクレオチドは次の通りである：

（１）ＤＮＡ３３０９４－１１３１（ＰＲＯ２１７）

ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＴＣＡＧＡＡＡＡＧＣＧＣＡＡＣＡＧＡＧＡＡ－３’（配列番号：３４８）

ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＴＧＴＣＴＴＣＣＡＴＧＣＣＡＡＣＣＴＴＣ－３’（配列番号：３４９）

（２）ＤＮＡ３３２２３－１１３６（ＰＲＯ２３０）

ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＧＧＣＧＡＴＧＴＣＣＡＣＴＧＧＧＧＣＴＡＣ－３’（配列番号：３５０）

ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＣＧＡＧＧＡＡＧＡＴＧＧＧＣＧＧＡＴＧＧＴ－３’（配列番号：３５１）

（３）ＤＮＡ３４４３５－１１４０（ＰＲＯ２３２）

ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＡＣＣＣＡＣＧ

ＣＧＴＣＣＧＧＣＴＧＣＴＴ－３’（配列番号：３５２）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＣＧＧＧＧＧＡ
ＣＡＣＣＡＣＧＧＡＣＣＡＧＡ－３’（配列番号：３５３）
（４）ＤＮＡ３５６３９－１１７２（ＰＲＯ２４６）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＴＴＧＣＴＧＣ
ＧＧＴＴＴＴＴＧＴＴＣＣＴＧ－３’（配列番号：３５４）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＧＣＴＧＣＣＧ
ＡＴＣＣＣＡＣＴＧＧＴＡＴＴ－３’（配列番号：３５５）
（５）ＤＮＡ４９４３５－１２１９（ＰＲＯ５３３）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＧＧＡＴＣＣＴ
ＧＧＣＣＧＧＣＣＴＣＴＧ－３’（配列番号：３５６）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＧＣＣＣＧＧＧ
ＣＡＴＧＧＴＣＴＣＡＧＴＴＡ－３’（配列番号：３５７）
（６）ＤＮＡ３５６３８－１１４１（ＰＲＯ２４５）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＧＧＧＡＡＧＡ
ＴＧＧＣＧＡＧＧＡＧＧＡＧ－３’（配列番号：３５８）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＣＣＡＡＧＧＣ
ＣＡＣＡＡＡＣＧＧＡＡＡＴＣ－３’（配列番号：３５９）
（７）ＤＮＡ３３０８９－１１３２（ＰＲＯ２２１）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＴＧＴＧＣＴＴ
ＴＣＡＴＴＣＴＧＣＣＡＧＴＡ－３’（配列番号：３６０）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＧＧＧＴＡＣＡ
ＡＴＴＡＡＧＧＧＧＴＧＧＡＴ－３’（配列番号：３６１）
（８）ＤＮＡ３５９１８－１１７４（ＰＲＯ２５８）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＣＣＧＣＣＴＣ
ＧＣＴＣＣＴＧＣＴＣＣＴＧ－３’（配列番号：３６２）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＧＧＡＴＴＧＣ
ＣＧＣＧＡＣＣＣＴＣＡＣＡＧ－３’（配列番号：３６３）
（９）ＤＮＡ３２２８６－１１９１（ＰＲＯ２１４）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＣＣＣＴＣＣＴ
ＧＣＣＴＴＣＣＣＴＧＴＣＣ－３’（配列番号：３６４）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＧＴＧＧＴＧＧ
ＣＣＧＣＧＡＴＴＡＴＣＴＧＣ－３’（配列番号：３６５）
（１０）ＤＮＡ３３２２１－１１３３（ＰＲＯ２２４）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＧＣＡＧＣＧＡ
ＴＧＧＣＡＧＣＧＡＴＧＡＧＧ－３’（配列番号：３６６）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＣＡＧＡＣＧＧ
ＧＧＣＡＧＡＧＧＧＡＧＴＧ－３’（配列番号：３６７）
（１１）ＤＮＡ３５５５７－１１３７（ＰＲＯ２３４）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＣＡＧＧＡＧＧ
ＣＧＴＧＡＧＧＡＧＡＡＡＣ－３’（配列番号：３６８）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＡＡＧＡＣＡＴ
ＧＴＣＡＴＣＧＧＧＡＧＴＧＧ－３’（配列番号：３６９）
（１２）ＤＮＡ３３１００－１１５９（ＰＲＯ２２９）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＣＧＧＧＴＧＧ
ＡＧＧＴＧＧＡＡＣＡＧＡＡＡ－３’（配列番号：３７０）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＣＡＣＡＧＡＣ
ＡＧＡＧＣＣＣＣＡＴＡＣＧＣ－３’（配列番号：３７１）
（１３）ＤＮＡ３４４３１－１１７７（ＰＲＯ２６３）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＣＡＧＧＧＡＡ

ＡＴＣＣＧＧＡＴＧＴＣＴＣ－３’（配列番号：３７２）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＧＴＡＡＧＧＧＧＡＴＧＣＣＡＣＣＧＡＧＴＡ－３’（配列番号：３７３）
（１４）ＤＮＡ３８２６８－１１８８（ＰＲＯ２９５）
ｐ１　５’－ＧＧＡＴＴＣＴＡＡＴＡＣＧＡＣＴＣＡＣＴＡＴＡＧＧＧＣＣＡＧＣＴＡＣＣＣＧＣＡＧＧＡＧＧＡＧＧ－３’（配列番号：３７４）
ｐ２　５’－ＣＴＡＴＧＡＡＡＴＴＡＡＣＣＣＴＣＡＣＴＡＡＡＧＧＧＡＴＣＣＣＡＧＧＴＧＡＴＧＡＧＧＴＣＣＡＧＡ－３’（配列番号：３７５）
【０３１８】
Ｇ．結果
ここに開示した上記の２つのＤＮＡ配列についてインサイツ分析を実施した。
これらの分析からの結果は次の通りである：
（１）ＤＮＡ３３０９４－１１３１（ＰＲＯ２１７）
明らかな生物学的機能は示さない高度に特徴的な発現パターンである。ヒト胚では、ＧＩ路の外側平滑筋層、呼吸器軟骨、分岐呼吸器上皮、骨芽細胞、腱、生殖腺、視神経頭及び発達真皮に発現された。成体では、チンパンジーの舌の表皮ペグ、前立腺の基底上皮細胞／筋上皮細胞及び膀胱に発現が観察された。成体肺の歯槽内側細胞、ペニスの勃起組織に並列した間葉細胞及び大脳皮質（おそらくはグリア細胞）にも発現された。腎臓では、発現は疾患、甲状腺化腎尿細管を囲む細胞でのみ見られた。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤及び下肢。
調べた成人ヒト組織：腎臓（正常及び末期）、副腎、心筋層、大動脈、脾臓、リンパ節、胆嚢、膵臓、肺、皮膚、眼（網膜を含む）、前立腺、膀胱、肝臓（正常、硬変、急性不全）。
調べた非ヒト霊長類組織：
（ａ）チンパンジー組織：唾液腺、胃、甲状腺、副甲状腺、皮膚、胸腺、卵巣、リンパ節。
（ｂ）アカゲザル組織：大脳皮質、海馬、小脳、ペニス。
（２）ＤＮＡ３３２２３－１１３６（ＰＲＯ２３０）
切片は胎児の動脈及び静脈血管に付随した強いシグナルを示した。動脈では、シグナルは平滑筋／周皮細胞に制限されているようにみえた。シグナルは毛細血管と糸球体にもまた見られた。内皮細胞がこのｍＲＮＡを発現していたかどうか明らかではない。発現はまた胎児水晶体の上皮細胞にも観察された。強い発現がまた胎盤栄養膜絨毛内の細胞にも見られた；これらの細胞は胎盤とＨＧＦ未定組織形成を発現する線維芽細胞様細胞の間にある。成人では発現の証拠はなく、大動脈と殆どの血管の壁は陰性のようであった。しかし、発現は正常な前立腺と胆嚢の内側の上皮の血管チャネルにわたって見られた。インシュアラー発現は軟部組織肉腫と腎細胞ガンの血管に見られた。
要約すると、これは胎児並びに幾つかの成人の器官に比較的特異的な血管発現を示す分子である。発現はまた胎児水晶体と成人胆嚢に観察された。
二次スクリーンでは、血管発現は上で観察されたものと同様に、胎児ブロックで観察された。発現はまた内皮よりもむしろ血管平滑筋上であった。発現はまた発育中の食道の平滑筋にも見られ、よってこの分子は血管特異的ではない。４つの肺及び４つの乳ガンにおいて発現が試験された。大幅な発現は少なくとも３／４の肺ガンと２／４の乳ガンの血管平滑筋に見られた。また、１つの乳ガンでは、未定組織形成（おそらくは筋線維芽細胞）の腫瘍周囲間質細胞に発現が観察された。この実験では内皮細胞発現は観察されなかった。
（３）ＤＮＡ３４４３５－１１４０（ＰＲＯ２３２）
前立腺上皮と膀胱上皮に強い発現が、気管支上皮に低いレベルの発現があった。とりわけ、骨、血液、軟骨肉腫、成人心臓及び胎児肝臓を含む多くの部位に高バックグラウンド／低レベルの発現が見られた。部分的にはこのレベルのシグナルは血液に見られたので、こ

のレベルのシグナルがバックグラウンドを示すと考えられる。全ての他のシグナルは陰性である。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤、精巣及び下肢。
調べた成人ヒト組織：腎臓（正常及び末期）、副腎、脾臓、リンパ節、膵臓、肺、眼（網膜を含む）、膀胱、肝臓（正常、硬変、急性不全）。
調べた非ヒト霊長類組織：
チンパンジー組織：副腎。
アカゲザル組織：大脳皮質、海馬。
二次スクリーンでは、前立腺の上皮、膀胱の尿管皮（ｕｒｅｔｈｅｌｉｕｍ）の表層、腎盂の内側の尿管皮、輸尿管の尿管皮（２実験のうち１つ）に発現が観察された。アカゲザルの輸尿管は陰あるいは性であった；これが輸尿管による発現の実際の欠乏を表しているのかどうか、あるいはそれがアカゲザル組織とプローブが交差反応できなかったせいなのかは不明である。前立腺と膀胱における知見は、同位体検出法を使用してこれまでに記載されたものと同様である。この抗原に対するｍＲＮＡの発現は前立腺上皮特異的ではない。抗原は尿管皮由来組織に対する有用なマーカーとなりうる。尿路上位の表面の有糸分裂後の細胞における発現は、特異的幹細胞マーカーを提示することは、これが基底上皮に特異的に発現されることが予想されるので、ないでろうことをまた示唆している。
【０３１９】
（４）ＤＮＡ３５６３９－１１７２（ＰＲＯ２４６）
ＣＮＳの組織を含む胎児血管内皮に強い発現があった。ＣＮＳを含む成人脈管構造には低レベルの発現。腫瘍血管内皮におけるより高レベルの発現は明らかではなかった。シグナルはまた骨基質及び成人脾臓にも見られ、明らかには細胞付随ではなく、おそらくこれらの部位で非特異的バックグラウンドに関連している。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤、精巣及び下肢。
調べた成人ヒト組織：腎臓（正常及び末期）、副腎、脾臓、リンパ節、膵臓、肺、眼（網膜を含む）、膀胱、肝臓（正常、硬変、急性不全）。
調べた非ヒト霊長類組織：
チンパンジー組織：副腎。
アカゲザル組織：大脳皮質、海馬。
（５）ＤＮＡ４９４３５－１２１９（ＰＲＯ５３３）
胎児脳の皮質性ニューロンにわたって穏やかな発現。胎児網膜の内側側面にわたる発現があり、発育中の水晶体に可能な発現がある。胎児皮膚、軟骨、小腸、胎盤絨毛及び臍帯にわたる発現。成人組織では、胆嚢上皮にわたって極めて高レベルの発現がある。成体腎臓、胃及び結腸上皮にわたって中程度の発現。低レベルの発現は多くの組織の多くの細胞型にわたって見られ、これはプローブの粘着性に関連しているかもしれず、よってこれらのデータはある程度の注意を持って解釈されなければならない。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤、精巣及び下肢。
調べた成人ヒト組織：腎臓（正常及び末期）、副腎、脾臓、リンパ節、膵臓、肺、眼（網膜を含む）、膀胱、肝臓（正常、硬変、急性不全）。
調べた非ヒト霊長類組織：
チンパンジー組織：副腎。
アカゲザル組織：大脳皮質、海馬、小脳。
（６）ＤＮＡ３５６３８－１１４１（ＰＲＯ２４５）
胎児及び胎盤血管のサブセットの内側の内皮に発現が観察された。内皮発現はこれらの組

織ブロックに限られていた。発現はまた胎盤の中間栄養芽細胞にわたって観察された。全ての他の組織は陰性である。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤及び下肢。
調べた成体組織：肝臓、腎臓、副腎、心筋層、大動脈、脾臓、リンパ節、膵臓、肺、皮膚、大脳皮質（アカゲザル）、海馬（アカゲザル）、小脳（アカゲザル）、ペニス、眼、膀胱、胃、胃ガン、大腸、大腸ガン、甲状腺（チンパンジー）、副甲状腺（チンパンジー）、卵巣（チンパンジー）及び軟骨肉腫。アセトミノフェン誘導肝臓傷害及び肝硬変。
（７）ＤＮＡ３３０８９－１１３２（ＰＲＯ２２１）
胎児脳白質及び灰白質、並びに脊髄のニューロンにわたっての特異的発現。プローブはラットと交差反応するように思われる。成体アカゲザルの脳の小脳ニューロンに低レベルの発現。全ての他の組織は陰性。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤及び下肢。
調べた成体組織：肝臓、腎臓、副腎、心筋層、大動脈、脾臓、リンパ節、膵臓、肺、皮膚、大脳皮質（アカゲザル）、海馬（アカゲザル）、小脳（アカゲザル）、ペニス、眼、膀胱、胃、胃ガン、大腸、大腸ガン及び軟骨肉腫。アセトミノフェン誘導肝臓傷害及び肝硬変。
（８）ＤＮＡ３５９１８－１１７４（ＰＲＯ２５８）
神経系に強い発現。アカゲザルの脳では、発現は皮質、海馬及び小脳ニューロンに観察されている。脊椎の脊髄ニューロン、発達中の脳及び胎児網膜の内側面にわたって発現。発達中の後根及び自律神経節並びに腸神経にわたって発現。成人前立腺の神経節細胞にわたって発現が観察された。ラットでは、発達中の後脳と脊髄に強い発現がある。胎盤絨毛の間質細胞にわたる強い発現。全ての他の組織は陰性。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤及び下肢。
調べた成体組織：肝臓、腎臓、腎臓細胞ガン、副腎、大動脈、脾臓、リンパ節、膵臓、肺、心筋層、皮膚、大脳皮質（アカゲザル）、海馬（アカゲザル）、小脳（アカゲザル）、膀胱、前立腺、胃、胃ガン、大腸、大腸ガン、甲状腺（チンパンジー）、副甲状腺（チンパンジー）、卵巣（チンパンジー）及び軟骨肉腫。アセトミノフェン誘導肝臓傷害及び肝硬変。
【０３２０】
（９）ＤＮＡ３２２８６－１１９１（ＰＲＯ２１４）
胎児組織：間葉中にわたる低レベルの発現。甲状腺と膜組織の胎盤間質性細胞における中程度の発現。皮質ニューロンにおける低レベルの発現。成体組織：全て陰性。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤及び下肢。
調べた成体組織：肝臓、腎臓、副腎、心筋層、大動脈、脾臓、リンパ節、膵臓、肺及び皮膚。
（１０）ＤＮＡ３３２２１－１１３３（ＰＲＯ２２４）
発現は、胎児脾臓、成体脾臓、胎児肝臓、成体甲状腺及び成体リンパ節（チンパンジー）に限られた。発現の更なる部位は発達中の脊椎神経節である。全ての他の祖危機は陰性である。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤及び下肢。

調べた成人ヒト組織：腎臓（正常及び末期）、副腎、心筋層、大動脈、脾臓、リンパ節、膵臓、肺、皮膚、眼（網膜を含む）、膀胱、肝臓（正常、硬変、急性不全）。
調べた非ヒト霊長類組織：
チンパンジー組織：唾液腺、胃、甲状腺、副甲状腺、皮膚、胸腺、卵巣、リンパ節。
アカゲザル組織：大脳皮質、海馬、小脳、ペニス。
（１１）ＤＮＡ３５５５７－１１３７（ＰＲＯ２３４）
胎児脊髄の腹側側面の発達中（脊髄の前角に発達）の運動ニューロンにわたって特異的な発現がある。全ての他の組織は陰性。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤及び下肢。
調べた成体組織：肝臓、腎臓、副腎、心筋層、大動脈、脾臓、リンパ節、膵臓、肺、皮膚、大脳皮質（アカゲザル）、海馬（アカゲザル）、小脳（アカゲザル）、ペニス、眼、膀胱、胃、胃ガン、大腸、大腸ガン及び軟骨肉腫。アセトミノフェン誘導肝臓傷害及び肝硬変。
（１２）ＤＮＡ３３１００－１１５９（ＰＲＯ２２９）
胎児と成体の脾臓の単核食細胞（マクロファージ）、肝臓、リンパ節及び成体胸腺（可染体マクロファージ中）に顕著に発現。最も高い発現は脾臓中である。全ての他の組織は陰性である。局在化と相同性は細網内皮系の細胞に対するスカベンジャーレセプターとしての役割と完全に一致している。胎盤単核細胞にまた発現が観察された。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤及び下肢。
調べた成人ヒト組織：腎臓（正常及び末期）、副腎、心筋層、大動脈、脾臓、リンパ節、胆嚢、膵臓、肺、皮膚、眼（網膜を含む）、前立腺、膀胱、肝臓（正常、硬変、急性不全）。
調べた非ヒト霊長類組織：
チンパンジー組織：唾液腺、胃、甲状腺、副甲状腺、皮膚、胸腺、卵巣、リンパ節。
アカゲザル組織：大脳皮質、海馬、小脳、ペニス。
（１３）ＤＮＡ３４４３１－１１７７（ＰＲＯ２６３）
ヒト胎児組織及び胎盤において単核細胞にわたって、おそらくはマクロファージ＋／－リンパ球において広範な発現。胎児副腎皮質の上皮細胞にもまた強い発現が見られる。全ての成人組織は陰性であった。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、心臓、大血管、食堂、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤及び下肢。
調べた成体組織：肝臓、腎臓、副腎、脾臓、リンパ節、膵臓、肺、皮膚、大脳皮質（アカゲザル）、海馬（アカゲザル）、小脳（アカゲザル）、膀胱、胃、胃ガン、大腸及び大腸ガン。アセトミノフェン誘導肝臓傷害及び肝硬変。
二次スクリーンは間質単核細胞、おそらくは組織球にわたる発現を証明した。
（１４）ＤＮＡ３８２６８－１１８８（ＰＲＯ２９５）
ヒト胎児脊髄神経節の神経節細胞と発達中の脊髄の前角の大ニューロンにわたって高い発現。成体では、チンパンジーの副腎髄質（神経）、アカゲザル脳のニューロン（海馬［＋＋＋］及び小脳）及び正常な成人ヒト前立腺（神経節細胞を含む、すなわちこの細胞型での発現は前立腺には制限されていないと思われる）のニューロンに発現があった。全ての他の組織は陰性。
調べたヒト胎児組織（Ｅ１２－Ｅ１６週）は次のものを含む：胎盤、臍帯、肝臓、腎臓、副腎、甲状腺、肺、大血管、胃、小腸、脾臓、胸腺、膵臓、脳、眼、脊髄、体壁、骨盤、精巣及び下肢。
調べた成人ヒト組織：腎臓（正常及び末期）、副腎、リンパ節、膵臓、肺、皮膚、眼（網

膜を含む）、膀胱、肝臓（正常、硬変、急性不全）。
調べた非ヒト霊長類組織：
チンパンジー組織：副腎
アカゲザル組織：大脳皮質、海馬、小脳。
【０３２１】
表１０

```
/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int     _day[26][26] = {
/*        A  B  C  D  E  F  G  H  I  J  K  L  M  N  O  P  Q  R  S  T  U  V  W  X  Y  Z */
/* A */    { 2, 0,-2, 0, 0,-4, 1,-1,-1, 0,-1,-2,-1, 0,_M, 1, 0,-2, 1, 1, 0, 0,-6, 0,-3, 0},
/* B */    { 0, 3,-4, 3, 2,-5, 0, 1,-2, 0, 0,-3,-2, 2,_M,-1, 1, 0, 0, 0, 0,-2,-5, 0,-3, 1},
/* C */    {-2,-4,15,-5,-5,-4,-3,-3,-2, 0,-5,-6,-5,-4,_M,-3,-5,-4, 0,-2, 0,-2,-8, 0, 0,-5},
/* D */    { 0, 3,-5, 4, 3,-6, 1, 1,-2, 0, 0,-4,-3, 2,_M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 2},
/* E */    { 0, 2,-5, 3, 4,-5, 0, 1,-2, 0, 0,-3,-2, 1,_M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 3},
/* F */    {-4,-5,-4,-6,-5, 9,-5,-2, 1, 0,-5, 2, 0,-4,_M,-5,-5,-4,-3,-3, 0,-1, 0, 0, 7,-5},
/* G */    { 1, 0,-3, 1, 0,-5, 5,-2,-3, 0,-2,-4,-3, 0,_M,-1,-1,-3, 1, 0, 0,-1,-7, 0,-5, 0},
/* H */    {-1, 1,-3, 1, 1,-2,-2, 6,-2, 0, 0,-2,-2, 2,_M, 0, 3, 2,-1,-1, 0,-2,-3, 0, 0, 2},
/* I */    {-1,-2,-2,-2,-2, 1,-3,-2, 5, 0,-2, 2, 2,-2,_M,-2,-2,-2,-1, 0, 0, 4,-5, 0,-1,-2},
/* J */    { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */    {-1, 0,-5, 0, 0,-5,-2, 0,-2, 0, 5,-3, 0, 1,_M,-1, 1, 3, 0, 0, 0,-2,-3, 0,-4, 0},
/* L */    {-2,-3,-6,-4,-3, 2,-4,-2, 2, 0,-3, 6, 4,-3,_M,-3,-2,-3,-3,-1, 0, 2,-2, 0,-1,-2},
/* M */    {-1,-2,-5,-3,-2, 0,-3,-2, 2, 0, 0, 4, 6,-2,_M,-2,-1, 0,-2,-1, 0, 2,-4, 0,-2,-1},
/* N */    { 0, 2,-4, 2, 1,-4, 0, 2,-2, 0, 1,-3,-2, 2,_M,-1, 1, 0, 1, 0, 0,-2,-4, 0,-2, 1},
/* O */    {_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M, 0,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M},
/* P */    { 1,-1,-3,-1,-1,-5,-1, 0,-2, 0,-1,-3,-2,-1,_M, 6, 0, 0, 1, 0, 0,-1,-6, 0,-5, 0},
/* Q */    { 0, 1,-5, 2, 2,-5,-1, 3,-2, 0, 1,-2,-1, 1,_M, 0, 4, 1,-1,-1, 0,-2,-5, 0,-4, 3},
/* R */    {-2, 0,-4,-1,-1,-4,-3, 2,-2, 0, 3,-3, 0, 0,_M, 0, 1, 6, 0,-1, 0,-2, 2, 0,-4, 0},
/* S */    { 1, 0, 0, 0, 0,-3, 1,-1,-1, 0, 0,-3,-2, 1,_M, 1,-1, 0, 2, 1, 0,-1,-2, 0,-3, 0},
/* T */    { 1, 0,-2, 0, 0,-3, 0,-1, 0, 0, 0,-1,-1, 0,_M, 0,-1,-1, 1, 3, 0, 0,-5, 0,-3, 0},
/* U */    { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */    { 0,-2,-2,-2,-2,-1,-1,-2, 4, 0,-2, 2, 2,-2,_M,-1,-2,-2,-1, 0, 0, 4,-6, 0,-2,-2},
/* W */    {-6,-5,-8,-7,-7, 0,-7,-3,-5, 0,-3,-2,-4,-4,_M,-6,-5, 2,-2,-5, 0,-6,17, 0, 0,-6},
/* X */    { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */    {-3,-3, 0,-4,-4, 7,-5, 0,-1, 0,-4,-1,-2,-2,_M,-5,-4,-4,-3,-3, 0,-2, 0, 0,10,-4},
/* Z */    { 0, 1,-5, 2, 3,-5, 0, 2,-2, 0, 0,-2,-1, 1,_M, 0, 3, 0, 0, 0, 0,-2,-6, 0,-4, 4}
};
```

```
/*
 */
#include <stdio.h>
#include <ctype.h>

#define MAXJMP  16      /* max jumps in a diag */
#define MAXGAP  24      /* don't continue to penalize gaps larger than this */
#define JMPS            1024    /* max jmps in an path */
#define MX              4       /* save if there's at least MX-1 bases since last jmp */

#define DMAT            3       /* value of matching bases */
#define DMIS            0       /* penalty for mismatched bases */
#define DINS0           8       /* penalty for a gap */
#define DINS1           1       /* penalty per base */
#define PINS0           8       /* penalty for a gap */
#define PINS1           4       /* penalty per residue */

struct jmp {
        short           n[MAXJMP];      /* size of jmp (neg for dely) */
        unsigned short  x[MAXJMP];      /* base no. of jmp in seq x */
};                                      /* limits seq to 2^16 -1 */

struct diag {
        int             score;          /* score at last jmp */
        long            offset;         /* offset of prev block */
        short           ijmp;           /* current jmp index */
        struct jmp      jp;             /* list of jmps */
};

struct path {
        int     spc;            /* number of leading spaces */
        short   n[JMPS];        /* size of jmp (gap) */
        int     x[JMPS];        /* loc of jmp (last elem before gap) */
};

char            *ofile;                 /* output file name */
char            *namex[2];              /* seq names: getseqs() */
char            *prog;                  /* prog name for err msgs */
char            *seqx[2];                       /* seqs: getseqs() */
int             dmax;                   /* best diag: nw() */
int             dmax0;                  /* final diag */
int             dna;                    /* set if dna: main() */
int             endgaps;                        /* set if penalizing end gaps */
int             gapx, gapy;             /* total gaps in seqs */
int             len0, len1;             /* seq lens */
int             ngapx, ngapy;           /* total size of gaps */
int             smax;                   /* max score: nw() */
int             *xbm;                   /* bitmap for matching */
long            offset;                 /* current offset in jmp file */
struct  diag    *dx;                    /* holds diagonals */
struct  path    pp[2];                  /* holds path for seqs */

char            *calloc(), *malloc(), *index(), *strcpy();
char            *getseq(), *g_calloc();
```

```
/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 *   where file1 and file2 are two dna or two protein sequences.
 *   The sequences can be in upper- or lower-case an may contain ambiguity
 *   Any lines beginning with ';', '>' or '<' are ignored
 *   Max file length is 65535 (limited by unsigned short x in the jmp struct)
 *   A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 *   Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static _dbval[26] = {
        1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static _pbval[26] = {
        1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
        128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
        1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
        1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)                                                    main
        int     ac;
        char    *av[];
{
        prog = av[0];
        if (ac != 3) {
                fprintf(stderr,"usage: %s file1 file2¥n", prog);
                fprintf(stderr,"where file1 and file2 are two dna or two protein sequences.¥n");
                fprintf(stderr,"The sequences can be in upper- or lower-case¥n");
                fprintf(stderr,"Any lines beginning with ';' or '<' are ignored¥n");
                fprintf(stderr,"Output is in the file ¥"align.out¥"¥n");
                exit(1);
        }
        namex[0] = av[1];
        namex[1] = av[2];
        seqx[0] = getseq(namex[0], &len0);
        seqx[1] = getseq(namex[1], &len1);
        xbm = (dna)? _dbval : _pbval;

        endgaps = 0;                    /* 1 to penalize endgaps */
        ofile = "align.out";            /* output file */

        nw();           /* fill in the matrix, get the possible jmps */
        readjmps();     /* get the actual jmps */
        print();                /* print stats, alignment */

        cleanup(0);     /* unlink any tmp files */
}
```

```
/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()                                                                    nw
{
        char            *px, *py;               /* seqs and ptrs */
        int             *ndely, *dely;  /* keep track of dely */
        int             ndelx, delx;    /* keep track of delx */
        int             *tmp;           /* for swapping row0, row1 */
        int             mis;            /* score for each type */
        int             ins0, ins1;     /* insertion penalties */
        register                id;             /* diagonal index */
        register                ij;             /* jmp index */
        register                *col0, *col1;   /* score for curr, last row */
        register                xx, yy;         /* index into seqs */

        dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

        ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
        dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
        col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
        col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
        ins0 = (dna)? DINS0 : PINS0;
        ins1 = (dna)? DINS1 : PINS1;

        smax = -10000;
        if (endgaps) {
                for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
                        col0[yy] = dely[yy] = col0[yy-1] - ins1;
                        ndely[yy] = yy;
                }
                col0[0] = 0;    /* Waterman Bull Math Biol 84 */
        }
        else
                for (yy = 1; yy <= len1; yy++)
                        dely[yy] = -ins0;

        /* fill in match matrix
         */
        for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
                /* initialize first entry in col
                 */
                if (endgaps) {
                        if (xx == 1)
                                col1[0] = delx = -(ins0+ins1);
                        else
                                col1[0] = delx = col0[0] - ins1;
                        ndelx = xx;
                }
                else {
                        col1[0] = 0;
                        delx = -ins0;
                        ndelx = 0;
```

...nw

```
for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
        mis = col0[yy-1];
        if (dna)
                mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
        else
                mis += _day[*px-'A'][*py-'A'];

        /* update penalty for del in x seq;
         * favor new del over ongong del
         * ignore MAXGAP if weighting endgaps
         */
        if (endgaps || ndely[yy] < MAXGAP) {
                if (col0[yy] - ins0 >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else {
                        dely[yy] -= ins1;
                        ndely[yy]++;
                }
        } else {
                if (col0[yy] - (ins0+ins1) >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else
                        ndely[yy]++;
        }

        /* update penalty for del in y seq;
         * favor new del over ongong del
         */
        if (endgaps || ndelx < MAXGAP) {
                if (col1[yy-1] - ins0 >= delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else {
                        delx -= ins1;
                        ndelx++;
                }
        } else {
                if (col1[yy-1] - (ins0+ins1) >= delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else
                        ndelx++;
        }

        /* pick the maximum score; we're favoring
         * mis over any del and delx over dely
         */
```

...nw

```
                id = xx - yy + len1 - 1;
                if (mis >= delx && mis >= dely[yy])
                        col1[yy] = mis;
                else if (delx >= dely[yy]) {
                        col1[yy] = delx;
                        ij = dx[id].ijmp;
                        if (dx[id].jp.n[0] && (!dna  || (ndelx >= MAXJMP
                        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
                                dx[id].ijmp++;
                                if (++ij >= MAXJMP) {
                                        writejmps(id);
                                        ij = dx[id].ijmp = 0;
                                        dx[id].offset = offset;
                                        offset += sizeof(struct jmp) + sizeof(offset);
                                }
                        }
                        dx[id].jp.n[ij] = ndelx;
                        dx[id].jp.x[ij] = xx;
                        dx[id].score = delx;
                }
                else {
                        col1[yy] = dely[yy];
                        ij = dx[id].ijmp;
if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
                        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
                                dx[id].ijmp++;
                                if (++ij >= MAXJMP) {
                                        writejmps(id);
                                        ij = dx[id].ijmp = 0;
                                        dx[id].offset = offset;
                                        offset += sizeof(struct jmp) + sizeof(offset);
                                }
                        }
                        dx[id].jp.n[ij] = -ndely[yy];
                        dx[id].jp.x[ij] = xx;
                        dx[id].score = dely[yy];
                }
                if (xx == len0 && yy < len1) {
                        /* last col
                         */
                        if (endgaps)
                                col1[yy] -= ins0+ins1*(len1-yy);
                        if (col1[yy] > smax) {
                                smax = col1[yy];
                                dmax = id;
                        }
                }
        }
        if (endgaps && xx < len0)
                col1[yy-1] -= ins0+ins1*(len0-xx);
        if (col1[yy-1] > smax) {
                smax = col1[yy-1];
                dmax = id;
        }
        tmp = col0; col0 = col1; col1 = tmp;
}

        (void) free((char *)ndely);
        (void) free((char *)dely);
        (void) free((char *)col0);
        (void) free((char *)col1);
}
```

```
/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() - -put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC     3
#define P_LINE  256     /* maximum output line */
#define P_SPC   3       /* space between name or num and seq */

extern  _day[26][26];
int     olen;           /* set output line length */
FILE    *fx;            /* output file */

print()                                                          print
{
        int     lx, ly, firstgap, lastgap;      /* overlap */

        if ((fx = fopen(ofile, "w")) == 0) {
                fprintf(stderr,"%s: can't write %s¥n", prog, ofile);
                cleanup(1);
        }
        fprintf(fx, "<first sequence: %s (length = %d)¥n", namex[0], len0);
        fprintf(fx, "<second sequence: %s (length = %d)¥n", namex[1], len1);
        olen = 60;
        lx = len0;
        ly = len1;
        firstgap = lastgap = 0;
        if (dmax < len1 - 1) {  /* leading gap in x */
                pp[0].spc = firstgap = len1 - dmax - 1;
                ly -= pp[0].spc;
        }
        else if (dmax > len1 - 1) {     /* leading gap in y */
                pp[1].spc = firstgap = dmax - (len1 - 1);
                lx -= pp[1].spc;
        }
        if (dmax0 < len0 - 1) { /* trailing gap in x */
                lastgap = len0 - dmax0 -1;
                lx -= lastgap;
        }
        else if (dmax0 > len0 - 1) {    /* trailing gap in y */
                lastgap = dmax0 - (len0 - 1);
                ly -= lastgap;
        }
        getmat(lx, ly, firstgap, lastgap);
        pr_align();
}
```

```
/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)                                    getmat
        int     lx, ly;                 /* "core" (minus endgaps) */
        int     firstgap, lastgap;              /* leading trailing overlap */
{
        int             nm, i0, i1, siz0, siz1;
        char            outx[32];
        double          pct;
        register                n0, n1;
        register char  *p0, *p1;

        /* get total matches, score
         */
        i0 = i1 = siz0 = siz1 = 0;
        p0 = seqx[0] + pp[1].spc;
        p1 = seqx[1] + pp[0].spc;
        n0 = pp[1].spc + 1;
        n1 = pp[0].spc + 1;

        nm = 0;
        while ( *p0 && *p1 ) {
                if (siz0) {
                        p1++;
                        n1++;
                        siz0--;
                }
                else if (siz1) {
                        p0++;
                        n0++;
                        siz1--;
                }
                else {
                        if (xbm[*p0-'A']&xbm[*p1-'A'])
                                nm++;
                        if (n0++ == pp[0].x[i0])
                                siz0 = pp[0].n[i0++];
                        if (n1++ == pp[1].x[i1])
                                siz1 = pp[1].n[i1++];
                        p0++;
                        p1++;
                }
        }

        /* pct homology:
         * if penalizing endgaps, base is the shorter seq
         * else, knock off overhangs and take shorter core
         */
        if (endgaps)
                lx = (len0 < len1)? len0 : len1;
        else
                lx = (lx < ly)? lx : ly;
        pct = 100.*(double)nm/(double)lx;
        fprintf(fx, "¥n");
        fprintf(fx, "<%d match%s in an overlap of %d: %.2f percent similarity¥n",



        nm, (nm == 1)? "" : "es", lx, pct);
```

```
        fprintf(fx, "<gaps in first sequence: %d", gapx);                    ...getmat
        if (gapx) {
                (void) sprintf(outx, " (%d %s%s)",
                        ngapx, (dna)? "base":"residue", (ngapx == 1)? "":"s");
                fprintf(fx,"%s", outx);

        fprintf(fx, ", gaps in second sequence: %d", gapy);
        if (gapy) {
                (void) sprintf(outx, " (%d %s%s)",
                        ngapy, (dna)? "base":"residue", (ngapy == 1)? "":"s");
                fprintf(fx,"%s", outx);
        }
        if (dna)
                fprintf(fx,
                "¥n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)¥n",
                smax, DMAT, DMIS, DINSO, DINS1);
        else
                fprintf(fx,
                "¥n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)¥n",
                smax, PINSO, PINS1);
        if (endgaps)
                fprintf(fx,
                "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s¥n",
                firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
                lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
        else
                fprintf(fx, "<endgaps not penalized¥n");
}

 static        nm;             /* matches in core -- for checking */
 static        lmax;           /* lengths of stripped file names */
 static        ij[2];          /* jmp index for a path */
 static        nc[2];          /* number at start of current line */
 static        ni[2];          /* current elem number -- for gapping */
 static        siz[2];
 static char   *ps[2];         /* ptr to current element */
 static char   *po[2];         /* ptr to next output char slot */
 static char   out[2][P_LINE]; /* output line */
 static char   star[P_LINE];   /* set by stars() */

/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()                                                              pr_align
{
        int             nn;     /* char count */
        int             more;
        register                i;

        for (i =  0, lmax = 0; i < 2; i++) {
                nn = stripname(namex[i]);
                if (nn > lmax)
                        lmax = nn;

                nc[i] = 1;


         ni[i] = 1;
         siz[i] = ij[i] = 0;
         ps[i] = seqx[i];
         po[i] = out[i];
}
```

...pr_align

```
	for (nn = nm = 0, more = 1; more; ) {
		for (i = more = 0; i < 2; i++) {
			/*
			 * do we have more of this sequence?
			 */
			if (!*ps[i])
				continue;

			more++;

			if (pp[i].spc) {	/* leading space */
				*po[i]++ = ' ';
				pp[i].spc--;
			}
			else if (siz[i]) {	/* in a gap */
				*po[i]++ = '-';
				siz[i]--;
			}
			else {		/* we're putting a seq element
					 */
				*po[i] = *ps[i];
				if (islower(*ps[i]))
					*ps[i] = toupper(*ps[i]);
				po[i]++;
				ps[i]++;

				/*
				 * are we at next gap for this seq?
				 */
				if (ni[i] == pp[i].x[ij[i]]) {
					/*
					 * we need to merge all gaps
					 * at this location
					 */
					siz[i] = pp[i].n[ij[i]++];
					while (ni[i] == pp[i].x[ij[i]])
						siz[i] += pp[i].n[ij[i]++];
				}
				ni[i]++;
			}
		}
		if (++nn == olen || !more && nn) {
			dumpblock();
			for (i = 0; i < 2; i++)
				po[i] = out[i];
			nn = 0;
		}
	}
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()
```

dumpblock

```
{

 register	i;

 for (i = 0; i < 2; i++)
	*po[i]-- = '¥0';
```

...dumpblock

```
        (void) putc('¥n', fx);
        for (i = 0; i < 2; i++) {
                if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
                        if (i == 0)
                                nums(i);
                        if (i == 0 && *out[1])
                                stars();
                        putline(i);
                        if (i == 0 && *out[1])
                                fprintf(fx, star);
                        if (i == 1)
                                nums(i);
                }
        }
}

/*
 * put out a number line: dumpblock()
 */
static
nums(ix)                                                        nums
        int     ix;     /* index in out[] holding seq line */
{
        char            nline[P_LINE];
        register                i, j;
        register char   *pn, *px, *py;

        for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
                *pn = ' ';
        for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
                if (*py == ' ' || *py == '-')
                        *pn = ' ';
                else {
                        if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                                j = (i < 0)? -i : i;
                                for (px = pn; j; j /= 10, px--)
                                        *px = j%10 + '0';
                                if (i < 0)
                                        *px = '-';
                        }
                        else
                                *pn = ' ';
                        i++;
                }
        }
        *pn = '¥0';
        nc[ix] = i;
        for (pn = nline; *pn; pn++)
                (void) putc(*pn, fx);
        (void) putc('¥n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */

static
putline(ix)                                                     putline
        int     ix;
{
```

...putline

```
        int             i;
        register char  *px;

        for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
                (void) putc(*px, fx);
        for (; i < lmax+P_SPC; i++)
                (void) putc(' ', fx);

        /* these count from 1:
         * ni[] is current element (from 1)
         * nc[] is number at start of current line
         */
        for (px = out[ix]; *px; px++)
                (void) putc(*px&0x7F, fx);
        (void) putc('¥n', fx);
}


/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()
{
        int             i;
        register char  *p0, *p1, cx, *px;

        if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
            !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
                return;
        px = star;
        for (i = lmax+P_SPC; i; i--)
                *px++ = ' ';

        for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
                if (isalpha(*p0) && isalpha(*p1)) {

                        if (xbm[*p0-'A']&xbm[*p1-'A']) {
                                cx = '*';
                                nm++;
                        }
                        else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                                cx = '.';
                        else
                                cx = ' ';
                }
                else
                        cx = ' ';
                *px++ = cx;
        }
        *px++ = '¥n';
        *px = '¥0';
}
```

stars

stripname

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)
        char    *pn;    /* file name (may be path) */
{
        register char  *px, *py;

        py = 0;
        for (px = pn; *px; px++)
                if (*px == '/')
                        py = px + 1;
        if (py)
                (void) strcpy(pn, py);
        return(strlen(pn));

}
```

```
/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 * readjmps() -- get the good jmps, from tmp file if necessary
 * writejmps() -- write a filled array of jmps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";             /* tmp file for jmps */
FILE    *fj;

int     cleanup();                              /* cleanup tmp file */
long    lseek();

/*
 * remove any tmp file if we blow
 */
cleanup(i)                                                          cleanup
        int     i;
{
        if (fj)
                (void) unlink(jname);
        exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)                                                   getseq
        char    *file;  /* file name */
        int     *len;   /* seq len */
{
        char            line[1024], *pseq;
        register char   *px, *py;
        int             natgc, tlen;
        FILE            *fp;

        if ((fp = fopen(file,"r")) == 0) {
                fprintf(stderr,"%s: can't read %s¥n", prog, file);
                exit(1);
        }
        tlen = natgc = 0;
        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '¥n'; px++)
                        if (isupper(*px) || islower(*px))
                                tlen++;
        }
        if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
                fprintf(stderr,"%s: malloc() failed to get %d bytes for %s¥n", prog, tlen+6, file);
                exit(1);

}
pseq[0] = pseq[1] = pseq[2] = pseq[3] = '¥0';
```

...getseq

```
        py = pseq + 4;
        *len = tlen;
        rewind(fp);

        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '¥n'; px++) {
                        if (isupper(*px))
                                *py++ = *px;
                        else if (islower(*px))
                                *py++ = toupper(*px);
                        if (index("ATGCU",*(py-1)))
                                natgc++;
                }
        }
        *py++ = '¥0';
        *py = '¥0';
        (void) fclose(fp);
        dna = natgc > (tlen/3);
        return(pseq+4);
}

char    *
g_calloc(msg, nx, sz)
        char    *msg;           /* program, calling routine */
        int     nx, sz;         /* number and size of elements */
{
        char            *px, *calloc();

        if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
                if (*msg) {
                        fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)¥n", prog, msg, nx,
sz);
                        exit(1);
                }
        }
        return(px);
}

/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()
{
        int             fd = -1;
        int             siz, i0, i1;
        register        i, j, xx;

        if (fj) {
                (void) fclose(fj);
                if ((fd = open(jname, O_RDONLY, 0)) < 0) {
                        fprintf(stderr, "%s: can't open() %s¥n", prog, jname);
                        cleanup(1);
                }
        }

   for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
           while (1) {
                   for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                           ;
```

g_calloc

readjmps

...readjmps

```
                if (j < 0 && dx[dmax].offset && fj) {
                        (void) lseek(fd, dx[dmax].offset, 0);
                        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
                        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
                        dx[dmax].ijmp = MAXJMP-1;
                }
                else
                        break;
        }
        if (i >= JMPS) {
                fprintf(stderr, "%s: too many gaps in alignment¥n", prog);
                cleanup(1);
        }
        if (j >= 0) {
                siz = dx[dmax].jp.n[j];
                xx = dx[dmax].jp.x[j];
                dmax += siz;
                if (siz < 0) {          /* gap in second seq */
                        pp[1].n[i1] = -siz;
                        xx += siz;
                        /* id = xx - yy + len1 - 1
                         */
                        pp[1].x[i1] = xx - dmax + len1 - 1;
                        gapy++;
                        ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
                        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
                        i1++;
                }
                else if (siz > 0) {     /* gap in first seq */
                        pp[0].n[i0] = siz;
                        pp[0].x[i0] = xx;
                        gapx++;
                        ngapx += siz;
/* ignore MAXGAP when doing endgaps */
                        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
                        i0++;
                }
        }
        else
                break;
    }

    /* reverse the order of jmps
     */
    for (j = 0, i0--; j < i0; j++, i0--) {
            i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
            i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
    }
    for (j = 0, i1--; j < i1; j++, i1--) {
            i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
            i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
    }
    if (fd >= 0)
            (void) close(fd);
    if (fj) {

            (void) unlink(jname);
            fj = 0;
            offset = 0;
    }
}
```

```
/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejmps(ix)                                                    writejmps
        int     ix;
{
        char    *mktemp();

        if (!fj) {
                if (mktemp(jname) < 0) {
                        fprintf(stderr, "%s: can't mktemp() %s¥n", prog, jname);
                        cleanup(1);
                }
                if ((fj = fopen(jname, "w")) == 0) {
                        fprintf(stderr, "%s: can't write %s¥n", prog, jname);
                        exit(1);
                }
        }
        (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
        (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}
```

【０３２２】

表１１

PRO	XXXXXXXXXXXXXXX	(Length = 15 アミノ酸)
比較タンパク質	XXXXXYYYYYYY	(Length = 12 アミノ酸)

% アミノ酸配列同一性 =

(ALIGN-2で決定したときの２つのポリペプチド配列間で同一に一致するアミノ酸残基の数)を(PROポリペプチドの全アミノ酸残基の数)で除する =

5を15で除する = 33.3%

【０３２３】

表１２

PRO	XXXXXXXXXX	(Length = 10 アミノ酸)
比較タンパク質	XXXXXYYYYYYZZYZ	(Length = 15 アミノ酸)

% アミノ酸配列同一性 =

(ALIGN-2で決定したときの２つのポリペプチド配列間で同一に一致するアミノ酸残基の数)を(PROポリペプチドの全アミノ酸残基の数)で除する=

5 を10 で除する= 50%

【０３２４】

表１３

PRO-DNA	NNNNNNNNNNNNN	(長さ = 14 ヌクレオチド)
比較DNA	NNNNNNLLLLLLLLL	(長さ = 16 ヌクレオチド)

% 核酸配列同一性 =

(ALIGN-2で決定したときの２つの核酸配列間で同一に一致するヌクレオチドの数)を(PRO-ＤＮＡ核酸配列の全ヌクレオチドの数)で除する=

6 を14で除する= 42.9%

【０３２５】

表１４

PRO-DNA	NNNNNNNNNNN	(長さ = 12 ヌクレオチド)
比較DNA	NNNNLLLVV	(長さ = 9 ヌクレオチド)

% 核酸配列一致性 =

(ALIGN-2で決定したときの２つの核酸配列間で同一に一致するヌクレオチドの数) を (PRO-ＤＮＡ核酸配列の全ヌクレオチドの数)で除する=

4 を12 で除する= 33.3%

【０３２６】

材料の寄託

次の材料をアメリカン・タイプ・カルチャー・コレクション，１２３０１ パークローンドライブ，ロックヴィル，ＭＤ，ＵＳＡ（ＡＴＣＣ）に寄託した：

材料	ATCC寄託番号	寄託日
DNA32292-1131	ATCC 209258	1997年9月16日
DNA33094-1131	ATCC 209256	1997年9月16日
DNA33223-1136	ATCC 209264	1997年9月16日
DNA34435-1140	ATCC 209250	1997年9月16日
DNA27864-1155	ATCC 209375	1997年10月16日
DNA36350-1158	ATCC 209378	1997年10月16日
DNA32290-1164	ATCC 209384	1997年10月16日
DNA35639-1172	ATCC 209396	1997年10月17日
DNA33092-1202	ATCC 209420	1997年10月28日
DNA49435-1219	ATCC 209480	1997年11月21日
DNA35638-1141	ATCC 209265	1997年9月16日
DNA32298-1132	ATCC 209257	1997年9月16日
DNA33089-1132	ATCC 209262	1997年9月16日
DNA33786-1132	ATCC 209253	1997年9月16日
DNA35918-1174	ATCC 209402	1997年10月17日
DNA37150-1178	ATCC 209401	1997年10月17日
DNA38260-1180	ATCC 209397	1997年10月17日
DNA39969-1185	ATCC 209400	1997年10月17日
DNA32286-1191	ATCC 209385	1997年10月16日
DNA33461-1199	ATCC 209367	1997年10月15日
DNA40628-1216	ATCC 209432	1997年11月7日
DNA33221-1133	ATCC 209263	1997年9月16日
DNA33107-1135	ATCC 209251	1997年9月16日
DNA35557-1137	ATCC 209255	1997年9月16日
DNA34434-1139	ATCC 209252	1997年9月16日
DNA33100-1159	ATCC 209373	1997年10月16日
DNA35600-1162	ATCC 209370	1997年10月16日
DNA34436-1238	ATCC 209523	1997年12月10日

DNA33206-1165	ATCC 209372	1997年10月16日
DNA35558-1167	ATCC 209374	1997年10月16日
DNA35599-1168	ATCC 209373	1997年10月16日
DNA36992-1168	ATCC 209382	1997年10月16日
DNA34407-1169	ATCC 209383	1997年10月16日
DNA35841-1173	ATCC 209403	1997年10月17日
DNA33470-1175	ATCC 209398	1997年10月17日
DNA34431-1177	ATCC 209399	1997年10月17日
DNA39510-1181	ATCC 209392	1997年10月17日
DNA39423-1182	ATCC 209387	1997年10月17日
DNA40620-1183	ATCC 209388	1997年10月17日
DNA40604-1187	ATCC209394	1997年10月17日
DNA38268-1188	ATCC 209421	1997年10月28日
DNA37151-1193	ATCC 209393	1997年10月17日
DNA35673-1201	ATCC 209418	1997年10月28日
DNA40370-1217	ATCC 209485	1997年11月21日
DNA42551-1217	ATCC 209483	1997年11月21日
DNA39520-1217	ATCC 209482	1997年11月21日
DNA41225-1217	ATCC 209491	1997年11月21日
DNA43318-1217	ATCC 209481	1997年11月21日
DNA40587-1231	ATCC 209438	1997年11月7日
DNA41338-1234	ATCC 209927	1998年6月2日
DNA40981-1234	ATCC 209439	1997年11月7日
DNA37140-1234	ATCC 209489	1997年11月21日
DNA40982-1235	ATCC 209433	1997年11月7日
DNA41379-1236	ATCC 209488	1997年11月21日
DNA44167-1243	ATCC 209434	1997年11月7日
DNA39427-1179	ATCC 209395	1997年10月17日
DNA40603-1232	ATCC 209486	1997年11月21日
DNA43466-1225	ATCC 209490	1997年11月21日
DNA43046-1225	ATCC 209484	1997年11月21日
DNA35668-1171	ATCC 209371	1997年10月16日

【０３２７】

これらの寄託は、特許手続き上の微生物の寄託の国際的承認に関するブダペスト条約及びその規則（ブダペスト条約）の規定に従って行われた。これは、寄託の日付から３０年間、寄託の生存可能な培養が維持されることを保証するものである。寄託物はブダペスト条約の条項に従い、またジェネンテク社とＡＴＣＣとの間の合意に従い、ＡＴＣＣから入手することができ、これは、どれが最初であろうとも、関連した米国特許の発行時又は任意の米国又は外国特許出願の公開時に、寄託培養物の後代を永久かつ非制限的に入手可能とすることを保証し、米国特許法第１２２条及びそれに従う特許庁長官規則（特に参照番号８８６ＯＧ６３８の３７ＣＦＲ第１．１４条を含む）に従って権利を有すると米国特許庁長官が決定した者に子孫を入手可能とすることを保証するものである。

本出願の譲受人は、寄託した培養物が、適切な条件下で培養されていた場合に死亡もしくは損失又は破壊されたならば、材料は通知時に同一の他のものと即座に取り替えることに同意する。寄託物質の入手可能性は、特許法に従いあらゆる政府の権限下で認められた権利に違反して、本発明を実施するライセンスであるとみなされるものではない。

上記の文書による明細書は、当業者に本発明を実施できるようにするために十分であると考えられる。寄託した態様は、本発明のある側面の一つの説明として意図されており、機能的に等価なあらゆる作成物がこの発明の範囲内にあるため、寄託された作成物により、本発明の範囲が限定されるものではない。ここでの物質の寄託は、ここに含まれる文書に

よる説明が、そのベストモードを含む、本発明の任意の側面の実施を可能にするために不十分であることを認めるものではないし、それが表す特定の例証に対して請求の範囲を制限するものと解釈されるものでもない。実際、ここに示し記載したものに加えて、本発明を様々に改変することは、前記の記載から当業者にとっては明らかなものであり、添付の請求の範囲内に入るものである。

【図面の簡単な説明】

【Ｆｉｇ１】天然配列ＰＲＯ２１１ｃＤＮＡのヌクレオチド配列（配列番号：１）を示す。配列番号：１は、ここで「ＤＮＡ３２２９２－１１３１」と命名されるクローンである。

【Ｆｉｇ２】Ｆｉｇ１に示した配列番号：１のコード化配列から誘導されたアミノ酸配列（配列番号：２）を示す。

【Ｆｉｇ３】天然配列ＰＲＯ２１７ｃＤＮＡのヌクレオチド配列（配列番号：３）を示す。配列番号：３は、ここで「ＤＮＡ３３０９４－１１３１と命名されるクローンである。

【Ｆｉｇ４】Ｆｉｇ３に示した配列番号：３のコード化配列から誘導されたアミノ酸配列（配列番号：４）を示す。

【Ｆｉｇ５】天然配列ＰＲＯ２３０ｃＤＮＡのヌクレオチド配列（配列番号：１１）を示す。配列番号：１１は、ここで「ＤＮＡ３３２２３－１１３６」と命名されるクローンである。

【Ｆｉｇ６】Ｆｉｇ５に示した配列番号：１１のコード化配列から誘導されたアミノ酸配列（配列番号：１２）を示す。

【Ｆｉｇ７】ここでＤＮＡ２００８８（配列番号：１３）と命名されるヌクレオチド配列を示す。

【Ｆｉｇ８】天然配列ＰＲＯ２３２ｃＤＮＡのヌクレオチド配列（配列番号：１７）を示す。配列番号：１７は、ここで「ＤＮＡ３４４３５－１１４０」と命名されるクローンである。

【Ｆｉｇ９】Ｆｉｇ８に示した配列番号：１７のコード化配列から誘導されたアミノ酸配列（配列番号：１８）を示す。

【Ｆｉｇ１０】天然配列ＰＲＯ１８７ｃＤＮＡのヌクレオチド配列（配列番号：２２）を示す。配列番号：２２は、ここで「ＤＮＡ２７８６４－１１５５」と命名されるクローンである。

【Ｆｉｇ１１】Ｆｉｇ１０に示した配列番号：２２のコード化配列から誘導されたアミノ酸配列（配列番号：２３）を示す。

【Ｆｉｇ１２】天然配列ＰＲＯ２６５ｃＤＮＡのヌクレオチド配列（配列番号：２７）を示す。配列番号：２７は、ここで「ＤＮＡ３６３５０－１１５８」と命名されるクローンである。

【Ｆｉｇ１３】Ｆｉｇ１２に示した配列番号：２７のコード化配列から誘導されたアミノ酸配列（配列番号：２８）を示す。

【Ｆｉｇ１４】天然配列ＰＲＯ２１９ｃＤＮＡのヌクレオチド配列（配列番号：３３）を示す。配列番号：３３は、ここで「ＤＮＡ３２２９０－１１６４」と命名されるクローンである。

【Ｆｉｇ１５】Ｆｉｇ１４に示した配列番号：３３のコード化配列から誘導されたアミノ酸配列（配列番号：３４）を示す。

【Ｆｉｇ１６】天然配列ＰＲＯ２４６ｃＤＮＡのヌクレオチド配列（配列番号：３８）を示す。配列番号：３８は、ここで「ＤＮＡ３５６３９－１１７２」と命名されるクローンである。

【Ｆｉｇ１７】Ｆｉｇ１６に示した配列番号：３８のコード化配列から誘導されたアミノ酸配列（配列番号：３９）を示す。

【Ｆｉｇ１８】天然配列ＰＲＯ２２８ｃＤＮＡのヌクレオチド配列（配列番号：４８）を示す。配列番号：４８は、ここで「ＤＮＡ３３０９２－１２０２」と命名されるクローンである。

【Ｆｉｇ１９】Ｆｉｇ１８に示した配列番号：４８のコード化配列から誘導されたアミノ酸配列（配列番号：４９）を示す。
【Ｆｉｇ２０】ここでＤＮＡ２１９５１（配列番号：５０）と命名されるヌクレオチド配列を示す。
【Ｆｉｇ２１】天然配列ＰＲＯ５３３ｃＤＮＡのヌクレオチド配列（配列番号：５８）を示す。配列番号：５８は、ここで「ＤＮＡ４９４３５－１２１９」と命名されるクローンである。
【Ｆｉｇ２２】Ｆｉｇ２１に示した配列番号：５８のコード化配列から誘導されたアミノ酸配列（配列番号：５９）を示す。
【Ｆｉｇ２３】天然配列ＰＲＯ２４５ｃＤＮＡのヌクレオチド配列（配列番号：６３）を示す。配列番号：６３は、ここで「ＤＮＡ３５６３８－１１４１」と命名されるクローンである。
【Ｆｉｇ２４】Ｆｉｇ２３に示した配列番号：６３のコード化配列から誘導されたアミノ酸配列（配列番号：６４）を示す。
【Ｆｉｇ２５】天然配列ＰＲＯ２２０ｃＤＮＡのヌクレオチド配列（配列番号：６８）を示す。配列番号：６８は、ここで「ＤＮＡ３２２９８－１１３２」と命名されるクローンである。
【Ｆｉｇ２６】Ｆｉｇ２５に示した配列番号：６８のコード化配列から誘導されたアミノ酸配列（配列番号：６９）を示す。
【Ｆｉｇ２７】天然配列ＰＲＯ２２１ｃＤＮＡのヌクレオチド配列（配列番号：７０）を示す。配列番号：７０は、ここで「ＤＮＡ３３０８９－１１３２」と命名されるクローンである。
【Ｆｉｇ２８】Ｆｉｇ２７に示した配列番号：７０のコード化配列から誘導されたアミノ酸配列（配列番号：７１）を示す。
【Ｆｉｇ２９】天然配列ＰＲＯ２２７ｃＤＮＡのヌクレオチド配列（配列番号：７２）を示す。配列番号：７２は、ここで「ＤＮＡ３３７８６－１１３２」と命名されるクローンである。
【Ｆｉｇ３０】Ｆｉｇ２９に示した配列番号：７２のコード化配列から誘導されたアミノ酸配列（配列番号：７３）を示す。
【Ｆｉｇ３１】天然配列ＰＲＯ２５８ｃＤＮＡのヌクレオチド配列（配列番号：８３）を示す。配列番号：８３は、ここで「ＤＮＡ３５９１８－１１７４」と命名されるクローンである。
【Ｆｉｇ３２】Ｆｉｇ３１に示した配列番号：８３のコード化配列から誘導されたアミノ酸配列（配列番号：８４）を示す。
【Ｆｉｇ３３】天然配列ＰＲＯ２６６ｃＤＮＡのヌクレオチド配列（配列番号：９０）を示す。配列番号：９０は、ここで「ＤＮＡ３７１５０－１１７８」と命名されるクローンである。
【Ｆｉｇ３４】Ｆｉｇ３３に示した配列番号：９０のコード化配列から誘導されたアミノ酸配列（配列番号：９１）を示す。
【Ｆｉｇ３５】天然配列ＰＲＯ２６９ｃＤＮＡのヌクレオチド配列（配列番号：９５）を示す。配列番号：９５は、ここで「ＤＮＡ３８２６０－１１８０」と命名されるクローンである。
【Ｆｉｇ３６】Ｆｉｇ３５に示した配列番号：９５のコード化配列から誘導されたアミノ酸配列（配列番号：９６）を示す。
【Ｆｉｇ３７】天然配列ＰＲＯ２８７ｃＤＮＡのヌクレオチド配列（配列番号：１０３を示す。配列番号：１０３は、ここで「ＤＮＡ３９９６９－１１８５」と命名されるクローンである。
【Ｆｉｇ３８】Ｆｉｇ３７に示した配列番号：１０３のコード化配列から誘導されたアミノ酸配列（配列番号：１０４）を示す。
【Ｆｉｇ３９】天然配列ＰＲＯ２１４ｃＤＮＡのヌクレオチド配列（配列番号：１０８を

示す。配列番号：１０８は、ここで「ＤＮＡ３２２８６－１１９１」と命名されるクローンである。
【Ｆｉｇ４０】Ｆｉｇ３９に示した配列番号：１０８のコード化配列から誘導されたアミノ酸配列（配列番号：１０９）を示す。
【Ｆｉｇ４１】天然配列ＰＲＯ３１７ｃＤＮＡのヌクレオチド配列（配列番号：１１３）を示す。配列番号：１１３は、ここで「ＤＮＡ３３４６１－１１９９」と命名されるクローンである。
【Ｆｉｇ４２】Ｆｉｇ４１に示した配列番号：１１３のコード化配列から誘導されたアミノ酸配列（配列番号：１１４）を示す。
【Ｆｉｇ４３】天然配列ＰＲＯ３０１ｃＤＮＡのヌクレオチド配列（配列番号：１１８）を示す。配列番号：１１８は、ここで「ＤＮＡ４０６２８－１２１６」と命名されるクローンである。
【Ｆｉｇ４４】Ｆｉｇ４３に示した配列番号：１１８のコード化配列から誘導されたアミノ酸配列（配列番号：１１９）を示す。
【Ｆｉｇ４５】天然配列ＰＲＯ２２４ｃＤＮＡのヌクレオチド配列（配列番号：１２６）を示す。配列番号：１２６は、ここで「ＤＮＡ３３２２１－１１３３」と命名されるクローンである。
【Ｆｉｇ４６】Ｆｉｇ４５に示した配列番号：１２６のコード化配列から誘導されたアミノ酸配列（配列番号：１２７）を示す。
【Ｆｉｇ４７】天然配列ＰＲＯ２２２ｃＤＮＡのヌクレオチド配列（配列番号：１３１）を示す。配列番号：１３１は、ここで「ＤＮＡ３３１０７－１１３５」と命名されるクローンである。
【Ｆｉｇ４８】Ｆｉｇ４７に示した配列番号：１３１のコード化配列から誘導されたアミノ酸配列（配列番号：１３２）を示す。
【Ｆｉｇ４９】天然配列ＰＲＯ２３４ｃＤＮＡのヌクレオチド配列（配列番号：１３６）を示す。配列番号：１３６は、ここで「ＤＮＡ３５５５７－１１３７」と命名されるクローンである。
【Ｆｉｇ５０】Ｆｉｇ４９に示した配列番号：１３６のコード化配列から誘導されたアミノ酸配列（配列番号：１３７）を示す。
【Ｆｉｇ５１】天然配列ＰＲＯ２３１ｃＤＮＡのヌクレオチド配列（配列番号：１４１）を示す。配列番号：１４１は、ここで「ＤＮＡ３４４３４－１１３９」と命名されるクローンである。
【Ｆｉｇ５２】Ｆｉｇ５１に示した配列番号：１４１のコード化配列から誘導されたアミノ酸配列（配列番号：１４２）を示す。
【Ｆｉｇ５３】天然配列ＰＲＯ２２９ｃＤＮＡのヌクレオチド配列（配列番号：１４７）を示す。配列番号：１４７は、ここで「ＤＮＡ３３１００－１１５９」と命名されるクローンである。
【Ｆｉｇ５４】Ｆｉｇ５３に示した配列番号：１４７のコード化配列から誘導されたアミノ酸配列（配列番号：１４８）を示す。
【Ｆｉｇ５５】天然配列ＰＲＯ２３８ｃＤＮＡのヌクレオチド配列（配列番号：１５２）を示す。配列番号：１５２は、ここで「ＤＮＡ３５６００－１１６２」と命名されるクローンである。
【Ｆｉｇ５６】Ｆｉｇ５５に示した配列番号：１５２のコード化配列から誘導されたアミノ酸配列（配列番号：１５３）を示す。
【Ｆｉｇ５７】天然配列ＰＲＯ２３３ｃＤＮＡのヌクレオチド配列（配列番号：１５８）を示す。配列番号：１５８は、ここで「ＤＮＡ３４４３６－１２３８」と命名されるクローンである。
【Ｆｉｇ５８】Ｆｉｇ５７に示した配列番号：１５８のコード化配列から誘導されたアミノ酸配列（配列番号：１５９）を示す。
【Ｆｉｇ５９】天然配列ＰＲＯ２２３ｃＤＮＡのヌクレオチド配列（配列番号：１６３）

を示す。配列番号：１６３は、ここで「ＤＮＡ３３２０６－１１６５」と命名されるクローンである。
【Ｆｉｇ６０】Ｆｉｇ５９に示した配列番号：１６３のコード化配列から誘導されたアミノ酸配列（配列番号：１６４）を示す。
【Ｆｉｇ６１】天然配列ＰＲＯ２３５ｃＤＮＡのヌクレオチド配列（配列番号：１６９）を示す。配列番号：１６９は、ここで「ＤＮＡ３５５５８－１１６７」と命名されるクローンである。
【Ｆｉｇ６２】Ｆｉｇ６１に示した配列番号：１６９のコード化配列から誘導されたアミノ酸配列（配列番号：１７０）を示す。
【Ｆｉｇ６３】天然配列ＰＲＯ２３６ｃＤＮＡのヌクレオチド配列（配列番号：１７４）を示す。配列番号：１７４は、ここで「ＤＮＡ３５５９９－１１６８」と命名されるクローンである。
【Ｆｉｇ６４】Ｆｉｇ６３に示した配列番号：１７４のコード化配列から誘導されたアミノ酸配列（配列番号：１７５）を示す。
【Ｆｉｇ６５】天然配列ＰＲＯ２６２ｃＤＮＡのヌクレオチド配列（配列番号：１７６）を示す。配列番号：１７６は、ここで「ＤＮＡ３６９９２－１１６８」と命名されるクローンである。
【Ｆｉｇ６６】Ｆｉｇ６５に示した配列番号：１７６のコード化配列から誘導されたアミノ酸配列（配列番号：１７７）を示す。
【Ｆｉｇ６７】天然配列ＰＲＯ２３９ｃＤＮＡのヌクレオチド配列（配列番号：１８４）を示す。配列番号：１８４は、ここで「ＤＮＡ３４４０７－１１６９」と命名されるクローンである。
【Ｆｉｇ６８】Ｆｉｇ６７に示した配列番号：１８４のコード化配列から誘導されたアミノ酸配列（配列番号：１８５）を示す。
【Ｆｉｇ６９】天然配列ＰＲＯ２５７ｃＤＮＡのヌクレオチド配列（配列番号：１８９）を示す。配列番号：１８９は、ここで「ＤＮＡ３５８４１－１１７３」と命名されるクローンである。
【Ｆｉｇ７０】Ｆｉｇ６９に示した配列番号：１８９のコード化配列から誘導されたアミノ酸配列（配列番号：１９０）を示す。
【Ｆｉｇ７１】天然配列ＰＲＯ２６０ｃＤＮＡのヌクレオチド配列（配列番号：１９４）を示す。配列番号：１９４は、ここで「ＤＮＡ３３４７０－１１７５」と命名されるクローンである。
【Ｆｉｇ７２】Ｆｉｇ７１に示した配列番号：１９４のコード化配列から誘導されたアミノ酸配列（配列番号：１９５）を示す。
【Ｆｉｇ７３】天然配列ＰＲＯ２６３ｃＤＮＡのヌクレオチド配列（配列番号：２００）を示す。配列番号：２００は、ここで「ＤＮＡ３４４３１－１１７７」と命名されるクローンである。
【Ｆｉｇ７４】Ｆｉｇ７３に示した配列番号：２００のコード化配列から誘導されたアミノ酸配列（配列番号：２０１）を示す。
【Ｆｉｇ７５】天然配列ＰＲＯ２７０ｃＤＮＡのヌクレオチド配列（配列番号：２０６）を示す。配列番号：２０６は、ここで「３９５１０－１１８１」と命名されるクローンである。
【Ｆｉｇ７６】Ｆｉｇ７５に示した配列番号：２０６のコード化配列から誘導されたアミノ酸配列（配列番号：２０７）を示す。
【Ｆｉｇ７７】天然配列ＰＲＯ２７１ｃＤＮＡのヌクレオチド配列（配列番号：２１２）を示す。配列番号：２１２は、ここで「ＤＮＡ３９４２３－１１８２」と命名されるクローンである。
【Ｆｉｇ７８】Ｆｉｇ７７に示した配列番号：２１２のコード化配列から誘導されたアミノ酸配列（配列番号：２１３）を示す。
【Ｆｉｇ７９】天然配列ＰＲＯ２７２ｃＤＮＡのヌクレオチド配列（配列番号：２２０）

を示す。配列番号：２２０は、ここで「ＤＮＡ４０６２０－１１８３」と命名されるクローンである。
【Ｆｉｇ８０】Ｆｉｇ７９に示した配列番号：２２０のコード化配列から誘導されたアミノ酸配列（配列番号：２２１）を示す。
【Ｆｉｇ８１】天然配列ＰＲＯ２９４ｃＤＮＡのヌクレオチド配列（配列番号：２２６）を示す。配列番号：２２６は、ここで「ＤＮＡ４０６０４－１１８７」と命名されるクローンである。
【Ｆｉｇ８２】Ｆｉｇ８１に示した配列番号：２２６のコード化配列から誘導されたアミノ酸配列（配列番号：２２７）を示す。
【Ｆｉｇ８３】天然配列ＰＲＯ２９５ｃＤＮＡのヌクレオチド配列（配列番号：２３５）を示す。配列番号：２３５は、ここで「ＤＮＡ３８２６８－１１８８」と命名されるクローンである。
【Ｆｉｇ８４】Ｆｉｇ８３に示した配列番号：２３５のコード化配列から誘導されたアミノ酸配列（配列番号：２３６）を示す。
【Ｆｉｇ８５】天然配列ＰＲＯ２９３ｃＤＮＡのヌクレオチド配列（配列番号：２４４）を示す。配列番号：２４４は、ここで「ＤＮＡ３７１５１－１１９３」と命名されるクローンである。
【Ｆｉｇ８６】Ｆｉｇ８５に示した配列番号：２４４のコード化配列から誘導されたアミノ酸配列（配列番号：２４５）を示す。
【Ｆｉｇ８７】天然配列ＰＲＯ２４７ｃＤＮＡのヌクレオチド配列（配列番号：２４９）を示す。配列番号：２４９は、ここで「ＤＮＡ３５６７３－１２０１」と命名されるクローンである。
【Ｆｉｇ８８】Ｆｉｇ８７に示した配列番号：２４９のコード化配列から誘導されたアミノ酸配列（配列番号：２５０）を示す。
【Ｆｉｇ８９】天然配列ＰＲＯ３０２ｃＤＮＡのヌクレオチド配列（配列番号：２５４）を示す。配列番号：２５４は、ここで「ＤＮＡ４０３７０－１２１７」と命名されるクローンである。
【Ｆｉｇ９０】Ｆｉｇ８９に示した配列番号：２５４のコード化配列から誘導されたアミノ酸配列（配列番号：２５５）を示す。
【Ｆｉｇ９１】天然配列ＰＲＯ３０３ｃＤＮＡのヌクレオチド配列（配列番号：２５６）を示す。配列番号：２５６は、ここで「ＤＮＡ４２５５１－１２１７」と命名されるクローンである。
【Ｆｉｇ９２】Ｆｉｇ９１に示した配列番号：２５６のコード化配列から誘導されたアミノ酸配列（配列番号：２５７）を示す。
【Ｆｉｇ９３】天然配列ＰＲＯ３０４ｃＤＮＡのヌクレオチド配列（配列番号：２５８）を示す。配列番号：２５８は、ここで「ＤＮＡ３９５２０－１２１７」と命名されるクローンである。
【Ｆｉｇ９４】Ｆｉｇ９３に示した配列番号：２５８のコード化配列から誘導されたアミノ酸配列（配列番号：２５９）を示す。
【Ｆｉｇ９５】天然配列ＰＲＯ３０７ｃＤＮＡのヌクレオチド配列（配列番号：２６０）を示す。配列番号：２６０は、ここで「ＤＮＡ４１２２５－１２１７」と命名されるクローンである。
【Ｆｉｇ９６】Ｆｉｇ９５に示した配列番号：２６０のコード化配列から誘導されたアミノ酸配列（配列番号：２６１）を示す。
【Ｆｉｇ９７】天然配列ＰＲＯ３４３ｃＤＮＡのヌクレオチド配列（配列番号：２６２）を示す。配列番号：２６２は、ここで「ＤＮＡ４３３１８－１２１７」と命名されるクローンである。
【Ｆｉｇ９８】Ｆｉｇ９７に示した配列番号：２６２のコード化配列から誘導されたアミノ酸配列（配列番号：２６３）を示す。
【Ｆｉｇ９９】天然配列ＰＲＯ３２８ｃＤＮＡのヌクレオチド配列（配列番号：２８４）

を示す。配列番号：２８４は、ここで「ＤＮＡ４０５８７－１２３１」と命名されるクローンである。
【Ｆｉｇ１００】Ｆｉｇ９９に示した配列番号：２８４のコード化配列から誘導されたアミノ酸配列（配列番号：２８５）を示す。
【Ｆｉｇ１０１】天然配列ＰＲＯ３３５ｃＤＮＡのヌクレオチド配列（配列番号：２８９）を示す。配列番号：２８９は、ここで「ＤＮＡ４１３８８－１２３４」と命名されるクローンである。
【Ｆｉｇ１０２】Ｆｉｇ１０１に示した配列番号：２８９のコード化配列から誘導されたアミノ酸配列（配列番号：２９０）を示す。
【Ｆｉｇ１０３】天然配列ＰＲＯ３３１ｃＤＮＡのヌクレオチド配列（配列番号：２９１）を示す。配列番号：２９１は、ここで「ＤＮＡ４０９８１－１２３４」と命名されるクローンである。
【Ｆｉｇ１０４】Ｆｉｇ１０３に示した配列番号：２９１のコード化配列から誘導されたアミノ酸配列（配列番号：２９２）を示す。
【Ｆｉｇ１０５】天然配列ＰＲＯ３２６ｃＤＮＡのヌクレオチド配列（配列番号：２９３）を示す。配列番号：２９３は、ここで「ＤＮＡ３７１４０－１２３４」と命名されるクローンである。
【Ｆｉｇ１０６】Ｆｉｇ１０５に示した配列番号：２９３のコード化配列から誘導されたアミノ酸配列（配列番号：２９４）を示す。
【Ｆｉｇ１０７】天然配列ＰＲＯ３３２ｃＤＮＡのヌクレオチド配列（配列番号：３０９）を示す。配列番号：３０９は、ここで「ＤＮＡ４０９８２－１２３５」と命名されるクローンである。
【Ｆｉｇ１０８】Ｆｉｇ１０７に示した配列番号：３０９のコード化配列から誘導されたアミノ酸配列（配列番号：３１０）を示す。
【Ｆｉｇ１０９】天然配列ＰＲＯ３３４ｃＤＮＡのヌクレオチド配列（配列番号：３１４）を示す。配列番号：３１４は、ここで「ＤＮＡ４１３７９－１２３６」と命名されるクローンである。開始及び停止コドンは太字及び下線フォントで示した。
【Ｆｉｇ１１０】Ｆｉｇ１０９に示した配列番号：３１４のコード化配列から誘導されたアミノ酸配列（配列番号：３１５）を示す。
【Ｆｉｇ１１１】天然配列ＰＲＯ３４６ｃＤＮＡのヌクレオチド配列（配列番号：３１９）を示す。配列番号：３１９は、ここで「ＤＮＡ４４１６７－１２４３」と命名されるクローンである。
【Ｆｉｇ１１２】Ｆｉｇ１１１に示した配列番号：３１９のコード化配列から誘導されたアミノ酸配列（配列番号：３２０）を示す。
【Ｆｉｇ１１３】天然配列ＰＲＯ２６８ｃＤＮＡのヌクレオチド配列（配列番号：３２４）を示す。配列番号：３２４は、ここで「ＤＮＡ３９４２７－１１７９」と命名されるクローンである。
【Ｆｉｇ１１４】Ｆｉｇ１１３に示した配列番号：３２４のコード化配列から誘導されたアミノ酸配列（配列番号：３２５）を示す。
【Ｆｉｇ１１５】天然配列ＰＲＯ３３０ｃＤＮＡのヌクレオチド配列（配列番号：３３１）を示す。配列番号：３３１は、ここで「ＤＮＡ４０６０３－１２３２」と命名されるクローンである。
【Ｆｉｇ１１６】Ｆｉｇ１１５に示した配列番号：３３１のコード化配列から誘導されたアミノ酸配列（配列番号：３３２）を示す。
【Ｆｉｇ１１７】天然配列ＰＲＯ３３９ｃＤＮＡのヌクレオチド配列（配列番号：３３８）を示す。配列番号：３３８は、ここで「ＤＮＡ４３４６６－１２２５」と命名されるクローンである。
【Ｆｉｇ１１８】Ｆｉｇ１１７に示した配列番号：３３８のコード化配列から誘導されたアミノ酸配列（配列番号：３３９）を示す。
【Ｆｉｇ１１９】天然配列ＰＲＯ３１０ｃＤＮＡのヌクレオチド配列（配列番号：３４０

）を示す。配列番号：３４０は、ここで「ＤＮＡ４３０４６－１２２５」と命名されるクローンである。

【Ｆｉｇ１２０】Ｆｉｇ１１９に示した配列番号：３４０のコード化配列から誘導されたアミノ酸配列（配列番号：３４１）を示す。

【Ｆｉｇ１２１】天然配列ＰＲＯ２４４ｃＤＮＡのヌクレオチド配列（配列番号：３７６）を示す。配列番号：３７６は、ここで「ＤＮＡ３５６６８－１１７１」と命名されるクローンである。

【Ｆｉｇ１２２】Ｆｉｇ１２１に示した配列番号：３７６のコード化配列から誘導されたアミノ酸配列（配列番号：３７７）を示す。

【図１】

ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA
CCCACGCGTCCGGGCCGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCGTCTTCCCG
CAGCGCTACCCGCC**ATG**CGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTG
CTGCTGCCGCCCGCGCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCT
GGTGGACAAGTTTAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACA
CGGCTTGGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTCGCCTGCTGGAGATC
CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA
GCACCTGGAGGCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGC
CAGGGCGGATCCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA
GGGCGACGGGTCCTGCCGGTGCCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGG
ACGGCTACTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC
TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT
GGACGAGGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTG
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTG
GGCTGCACAGGGGAAGGCCCAGGAAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA
CGGACAGTGTGCAGATGTGGACGAGTGCTCACTAGCAGAAAAAACCTGTGTGAGGAAAAACG
AAAACTGCTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCCTGACGGCTTCGAAGAAACG
GAAGATGCCTGTGTGCCGCCGGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT
GCCCTCCCGCGAAGACCTG**TAA**TGTGCCGGACTTACCCTTTAAATTATTCAGAAGGATGTCC
CGTGGAAAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGC
CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTA
AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG
AGGAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGC
TTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA
TCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTC
ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTCGGCGCAGCACCATGGCCTGAAAT
AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG
TGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAGCATGC
ATCTCAATTAGTCAGCAACCCAGTTTT

【図２】

><subunit 1 of 1, 353 aa, 0 stop
><MW: 38192, pI: 4.53, NX(S/T): 2
MRLPRRAALGLLPLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEEKTLSKYESSEIRL
LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLKVCCSPGTYGPDCLACQGGSQRPCSG
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESCKTCSGLTNRDCGECEVGWVLDE
GACVDVDECAAEPPPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYARERHGQCADVDECSLAEKT
CVRKNENCYNTPGSYVCVCPDGFEETEDACVPPAEAEATEGESPTQLPSREDL

シグナルペプチド：
アミノ酸 1-24

N-グリコシル化部位
アミノ酸 190-194 及び 251-255

グリコサミノグリカン付着部位
アミノ酸 149-153 及び 155-159

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位
アミノ酸 26-30

カゼインキナーゼIIリン酸化部位
アミノ酸 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343 及び 349-353

チロシンキナーゼリン酸化部位
アミノ酸 303-310

N-ミリストイル化部位
アミノ酸 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258及び 313-319

アスパラギン酸及びアスパラギンヒドロキシル化部位
アミノ酸 308-320

EGF-様ドメインシステインパターンシグネチャー
アミノ酸 166-178

ロイシンジッパーパターン
アミノ酸 94-116

【図３】

CAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTC
GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCGCCCAGCCGTCTAAACGGGAACA
GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCG
GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGCATGGCCCGGAGGAGCGCCTTC
CCTGCCGCCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGC
CGGGCCGCCGCAGGAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA
TAGGATTTGAAGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGAT
TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCCTGTCAATATCCATTCCATGAATTTTAC
CTGGCAAGCTGCAGGGCAGGCAGAATACTTCTATGAATTCCTGTCCTTGCGCTCCCTGGATA
AAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCA
TCAGTTGTTCAAGTTGGTTTCCCATGTCTTGGAAAACAGGATGGGGTGGCAGCATTTGAAGT
GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCT
TCTTTAAAACATGTCAACAAGCTGAGTGCCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAAT
GAAAGACGCATCTGCGAGTGTCCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTG
TACCCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCCACCTG
GATTCTATGGAGTGAACTGTGACAAAGCAAACTGCTCAACCACCTGCTTTAATGGAGGGACC
TGTTTCTACCCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG
CAAATGCCCACAACCCTGTCGAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTT
CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACAT
GGAACCTGCCATGAACCCAACAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAA
TAAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGCAGGCGCCCAGCTCAGGCAGC
ACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG
TGAACTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTCA
TGTGTTGAATGTTCAAATAATGTTCATTACACTTAAGAATACTGGCCTGAATTTTATTAGCT
TCATTATAAATCACTGAGCTGATATTTACTCTTCCTTTTAAGTTTCTAAGTACGTCTGTAG
CATGATGGTATAGATTTTCTTGTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGA
TCAGGTTAAAATTTTCAGTGTGTAGTTGGCAGATATTTTCAAAATTACAATGCATTTATGGT
GTCTGGGGGCAGGGGAACATCAGAAAGGTTAAATTGGGCAAAAATGCGTAAGTCACAAGAAT
TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGAT
GTTTGTTACATTTTTAAAAATTGCTCTTAATTTTTAAACTCTCAATACAATATATTTTGACC
TTACCATTATTCCAGAGATTCAGTATTAAAAAAAAAAAAAATTACACTGTGGTAGTGGCATTT
AAACAATATAATATATTCTAAACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCAT
TGGCTTGAAGCAATATAATATATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTAT
ACTGTTTGTATGTATAAAATAAAGGTGCTGCTTTAGTTTTTTGGAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGC
CGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

【図４】

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094
><subunit 1 of 1, 379 aa, 0 stop
><MW: 41528, pI: 7.97, NX(S/T): 2
MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFLSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPGGC
RNGGFCNERRICECPDGFHGPHCEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEERRDP
PESNYIW

シグナルペプチド:

アミノ酸 1-28

N-グリコシル化部位

アミノ酸 88-92, 245-249

カゼインキナーゼIIリン酸化部位

アミノ酸 319-323

チロシンキナーゼリン酸化部位

アミノ酸 370-378

N-ミリストイル化部位

アミノ酸 184-190, 185-191, 189-195, 315-321

ATP/GTP-結合部位モチーフ(P-ループ)

アミノ酸 285-293

EGF-様ドメインシステインパターンシグネチャー

アミノ酸 198-210, 230-242, 262-274, 294-306, 326-338

【図５】

CGGACGCGTGGGCGTCCGGCGGTCGCAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGG
CCCCAGCCCACACCTTCACCAGGGCCCAGGAGCCACCATGTGGCGATGTCCACTGGGGCTAC
TGCTGTTGCTGCCGCTGGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTCGTGGGCGCCGG
GAGCTAGCACCGGGTCTGCACCTGCGGGGCATCCGGGACGCGGGAGGCCGGTACTGCCAGGA
GCAGGACCTGTGCTGCCGCGGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT
GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTC
TGCCTCGGCGTGCCACCCCCTTTCCCCCGATCCAAGGATGTATGCATGGAGGTCGTATCTA
TCCAGTCTTGGGAACGTACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGT
GGCATGGTGGATCCAGACATGATCAAAGCCATCAACCAGGGCAACTATGGCTGGCAGGCTGG
GAACCACAGCGCCTTCTGGGGCATGACCCTGGATGAGGGCATTCGCTACCGCCTGGGCACCA
TCCGCCCATCTTCCTCGGTCATGAACATGCATGAAATTTATACAGTGCTGAACCCAGGGGAG
GTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCATGAGCCTCTTGA
CCAAGGCAACTGTGCAGGCTCCTGGGCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT
CAATCCATTCTCTGGGACACATGACGCCTGTCCTGTGCGCCCCAGAACCTGCTGTCTTGTGAC
ACCCACCAGCAGCAGGGCTGCCGCGGTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCG
CCGAGGGGTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCC
CTGCGCCCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCACTGCC
CACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCT
CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA
TGGAGGTGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGC
CTTGGGAGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG
AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCCTGGGGCC
CAGCCTGGGGCGAGAGGGCCACTTCCGCATCGTGCGCGGCGTCAATGAGTGCGACATCGAG
AGCTTCGTGCTGGGCGTCTGGGGCCGCGTGGGCATGGAGGACATGGGTCATCACTGAGGCTG
CGGGCACCACGCGGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCCAATG
GGGCGGTGACCCCAGCCTCGCCCGACAGAGCCCGGGGCGCAGGCGGGCGCCAGGGCGCTAAT
CCCGGCGCGGGTTCCGCTGACGCAGCGCCCCGCCTGGGAGCCGCGGGCAGGCGAGACTGGCG
GAGCCCCCAGACCTCCCAGTGGGGACGGGGCAGGGCCTGGCCTGGGAAGAGCACAGCTGCAG
ATCCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGC
CCCAATACCCCACCCCAATCCCGTATTCTTTTTTTTTTTTTTAGACAGGGTCTTGCTCCG
TTGCCCAGGTTGGAGTGCAGTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCA
AGTGACCCTCCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCACCACACCTGGC
TAATTTTTGTATTTTTTGTAAAGAGGGGGGTCTCACTGTGTTGCCCAGGCTGGTTTCGAACT
CCTGGGCTCAAGCGGTCCACCTGCCTCCGCCTCCCAAAGTGCTGGGATTGCAGGCATGAGCC
ACTGCACCCAGCCCTGTATTCTTATTCTTCAGATATTTATTTTTCTTTTCACTGTTTTAAAA
TAAAACCAAAGTATTGATAAAAAAAAA

【図６】

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223
><subunit 1 of 1, 164 aa, 1 stop
><MW: 18359, pI: 7.45, NX(S/T): 1
MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQDLCCRGRADDC
ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPPPPPIQGCMHGGRIYPVLGTYWDNCNR
CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

N-グリコシル化部位

アミノ酸 78-82, 161-165

カゼインキナーゼIIリン酸化部位

アミノ酸 80-84, 117-121, 126-130, 169-173, 205-209, 296-300, 411-415

N-ミリストイル化部位

アミノ酸 21-27, 39-45, 44-50, 104-110, 160-164, 224-230, 269-275, 378-384, 442-448

アミド化部位

アミノ酸 26-30, 318-322

真核生物チオール(システイン)プロテアーゼヒスチジン活性部位

アミノ酸 398-409

【図７】

AGGCTCCTTGGCCCTTTTCCACAGCAAGCTTNTGCNATCCCGATTCGTTGTCTCAAATCCA
ATTCTCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCGCCGAGG
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC
CCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCACTGCCCACTGC
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG
TGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG
AGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAG

【図８】

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG
GAGCAGTGCTGGACCGCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG
CAGCTTGAACTGCGTGGATGACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT
GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT
GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCCGGCCAGCTATAGGCTCTGGGGGG
CCCCGCTGCAGCCCACACTGGGTGTGGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG
GCCCAGTGGGAGCCTGTCCTGGTTCCTGAGGCACATCCTAACGCAAGTCTGACCATGTATGT
CTGCACCCCTGTCCCCCACCCTGACCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA
TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCTCCAACCCTCTCTGCTGCTGTTTC
CATGGCCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT
CCCTGCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA
GGGGACAGGCACTCAGGAGGGCCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA
GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA
GGGGCCAGGCCTCACATTCGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT
AATAAACACCTGTTGGATAAGCCAAAAAAA

【図９】

MTHRTTTWARRTSRAVTPTCATPAGPMPCSRLPPSLRCSLHSACCSGDPASYRLWGAPLQPT
LGVVPQASVPLLTDLAQWEPVLVPEAHPNASLTMYVCTPVPHPDPPMALSRTPTRQISSSDT
DPPADGPSNPLCCCFHGPAFSTLNPVLRHLFPQEAFPAHPIYDLSQVWSVVSPAPSRGQALR
RAQ

シグナルペプチド：

アミノ酸 1-47

N-グリコシル化部位

アミノ酸 31-35, 74-78, 84-88

カゼインキナーゼIIリン酸化部位

アミノ酸 22-26, 76-80

N-ミリストイル化部位

アミノ酸 56-60

アミド化部位

アミノ酸 70-74

【図１０】

CCCACGCGTCCGAACCTCTCCAGCGATGGGAGCCGCCCGCCTGCTGCCCAACCTCACTCTGT
GCTTACAGCTGCTGATTCTCTGCTGTCAAACTCAGTACGTGAGGGACCAGGGCGCCATGACC
GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA
CGTGCAGGTCACCGGGCGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC
TCATAGTGGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGGCTGAGAGTGAGAAG
TACATCTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTG
CGTGTTCACGGAGATCGTGCTGGAGAACAACTATACGGCCTTCCAGAACGCCCGGCACGAGG
GCTGGTTCATGGCCTTCACGCGGCAGGGGCGGCCCCGCCAGGCTTCCCGCAGCCGCCAGAAC
CAGCGCGAGGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCACGC
CGAGAAGCAGAAGCAGTTCGAGTTTGTGGGCTCCGCCCCCACCCGCCGGACCAAGCGCACAC
GGCGGCCCCAGCCCCTCACGTAGTCTGGGAGGCAGGGGGCAGCAGCCCCTGGGCCGCCTCCC
CACCCCTTTCCCTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCC
GAGGGAGGACCCTGAGGGCCGCGAAGCATCCGAGCCCCCAGCTGGGAAGGGGCAGGCCGGTG
CCCCAGGGGCGGCTGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACT
GAGTGTCACCCTGATCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCC
CGCTGAAAGGTCAGCGACTGAAGGCCTTGCAGACAACCGTCTGGAGGTGGCTGTCCTCAAAA
TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCCAGCCCCCAAACTCCTCCTGGCTAGACTGTA
GGAAGGGACTTTTGTTTGTTTGTTTGTTTCAGGAAAAAAGAAAGGGAGAGAGGAAAATAG
AGGGTTGTCCACTCCTCACATTCCACGACCCAGGCCTGCACCCCACCCCCAACTCCCAGCCC
CGGAATAAAACCATTTTCCTGC

【図１１】

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRI
SATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLIGKPSGKSKDCVFTEIVLE
NNYTAFQNARHEGWFMAFTRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQKQFEF
VGSAPTRRTKRTRRPQPLT

シグナルペプチド：

アミノ酸 1-22

N-グリコシル化部位

アミノ酸 9-13, 126-130

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 60-64

カゼインキナーゼIIリン酸化部位

アミノ酸 65-69

チロシンキナーゼリン酸化部位

アミノ酸 39-48, 89-97

N-ミリストイル化部位

アミノ酸 69-75, 188-194

アミド化部位

アミノ酸 58-62

HBGF/FGFファミリーシグネイチャー

アミノ酸 103-128

【図１２】

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG
GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT
CATTTTGATTTTGCTGTTTATTTTTTTTTCTTTTTCTTTTTCCCACCACATTGTATTTTAT
TTCCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCCT
GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC
CTAGTGTGTGCCGCTGCGACAGGAACTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC
TGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC
AACTGGACGAATTCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC
AATATTCAGACCATTTCACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCACCT
GGATGACAACTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCC
TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGAC
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCATATCCGACATGGCCTTCCAGAA
TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG
AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC
CACCCTCCTCCCGATCTCCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT
AAACCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA
ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC
ACTGCTCGGAATAACCCTTGGTTTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAA
ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGG
GGATGGCCGTCAGGGAATTAAATATGAATCTTTTGTCCTGTCCCACCACGACCCCCGGCCTG
CCTCTCTTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT
TCCAAACCCTAGCAGAAGCTACACGCCTCCAACTCCTACCACATCGAAACTTCCCACGATTC
CTGACTGGGATGGCAGAGAAAGAGTGACCCCACCTATTTCTGAACGGATCCAGCTCTCTATC
CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTCACCGTGATGGCATA
CAAACTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTCAGGAGCGCATAG
TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT
TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGC
CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA
CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT
GTGCTGGTGGTCTTGCTCAGCGTCTTTTGCTGGCATATGCACAAAAAGGGGCGCTACACCTC
CCAGAAGTGGAAATACAACCGGGGCCGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA
AGGACAACTCCATCCTGGAGATGACAGAAACCAGTTTTCAGATCGTCTCCTTAAATAACGAT
CAACTCCTTAAAGGAGATTTCAGACTGCAGCCCATTTACACCCCAAATGGGGGCATTAATTA
CACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC
ACTGCCATACGTGACAGCCAGAGGCCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA
CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT
TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAAGTG
CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTAAATCTT

【図１３】

MGLQTTKWPSHGAPFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP
EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI
SRAALAQLLKLEELHLDDNSISTVGVEDGAFREAISLKLLFLSKNHLSSVPVGLPVDLQELR
VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD
LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN
PWFCDCSIKWVTEWLKYIPSSLNVRGFMCQGPEQVRGMAVRELNMNLLSCPTTTPGLPLFTP
APSTASPTTQPPTTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVND
TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL
DAFNYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVL
LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG
DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

シグナルペプチド：

アミノ酸 1-42

膜貫通ドメイン：

アミノ酸 542-561

N-グリコシル化部位

アミノ酸 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

カゼインキナーゼIIリン酸化部位

アミノ酸 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

チロシンキナーゼリン酸化部位

アミノ酸 319-328

N-ミリストイル化部位

アミノ酸 2-8, 60-66, 149-155, 213-219, 220-226, 294-300, 522-528, 545-551, 633-639

アミド化部位

アミノ酸 581-585

ロイシンジッパーパターン

アミノ酸 164-186

ホスホリパーゼA2アスパラギン酸活性部位

アミノ酸 39-50

【図１４】

ACTTGGAGCAAGCGGCGGCGGCGGAGACAGAGGCAGAGGCAGAAGCTGGGGCTCCGTCCTCGCCTCCCACGAGCG
ATCCCCGAGGAGAGCCGCGGCCCTCGGCGAGGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCGCCCCTGCC
TCGCTTCCCAGGGCCGCGGCGGCTGCAGCCTTGCCCCTCTTGCTCGCCTTGAAAATGGAAAAGATGCTCGCAGGCT
GCTTTCTGCTGATCCTCGGACAGATCGTCCTCCTCCCTGCCGAGGCCAGGGAGCGGTCACGTGGGAGGTCCATCT
CTAGGGGCAGACACGCTCGGACCCACCCGCAGACGGCCCTTCTGCACAGTTCCTGTGAGAACAAGCGGCCAGACC
TGGTTTTCATCATTGACAGCTCTCGCAGTGTCAACACCCATGACTATGCAAAGGTCAAGGAGTTCATCGTGGACA
TCTTGCAATTCTTGGACATTGGTCCTGATGTCACCCGAGTGGGCCTGCTCCAATATGGCAGCACTGTCAAGAATG
AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGCTGTCAAGAGGATGCGGCATCTGTCCACGG
GCACCATGACTGGGCTGGCCATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGACGGGGCCCGCCCCCTGA
GGGAGAATGTGCCACGGGTCATAATGATCGTGACAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGGCTGCTA
AGGCACGGGACACGGGCATCCTAATCTTTGCCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATTG
GGAGTGAGCCCCATGAGGACCATGTCTTCCTTGTGGCCAATTTCAGCCAGATTGAGACGCTGACCTCCGTGTTCC
AGAAGAAGTTGTGCACGGCCCACATGTGCAGCACCCTGGAGCATAACTGTGCCCACTTCTGCATCAACATCCCTG
GCTCATACGTCTGCAGGTGCAAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTGT
CTGCCATGGAGGACCACAACTGTCAGCAGCTCTGTGTGAATGTGCCGGCTCGTCGTCGTCAGTGCTACAGTG
GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCCTCAGAAAACCACGGATGTGAAC
ATGAGTGTGTAAATGCTGATGGCTCCTACCTTTGCCAGTGCCATGAAGGATTTGCTCTTAACCCAGATGAAAAAA
CGTGCACAAGGATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGCGTCAACATGGAGGAGAGCT
ACTACTGCCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCAGCCGAGTGGACCACTGTGCAC
AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCCTTCGTCTGCCAGTGCTCAGAAGGCTTCC
TCATCAACGAGGACCTCAAGACCTGCTCCCGGGTGGATTACTGCCTGCTGAGTGACCATGGTTGTGAATACTCCT
GTGTCAACATGGACAGATCCTTTGCCTGTCAGTGTCCTGAGGGACACGTGCTCCGCAGCGATGGGAAGACGTGTG
CAAAATTGGACTCTTGTCCTCTGGGGGACCACGGTTGTGAACATTCGTGTGTAAGCAGTGAAGATTCGTTTGTGT
GCCAGTGCTTTGAAGGTTATATACTCCGTGAAGATGSAAAAACCTGCAGAAGGAAAGATGTCTGCCAAGCTATAG
ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACACGTGCGAGTGCTTGGAGGGATTCCGGCTCG
CTGAGGATGGGAAACGCTGCCGAAGGAAGGATGTCTGCAAATCAACCCACCACTGGCTGCGAACACATTFGTGTTA
ATAATGGGAATTCCTACATCTGCAAATGTCAGAGGGATTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAAT
GCACTGAAGCCCCAATTGACCTGCTCTTTGTGATCGATGGATCCAAGAGTCTTGGAGAAGAGAATTTTGAGGTCG
TGAAGCAGTTTGTCACTGGAATTATAGATTCCTTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCTCCAGT
ATTCCACACAGGTCCACACAGAGTTCACTCTGAGAAACTTCAACTCAGCCAAAGACATGAAAAAAGCCGTGGCCC
ACATGAAATACATGGGAAAGGGCTCTATGACTGGGCTGCCCCTGAAACACATGTTTGAGAGAAGTTTTACCCAAG
GAGAAGGGGCCAGGCCCCTTCCACAAGGGTGCCCAGAGCAGCCATTGTGTTCACCGACGGACGGGCTCAGGATG
ACGTCTCCGAGTGGGCCAGTAAAGCCAAGGCCAATGGTATCACTATGTATGCTGTTGGGGTAGGAAAAGCCATTG
AGGAGGAACTACAAGAGATTGCCTCTGAGCCCACAAACAAGCATCTCTTCTATGCCGAAGACTTCAGCACAATGG
ATGAGATAAGTGAAAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAG
CAGGGGAACTGCCAAAACGTCCACACAGCCAACAGAATCTGAGCCACTCACATAAATATCCAGACTCTACTTT
CCTGTTCTAATTTGCAGTGCAACACAGATATCTGTTGAAGAAGACAATCTTTACGGTCTACACAAAAGCTTT
CCCATTCAACAAAACCTTCAGGAAGCCCTTTGGAAGAAAACACGATCAATGCAAATGTGAAAACCTTATAATGT
TCCAGAACCTTGCAAACGAAGAAGTAAGAAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC
TGGAAAATCGCCTGAGATACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCATTGTATCACGGATTACAAT
GAACCGAGTGCAGAGCCCCAAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAAACAATCAGTACTGA
GAAACCTGGTTTGCCACAGAACAAAGACAAGAAGTATACACTAACTTGTATAAATTTATCTAGGAAAAAAATCCT
TCAGAATTCTAAGATGAATTTACCAGGTGAGAATGAATAAGCTATGCAAGGTATTTGTAATATACTGTGGACAC
AACTTGCTTCTGCCTCATCCTGCCTTAGTGTGCAATCTCATTTGACTATACGATAAAGTTTGCACAGTCTTACTT
CTGTAGAACACTGGCCATAGGAAATGCTGTTTTTTTGTACTGGACTTTACCTTGATATATGTATATGGATGTATG
CATAAAATCATAGGACATATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAAAATTCACCACTTCAG

【図１５】

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPQTALLESSCENKRADLVFII
DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV
KRMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD
TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHN
CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA
LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKPGC
EHECVNMEESYYCRCHRGYTLDPNGKTCSRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLI
NEDLKTCSRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTCAKLDSCALGDHGCE
HSCVSSEDSFVCQCPEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLA
EDGKRCRRKDVCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVID
GSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVA
HMKYMGKGSMTGLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN
GITMYAVGVGKAIPEELQEIASEPTNKHLFYAEDFSTMDEISEKLKKGICEALEDSDGRQDS
PAGELPKTVQQPTESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPL
EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMEALENRLRYR

シグナルペプチド：

アミノ酸 1-23

N-グリコシル化部位

アミノ酸 221-225

cAMP- 及び cGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 115-119, 606-610, 892-896

カゼインキナーゼIIリン酸化部位

アミノ酸 49-53, 118-122, 149-153, 176-180, 223-227, 243-247, 401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784, 781-785, 819-823, 866-870

N-ミリストイル化部位

アミノ酸 133-139, 258-264, 299-305, 340-346, 453-459, 494-500, 639-645, 690-696, 752-758, 792-798

アミド化部位

アミノ酸 314-318, 560-564, 601-605

アスパラギン酸及びアスパラギンヒドロキシル化部位

アミノ酸 253-265, 294-306, 335-347, 376-388, 417-423, 458-464, 540-546, 581-587

【図１６】

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCG
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTTCCCTCCCGGGGCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGCACGGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT
TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG
GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCCAC
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT
TTCCTCTGTCACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT
TGACCCCCACGCCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCCACGACAGAT
GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCG
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCCAGAGTCAAGCTGGCTCTCTGGTATGACTGAC
CCCACCACTCATTGGCTAAAGGATTTGGGGTCTCTCCTTCCTATAAGGGTCACCTCTAGCAC
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCCACCTCTC
TTTACTGTCGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTCACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCCAGGC
CCCCTTGATCTGTACCCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC
TTATTAAAACTAACATGAAATATGTGTTGTTTTCATTTGCAAATTTAAATAAAGATACATAA
TGTTTGTATGAAAAA

【図１７】

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS
SQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY
SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSCQSPRSKPAVQYQ
WDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA
VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLPWPKSSDTISKNGTL
SSVTSARALRPPHGPPRPGALTFTPSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSR
MGAVPVMVPAQSQAGSLV

シグナルペプチド：

アミノ酸 1-29

膜貫通ドメイン：

アミノ酸 245-267

N-グリコシル化部位

アミノ酸 108-112, 169-173, 213-217, 236-240, 307-311

N-ミリストイル化部位

アミノ酸 90-96, 167-173, 220-226, 231-237, 252-258, 256-262, 262-268, 308-314, 363-369, 364-370

原核生物膜リポタンパク質脂質付着部位

アミノ酸 164-175

【図１８】

CGCCACCACTGCGGCCACCGCCA**ATG**AAACGCCTCCCGCTCCTAGTGGTTTTTCCACTTTG
TTGAATTGTTCCTATACTCAAAATTGCACCAAGACACCTTGTCTCCCAAATGCAAAATGTGA
AATACGCAATGGAATTGAAGCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTGTCACAA
TTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGTCCTGTGGCGAAAATGCTAATTGC
ACTAACACAGAAGGAAGTTATTATTGTATGTGTGTACCTGGCTTCAGATCCAGCAGTAACCA
AGACAGGTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCATT
TAGATAATGTCTGTATAGCTGCAAATATTAATAAAACTTTAACAAAAATCAGATCCATAAAA
GAACCTGTGGCTTTGCTACAAGAAGTCTATAGAAATTCTGTGACAGATCTTTCACCAACAGA
TATAATTACATATATAGAAATATTAGCTGAATCATCTTCATTACTAGGTTACAAGAACAACA
CTATCTCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTGAATTTGTAAAAACCGTGAAT
AATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAGAAC
ACATCTTACAAAACTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCCAGAGCTTCC
AAAAGACCACAGAGTTTGATACAAATTCAACGGATATAGCTCTCAAAGTTTCTTTTTTGAT
TCATATAACATGAAACATATTCATCCTCATATGAATATGGATGGAGACTACATAAATATATT
TCCAAAGACAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTATATTATA
AGAGTATTGGTCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAAAATTATGAT
AATTCTGAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATGAGCTCAAACCC
ACCCACATTATATGAA**CTT**GAAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATA
GGTATAGGAGTCTATGTGCATTTTGGAATTACTCACCTGATACCATGAATGGCAGCTGGTCT
TCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCT
GACACATTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTC
TTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTT
ACCTTCTGGTTCTTCAGTGAAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTG
TAGCCTATTTCTTGCTGAACTTGTTTTTCTTGTTGGGATCAATACAAATACTAATAAGCTCT
TCTGTTCAATCATTGCCGGACTGCTACACTACTTCTTTTTAGCTGC**TTT**GCATGGATGTGC
ATTGAAGGCATACATCTCTATCTCATTGTTGTGGGTGTCATCTACAACAAGGGATTTTTGCA
CAAGAATTTTTATATCTTTGGCTATCTAAGCCCAGCCGTGGTAGTTGGATTTTCGGCAGCAC
TAGGATACAGATATTATGGCACAACCAAAGTATGTTGGCTTAGCACCGAAAACAACTTTATT
TGGAGTTTTATAGGACCAGCATGCCTAATCATTCTTGTTAATCTCTTGGC**TTT**GGAGTCAT
CATATACAAAGTTTTCGTCACACTGCAGGGTTGAAACCAGAAGTTAGTTGCTTTGAGAACA
TAAGGTCTTGTGCAAGAGGAGCCCTCGCTCTTCTGTTCCTTCTCGGCACCACCTGGATCTTT
GGGGTTCTCCATGTTGTGCACGCATCAGTGGTTACAGCTTACCTCTTCACAGTCAGCAATGC
TTTCCAGGGGATGTTCATTTTTTTATTCCTGTGTGTTTTATCTAGAAAGATTCAAGAAGAAT
ATTACAGATTGTTCAAAAATGTCCCCTGTTGTTTTGGATGTTTAAG**GTAA**ACATAGAGAATG
GTGGATAATTACAACTGCACAAAAATAAAAATTCCAAGCTGTGGATGACCAATGTATAAAAA
TGACTCATCAAATTATCCAATTATTAACTACTAGACAAAAAGTATTTTAAATCAGTTTTTCT
GTTTATGCTATAGGAACTGTAGATAATAAGGTAAAATTATGTATCATATAGATATACTATGT
TTTTCTATGTGAAATAGTTCTGTCAAAAATAGTATTGCAGATATTTGGAAAGTAATTGGTTT
CTCAGGAGTGATATCACTGCACCCAAGGAAAGATTTTCTTTCTAACACGAGAAGTATATGAA
TGTCCTGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGTGTTGCCTTTGAAACTAGTCC
CCTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGGCAGA
ATATCAAACAGTGAAAAGGGAATGATAAGATGTATTTTGAATGAACTGTTTTTTCTGTAGAC
TAGCTGAGAAATTGTTGACATAAAATAAAGAATTGAAGAAACACATTTTACCATTTTGTGAA
TTGTTCTGAACTTAAATGTCCACTAAAACAACTTAGACTTCTGTTTGCTAAATCTGTTTCTT
TTTCTAATATTCTAAAAAAAAAAAAAAGGTTTACCTCCACAAATTGAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

【図１９】

MKRLPLLVVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC
GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNVCIAA
NINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSSLLGYKNNTISAKDTL
SNSTLTEFVKTVNNFVQRDTPVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSFQKTTEFDT
NSTDIALKVFFFDSYNMKHIHPHMNMDGDYINIFPPKRKAAVDSNGNVAVAPLYYKSIGPLLS
SSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTDRYRSLCAF
WNYSPDTMNGSWSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG
IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLPLAELVFLVGINTNTNKLFCSIIAGL
LHYFFLAAFAWMCIEGIHLYLIVVGVIYNKGFLHKNFYIFGYLSPAVVVGFSAALGYRYYGT
TKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGA
LALLFLLGTTWIFGVLHVVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEYYRLFKNV
PCCFGCLR

シグナルペプチド：

アミノ酸 1-19

膜貫通ドメイン：

アミノ酸 430-450, 465-486, 499-513, 535-549, 573-593, 619-636, 648-664

N-グリコシル化部位

アミノ酸 15-19, 21-25, 64-68, 74-78, 127-131, 177-181, 188-192, 249-253, 381-385, 395-399

グリコサミノグリカン付着部位

アミノ酸 49-53

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 360-364

カゼインキナーゼIIリン酸化部位

アミノ酸 54-58, 68-72, 76-80, 94-98, 135-139, 150-154, 155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329, 346-350, 608-612

チロシンキナーゼリン酸化部位

アミノ酸 36-44, 669-677, 670-678

N-ミリストイル化部位

アミノ酸 38-44, 50-56, 52-58, 80-86, 382-388, 388-394, 434-440, 480-486, 521-527

アスパラギン酸及びアスパラギンヒドロキシル化部位

アミノ酸 75-87

【図２０】

TGGAAACATATCCTCCCTCATATGAATATGGATGGAGACTACATAAATATATTTCCAAAGNG
AAAAGCCGGCATATGGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATAAGAGTAT
TGGTCCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAAAATTATGATAATTCT
GAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATGAGCTCAAACCCACCCAC
ATTATATGAACTTGAAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATAGGTATA
GGAGTCTATGTGGCATTTTGGAATACTCACCTGATACCATGAATGGCAGCTGGTCTTCAGAG
GGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACA
TTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTCTTACAA
GGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTTACCTTC
TGGTTCTTCAGTGAAATTCAAAGCACCAGGA

【図２１】

GCTCCCAGCCAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG
CTAAGCGAGGCCTCCTCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTCACCCCGGCTGGGA
CAAGAAGCCGCCGCCTGCCTGCCCGGGCCCGGGGAGGGGGCTGGGGCTGGGGCCGGAGGCGG
GGTGTGAGTGGGTGTGTGCGGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
TGTCTTGGGCACCTACCCGTGGGGCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGGAG
CCGCCGCGCCGTCAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC
GGCACTCACAGCCCCGCAGCGCATCCCGGTCGCCGCCCAGCCTCCCGCACCCCCATCGCCGG
AGCTGCGCCGAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGTGGTGGTCCACGTATGG
ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCC
CCACGTGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCCC
ACGGGCTCTCCAGCTGCTTCCTGCGCATCCGTGCCGACGGCGTCGTGGACTGCGCGCGGGGC
CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG
CGTGCACAGCGTGCGGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT
ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGA
TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTCCAAACAGCGGCAGCTGTACAAGAA
CAGAGGCTTTCTTCCACTCTCTCATTTCCTGCCCATGCTGCCCATGGTCCCAGAGGAGCCTG
AGGACCTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCCTGGAGACCGACAGCATG
GACCCATTTGGGCTTGTCACCGGACTGGAGGCCGTGAGGAGTCCCAGCTTTGAGAAGTAACT
GAGACCATGCCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG
TGCTTCTACAAGAACAGTCCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAA
GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT
CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCTGCCTGGGCCCCCATTCTGCTCCCTCGA
GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC
TCACTTCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC
CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTCAGGAACAGGTGATCCACTCTGTA
AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG
ACCATTTGCCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCACCAG
GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTCCCC
CTGAGGCCAGTTCTGTCATGGATGCTGTCCTGAGAATAACTTGCTGTCCCGGTGTCACCTGC
TTCCATCTCCCAGCCCACCAGCCCTCTGCCCACCTCACATGCCTCCCCATGGATTGGGGCCT
CCCAGGCCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT
TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC
CCTTTCCCCAGCACTTGGTTTTCCAACATGATATTTATGAGTAATTTATTTTGATATGTACA
TCTCTTATTTTCTTACATTATTTATGCCCCCAAATTATATTTATGTATGTAAGTGAGGTTTG
TTTTGTATATTAAAATGGAGTTTGTTTGT

【図２２】

MRSGCVVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDPIRLRHLYTSGPHGLSSCFLRI
RADGVVDCARGQSAHSLLEIKAVALRTVAIKGVHSVRYLCMGADGKMQGLLQYSEEDCAFEE
EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD
MFSSPLETDSMDPFGLVTGLEAVRSPSFEK

シグナルペプチド：

アミノ酸 1-22

カゼインキナーゼII リン酸化部位

アミノ酸 78-82, 116-120, 190-194, 204-208

N-ミリストイル化部位

アミノ酸 15-21, 54-60, 66-72, 201-207

原核生物膜リポタンパク質脂質付着部位

アミノ酸 48-59

【図２３】

CCCAGAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCA
GAGCAGCCGGCTGCCGCCCCGGGAAGATGGCGAGGAGGAGCCGCCACCGCCTCCTCCTGCTG
CTGCTGCGCTACCTGGTGGTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA
AGAAGACTGTTTCCTCCAGATTAGAGTGGAAGAAACTGGGTCGGAGTGTCTCCTTTGTCTAC
TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCG
GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG
AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT
CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAACTGTGGTAGAGCTACGATGTCAAGA
CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA
ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATACACAATGAATACAAAAACTGGAACT
CTGCAATTTAATACTGTTTCCAAACTGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC
TGTTGGATATCGCAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA
TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTAT
GCTCAGAGGAAAGGCTACTTTTCAAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA
AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGG
CCGCGGCGGGCGGATCACGAGGTCAGGAGTTCTAGACCAGTCTGGCCAATATGGTGAAACCC
CATCTCTACTAAAATACAAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGC
TTGGGAGACAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC
CACTGCAGTCCAGCCTGGGTAACAGAGCAAGATTCCATCTCAAAAAATAAAATAAATAAATA
AATAAATACTGGTTTTTACCTGTAGAATTCTTACAATAAATATAGCTTGATATTC

【図２４】

MARRSRHRLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLE
WKKLGRSVSFVYYQQTLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQNLEED
TVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQST
NSSYTMNTKTGTLQFNTVSKLDTGEYSCEARNSVGYRRCPGKRMQVDDLNISGIIAAVVVVA
LVISVCGLGVCYAQRKGYFSKETSFQKSNSSSKATTMSENVQWLTPVIPALWKAAAGGSRGQ
EF

シグナルペプチド：

アミノ酸 1-20

膜貫通ドメイン：

アミノ酸 130-144, 238-258

N-グリコシル化部位

アミノ酸 98-102, 187-191, 236-240, 277-281

カゼインキナーゼII リン酸化部位

アミノ酸 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

N-ミリストイル化部位

アミノ酸 182-188, 239-245, 255-261, 257-263, 305-311

アミド化部位

アミノ酸 226-230

【図２５】

GACATCGGAGGTGGGCTAGCACTGAAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAG
AAAGAAGAGGAAGATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCAT
CATGCTGCTATTCCTGCAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA
ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA
TTATATCATTAAGGAAATAGTAACCTTCTCTTCTCCAATATGCATGACATTTTTGGACAATG
CAATTGTGGCACTGGCACTTATTTCAGTGAAGAAAAACTTTGTGGTTCTATGGCATTCATCA
TTTGACAAATGCAAGCATCTTCCTTATCAATCAGCTCCTATTGAACTTACTAGCACTGACTG
TGGAATCCTTAAGGGCCCATTACATTTCTGAAGAAGAAAGCTAAGATGAAGGACATGCCACT
CCGAATTCATGTGCTACTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG
TGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTTGGTTTACACCCAGATCCATTTAT
ATGGAAGCATCTACAGTGGATTGTAATGATTTAGGTCTTTTAACTTTCCCAGCCAGATTGCC
AGCTAACACACAGATTCTTCTCCTACAGACTAACAATATTGCAAAAATTGAATACTCCACAG
ACTTTCCAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAAT
ATTAATGTAAAAAAGATGCCTCAGCTCCTTTCTGTGTACCTAGAGGAAAACAAACTTACTGA
ACTGCCTGAAAAATGTCTGTCCGAACTGAGCAACTTACAAGAACTCTATATTAATCACAACT
TGCTTTCTACAATTTCACCTGGAGCCTTTATTGGCCTACATAATCTTCTTCGACTTCATCTC
AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGAT
TCTGATGATTGGGGAAAATCCAATTATCAGAATCAAAGACATGAACTTTAAGCCTCTTATCA
ATCTTCGCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTT
GGACTGGAAAACTTAGAAAGCATCTCTTTTTACGATAACAGGCTTATTAAAGTACCCCATGT
TGCTCTTCAAAAAGTTGTAAATCTCAAATTTTTGGATCTAAATAAAAATCCTATTAATAGAA
TACGAAGGGGTGATTTTAGCAATATGCTACACTTAAAAGAGTTGGGGATAAATAATATGCCT
GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC
TACTAACAACCCTAGATTGTCTTACATTCACCCAATGCATTTTTCAGACTCCCCAAGCTGG
AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTG
CCAAACCTCAAGGAAATCAGCATACACAGTAACCCCATCAGGTGTGACTGTGTCATCCGTTG
GATGAACATGAACAAAACCAACATTCGATTCATGGAGCCAGATTCACTGTTTTGCGTGGACC
CACCTGAATTCCAAGGTCAGAATGTTCGGCAAGTGCATTTCAGGGACATGATGGAAATTTGT
CTCCCTCTTATAGCTCCTGAGAGCTTTCCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT
TTCCTTTCACTGTAGAGCTACTGCAGAACCACAGCCTGAAATCTACTGGATAACACCTTCTG
GTCAAAAACTCTTGCCTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTA
GATATAAATGGCGTAACTCCCAAAGAAGGGGGTTTATATACTTGTATAGCAACTAACCTAGT
TGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGATCTTTTCCACAAGATAACAATG
GCTCTTTGAATATTAAAATAAGAGATATTCAGGCCAATTCAGTTTTGGTGTCCTGGAAAGCA
AGTTCTAAAATTCTCAAATCTAGTGTTAAATGGACAGCCTTTGTCAAGACTGAAAATTCTCA
TGCTGCGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC
CATCAACTGAGTATAAAATTTGTATTGATATTCCCACCATCTATCAGAAAAACAGAAAAAAA
TGTGTAAATGTCACCACCAAAGGTTTGCACCCTGATCAAAAAGAGTATGAAAAGAATAATAC
CACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGATTATTGGTGTGATATGTCTTATCA
GCTGCCTCTCTCCAGAAATGAACTGTGATGGTGGACACAGCTATGTGAGGAATTACTTACAG
AAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAA
AGAAAAAAGTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCCT
AAAAACCACCAAGGAAACCTACTCCAAAAATGAAC

【図２６】

MKDMPLRIHVLLGLAITTLVQAVDKKVDCPRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT
FPARLPANTQILLLQTNNIAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE
ENKLTELPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLHLNSNRLQMINSKWFDA
LPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRL
IKVPHVALQKVVNLKFLDLNKNPINRIRRGDFSNMLHLKELGINNMPELISIDSLAVDNLPD
LRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSNPIRC
DCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGQNVRQVHFRDMMEICLPLIAPESFPSNLNV
EAGSYVSPHCRATAEPQPEIYWITPSGQKLLPNTLTDKFYVHSEGTLDINGVTPKEGGLYTC
IATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFV
KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE
YEKNNTTTLMACLGGLLGIIGVICLISCLSPEMNCDGGHSYVRNYLQKPTFALGELYPPLIN
LWEAGKEKSTSLKVKATVIGLPTNMS

シグナル配列：

アミノ酸 1-22

膜貫通ドメイン：

アミノ酸 633-650

N-グリコシル化部位

アミノ酸 93-97, 103-107, 223-227, 382-386, 522-526, 579-583, 608-612, 624-628, 625-629

カゼインキナーゼII リン酸化部位

アミノ酸 51-55, 95-99, 242-246, 468-472, 487-491

チロシンキナーゼリン酸化部位

アミノ酸 570-579

N-ミリストイル化部位

アミノ酸 13-19, 96-102, 158-164, 221-227, 352-358, 437-443, 491-497, 492-498, 634-640, 702-708

細胞付着配列

アミノ酸 277-280

【図２７】

GCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAG
CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC
CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT
GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTA
TGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG
GGTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGG
ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT
GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG
TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC
ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC
AACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGC
CAACGACGCTGACCTTTGTAACCTCCCTAAAAAAACTACCGATTATGCCATGCTGGTCACCA
TGTTTGGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG
GATGCCCGGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA
ACCTGATGATATTAGCACTGTGGTATAGTGTCCAAACTGACTGTCATTGAGAAAGAAAGAAA
GTAGTTTGCGATTGCAGTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAA
ACTTTGTATTTCAGTTTTTTTTGAATTATGCCACTGCTGAACTTTTAACAAACACTACAACA
TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCCGATGGTATATTTCTGAGT
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTT
AATTTAAAAGCAAATAAAAGCTTAACTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAAAA
ACA

【図２８】

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDL
PPETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDLSDNR
IQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFL
NAANDADLCNLPKKTTDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKSLPSRQKK
ADEPDDISTVV

シグナル配列：

アミノ酸 1-33

膜貫通ドメイン：

アミノ酸 205-220

N-グリコシル化部位

アミノ酸 47-51, 94-98

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 199-203

カゼインキナーゼII リン酸化部位

アミノ酸 162-166, 175-179

N-ミリストイル化部位

アミノ酸 37-43, 45-51, 110-116

【図２９】

【図３０】

MQVSKRMLAGGVRSMPSPLLACWQPILLLVLGSVLSGSATGCPPRCECSAQDRAVLCHRKCF
VAVPEGIPTETRLLDLGKNRIKTLNQDEFASPPHLEELELNENIVSAVEPGAFNNLFNLRTL
GLRSNRLKLIPLGVFTGLSNLTKQDISENKIVILLDYMFQDLYNLKSLEVGDNDLVYISHRA
FSGLNSLEQLTLEKCNLTSIPTEALSHLHGLIVLRLRHLNINAIRDYSFKRLYRLKVLEISH
WPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLSYNPISTIEGSMLHEL
LRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLEESVFHSVGNLETLILDSNPLA
CDCRLLWVFRRRWRLNFNRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV
FVDEGHTVQFVCRADGDPPPAILWLSPRKHLVSAKSNGRLTVFPDGTLEVRYAQVQDNGTYL
CIAANAGGNDSMPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI
IATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPRKSDAGISSADAPRKFNMKMI

シグナル配列：

アミノ酸 1-41

膜貫通ドメイン：

アミノ酸 556-578

N-グリコシル化部位

アミノ酸 144-148, 202-206, 264-268, 274-278, 293-297, 341-345, 492-496, 505-509, 526-530, 542-546

カゼインキナーゼIIリン酸化部位

アミノ酸 49-53, 108-112, 146-150, 300-304, 348-352, 349-353, 607-611

チロシンキナーゼリン酸化部位

アミノ酸 590-598

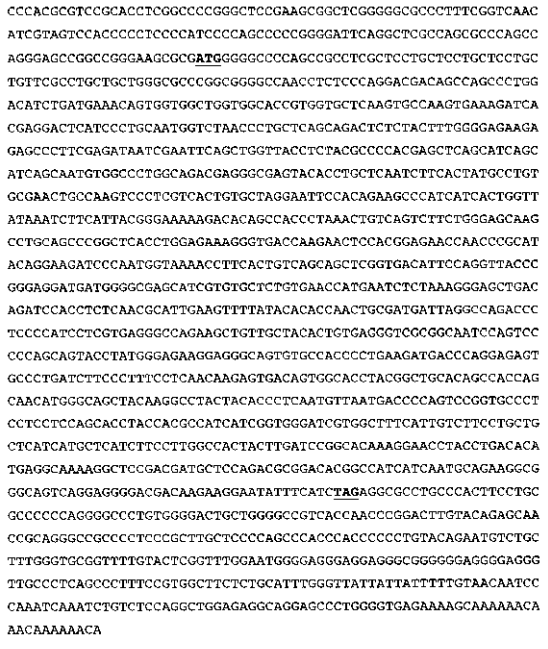
N-ミリストイル化部位

アミノ酸 10-16, 32-38, 37-43, 113-119, 125-131, 137-143, 262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

原核生物膜リポタンパク質脂質付着部位

アミノ酸 32-43

【図３１】



【図３２】

MGAPAASLLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLKCQVKDHEDSSLQW
SNPAQQTLYFGEKRALRDNRIQLVTSTPHELSISISNVALADEGEYTCSIFTMPVRTAKSLV
TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQEDPNGK
TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQ
KLLLHCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKA
YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIPLGHYLIRHKGTYLTHEAKGSDD
APDADTAIINAEGGQSGGDDKKEYFI

シグナル配列：

アミノ酸 1-20

膜貫通ドメイン：

アミノ酸 331-352

N-グリコシル化部位

アミノ酸 25-29, 290-294

カゼインキナーゼIIリン酸化部位

アミノ酸 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-ミリストイル化部位

アミノ酸 2-8, 23-29, 156-162, 218-224, 295-301, 298-304, 306-310, 334-340, 360-364, 385-389, 386-390

原核生物膜リポタンパク質脂質付着部位

アミノ酸 7-18

【図３３】

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCCAAACCCTTTTCTTCTCCTTTCCTGG
CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG
TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAA**ATG**CTGCTTTGGATTCTGTT
GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA
TCTGTTCCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAAGGGCTTCACA
AGTCTGCAGCGTTTCACTGCCCCGACTTCCCAGTTTTACCATTTATTTCTGCATGGCAATTC
CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG
AAAACAATGGCTTGCATGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGG
CTGCACATCAACAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA
TCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATATAGACCCGGGGGCCTTCCAGG
ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCCTACCTGCCAAC
GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC
CTATGAGGAGGTCTTGGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCCTT
GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCC
CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC
CACCGAACAGGACTTGTGTCCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCCTG
CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG
GATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT
CAAAATCAGACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCCTTAGCTAACA
GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTCGGGTTTAAAGATGAAC
TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA
GCTTTTCCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCCACTTTGTGGATTACAAGA
ACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG
AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA
GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA
TCCTCCCGGGCACTTTCAATGCCATGCCCAAACTGAGGATCCTCATTCTCAACAACAACCTG
CTGAGGTCCCTGCCTGTGGACGTGTTCGCTGGGGTCTCGCTCTCTAAACTCAGCCTGCACAA
CAATTACTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA
TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCA
GAACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACTTCTT
TAGAAAGGATTTCATGCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT
CGCCCACGTTAACTTCGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCGGGACGCACTCC
AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT
GTTTGTCACCTCCGCCTTCACCGTGGTGGGCATGCTCGTGTTTATCCTGAGGAACCGAAAGC
GGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATTCCCTACAGACAGTCTGT
GACTCTTCCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCCACAGAGTGTATGACTG
TGGCTCTCACTCGCTCTCAGAC**TAA**GACCCCAACCCCAATAGGGGAGGGCAGAGGG**AA**GGCG
ATACATCCTTCCCCACCGCAGGCACCCCGGGGGCTGGAGGGGCGTGTACCCAAATCCCCGCG
CCATCAGCCTGGATGGGCATAAGTAGATAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT
GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA
GCCAGCTCGCTCTTTGCTGAGAGCCCCTTTTGACAGAAAGCCCAGCACGACCCTGCTGGAAG
AACTGACAGTGCCCCTCGCCCTCGGCCCCGGGGCCTGTGGGGTTGGATGCCGCGGTTCTATAC
ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTCCCCTGTGGATTAG
CCCCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGGCAGTTGCACGAAGGCATGAATGTAT
TGTAAATAAGTAACTTTGACTTCTGAC

【図３４】

MLLWILLLETSLCFAAGNVTGDVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYH
LFLHGNSLTRLFPNEFANFYNAVSLHMENNGLHEIVPGAFLGLQLVKRLHINNNKIKSFRKQ
TFLGLDDLEYLQADFNLLRDIDPGAFQDLNKLEVLILNDNLISTLPANVFQYVPITHLDLRG
NRLKTLPYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQ
GKDLNETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFKTNGQEDHATPGSAPNGGTK
IPGNWQIKIRPTAAIATGSSRNKPLANSLPCPGGCSCDHIPGSGLKMNCNNRNVSSLADLKP
KLSNVQELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLLDLRWLYMDSNY
LDTLSREKFAGLQNLEYLNVEYNAIQLILPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSL
SKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGNPWECSCTIVPFKQWAERLGSEVLMSDLKC
ETPVNFFRKDFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL
VPGLLLVFVTSAFTVVGMLVFILRNRKRSKRRDANSSASEINSLQTVCDSSYWHNGPYNADG
AHRVYDCGSHSLSD

シグナル配列：

アミノ酸 1-15

膜貫通ドメイン：

アミノ酸 618-638

N-グリコシル化部位

アミノ酸 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 122-126, 646-650

カゼインキナーゼIIリン酸化部位

アミノ酸 30-34, 180-184, 222-226, 256-260, 366-370, 573-577, 608-612, 657-661, 666-670, 693-697

N-ミリストイル化部位

アミノ酸 17-23, 67-73, 100-106, 302-308, 328-334, 343-349, 354-360, 465-471, 493-499, 598-604, 603-609

真核生物膜リポタンパク質脂質付着部位

アミノ酸 337-348

【図３５】

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCCTGACCCCAGAATAACTCAGGGC
TGCACCGGGCCTGGCAGCGCTCCGCACACATTCCTGTCGCGGCCTAAGGGAAACTGTTGGC
CGCTGGGCCCGCGGGGGGATTCTTGGCAGTTGGGGGGTCCGTCGGGAGCGAGGGCGGAGGGG
AAGGGAGGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC
AGCTCTGCGTCCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGCGGGGCCTCAG
AGA**ATG**AGGCCGGCGTTCGCCCTGTGCCTCCTCTGGCAGGCGCTCTGGCCCGGGCCGGGCGG
CGGCGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCTCGGGGGCCTGCTACAGCCTGC
ACCACGCTACCATGAAGCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCCCTC
AGCACCGTGCGTGCGGGCGCCGAGCTGCGCGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGG
GCCCGGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTCGCACTGGAGCGCAGGCGTTCCCACT
GCACCCTGGAGAACGAGCCTTTGCGGGGTTTCTCCTGGCTGTCCTCCGACCCCGGCGGTCTC
GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCCGCGGAGATGCGC
GGTACTCCAGGCCACCGGTGGGGTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC
GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTTGTCCTGCGCCGCGCCCCGGG
GCCGCCTCTAACTTGAGCTATCGCGCGCCCTTCCAGCTGCACAGCGCCGCTCTGGACTTCAG
TCCACCTGGGACCGAGGTGAGTGCGCTCTGCCGGGGACAGCTCCCGATCTCAGTTACTTGCA
TCGCGGACGAAATCGGCGCTCGCTGGGACAAACTCTCGGGCGATGTGTTGTGTCCCTGCCCC
GGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG
CTTTGCCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCA
GTGGGGAAGGACAGCCGACCCTTGGGGGGACCGGGGTGCCCACCAGGCGCCCGCCGGCCACT
GCAACCAGCCCCGTGCCGCAGAGAACATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC
ACCACTTGTCCCTGAACAAGACAATTCAGTAACATCTATTCCTGAGATTCCTCGATGGGGAT
CACAGAGCACGATGTCTACCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC
CCATCAGGGAGCGTGATTTCCAAGTTTAATTCTACGACTTCCTCTGCCACTCCTCAGGCTTT
CGACTCCTCCTCTGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTTGGTGATCT
TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTCACGAAAGCCCCTCTTCCCAGCCA
AGGAAGGAGTCTATGGGCCCGCCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTC
CAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG
CAGAGGGTGCCTTGCTGGCGGAGTCCCCTCTTGGCTCTAGTGATGCA**TAG**GGAAACAGGGGA
CATGGGCACTCCTGTGAACAGTTTTCACTTTTGATGAAACGGGGAA**CC**AAGAGGAACTTAC
TTGTGTAACTGACAATTTCTGCAGAAATCCCCCTTCCTCTAAATTCCCTTTACTCCACTGAG
GAGCTAAATCAGAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGA
TGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTCGGAGAA
TTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATATAATTTACA
TTAAAAAATAATTTCTACCAAAATGGAAAGGAAATGTTCTATGTTGTTCAGGCTAGGAGTAT
ATTGGTTCGAAATCCCAGGGAAAAAAATAAAAATAAAAAATTAAAGGATTGTTGAT

【図３６】

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALS
TVRAGAELRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGPSWLSSDPGGLE
SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQPEVLCPAPRPGA
ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG
RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSGEGQPTLGGTGVPTRRPPATA
TSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP
SGSVISKFNSTTSSATPQAFDSSSAVVFIFVSTAVVVLVILTMTVLGLVKLCFHESPSSQPR
KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

シグナル配列：

アミノ酸 1-16

膜貫通ドメイン：

アミノ酸 399-418

N-グリコシル化部位

アミノ酸 189-193, 381-385

グリコサミノグリカン付着部位

アミノ酸 289-293

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 98-102, 434-438

カゼインキナーゼIIリン酸化部位

アミノ酸 275-279, 288-292, 342-346, 445-449

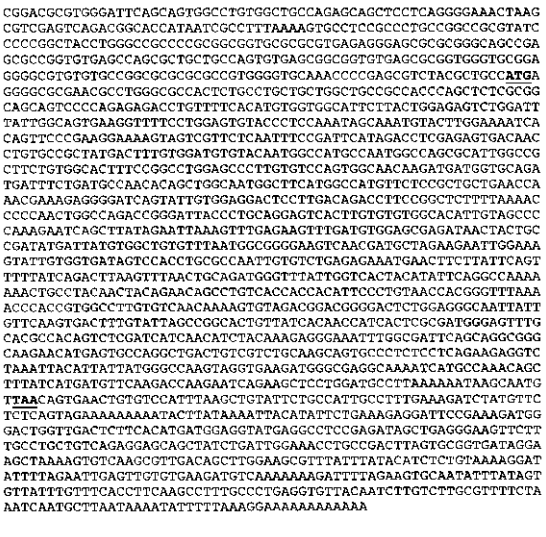
N-ミリストイル化部位

アミノ酸 30-36, 35-41, 58-64, 59-65, 121-127, 151-157, 185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469, 477-483

アスパラギン酸及びアスパラギンヒドロキシル化部位

アミノ酸 262-274

【図３７】



【図３８】

MRGANAWAPLCLLLAAATQLSRQQSPERPVFTCGGILTGESGFIGSEGFPGVYPPNSKCTWK
ITVPEGKVVVLNFRFIDLESDNLCRYDFVDVYNGHANGQRIGRFCGTFRPGALVSSGNKMMV
QMISDANTAGNGFMAMFSAAEPNERGDQYCGGLLDRPSGSFKTPNWPDRDYPAGVTCVWHIV
APKNQLIELKFEKFDVERDNYCRYDYVAVFNGGEVNDARRIGKYCGDSPPAPIVSERNELLI
QPLSDLSLTADGFIGHYIFRPKKLPTTTEQPVTTTPPVTTGLKPTVALCQQKCRRTGTLEGN
YCSSDFVLAGTVITTITRDGSLHATVSIINIYKEGNLAIQQAGKNMSARLTVVCKQCPLLRR
GLNYIIMGQVGEDGRGKIMPNSFIMMFKTKNQKLLDALKNKQC

シグナル配列：

アミノ酸 1-23

N-グリコシル化部位

アミノ酸 355-359

カゼインキナーゼIIリン酸化部位

アミノ酸 64-68, 142-146, 274-278

チロシンキナーゼリン酸化部位

アミノ酸 199-208

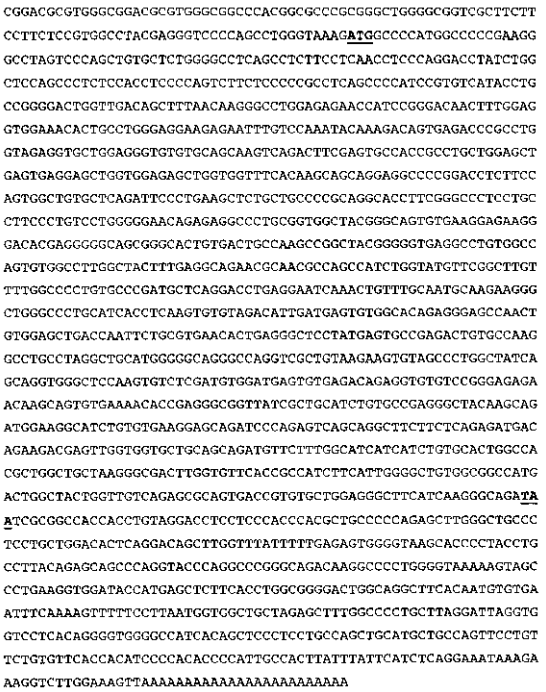
N-ミリストイル化部位

アミノ酸 34-40, 35-41, 100-106, 113-119, 218-224, 289-295, 305-311, 309-315, 320-326, 330-336

細胞付着配列

アミノ酸 149-152

【図３９】



【図４０】

MAPWPPKGLVPAVLWGLSLFLNLPGPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLER
TIRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQ
QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAG
YGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPEESNCLQCKKGWALHHLKCVDIDE
CGTEGANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECE
TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDELVVLQQMFFG
IIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

シグナル配列：

アミノ酸 1-29

膜貫通ドメイン：

アミノ酸 372-395

N-グリコシル化部位

アミノ酸 79-83, 205-209

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 290-294

カゼインキナーゼIIリン酸化部位

アミノ酸 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

N-ミリストイル化部位

アミノ酸 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179, 177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289, 326-332, 372-378, 395-401

アスパラギン酸及びアスパラギンヒドロキシル化部位

アミノ酸 321-333

EGF-様ドメインシステインパターンシグネチャー

アミノ酸 181-193

【図４１】

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA
GCACCATGCAGCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCCTGGCCAGCCCCGGG
GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGT
GCCCACCCTGGACAGGGCCGACATGGAGGAGCTGGTCATCCCCACCCACGTGAGGGCCCAGT
ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGGAAAGAGGTTCAGCCAGAGC
TTCCGAGAGGTGGCCGGCAGGTTCCTGGCGTTGGAGGCCAGCACACACCTGCTGGTGTTCGG
CATGGAGCAGCGGCTGCCGCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG
AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGCTGTCCCCGCGCAGCGCCCGGGCC
CGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACCTCCCTCATCGA
CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA
ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTGCTACAGGTGTCGGTGCAGAGG
GAGCATCTGGGCCCGCTGGCGTCCGGCGCCCACAAGCTGGTCCGCTTTGCCTCGCAGGGGGC
GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCCTGGACCTTGGGGACTATGGAG
CTCAGGGCGACTGTGACCCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG
ATGTACATTGACCTGCAGGGGATGAAGTGGGCCGAGAACTGGGTGCTGGAGCCCCCGGGCTT
CCTGGCTTATGAGTGTGTGGGCACCTGCCGGCAGCCCCCGGAGGCCCTGGCCTTCAAGTGGC
CGTTTCTGGGGCCTCGACAGTGCATCGCCTCGGAGACTGACTCGCTGCCCATGATCGTCAGC
ATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGTGCAGAA
GTGCAGCTGTGCCTCGGATGGTGCGCTCGTGCCAAGGAGGCTCCAGCCATAGGCGCCTAGTG
TAGCCATCGAGGGACTTGACTTGTGTGTGTTTCTGAAGTGTTCGAGGGTACCAGGAGAGCTG
GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTT
CCTCTGACAAGTTACCTCACCTAATTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA
GAGCCCTTGCTCAGTTTTCTCTATTCTTATTATTCACTGCACTATATTCTAAGCACTTACAT
GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCANTGTGTCATTGTTTACTTGTCCTGTCAC
TGGATCTGGGCTAAAGTCCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT
TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCT
AAAA

【図４２】

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV
ALLQRSHGDRSRGKRFSQSFREVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP
VPKAALHRHGRLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF
WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ
GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQPPEALAFKWPF
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSCASDGALVPRRLQP

シグナル配列：
アミノ酸 1-18

N-グリコシル化部位
アミノ酸 158-162

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位
アミノ酸 76-80

カゼインキナーゼIIリン酸化部位
アミノ酸 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-ミリストイル化部位
アミノ酸 19-25, 156-162, 225-231, 260-266, 274-280

アミド化部位
アミノ酸 74-78

TGF-. ファミリーシグネイチャー
アミノ酸 282-298

【図４３】

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCGATGGGGACAAA
GGCGCAAGTCGAGAGGAAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
CATTGGGCAGTGTTACAGTGCACTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
TGACCTTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGG
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT
GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCT
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGGTCATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT
CTTGGTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTCCTGGTGTGAGCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT
CAGGTGCTACCGGACTCTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC
TACACCCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC
ATCCTCCTTCATGCCCTCCCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACTTGTTTAAA
GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGCAAAAATGGCGGGGGTCGCAGGAATCTGCACTCAACTGCCCACCTGGC
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG
TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGTCCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAACTGGAGGCTGGGCGCAGTGGCTCACGCCTG
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA
ACATGGAGAAACCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAAACTCCAGCTCAAAAAAAAAAAAAAAA

【図４４】

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW
KFDQGDTTRLVCYNNKITASYEDRVTFLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWFKDGIVMPTNPKSTRAFSNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVAAVLVTLI
LLGILVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

シグナル配列：
アミノ酸 1-27

膜貫通ドメイン：
アミノ酸 238-255

N-グリコシル化部位
アミノ酸 185-189

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位
アミノ酸 270-274

カゼインキナーゼIIリン酸化部位
アミノ酸 34-38, 82-86, 100-104, 118-122, 152-156, 154-158, 193-197, 203-207, 287-291

N-ミリストイル化部位
アミノ酸 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

【図４５】

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGC**ATG**AGCGGCGGTTGGATGGCGCAGGTTGGA
GCGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTCGGCCTCGGACTAGGCCT
GGAGGCCGCCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCAGGCCCCAGCTCAG
GCTCGTGCCCACCCACCAAGTTCCAGTGCCGCACCAGTGGCTTATGCGTGCCCCTCACCTGG
CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCC
ATGTACCCAGAAAGGGCAATGCCCACCGCCCCCTGGCCTCCCCTGCCCCTGCACCGGCGTCA
GTGACTGCTCTGGGGGAACTGACAAGAAACTGCGCAACTGCAGCCGCCTGGCCTGCCTAGCA
GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCA
CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG
GGGATGCCACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCACCTCTCTCAGGAATGCC
ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCCTCTGTCGGGAATGCCACATCCTC
CTCTGCCGGAGACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC
TCAGTGCAAGCCTGGTCACCGCCACCCTCCTCCTTTTGTCCTGGCTCCGAGCCCAGGAGCGC
CTCCCCCCACTGGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC
CTCGCTGCCC**TGA**GGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA
GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCC
TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGAACTGAG
GGGCTGGCCCCAGGCAGCTCCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG
CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

【図４６】

MSGGWMAQVGAWRTGALGLALLLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR
TSGLCVPLTWRCDRDLDCSDGSDEEECRIEPCTQKGQCPPPPGLPCPCTGVSDCSGGTDKKL
RNCSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVT
LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAAVLSASLVTATLL
LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

シグナル配列：

アミノ酸 1-30

膜貫通ドメイン：

アミノ酸 230-246

N-グリコシル化部位

アミノ酸 126-130, 195-199, 213-217

カゼインキナーゼIIリン酸化部位

アミノ酸 84-88, 140-144, 161-165, 218-222

N-ミリストイル化部位

アミノ酸 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218, 224-230, 230-236, 263-269

真核生物膜リポタンパク質脂質付着部位

アミノ酸 44-55

ロイシンジッパーパターン

アミノ酸 17-39

【図４７】

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG
GTTAGACTGGCGGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGA
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAG**ATG**TATCATGGAATGAACCCGAGCAATG
GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCCAGAGACTCTTG
GCCGTGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTT
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG
GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG
GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA
TAATTCCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA
ACAAGACATATAGACATGGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG
TACCCCGACCTACACAATATGGTTTCATTATGTCGCGATGATGGAACGTGGAATAATCTGCC
CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC
TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCCGGTG
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCGTCT
GCCACCCGCGGCCTTGTGAGCGCTACAACCACGGAACTGTGGTGGAGTTTTACTGCGATCCT
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCCTTC
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA
CCACGTGGAAGATTGTGGCGTTCACGGCAACCAGTGTGCTGCTGGTGCTGCTGCTCGTCATC
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCCAGGGGGCCTCCCCGGAG
TTCCAGCAGTGACCCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG
ACGAAGCTGTGAGTGGCGGCTTGAGTGCCTTAGGCCCCGGGTACATGGCCTCTGTGGGCCAG
GGCTGCCCCTTACCCGTGGACGACCAGAGCCCCCCAGCATACCCCGGCTCAGGGGACACGGA
CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA
GTCTGTATTCACCTCCCAGGTGCCAAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATA
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCCAGGCATCCATCATGCCCACTGGGT
GTTGTTCCTAAGAAAC**TGA**TTGATTAAAAAATTTCCCAAAGTGTCCTGAAGTGTCTCTTCAA
ATACATGTTGATCTGTGGAGTTGATTCCTTTCCTTCTCTTGGTTTTAGACAAATGTAAACAA
AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC
CTGTTTCTTCTTGACACAGACTGATTAAAAATTAAAAGNAAAAAA

【図４８】

MKFLLDILLLLPLLIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK
LVLWDINKHGLEETAAKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAEIGDVSILVNNAGVV
YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC
SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVVNRLMH
GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMKAQ

シグナル配列：

アミノ酸 1-19

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 30-34, 283-287

カゼインキナーゼIIリン酸化部位

アミノ酸 52-56, 95-99, 198-202, 267-271

N-ミリストイル化部位

アミノ酸 43-49, 72-78, 122-128, 210-216

【図４９】

CCCACGCGTCCGCTCCGCGCCCTCCCCCCCCGCCTCCCGTGCGGTCCGTCGGTGGCCTAGAGA
TGCTGCTGCCGCGGTTGCAGTTGTCGCGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGT
AGCGCCCGAGTGTCGGGGGGCGCACCCGAGTCGGGCC**ATG**AGGCCGGGAACCGCGCTACAGG
CCGTGCTGCTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTCGCCTGCTGAGTGCC
TCGGATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCTTGTTA
TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT
GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAAACTGATAGAA
AAGTTCATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGA
GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATAT
CACAATTTAGGAACTGGTATGTGGATGAGCCGTCCTGCGGCAGCGAGGTCTGCGTGGTCATG
TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCCTACATGTTCCAGTGGAATGATGA
CCGGTGCAACATGAAGAACAATTTCATTTGCAAATATTCTGATGAGAAACCAGCAGTTCCTT
CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG
GAAGAAGATGCCAAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT
CCTAATCCCCAGCATTCCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTT
GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCCTAGCACAAAGAAGCAACACACCATC
TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA
AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCATTCCGAGTGTGTT
CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA
GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA
GTTCTCCCCAGACCAAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG
GTTAT**TAG**GACATATAAAAAACTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAAATC
CTCTTATTTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT
GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTAT
CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTA
GAGTCTCAATAAATGTCACTTGGTTGGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTG
GCAGTGATAAAGATGGGCTGTGGAGCTTGGAAAACCACCTCTGTTTTCCTTGCTCTATACAG
CAGCACATATTATCATACAGACAGAAAATCCAGAATCTTTTCAAAGCCCACATATGGTAGCA
CAG
GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTTCAAAGAATAAAATCAAATAAAGA
GCAGGAAAAAAAAA

【図５０】

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRL
NFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDL
YAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY
SDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV
VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRPDL
KNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKES
GWVENEIYGY

シグナル配列：

アミノ酸 1-21

膜貫通ドメイン：

アミノ酸 235-254

N-グリコシル化部位

アミノ酸 117-121, 312-316

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 296-300

カゼインキナーゼIIリン酸化部位

アミノ酸 28-32, 30-34, 83-87, 100-104, 214-218, 222-226, 299-303, 306-310, 323-327

N-ミリストイル化部位

アミノ酸 18-24, 37-43, 76-82, 146-152

【図５１】

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGC
ATCCGCAGGTTCCCGCGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT
GTTTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACC**ATG**ATCACTGGTGT
GTTCAGCATGCGCTTGTGGACCCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACC
AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC
CTGCTGAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCC
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAACTC
AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT
CAATACCATGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT
GCAGCAAATGTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC
TTTCACCAACCTTCAACCCACAGGAGGTCTTTATTCGTTCCACTAACATTTTTCGGAATCTG
GAGTCCACCCGTTGTTTGCTGGCTGGGCTTTTCCAGTGTCAGAAAGAAGGACCCATCATCAT
CCACACTGATGAAGCAGATTCAGAAGTCTTGTATCCCAACTACCAAAGCTGCTGGAGCCTGA
GGCAGAGAACCAGAGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTG
AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTTG
CACGGATGATCGAACAGAGAGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGG
GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC
CATGGACTCTGCCACTGCCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG
TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTGACCACAAATGGCCACCGTTTGCT
GTTGACCTGACCATGGAACTTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA
TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGT
TCTTGAATGCCATGTCAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA
ACTCAGGTGATGGAAGTTGGAAATGAAGAG**TAA**CTGATTTATAAAAGCAGGATGTGTTGATT
TTAAAATAAAGTGCCTTTATACAATG

【図５２】

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGAR
SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL
TKVGMQQMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKE
GPIIIHTDEADSEVLYPNYQSCWSLRQRTRGRRQTASLQPGISEDLKKVKDRMGIDSSDKVD
FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFLHILES
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEW
FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMSVYTLSPEKYHALCSQTQVMEVGNEE

シグナル配列：

アミノ酸 1-23

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 218-222

カゼインキナーゼIIリン酸化部位

アミノ酸 87-91, 104-108, 320-324

チロシンキナーゼリン酸化部位

アミノ酸 280-288

N-ミリストイル化部位

アミノ酸 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

アミド化部位

アミノ酸 216-220

ロイシンジッパーパターン

アミノ酸 10-32

ヒスチジン酸性ホスファターゼホスホヒスチジンシグネイチャー

アミノ酸 50-65

【図５３】

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT
TAAATTTCAGCTCATCACCTTCACCTGCCTTGGTC**ATG**GCTCTGCTATTCTCCTTGATCCTT
GCCATTTGCACCAGACCTGGATTCCTAGCGTCTCCATCTGGAGTGCGGCTGGTGGGGGGCCT
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG
ACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCA
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG
ATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCTTTCTCCCCA
GTCCCAGAGGGTGTCAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA
GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCCGCAAAGGTGG
TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCA
GGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTCG
AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGA
CCAGGTGGTATGCAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTCGTTGCTCAGGGGAGGAG
CAGTCCCTGGAGCAGTGCCAGCACAGATTTTGGGGGTTTCACGACTGCACCCACCAGGAAGA
TGTGGCTGTCATCTGCTCAGTG**TAG**GTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA
GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT
TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCCTGATTCTCAGGCCTTCAGAGTTGG
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA
CCACCTTTCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACACATA
CACCATTTGTCCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA
TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACTTTATTTACAATAATAAAGATAGCAC
TATGTGTTCAAA

【図５４】

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC
RELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEEVYDCSHDEDAGASC
ENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCQTGWSLRAAKVVCRQLGCGRAVL
TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDLRLVG
GDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL
DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

シグナル配列：

アミノ酸 1-15

カゼインキナーゼIIリン酸化部位

アミノ酸 47-51, 97-101, 115-119, 209-213, 214-218, 234-238, 267-271, 294-298, 316-320, 336-340

N-ミリストイル化部位

アミノ酸 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143, 180-186, 263-269, 286-292

アミド化部位

アミノ酸 196-200

スペクトレセプター反復ドメインシグネイチャー

アミノ酸 29-67, 249-287

【図５５】

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG
GTGAAGGCC**ATG**GACTTCATCACCTCCACAGCCATCCTGCCCCTGCTGTTCGGCTGCCTGGG
CGTCTTCGGCCTCTTCCGGCTGCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCGGAATGCTG
TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAACGAATGTGCAAAAGTCTTCTATGCT
GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA
ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC
TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC
GACATACTTGTCAACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA
TGTGGACAAGAGGGTCATGGAGACAAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC
TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG
ATGAGCATTCCTTTTCGATCAGCATATGCAGCCTCCAAGCACGCAACCCAGGCTTTCTTTGA
CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA
TCCACACCAACCTCTCTGTAAATGCCATCACCGCGGATGGATCTAGGTATGGAGTTATGGAC
ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTTCCTTGGCTGTTTATCTTCCAA
CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC
AAGAACTCC**TAG**TACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC
TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG
AAAGACTGAAGAAACACATCTCGTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA
AGCTTCTTCCCAGGGTGAGGGGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACACT
AAAAACTAGAAATAAACATCTCAAACAGTAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAG
AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTA
C

【図５６】

MDPITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA
KLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL
VNNAGISYRGTIMDTTVDVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAISSIQGKMSI
PFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDTTT
AQGRSPVEVAQDVLAAVGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSKNS

シグナル配列：

アミノ酸 1-21

膜貫通ドメイン：

アミノ酸 104-120, 278-292

N-グリコシル化部位

アミノ酸 228-232

グリコサミノグリカン付着部位

アミノ酸 47-51

カゼインキナーゼIIリン酸化部位

アミノ酸 135-139, 139-143, 253-257

チロシンキナーゼリン酸化部位

アミノ酸 145-153, 146-153

N-ミリストイル化部位

アミノ酸 44-50, 105-111, 238-244, 242-248, 291-297

アミド化部位

アミノ酸 265-269

真核生物膜リポタンパク質脂質付着部位

アミノ酸 6-17

【図５７】

CCCACGCGTCCGCTGGTGTTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA
AAAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGATGAAATTTCTTCTGGACATCCTC
CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCCTAA
GAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTACAGGAGCTGGGCATGGAATTGGGA
GACTGACTGCCTATGAATTTGCTAAACTTAAAAGCAAGCTGGTTCTCTGGGATATAAATAAG
CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTCATACCTTTGT
GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG
GAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTCTATACATCAGATTTGTTTGCTACA
CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA
GGCATTTCTTCCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG
CTGGACATGTCTCGGTCCCCTTCTTACTGGCTTACTGTTCAAGCAAGTTTGCTGCTGTTGGA
TTTCATAAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAAACAACATG
TCTGTGTCCTAATTTCGTAAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA
CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGGATTCTGACTGAGCAGAAGATG
ATTTTTATTCCATCTTCTATAGCTTTTTTAACAACATTGGAAAGGATCCTTCCTGAGCGTTT
CCTGGCAGTTTAAAACGAAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA
AAGCGCAATAAGCACCTAGTTTTCTGAAAACTGATTTACCAGGTTTAGGTTGATGTCATCTA
ATAGTGCCAGAATTTTAATGTTTGAACTTCTGTTTTTTCTAATTATCCCCATTTCTTCAATA
TCATTTTTGAGGCTTTGGCAGTCTTCATTTACTACCACTTGTTCTTTAGCCAAAAGCTGATT
ACATATGATATAAACAGAGAAATACCTTTAGAGGTGACTTTAAGGAAAATGAAGAAAAAGAA
CCAAAATGACTTTATTAAAATAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA
AATTTGTACCATAACCGTTTATTTAACATATATTTTTATTTTTGATTGCACTTAAATTTTGT
ATAATTTGTGTTTCTTTTTCTGTTCTACATAAAATCAGAAACTTCAAGCTCTCTAAATAAAA
TGAAGGACTATATCTAGTGGTATTTCACAATGAATATCATGAACTCTCAATGGGTAGGTTTC
ATCCTACCCATTGCCACTCTGTTTCCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT
GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG
GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

【図５８】

MKFLLDILLLLPLLIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK
LVLWDINKHGLEETAAKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAEIGDVSILVNNAGVV
YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC
SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVVNRLMH
GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMKAQ

シグナル配列：

アミノ酸 1-19

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 30-34, 283-287

カゼインキナーゼIIリン酸化部位

アミノ酸 52-56, 95-99, 198-202, 267-271

N-ミリストイル化部位

アミノ酸 43-49, 72-78, 122-128, 210-216

【図５９】

CCCACGCGTCCGCGGACGCGTGGGTCGACTAGTTCTAGATCGCGAGCGGCCGCCCGCGGCTC
AGGGAGGAGCACCGACTGCGCCGCACCCTGAGAGATGGTTGGTGCCATGTGGAAGGTGATTG
TTTCGCTGGTCCTGTTGATGCCTGGCCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT
GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC
TGGGAAGATCCAAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCCAGGACTGAACATGA
AGAGTTATGCCGGCTTCCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGGTTC
TTCCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG
AGGTTCATCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCACAAGTAACATGA
CCTTGCGTGACAGAGACTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA
GTGGGCACAGGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT
AGCACGGGATTTATACAGTGCACTAATTCAGTTTTTCCAGATATTTCCTGAATATAAAAATA
ATGACTTTTATGTCACTGGGGAGTCTTATGCAGGGAAATATGTGCCAGCCATTGCACACCTC
ATCCATTCCCTCAACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA
TGGATATTCTGATCCCGAATCAATTATAGGGGGCTATGCAGAATTCCTGTACCAAATTGGCT
TGTTGGATGAGAAGCAAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC
AGGAAGCAGAACTGGTTTGAGGCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC
AAGTGATCCTTCTTACTTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTTTGCGGT
GCACGGAACCTGAGGATCAGCTTTACTATGTGAAATTTTTGTCACTCCCAGAGGTGAGACAA
GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACTATAGTTGAAAAGTACTTGCGAGA
AGATACAGTACAGTCAGTTAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA
TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGC
ATGGACTGGAAAGGATCCCAGGAATACAAGAAGGCAGAAAAAAAAGTTTGGAAGATCTTTAA
ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGACTTCCATCAGGTAATTATTC
GAGGTGGAGGACATATTTACCCTATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGA
TTCATTTATGGAAAAGGATGGGATCCTTATGTTGGATAAACTACCTTCCCAAAAGAGAACAT
CAGAGGTTTTCATTGCTGAAAAGAAAATCGTAAAAACAGAAAATGTCATAGGAATAAAAAAA
TTATCTTTTCATATCTGCAAGATTTTTTTCATCAATAAAAATTATCCTTGAAACAAGTGAGC
TTTTGTTTTTGGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA
TTATTTAACTTAAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA
AATTTTAGGGTCTTGAATAGGAAGTTTTAATTTCTTCTAAGAGTAAGTGAAAAGTGCAGTTG
TAACAAACAAAGCTGTAACATCTTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT
GTTTGGAAATATTATTGGATAAGAATAGCTCAATTATCCCAAATAAATGGATGAAGCTATAA
TAGTTTTGGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA
AAAATATTATATATAAAAGTAAAAAAAAAA

【図６０】

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDSGQPLFLTPYIEAGKIQKGRELSL
VGPFPGLNMKSYAGFLTVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH
GPYVVTSNMTLRDRDFPWTTTLSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQF
FQIPPEYKNNDFYVTGESYAGKYVPAIAHLIHSLNPVREVKINLNGIAIGDGYSDPESIIGG
YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNVTG
CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT
EIMNNYKVLIYNGQLDIIVAAALTERSLMGMDWKGSQEYKKAEKKVWKIFKSDSEVAGYIRQ
AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

シグナル配列：

アミノ酸 1-22

N-グリコシル化部位

アミノ酸 81-85, 132-136, 307-311, 346-350

カゼインキナーゼIIリン酸化部位

アミノ酸 134-138, 160-164, 240-244, 321-325, 334-338, 348-352, 353-357, 424-428

チロシンキナーゼリン酸化部位

アミノ酸 423-432

N-ミリストイル化部位

アミノ酸 22-28, 110-116, 156-162, 232-238

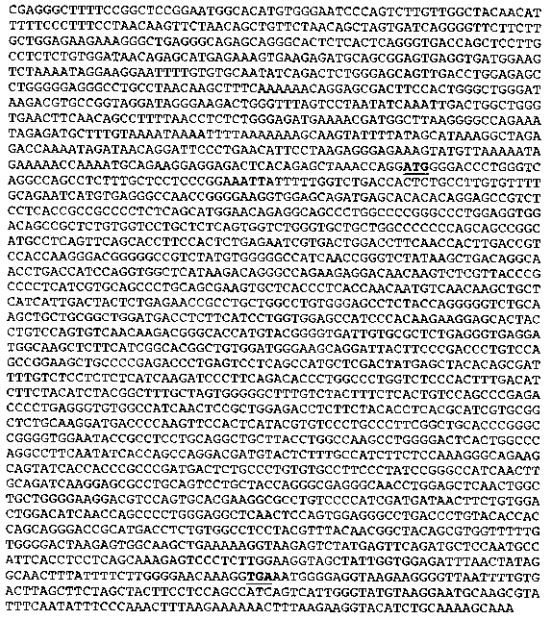
セリンカルボキシペプチダーゼ,セリン活性部位

アミノ酸 200-208

クリスタリン.及びγ・グリークキー・モチーフシグネイチャー

アミノ酸 375-391

【図６１】



【図６２】

MGTLGQASLFAPPGNYFWSDHSALCPAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP
RALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTPHSENRDWTFNHLTVHQGTGAVYVGAINRV
YKLTGNLTIQVAHKTGPEEDNKSRYPPLIVQPCSEVLTLTNNVNKLLIIDYSENRLLACGSL
YQGVCKLLRLDDLFILVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDY
FPTLSSRKLPRDPESSAMLDYELHSDFVSSLIKIPSDTLALVSHFDIFYIYGFASGGFVYFL
TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP
GDSLAQAFNITSQDDVLFAIFSKGQKQYHHPPDDSALCAFPIRAINLQIKERLQSCYQGEGN
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG
YSVVFVGTKSGKLKKVRVYEFRCSNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

シグナル配列：

アミノ酸 1-32

膜貫通ドメイン：

アミノ酸 71-87

N-グリコシル化部位

アミノ酸 130-134, 145-149, 217-221, 381-385

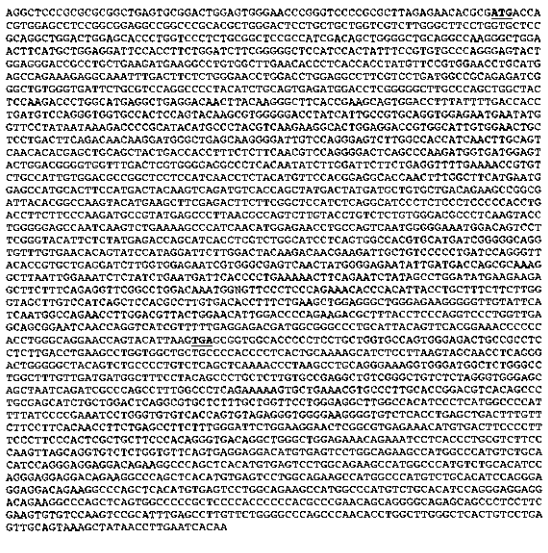
カゼインキナーゼII リン酸化部位

アミノ酸 139-143, 229-233, 240-244, 291-295, 324-328, 383-387, 384-388, 471-475, 481-485, 530-534

N-ミリストイル化部位

アミノ酸 220-226, 319-325, 353-359, 460-466, 503-509

【図６３】



【図６４】

MTTWSLRRRPARTLGLLLLVVLGPLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW
IFGGSIHYFRVPREYWRDRLLKMKACGLNTLTTYVPWNLHEPERGKFDFSGNLDLEAFVLMA
AEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMSRVVPLQ
YKRGGPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLLTSDNKDGLSKGIVQGVLAT
INLQSTHELQLLTTFLFNVQGTQPKMVMEYWTGWFDSWGGPHNILDSSEVLKTVSAIVDAGS
SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTAKYMKLRDFFGSISGIP
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSPGYILYE
TSITSSGILSGHVHDRGQVFVNTVSIGFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN
IDDQRKGLIGNLYLNDSPLKNFRIYSLDMKKSFFQRFGLDKWXSLPETPTLPAFFLGSLSIS
STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLSSGINQVIVFEETMAGPA
LQFTETPFHLGRNQYIK

シグナル配列：

アミノ酸 1-27

カゼインキナーゼII リン酸化部位

アミノ酸 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

N-ミリストイル化部位

アミノ酸 146-152, 236-242, 240-246, 244-250, 287-293, 309-315, 320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

【図６５】

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC
CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG
GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC
CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCCATGGCTCCC
AAGAAGCTGTCCTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCA
GGCAGACACTCGGTCGTTCGTAGTGGATAGGGGTCATGACCGGTTTCTCCTAGACGGGGCCC
CGTTCCGCTATGTGTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC
CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACTA
CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA
ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG
TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTC
AGATCCAGACTTCCTTGCCGCAGTGGACTCCTGGTTCAAGGTCTTGCTGCCCAAGATATATC
CATGGCTTTATCACAATGGGGGCAACATCATTAGCATTCAGGTGGAGAATGAATATGGTAGC
TACAGAGCCTGTGACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCCGTGCACTGCTAGG
AGAAAAGATCTTGCTCTTCACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGG
GACTCTATACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT
CGGAAGTATGAACCCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA
CTGGGGCCAGAATCACTCCACACGGTCTGTGTCAGCTGTAACCAAAGGACTAGAGAACATGC
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG
AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC
TATATCTGAAGCAGGGGACCCCACACCTAAGCTTTTTGCTCTTCGAGATGTCATCAGCAAGT
TCCAGGAAGTTCCTTTGGGACCTTTACCTCCCCGAGCCCCAAGATGATGCTTGGACCTGTG
ACTCTGCACCTGGTTGGGCATTTACTGGCTTTCCTAGACTTGCTTTGCCCCCGTGGGCCCAT
TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC
GAACCTATATGACCCATACCATTTTTGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC
CATGACCGTGCCTATGTGATGGTGGATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAG
AGACAAACTATTTTTGACGGGGAAACTGGGGTCCAAACTGGATATCTTGGTGGAGAACATGG
GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG
GGGCAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG
GTGGTTTCCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT
ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG
ACCAAGGGCCAAGTCTGGATCAATGGGTTTAACTTGGGCCGGTACTGGACAAAGCAGGGGCC
ACAACAGACCCTCTACGTGCCAAGATTCCTGCTGTTTCCTAGGGGAGCCCTCAACAAAATTA
CATTGCTGGAACTAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTTGGATAAGCCTATC
CTCAATAGCACTAGTACTTTGCACAGGACACATATCAATTCCCTTTCAGCTGATACACTGAG
TGCCTCTGAACCAATGGAGTTAAGTGGGCACTGAAAGGTAGGCCGGGCATGGTGGCTCATGC
CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTCAGGACTTCAAGA
CCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGCCGGGCGTG
ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC
AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA
GACACTCCATCTCAAAAAAAAAAAA

【図６６】

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDRGHDRFLLDGAPFRYVSGSLHYFRVPRVL
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYI
CAEWEMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE
YGSYRACDFSYMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF
TLLRKYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNMYMFHGGTNF
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPPSPKMML
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPPWVPN
NGVHDRAYVMVDGVFQGVVERNMRDKLPLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKP
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFYSKTFPILGSVGDTFLYL
PGWTKGQVWINGFNLGRYWTKQGPQQTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLD
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

シグナル配列：

アミノ酸 1-27

N-グリコシル化部位

アミノ酸 97-101, 243-247, 276-280, 486-490, 625-629

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 4-8

カゼインキナーゼIIリン酸化部位

アミノ酸 148-152, 234-238, 327-331, 423-427, 469-473, 550-554, 603-607, 644-648

チロシンキナーゼリン酸化部位

アミノ酸 191-198

N-ミリストイル化部位

アミノ酸 131-137, 176-182, 188-194, 203-209, 223-229, 227-233, 231-237, 274-280, 296-300, 307-313, 447-453, 484-490

【図６７】

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC
ACCCACAATATGGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT
TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT
CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT
GCGTTCCTTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT
CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG
AAAAACTCAGGCAGCACATTTCACGCAACGCCCAGGACAAGCAGGAGTTGCATCTGTTCATG
CTGTCGGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAACT
AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC
TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA
TGCCTTCACGTGAAGTTCACTGATGTGGCTGAAATTCCTGCCTGGGTGTATTTGCTCAAAAA
CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG
AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT
CCCTCCAACATTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTCATAATGACGGCAC
TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTCGCTGAGCTGGAACTCCAGA
ACTGTGAGCTAGAGAGAATCCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT
TTAAAGTCCAATAACATTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT
GACTTGTTTAAAATTATGGCATAACAAAATTGTTACTATTCCTCCCTCTATTACCCATGTCA
AAAACTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT
AGTTTACAGAAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA
AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTCTGC
CAAAACAATTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC
TCACTCCCAGAGAAAGTTGGTCAGCTCTCCCAGCTCACTCAGCTGGAGCTGAAGGGGAACTG
CTTGGACCGCCTGCCAGCCCAGCTGGGCCAGTGTCGGATGCTCAAGAAAAGCGGGCTTGTTG
TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA
AATATTCCCTTTGCAAATGGGATTTAAACTAAGATAATATATGCACAGTGATGTGCAGGAAC
AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG
ATACATCTTTTAAAATAAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT
GTTCAATGTTTGTAGGGTTTTAAGTCATTCATTTCCAAATCATTTTTTTTTTCTTTTGGGG
AAAGGGAAGGAAAAATTATAATCACTAATCTTGGTTCTTTTTAAATTGTTTGTAACTTGGAT
GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT
TTCTTACTAAAAAAAAAAAAAAAA

【図６８】

MAYMLKKLLISYISIICVYGFICLYTLPWLFRIPLKEYSPEKVREESSFSDIPDVKNDFAFL
LHMVDQYDQLYSKRFGVFLSEVSENKLREISLNHEWTFEKLRQHISRNAQDKQELHLFMLSG
VPDAVFDLTDLDVLKLELIPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRDHLRCLH
VKFTDVAEIPAWVYLLKNLRELYLIGNLNSENNKMIGLESLRELRHLKILHVKSNLTKVPSN
ITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAELELQNCELERIPHAIFSLSNLQELDLKS
NNIRTIEEIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLESLYFSNNKLESLPVAVFSLQ
KLRCLDVSYNNISMIPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTLNLGQNCITSLP
EKVGQLSQLTQLELKGNCLDRLPAQLGQCRMLKKSGLVVEDHLFDTLPLEVKEALNQDINIP
FANGI

シグナル配列：

アミノ酸 1-20

N-グリコシル化部位

アミノ酸 241-245, 248-252, 383-387

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 326-330

カゼインキナーゼIIリン酸化部位

アミノ酸 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

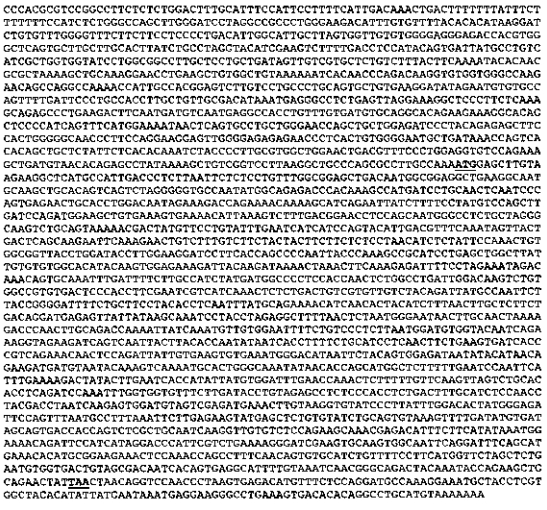
チロシンキナーゼリン酸化部位

アミノ酸 349-355, 375-381

N-ミリストイル化部位

アミノ酸 78-84, 124-130, 212-218, 392-398

【図６９】



【図７０】

MELVRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI
ERPENKSIRIIFSYVQLDPDGSCESENIKVFDGTSSNGPLLGQVCSKNDYVPVFESSSSTLT
FQIVTDSARIQRTVPVFYYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYCVWHIQV
EKDYKIKLNFKEIFLEIDKQCKPDFLAIYDGPSTNSGLIGQVCGRVTPTFESSSNSLTVVLS
TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIISKSYLEAFNSNGNNLQLKDPTCRP
KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNST
VEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN
LVVFLDTCRASPTSDPASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYL
QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDSIIGPIRLKRDRSASGNSGFQHETHA
EETPNQPFNSVHLFSPMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

シグナル配列：

アミノ酸 1-24

膜貫通ドメイン：

アミノ酸 571-586

N-グリコシル化部位

アミノ酸 29-33, 57-61, 67-71, 148-152, 271-275, 370-374, 394-398, 419-423

カゼインキナーゼIIリン酸化部位

アミノ酸 22-26, 108-112, 289-293, 348-352, 371-375, 379-383, 408-412, 463-467, 520-524, 556-560

チロシンキナーゼリン酸化部位

アミノ酸 172-180, 407-415, 407-416, 519-528

N-ミリストイル化部位

アミノ酸 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

真核生物膜リポタンパク質脂質付着部位

アミノ酸 7-18

【図７１】

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG
GGAC**ATG**CGGCCCCAGGAGCTCCCCAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGTTGCTGC
TGCTGCCGCCGCCGCCGTGCCCTGCCCACAGCGCCACGCGCTTCGACCCCACCTGGGAGTCC
CTGGACGCCCGCCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG
GGGAGTGTTTTCCGTGCCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCAAAAGGAAA
AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCCTCCTAGTTTCAAATATGAAGAT
TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAGTGGGCAGATATTTTTCAGGC
CTCTGGTGCCAAATACATTGTCTTAACTTCCAAACATCATGAAGGCTTTACCTTGTGGGGGT
CAGAATATTCGTGGAACTGGAATGCCATAGATGAGGGGCCCAAGAGGGACATTGTCAAGGAA
CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTGGACTGTACTATTCCCTTTTTGA
ATGGTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG
TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGG
TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT
ATATAATGAAAGCCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA
TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA
CATAAATGGGAAAACTGCATGACAATAGACAAACTGTCCTGGGGCTATAGGAGGGAAGCTGG
AATCTCTGACTATCTTACAATTGAACAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG
GAGGAAATCTTTTGATGAATATTGGGCCCACACTAGATGGCACCATTTCTGTAGTTTTTGAG
GAGCGACTGAGGCAAGTGGGGTCCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA
TACCTGGCGATCCCAGAATGACACTGTCACCCCAGATGTGTGGTACACATCCAAGCCTAAAG
AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCCACATCAGGACAGCTGTTCCTTGGCCAT
CCCAAAGCTATTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAACTG
GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTCATCAGATGC
CGTGTAAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATC**TAA**AGTGCAGCAGAGTGGCTG
ATGCTGCAAGTTATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGA
AAGCAATG**TAAA**CTGGATAAGAAAATTATTTGGCAGTTCAGCCCTTTCCCTTTTTCCCACTA
AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCCATTAAAGTC
TCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTCACATTATAGTAG
CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATG
ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAAT
TTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACA
CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG
AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG
GTACAGTAAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG
CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCTAGGACATTA
TTGAACACTGCCAGACGTTATAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAA
GTTTTTCTTTCTTCAATTATAAATTAACATAAGTGTACTGTAACTTTACAAACGTTTTAATT
TTTAAAACCTTTTTGGCTCTTTTGTAATAACACTTAGCTTAAAACATAAACTCATTGTGCAA
ATGTAA

【図７２】

MRPQELPRLAFPLLLLLLLLPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFIHWG
VFSVPSFGSEWFWWYWQKEKIPKYVEFMKDNYPPSFKYEDFGPLFTAKFFNANQWADIFQAS
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLFEW
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKWGWALALTNVI

シグナル配列：

アミノ酸 1-28

N-グリコシル化部位

アミノ酸 171-175, 239-243, 377-381

カゼインキナーゼIIリン酸化部位

アミノ酸 32-36, 182-186, 209-213, 227-231, 276-280, 315-319, 375-375

チロシンキナーゼリン酸化部位

アミノ酸 361-369, 389-397

N-ミリストイル化部位

アミノ酸 143-149, 178-184, 255-261, 272-278, 428-434

ロイシンジッパーパターン

アミノ酸 410-432

α-L-フコシダーゼ推定活性部位

アミノ酸 283-295

【図７３】

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT
TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC
TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACG**ATG**GCCAGGTGCTTCAGCCTGGTGTTG
CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT
TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGC
AGCTGAATTTCACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG
GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTTGGAGA
TGGATTCGTGGTCATCTCTAGGATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTG
TCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACTCATCTGAT
ACTTGGACTAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA
AACTGCAACACAAACAACAGAATTTATTGTCAGTGACAGTACCTACTCGGTGGCATCCCCTT
ACTCTACAATACCTGCCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG
AGAAAAAAATTGATTTGTGTCACAGAAGTTTTTATGGAAACTAGCACCATGTCTACAGAAAC
TGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCC
CCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGCTGCAGCTGGTCTTGGATTTTGC
TATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAATCAGCAGAAGGAAATGAT
CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA
AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCGATGCCTGGAA
GCTGAAGTT**TAG**ATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTTCATGCTCC
TTACCCTGCCCCAGCTGGGGAAATCAAAAGGGCCAAAGAACCAAAGAAGAAAGTCCACCCTT
GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC
CCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCC
TTTCTAGCCTGGCTATGTCCTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAA
GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG
TGGGTTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGAC
CCTTTCTTCAGCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTA
AGAGCAAAAGAATGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG
ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT
GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA
ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT
AGGAAATATACTTTTACAAGTAACAAAAATAAAAACTCTTATAAATTTCTATTTTTATCTGA
GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA
ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT
TGAATATTATTCCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTTTCA
GTTTTGATATTTCTAGCTTATCTACTTCCAAACTAATTTTTATTTTTGCTGAGACTAATCTT
ATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG
TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAACAAATGTATCACTA
GCCCTCCTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA
AGCATTTAGAAAACTT

【図７４】

MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACR
LLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNFKCGKNGVGVLIWKVPVSRQF
AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP
APASTSIPRRKKLICVTEVFMETSTMSTETEPPVENKAAFKNEAAGFGGVPTALLVLALLFF
GAAAGLGPCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSP
SKTTVRCLEAEV

シグナル配列：

アミノ酸 1-16

膜貫通ドメイン：

アミノ酸 235-254

N-グリコシル化部位

アミノ酸 53-57, 130-134, 289-293

カゼインキナーゼIIリン酸化部位

アミノ酸 145-149, 214-218

チロシンキナーゼリン酸化部位

アミノ酸 79-88

N-ミリストイル化部位

アミノ酸 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

【図７５】

AG**ATG**GCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTCGGTGCCGCGACTTTCACGATGG
CTCGCCCAACCTTACTACCTTCTGTCGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAA
ACTGCCGCCGCTCTGCCACGGTCTGCCCACCCAACGCGAAGACGGTAACCCGTGTGACTTTG
ACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA
TCCATCACTGTGGAGCAACATATAGGCAACATTTCATGTTTAGTAAAGTGGCCAACACAAT
TCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC
TGATGACGTGCAAACCCCCCCCTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAA
ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA
TTGGTCTAATGACTGCCAATCATTTGCCCCTATCTATGCTGACCTCTCCCTTAAATACAACT
GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC
AAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAA
GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGGCTGTCTCATGGACCTTCTCTG
AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAAACTATCA
AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTCAGA
TGGGGAAAACAAGAAGGATAAA**TAA**GATCCTCACTTTGGCAGTGCTTCCTCTCCTGTCAATT
CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGG
CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAG
GCACCCTACAGGAAGGCCTGCCATGCTGTGGCCAACTGTTTCACTGGAGCAAGAAAGAGATC
TCATAGGACGGAGGGGGAAATGGTTTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATC
AGCTATTCAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCCTTCTTAG
TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAA
CGCTAAGAATTTTCCCCCAAGGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTC
TTCATTAAAAGTATAAGCCTAACTTTGTCGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG
TTTTTATCATTGAAGACAATATTGAACAACCCCCTATTTTGTGGGGATTGAGAAGGGGTGAA
TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCAC
TAACCCTCTGACATACTCCCCACACCCAGTTGATGGCTTTCCGTAATAAAAAGATTGGGATT
TCCTTTTG

【図７６】

MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAPLLVRKLPPLCHGLPTQREDGNPCDFD
WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC1VFL
MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNC
TGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE
ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

シグナル配列：

アミノ酸 1-48

膜貫通ドメイン：

アミノ酸 111-125

N-グリコシル化部位

アミノ酸 165-169, 185-189

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 154-158, 265-269

カゼインキナーゼIIリン酸化部位

アミノ酸 51-55, 145-149, 245-249, 286-290, 288-292

N-ミリストイル化部位

アミノ酸 188-194, 225-231

Myb DNA-結合ドメイン反復シグネイチャー 1

アミノ酸 244-253

【図７７】

GGACAGCTCGCGGCCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGACGTTTGCCCTG
GGGCCCCAGCCTGGCCCGGGTCACCCTGGCATGAGGAGATGGCCTGTTGCTCCTGGTCCCA
TTGCTCCTGCTGCCCGGCTCCTACGGACTGCCCTTCTACAACGGCTTCTACTACTCCAACAG
CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG
TGGTGGAGACACCCGAGGAGACCCTGTTCACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC
CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAGCT
GTCGGAGAACGGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT
TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG
ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCATTGACGGGCTGGA
GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCGGGGTGTGGTCTTTCCTTACCAGTCCCCCA
ACGGGCGCTACCAGTTCAACTTCCACGAGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG
GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCTGGGAGGAGGGCCTGGACTGGTGCAACGCGGG
CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCCGGCAGCCCTGCGGTGGCC
CAGGCCTGGCACCTGGCGTGCGAAGCTACGGCCCCCGCCACCGCCGCCTGCACCGCTATGAT
GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCCTGAGAAGCT
GACGCTGACAGAGGCAAGGGAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC
AGCTCTTTGCCGCCTGGAAGTTCCATGGCCTGGACCGCTGCGACGCTGGCTGGCTGGCAGAT
GGCAGCGTCCGCTACCCTGTGGTTCACCCGCATCCTAACTGTGGGCCCCCAGAGCCTGGGGT
CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTGTACGGTGTTTACTGCTACCGCCAGC
ACTAGGACCTGGGGCCCTCCCCTGCCGCATTCCCTCACTGGCTGTGTATTTATTGAGTGGTT
CGTTTTCCCTTGTGGGTTGGAGCCATTTTAACTGTTTTTATACTTCTCAATTTAAATTTTCT
TTAAACATTTTTTTACTATTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG
GATGCCCCACTCCAGGAATCATGCTTGCTCCCCTGGGCCATTTGCGGTTTTGTGGGCTTCTG
GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTCCCTTAAGGAGGTTGGTGCCCAGAGTGGGC
GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCCT
CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGGAGCTGGGCTTTGGGCCTCTCCTGCC
CACCTCTACTTCTCTGTGAAGCCGCTGACCCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA
GCCAGTTCTAGGCTTCCAGGCGAAATCTGAGGGAAGGAAGAAACTCCCCTCCCCGTTCCCCT
TCCCCTCTCGGTTCCAAAGAATCTGTTTTGTTGTCATTTGTTTCTCCTGTTTCCCTGTGTGG
GGAGGGGCCCTCAGGTGTGTGTACTTTGGACAATAAATGGTGCTATGACTGCCTTCCGCCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
A

【図７８】

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNGHGKDLLNGVKLVVETPEETLFTYQ
GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQD
KEHDVSLEIQDLRLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ
VCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR
HRRLHRYDVFCFATALKGRVYYLEHPEKLTLTEAREACQEDDATIAKVGQLFAAWKFHGLDR
CDAGWLADGSVRYPVVHPHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

シグナル配列：

アミノ酸 1-17

カゼインキナーゼIIリン酸化部位

アミノ酸 29-33, 53-57, 111-115, 278-282

チロシンキナーゼリン酸化部位

アミノ酸 137-145

N-ミリストイル化部位

アミノ酸 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

【図７９】

GGAGAGCGGAGCGAAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC
TTCTGTTGCTACTGAGGCACGGGGCCCAGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC
CAGGGGAGGGTGCACCAGGCGGCCCCCCTGAGCGACGCTCCCCATGATGACGCCCACGGGAA
CTTCCAGTACGACCATGAGGCTTTCCTGGGACGGGAAGTGGCCAAGGAATTCGACCAACTCA
CCCCAGAGGAAAGCCAGGCCCGTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGAC
GGCGACGGCTGGGTGTCGCTGGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCA
CATACGGGACTCGGTGAGCGCGGCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGG
GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCGGTGAAGAATTTCATGAC
GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGGCGTTTCCGGGTGGC
CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGAGCTGACAGCCTTCCTGCACCCCGAGG
AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAAA
GATGGCTATGTCCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGA
GGAGCCGGCGTGGGTGCAGACGGAGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGG
ATGGGCACCTGGATGGGAGTGAGGTGGGCCACTGGGTGCTGCCCCCTGCCCAGGACCAGCCC
CTGGTGGAAGCCAACCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA
AGCGGAAATCCTGGGTAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGG
ACCTGACCCGGCACCACGATGAGCTGTGAGCACCGCGCACCTGCCACAGCCTCAGAGGCCCG
CACAATGACCGGAGGAGGGGCCGCTGTGGTCTGGCCCCCTCCCTGTCCAGGCCCCGCAGGAG
GCAGATGCAGTCCCAGGCATCCTCCTGCCCCTGGGCTCTCAGGGACCCCCTGGGTCGGCTTC
TGTCCCTGTCACACCCCAACCCCAGGGAGGGGCTGTCATAGTCCCAGAGGATAAGCAATAC
CTATTTCTGACTGAGTCTCCCAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTCAGCTCTAA
GAACCGCCCCAACCCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCCT
GGCCCCAGCCCTCTCCTGCCTGGCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCAATAA
AAGCCAGCGCCGGGACCTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

【図８０】

MMWRPSVLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDAPHDDAHGNFQYDHEAFLGR
EVAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT
YDTDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGDSMATRE
ELTAFLHPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPAWVQTERQQ
FRDFRDLNKDGHLDGSEVGHWVLPPAQDQPLVEANHLLHESDTDKDGRLSKAEILGNWNMFV
GSQATNYGEDLTRHHDEL

シグナル配列：

アミノ酸 1-20

N-グリコシル化部位

アミノ酸 140-144

カゼインキナーゼIIリン酸化部位

アミノ酸 72-76, 98-102, 127-131, 184-188, 208-212, 289-293, 291-295, 298-302

N-ミリストイル化部位

アミノ酸 263-269, 311-317

内質網状組織標的配列

アミノ酸 325-330

【図８１】

GGGGCCTTGCCTTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCCGAGCCCCGG
GCGGCGGGCGCGGGTGCGAGGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTCGCTCCCAG
CCTGTCTGTCGTCGTTTTGGCGCCCCCGCCTCCCCGCGGTGCGGGGTTGCACACCGATCCTG
GGCTTCGCTCGATTTGCCGCCGAGGCGCCTCCCAGACCTAGAGGGGCGCTGGCCTGGAGCAG
CGGGTCGTCTGTGTCCTCTCTCCTCTGCGCCGCGCCCGGGGATCCGAAGGGTGCGGGGCTCT
GAGGAGGTGACGCGCGGGGCCTCCCGCACCCTGGCCTTGCCCGCATTCTCCCTCTCTCCCAG
GTGTGAGCAGCCTATCAGTCACCATGTCCGCAGCCTGGATCCCGGCTCTCGGCCTCGGTGTG
TGTCTGCTGCTGCCGGGGCCCGCGGGCAGCGAGGGAGCCGCTCCCATTGCTATCACATG
TTTTACCAGAGGCTTGGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCCAGGGGGCTGCC
CTCTTGAGGAATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGG
GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC
TGGTCGAGAAAACTATTCCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGAT
GGTCTGCTTCTTTCACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA
GTGTCCACAGCACATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACTGG
CAATAAAGATTGTAAAGCAGACATTGCATTTCTGATTGATGGAAGCTTTAATATTGGGCAGC
GCCGATTTAATTTACAGAAGAATTTTGTTGGAAAAGTGGCTCTAATGTTGGGAATTGGAACA
GAAGGACCACATGTGGGCCTTGTTCAAGCCAGTGAACATCCCAAAATAGAATTTTACTTGAA
AAACTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGTA
ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCTGGA
GTAAGAAAAGGGATCCCCAAAGTGGTGGTGGTATTTATTGATGGTTGGCCTTCTGATGACAT
CGAGGAAGCAGGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCA
AGCCTATCCCTGAAGAACTGGGGATGGTTCAGGATGTCACATTTGTTGACAAGGCTGTCTGT
CGGAATAATGGCTTCTTCTCTTACCACATGCCCAACTGGTTTGGCACCACAAAATACGTAAA
GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT
CAGTGAACATTGCCTTTCTAATTGATGGCTCCAGCAGTGTTGGAGATAGCAATTTCCGCCTC
ATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGAT
AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCAGTTTCACTGACTATAGCACCA
AAGAGAATGTCCTAGCTGTCATCAGAAACATCCGCTATATGAGTGGTGGAACAGCTACTGGT
GATGCCATTTCCTTCACTGTTAGAAATGTGTTTGGCCCTATAAGGGAGAGCCCCAACAAGAA
CTTCCTAGTAATTGTCACAGATGGGCAGTCCTATGATGATGTCCAAGGCCCTGCAGCTGCTG
CACATGATGCAGGAATCACTATCTTCTCTGTTGGTGTGGCTTGGGCACCTCTGGATGACCTG
AAAGATATGGCTTCTAAACCGAAGGAGTCTCACGCTTTCTTCACAAGAGAGTTCACAGGATT
AGAACCAATTGTTTCTGATGTCATCAGAGGCATTTGTAGAGATTTCTTAGAATCCCAGCAAT
AATGGTAACATTTTGACAACTGAAAGAAAAAGTACAAGGGGATCCAGTGTGTAAATTGTATT
CTCATAATACTGAAATGCTTTAGCATACTAGAATCAGATACAAAACTATTAAGTATGTCAAC
AGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGT
ACTTTGTTAAAAACACTGCTGAGGCTTCATAATCATGGCTCTTAGAAACTCAGGAAAGAGGA
GATAATGTGGATTAAAACCTTAAGAGTTCTAACCATGCCTACTAAATGTACAGATATGCAAA
TTCCATAGCTCAATAAAAGAATCTGATACTTAGACCAAAAAAAAAAA

【図８２】

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPGGCPLEEFSVY
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASFTVT
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRRFNLQKN
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLKNFTSAKDVLFAIKEVGFRGGNSNTGKAL
KHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG
MVQDVTPVDKAVCRNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI
DGSSSVGDSNFRLMLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAHDAGITI
FSVGVAWAPLDDLKDMASKPKESHAFFTREFTGLEPIVSDVIRGICRDFLESQQ

シグナル配列：

アミノ酸 1-24

N-グリコシル化部位

アミノ酸 100-104, 221-225

カゼインキナーゼIIリン酸化部位

アミノ酸 102-106, 129-133, 224-228, 316-320, 377-381, 420-424, 425-429, 478-482, 528-532

N-ミリストイル化部位

アミノ酸 10-16, 23-29, 81-87, 135-141, 158-164, 205-211, 239-245, 240-246, 261-267, 403-409, 442-448, 443-449

アミド化部位

アミノ酸 145-149

【図８３】

CGCCGCGCTCCCGCACCCGCGGCCCGCCCACCGCGCCGCTCCCGCATCTGCACCCGCAGCCC
GGCGGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCG
GGGCGGCGGCTGCGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGC
TGCTGGCGGCGGCGGTCCCCACGGCCCCCGGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTC
AAGCCCGGCCCGGCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGA
GGTTGAGGAACTGATGGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGG
CAGAAGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCCAGCTAT
CACAATGAGACCAACACAGACACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAAT
TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTG
TGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGC
ATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCT
CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAA
TGGCCACCAGGGGCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTG
TGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA
GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATG
GAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG
GTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC
CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG
AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCGCT
GCACTGCTGGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA
ATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTACA
TCTTCTTCCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTCAGC
TCCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAAAC
TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAG
ACAGCCGTTTGTTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGCGAAACAATGTGG
AGTCTCCCTCTGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAA
CCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTG
TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTCACCCTGCATTACATGTGTTTATTCATCC
AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC
CCTCTCTCAGCACAGCCTGGGGAGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG
GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT
CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCCACACCAGCCTTGGTGCCACCAA
AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTTCTTGAGGCATGCACATCTGGAATTAAG
GTCAAACTAATTCTCACATCCCTCTAAAAGTAAACTACTGTTAGGAACAGCAGTGTTCTCAC
AGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGT
TGCATTAGTAACTTTGAAAGGTATATGACTGAGCGTAGCATACAGGTTAACCTGCAGAAACA
GTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAAC
TGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTT
GTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATGTTTTCAGGTGTCA
TGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTA
TAAGCATGCTTTCTTTGAGTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGC
ATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAAA

【図８４】

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQ
HKLRSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTG
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMLCTRDSECCG
DQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRL
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEEI

シグナル配列：

アミノ酸 1-19

N-グリコシル化部位

アミノ酸 96-100, 106-110, 121-125, 204-208

カゼインキナーゼIIリン酸化部位

アミノ酸 46-50, 67-71, 98-102, 135-139, 206-210, 312-316, 327-331

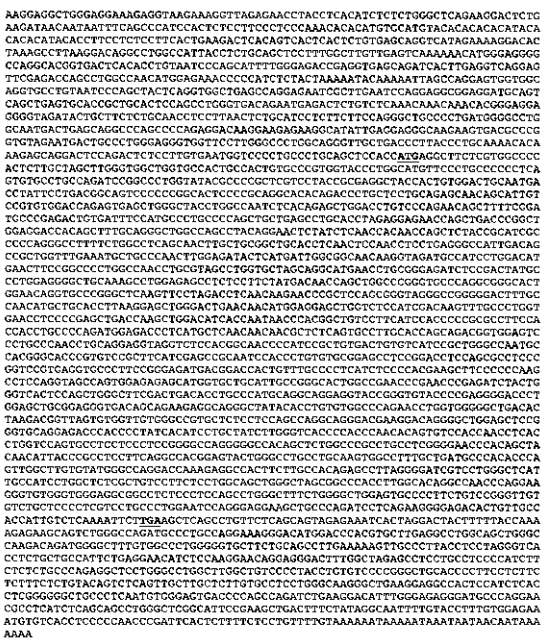
N-ミリストイル化部位

アミノ酸 202-208, 217-223

アミド化部位

アミノ酸 140-144

【図８５】



【図８６】

MRLLVAPLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLFLTA
VPPALPAGTQTLLLQSNSIVRVDQSELGYLANLTELDLSQNSFSDARDCDFHALPQLLSLHL
EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE
MLPNLEILMIGGNKVDAILDMNFRPLANLRSLVLAGMNLREISDYALEGLQSLESLSFYDNQ
LARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELVSIDKFALVNLP
ELTKLDITNNPRLSFIHPRAFHHLPQMETLMLNNNALSALHQQTVESLPNLQEVGLHGNPIR
CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVPFREMTDHCLPLISPRSFPPSLQ
VASGESMVLHCRALAEPPEPEIYWVTPAGLRLTPAHAGRRYRVYPEGTLELRRVTAEEAGLYT
CVAQNLVGADTKTVSVVVGRALLQPGRDEGQGLELRVQETHPYHILLSWVTPPNTVSTNLTW
SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS
CHRALGDRPGLIAILALAVLLLAAGLAAHLGTGQPRKGVGGRRPLPPAWAFWGWSAPSVRVV
SAPLVLPWNPGRKLPRSSEGETLLPPLSQNS

シグナル配列：

アミノ酸 1-18

膜貫通ドメイン：

アミノ酸 629-648

N-グリコシル化部位

アミノ酸 94-98, 381-385, 555-559, 583-587

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 485-489

カゼインキナーゼIIリン酸化部位

アミノ酸 46-50, 51-55, 96-100, 104-108, 130-134, 142-146, 243-247, 313-317, 488-492, 700-704

チロシンキナーゼリン酸化部位

アミノ酸 532-540

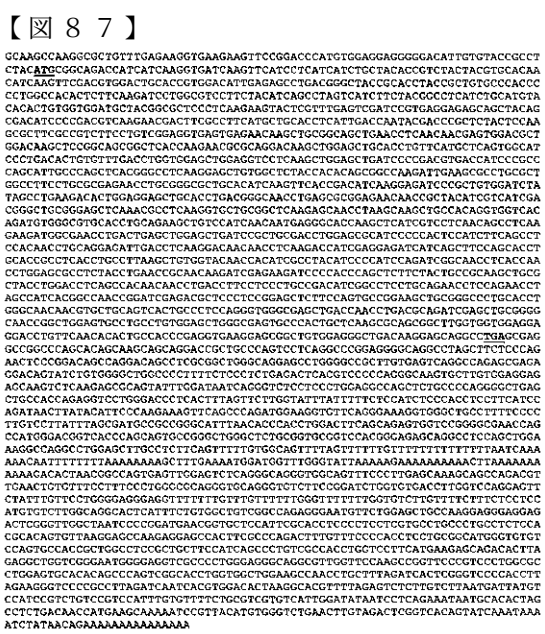
N-ミリストイル化部位

アミノ酸 15-21, 493-499, 566-572

アミド化部位

アミノ酸 470-474, 660-664, 692-696

【図８７】



【図８８】

MRQTIIKVIKFILIICYTVYYVHNIKFDVDCTVDIESLTGYRTYRCAHPLATLFKILASFYI
SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK
RFAVFLSEVSENKLRQLNLNNEWTLDKLRQRLTKNAQDKLELHLFMLSGIPDTVFDLVELEV
LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLRENLRALHIKFTDIKEIPLWI
YSLKTLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI
NNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKTIEEIISFQ
HLHRLTCLKLWYNHIAYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKLRYLDLSHNNLT
FLPADIGLLQNLQNLAITANRIETLPPELFQCRKLRALHLGNNVLQSLPSRVGELTNLTQIE
LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

膜貫通ドメイン：

アミノ酸 51-75 (type II)

N-グリコシル化部位

アミノ酸 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 85-89

カゼインキナーゼIIリン酸化部位

アミノ酸 91-95, 97-101, 177-181, 253-257, 330-334, 364-368, 398-402, 493-497

N-ミリストイル化部位

アミノ酸 173-179, 261-267, 395-401, 441-447

【図８９】

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCATGGAGCTGGCACTGCGGCGCTCTCCCGT
CCCGCGGTGGTTGCTGCTGCTGCCGCTGCTGCTGGGCCTGAACGCAGGAGCTGTCATTGACT
GGCCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG
TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCTGGTCAT
GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTGGAAACTTTGAGGAAATTGGGC
CCCTTGACAGTGATCTCAAACCACGGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT
GTGGATAATCCCGTGGGCACTGGGTTCAGTTATGTGAATGGTAGTGGTGCCTATGCCAAGGA
CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG
AATTCCAGACAGTTCCATTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGC
ATTGGTCTAGAGCTTTATAAGGCCATTCAGCGAGGGACCATCAAGTGCAACTTTGCGGGGGT
TGCCTTGGGTGATTCCTGGATCTCCCCTGTTGATTCGGTGCTCTCCTGGGGACCTTACCTGT
ACAGCATGTCTCTTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA
CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT
GATCATTGAACAGAACACAGATGGGGTGAACTTCTATAACATCTTAACTAAAAGCACTCCCA
CGTCTACAATGGAGTCGAGTCTAGAATTCACACAGAGCCACCTAGTTTGTCTTTGTCAGCGC
CACGTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAA
GCTCAAAATTATTCCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA
TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC
AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTCAGGAGGCCTG
GGTGCGGAAACTGAAGTGGCCAGAACTGCCTAAATTCAGTCAGCTGAAGTGGAAGGCCCTGT
ACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTC
TACTGGATTCTGAAAGCTGGTCATATGGTTCCTTCTGACCAAGGGGACATGGCTCTGAAGAT
GATGAGACTGGTGACTCAGCAAGAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT
TGGGGCACAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT
AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT
GGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAAACCTAAGATTTTTTAAAAAATTGAT
TTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

【図９０】

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC
KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSDLKPRKTTWLQAASLLFVDNPVGTGFSY
VNGSGAYAKDLAMVASDMMVLLKTFPSCHKEFQTVPFYIFSESYGGKMAAGIGLELYKAIQR
GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVNKGLYRE
ATELWGKAEMIIEQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQRDALS
QLMNGPIRKKLKIIPEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL
IVDTMGQEAWVRKLKWPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP
SDQGDMALKMMRLVTQQE

シグナル配列：

アミノ酸 1-25

N-グリコシル化部位

アミノ酸 64-68, 126-130, 362-366

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 101-105

カゼインキナーゼIIリン酸化部位

アミノ酸 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

N-ミリストイル化部位

アミノ酸 22-28, 76-82, 79-85, 80-86, 119-125, 169-175, 187-193, 195-201, 331-337, 332-338, 360-366

【図９１】

GGCCGCGGGAGAGGAGGCCATGGGCGCGCGCGGGGCGCTGCTGCTGGCGCTGCTGCTGGCTC
GGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGGCGGCGCCGTTATCAGGACCATGCGGCCGA
CGGGTCATCACGTCGCGCATCGTGGGTGGAGAGGACGCCGAACTCGGGCGTTGGCCGTGGCA
GGGGAGCCTGCGCCTGTGGGATTCCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGG
CACTCACGGCGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG
GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCG
TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCTATGACATTG
CCTTGGTGAAGCTGTCTGCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAG
GCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAA
AGAGGATGAGGCACTGCCATCTCCCCACACCCTCCAGGAAGTTCAGGTCGCCATCATAAACA
ACTCTATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG
GTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCTT
GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTGGGCTGTG
GTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAAG
CTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTCTTTTTCCCTCT
TCTCTGGGCTCTCCCACTCCTGGGGCCGGTCTGAGCCTACCTGAGCCCATGCAGCCTGGGGC
CACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTTTGGTAATAAACACATTCCAGTTGA
TGCCTTGCAGGGCATTCTTCAAAAAAAAAAAAAAAAAAAAAAAAAAA

【図９２】

MGARGALLLALLLARAGLRKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSLRLW
DSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI
YLSPRYLGNSPYDIALVKLSAPVTYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP
SPHTLQEVQVAIINNSMCNHLFLKYSFRKDIFGDMVCAGNAQGGKDACFGDSGGPLACNKNG
LWYQIGVVSWGVGCGRPNRPGVYTNISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL
LGPV

シグナル配列：

アミノ酸 1-18

N-グリコシル化部位

アミノ酸 167-171, 200-204, 273-277

カゼインキナーゼIIリン酸化部位

アミノ酸 86-90, 134-138, 161-165, 190-194, 291-295

N-ミリストイル化部位

アミノ酸 2-8, 44-50, 101-107, 225-231, 229-235, 239-245, 259-265, 269-275

アミド化部位

アミノ酸 33-37

真核生物膜リポタンパク質脂質付着部位

アミノ酸 252-263,

セリンプロテアーゼ,トリプシンファミリー,ヒスチジン活性部位

アミノ酸 78-84

【図９３】

CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGAATGGACTCCAAGCCTGCCTCCTAGGGCT
CTTTGCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGACCAGCGGAGGACGC
TGCCCCCAGGCTGGGTGTCCCTGGGCCGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTT
GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGCAGGCTGTGTCGGATCCCAG
CTCTCCTCAATACGGAAAATACCTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC
CACTGACCCTCCACACGGTGCAAAAATGGCTCTTGGCAGCCGGAGCCCAGAAGTGCCATTCT
GTGATCACACAGGACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCC
TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCCAC
ATCCCTACCAGCTTCCACAGGCCTTGGCCCCCCATGTGGACTTTGTGGGGGGACTGCACCGT
TTTCCCCCAACATCATCCCTGAGGCAACGTCCTGAGCCGCAGGTGACAGGGACTGTAGGCCT
GCATCTGGGGGTAACCCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAAGACGTGG
GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCCAGTTCCTGGAGCAGTATTTCCATGAC
TCAGACCTGGCTCAGTTCATGCGCCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGC
CCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCCGGGATTGAGGCCAGTCTAGATGTGCAGT
ACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG
GGACAGGAGCCCTTCCTGCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGT
GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA
ACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCCTGCTCTTCGCCTCAGGTGACAGT
GGGGCCGGGTGTTGGTCTGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAG
CCCCTATGTCACCACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCACAAATGAAA
TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCCACGGCCTTCATACCAGGAG
GAAGCTGTAACGAAGTTCCTGAGCTCTAGCCCCCACCTGCCACCATCCAGTTACTTCAATGC
CAGTGGCCGTGCCTACCCAGATGTGGCTGCACTTTCTGATGGCTACTGGGTGGTCAGCAACA
GAGTGCCCATTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTTTGGGGGGATCCTA
TCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCCTCTTGGCTTTCTCAACCCAAG
GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTC
TGGATGAAGAGGTAGAGGGCCAGGGTTTCTGCTCTGGTCCTGGCTGGGATCCTGTAACAGGC
TGGGGAACACCAACTTCCCAGCTTTGCTGAGACTCTACTCAACCCCTGACCCTTTCCTATC
AGGAGAGATGGCTTGTCCCCTGCCCTGAAGCTGGCAGTTCAGTCCCTTATTCTGCCCTGTTG
GAAGCCCTGCTGAACCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCTGAAA
TGCTGTGAGCTTGACTTGACTCCCAACCCTACCATGCTCCATCATACTCAGGTCTCCCTACT
CCTGCCTTAGATTCCTCAATAAGATGCTGTAACTAGCATTTTTTGAATGCCTCTCCCTCCGC
ATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAGACTCTGTGCACTA
TTTCACTTGATATTCATTCCCCAATTCACTGCAAGGAGACCTCTACTGTCACCGTTTACTCT
TTCCTACCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATG
GCCTTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACTTCTCTG
ACTACTCTTGTCTTCCTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGC
TCCATTTGTAGATTTTTGCTCTTCTCAGTTTACTCATTGTCCCCTGGAACAAATCACTGACA
TCTACAACCATTACCATCTCACTAAATAAGACTTTCTATCCAATAATGATTGATACCTCAAA
TGTAAAAAA

【図９４】

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQVGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWL
SIRQAELLLPGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGGLHRFPPTSSLRQRP
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG
RPPLGFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPTSQLC

シグナル配列：

アミノ酸 1-16

N-グリコシル化部位

アミノ酸 210-214, 222-226, 286-290, 313-317, 443-447

グリコサミノグリカン付着部位

アミノ酸 361-365, 408-412, 538-542

カゼインキナーゼIIリン酸化部位

アミノ酸 212-216, 324-328, 392-396, 420-424, 525-529

N-ミリストイル化部位

アミノ酸 2-8, 107-113, 195-201, 199-205, 217-223, 219-225, 248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488, 521-527, 533-539, 549-555

【図９５】

GCCGCGCGCTCTCTCCCGGCGCCCACACCTGTCTGAGCGGCGCAGCGAGCCCGCGGCCCGGGC
GGGCTGCTCGGCGCGGAACAGTGCTCGGCATGGCAGGGATTCCAGGGCTCCTCTTCCTTCTC
TTCTTTCTGCTCTGTGCTGTTGGGCAAGTGAGCCCTTACAGTGCCCCCTGGAAACCCACTTG
GCCTGCATACCGCCTCCCTGTCGTCTTGCCCCAGTCTACCCTCAATTTAGCCAAGCCAGACT
TTGGAGCCGAAGCCAAATTAGAAGTATCTTCTTCATGTGGACCCCAGTGTCATAAGGGAACT
CCACTGCCCACTTACGAAGAGGCCAAGCAATATCTGTCTTATGAAACGCTCTATGCCAATGG
CAGCCGCACAGAGACGCAGGTGGGCATCTACATCCTCAGCAGTAGTGGAGATGGGGCCCAAC
ACCGAGACTCAGGGTCTTCAGGAAAGTCTCGAAGGAAGCGGCAGATTTATGGCTATGACAGC
AGGTTCAGCATTTTTGGGAAGGACTTCCTGCTCAACTACCCTTTCTCAACATCAGTGAAGTT
ATCCACGGGCTGCACCGGCACCCTGGTGGCAGAGAAGCATGTCCTCACAGCTGCCCACTGCA
TACACGATGGAAAAACCTATGTGAAAGGAACCCAGAAGCTTCGAGTGGGCTTCCTAAAGCCC
AAGTTTAAAGATGGTGGTCGAGGGGCCAACGACTCCACTTCAGCCATGCCCGAGCAGATGAA
ATTTCAGTGGATCCGGGTGAAACGCACCCATGTGCCCAAGGGTTGGATCAAGGGCAATGCCA
ATGACATCGGCATGGATTATGATTATGCCCTCCTGGAACTCAAAAAGCCCCACAAGAGAAAA
TTTATGAAGATTGGGGTGAGCCCTCCTGCTAAGCAGCTGCCAGGGGGCAGAATTCACTTCTC
TGGTTATGACAATGACCGACCAGGCAATTTGGTGTATCGCTTCTGTGACGTCAAAGACGAGA
CCTATGACTTGCTCTACCAGCAATGCGATGCCCAGCCAGGGGCCAGCGGGTCTGGGGTCTAT
GTGAGGATGTGGAAGAGACAGCAGCAGAAGTGGGAGCGAAAAATTATTGGCATTTTTTCAGG
GCACCAGTGGGTGGACATGAATGGTTCCCCACAGGATTTCAACGTGGCTGTCAGAATCACTC
CTCTCAAATATGCCCAGATTTGCTATTGGATTAAAGGAAACTACCTGGATTGTAGGGAGGGG
TGACACAGTGTTCCCTCCTGGCAGCAATTAAGGGTCTTCATGTTCTTATTTTAGGAGAGGCC
AAATTGTTTTTTGTCATTGGCGTGCACACGTGTGTGTGTGTGTGTGTGTGTAAGGTGT
CTTATAATCTTTTACCTATTTCTTACAATTGCAAGATGACTGGCTTTACTATTTGAAAACTG
GTTTGTGTATCATATCATATATCATTTAAGCAGTTTGAAGGCATACTTTTGCATAGAAATAA
AAAAAATACTGATTTGGGGCAATGAGGAATATTTGACAATTAAGTTAATCTTCACGTTTTTG
CAAACTTTGATTTTTATTTCATCTGAACTTGTTTCAAAGATTTATATTAAATATTTGGCATA
CAAGAGATATGAAAAAAAAAAAAAA

【図９６】

MAGIPGLLFLLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS
SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIYILSSSGDGAQHRDSGSSGKS
RRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG
TQKLRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYA
LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD
AQPGASGSGVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW
IKGNYLDCREG

シグナル配列：

アミノ酸 1-19

N-グリコシル化部位

アミノ酸 93-97, 207-211

グリコサミノグリカン付着部位

アミノ酸 109-113, 316-320

カゼインキナーゼIIリン酸化部位

アミノ酸 77-81, 95-99, 108-112, 280-284, 351-355

N-ミリストイル化部位

アミノ酸 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

セリンプロテアーゼ,トリプシンファミリー, ヒスチジン活性部位

アミノ酸 171-177

【図９７】

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTGCTCCCCCGCCCCACCAGCCATGGTGGTTT
CTGGAGCGCCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG
GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA
GCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA
GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC
ACTGCTGCCCACTGTTTCAAGGACAACCTGAACAAACCATACCTGTTCTCTGTGCTGCTGGG
GGCCTGGCAGCTGGGGAACCCTGGCTCTCGGTCCCAGAAGGTGGGTGTTGCCTGGGTGGAGC
CCCACCCTGTGTATTCCTGGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCGTCTCGAG
CGCTCCATACAGTTCTCAGAGCGGGTCCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT
CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC
CCCACCCTCAGACCCTGCAGAAGCTGAAGGTTCCTATCATCGACTCGGAAGTCTGCAGCCAT
CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT
GGAGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG
GCGCCTGGCTGCTGGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCC
GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA
GCTCCGCGGGCGCGCTCAGGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG
CCGCGCGCTCCTAGGGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGGCGGCCAGATCCACA
TCTGGATCTGGATCTGCGGCGGCCTCGGGCGGTTTCCCCCGCCGTAAATAGGCTCATCTACC
TCTACCTCTGGGGGCCCGGACGGCTGCTGCGGAAAGGAAACCCCCTCCCCGACCCGCCCGAC
GGCCTCAGGCCCCCCTCCAAGGCATCAGGCCCCGCCCAACGGCCTCATGTCCCCGCCCCCAC
GACTTCCGGCCCCGCCCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTAT
AGGTATTTGTAACCCTGCCCACATATCTTATTTATTCCTCCAATTTCAATAAATTATTTATT
CTCCAAAAAAAAAA

【図９８】

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318
><subunit 1 of 1, 317 aa, 1 stop
><MW: 33732, pI: 7.90, NX(S/T): 1
MVVSGAPPALGGGCLGTFTSLLLLASTAILNAARIPVPPACGKPQQLNRVVGGEDSTDSEWP
WIVSIQKNGTHHCAGSLLTSRWVITAAHCFKDNLNKPYLPSVLLGAWQLGNPGSRSQKVGVA
WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPPNTHCWISGWGSIQDG
VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAQQGPITEDMLCAGYLEGERDACLGDSGGPLMC
QVDGAWLLAGIISWGEGCAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG
SGAAARS

シグナル配列：

アミノ酸 1-32

N-グリコシル化部位

アミノ酸 62-66, 96-100, 214-218, 382-386, 409-413, 455-459, 628-632, 669-673, 845-849, 927-931, 939-943, 956-960

グリコサミノグリカン付着部位

アミノ酸 826-830

カゼインキナーゼIIリン酸化部位

アミノ酸 17-21, 39-43, 120-124, 203-207, 254-258, 264-268, 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653, 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962, 1013-1017, 1019-1023, 1021-1025

チロシンキナーゼリン酸化部位

アミノ酸 607-615

N-ミリストイル化部位

アミノ酸 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

【図９９】

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTA
CTGCTGGTGGCCACCACAGGCCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGT
GGAGCTGCACAACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA
GATGGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCGTGTGGGGCCAC
AACAAGGAGCGCGGGCGCCGCGGCGAGAATCTGTTCGCCATCACAGACGAGGGCATGGACGT
GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT
GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC
GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACT
GGTGTGCAACTATGAGCCTCCGGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC
CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC
CCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCATCCTTCCGGGCGACTGA
AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT
TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAAACCCAGGCC
CCAACTTCCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC
AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA
CCTTCCCCAAATCGACCCATGTTCCTATCCCAAAATCAGCAGACAAAGTGACAGACAAAACA
AAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG
GGAACTCCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG
AGGTCTTGGCCTCAGTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC
CACACGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAA
TGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCGTCCTTGCCAGGTGCAGAGGGCCCTGACA
AGCCTAGCGTTGTGTCAGGGCTGAACTCGGGCCCTGGTCATGTGTGGGGCCCTCTCCTGGGA
CTACTGCTCCTGCCTCCTCTGGTGTTGGCTGGAATCTTCTGAATGGGATACCACTCAAAGGG
TGAAGAGGTCAGCTGTCCTCCTGTCATCTTCCCCACCCTGTCCCCAGCCCCTAAACAAGATA
CTTCTTGGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC
ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGCTCAGGAGGCCGCCTGAGGACTGCACACC
GGGCCCACACCTCTCCTGCCCCTCCCTCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCT
CACTGCCTACCTGGCCTGGGGCTGTCTGCCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG
TGTAGCTGGGGATGGGGATTCCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC
TTTGAGTGGGGGAGGCAGGGACGAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAAACCT
GTCCAACCTGTGAAA

【図１００】

MHGSCSFLMLLLPLLLLLVATTGPVGALTDEEKRLMVELHNLYRAQVSPTASDMLHMRWDEE
LAAFAKAYARQCVWGHNKERGRRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ
MCGHYTQVVWAKTERIGCGSHFCEKLQGVEETNIELLVCNYEPPGNVKGKRPYQEGTPCSQC
PSGYHCKNSLCEPIGSPEDAQDLPYLVTEAPSFRATEASDSRKMGTPSSLATGIPAFLVTEV
SGSLATKALPAVETQAPTSLATKDPPSMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKS
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEAELPPSSEVLAS
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV
SGLNSGPGHVWGPLLGLLLLPPLVLAGIF
シグナル配列：

アミノ酸 1-22

N-グリコシル化部位

アミノ酸 114-118, 403-407, 409-413

グリコサミノグリカン付着部位

アミノ酸 439-443

カゼインキナーゼIIリン酸化部位

アミノ酸 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-ミリストイル化部位

アミノ酸 123-129, 143-149, 152-158, 169-175, 180-186, 231-237, 250-256

アミド化部位

アミノ酸 82-86, 172-176

ベルオキシダーゼ近位ヘム-リグ及びシグネイチャー

アミノ酸 287-298

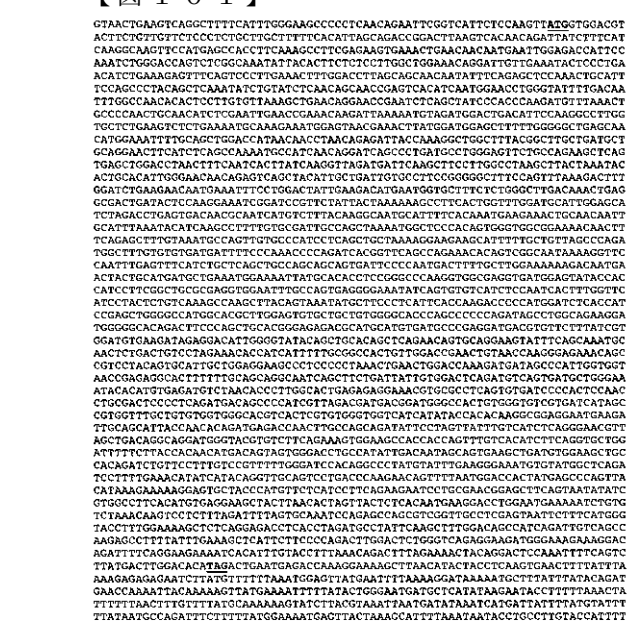
細胞外タンパク質 SCP/Tpx-1/Ag5/PR-1/Sc7 シグネイチャー 1

アミノ酸 127-138

細胞外タンパク質 SCP/Tpx-1/Ag5/PR-1/Sc7 シグネイチャー 2

アミノ酸 160-172

【図１０１】



【図１０２】

MVDVLLLFSLCLLFHISRPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSAN
ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEP
GYFDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKM
QRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWE
PCQKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNE
ISWTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ
MKKLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCD
DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQG
GEVMEYTTILRLREVEFASEGKYQCVISNHFPGSSYSVKAKLTVNMLPSPTKTPMDLTIRAGA
MARLECAAVGHPAPQIAWQKDGGTDFPAARERMHVMPEDDVFFIVDVKIEDIGVYSCTAQN
SAGSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTER
HFFAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDG
WATVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTLAD
RQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGP
MYLKGNVYGSDPPETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISW
PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDA
YSSFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNPQS
YDLDT

シグナル配列：

アミノ酸 1-19

膜貫通ドメイン：

アミノ酸 746-765

N-グリコシル化部位

アミノ酸 62-66, 96-100, 214-220, 382-386, 409-413, 455-459, 628-632, 669-673, 845-849, 927-931, 939-943, 956-960

グリコサミノグリカン付着部位

アミノ酸 826-830

カゼインキナーゼIIリン酸化部位

アミノ酸 17-21, 39-43, 120-124, 203-207, 254-258, 264-268, 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653, 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962, 1013-1017, 1019-1023, 1021-1025

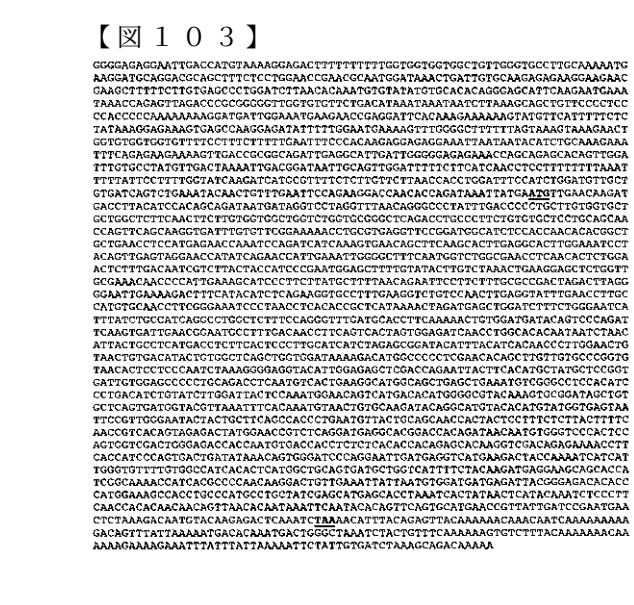
チロシンキナーゼリン酸化部位

アミノ酸 607-615

N-ミリストイル化部位

アミノ酸 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

【図１０３】



【図１０４】

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGISTNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSNSVGN
TTASATLNVTAATTTPPSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

シグナル配列：

アミノ酸 1-44

膜貫通ドメイン：

アミノ酸 523-543

N-グリコシル化部位

アミノ酸 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446, 488-492, 606-610

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 183-187

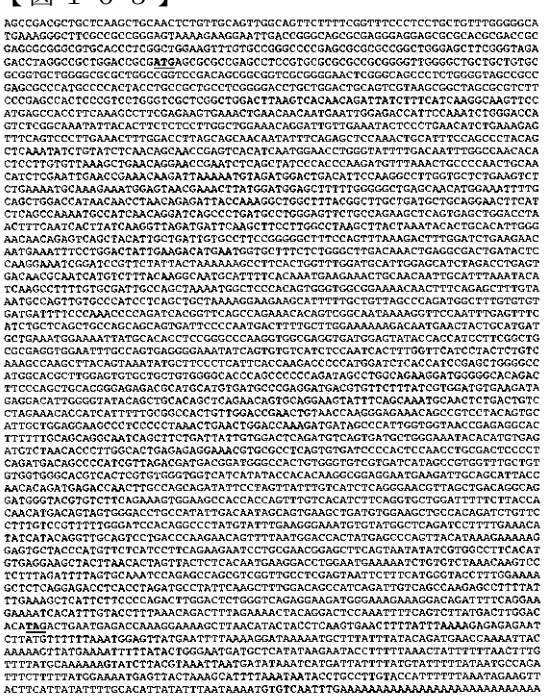
カゼインキナーゼIIリン酸化部位

アミノ酸 268-272, 417-421, 465-469, 579-583, 620-624

N-ミリストイル化部位

アミノ酸 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397, 422-428, 433-439, 531-537

【図１０５】



【図１０６】

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAAERPCPTTCRCLGDLLDCSR
KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT
LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY
FDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKMQR
NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWEFC
QKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEIS
WTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK
KLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAMPQLLKGRSIFAVSPDGFVCDDF
PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQGGE
VMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGAMA
RLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQNSA
GSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPFPKLNWTKDDSPLVVTERHF
FAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDGWA
TVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTLADRQ
DGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGPMY
LKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISWPS
HVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTPGKALRRPHLDAYS
SFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQSYD
LDT

シグナル配列：

アミノ酸 1-27

膜貫通ドメイン：

アミノ酸 808-828

N-グリコシル化部位

アミノ酸 122-126, 156-160, 274-278, 442-446, 469-473, 515-519, 688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

グリコサミノグリカン付着部位

アミノ酸 886-890

カゼインキナーゼIIリン酸化部位

アミノ酸 99-103, 180-184, 263-267, 314-318, 324-328, 374-378, 383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735, 799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022, 1073-1077, 1079-1083, 1081-1085

チロシンキナーゼリン酸化部位

アミノ酸 667-675

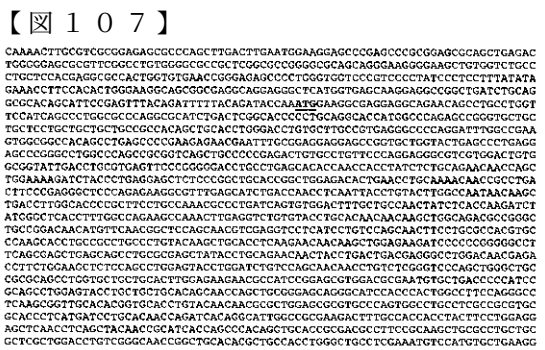
N-ミリストイル化部位

アミノ酸 14-20, 36-42, 239-245, 257-263, 380-386, 427-433, 513-519, 588-594, 672-678, 683-687, 774-780, 933-939

ロイシンジッパーパターン

アミノ酸 58-80, 65-87

【図１０７】



【図１０８】

MEGEEAEQPAWFHQPWRPGASDSAPPAGTMAQSRVLLLLLLLPPQLHLGPVLAVRAPGFGRS
GGHSLSPEENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCGGIDLREFPGDLP
EHTNHLSLQNNQLEKIYPEELSRLHRLETLNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNK
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNV
EVLILSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN
ETFWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLLHSN
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMILHNQITGIGREDFATTYF
LEELNLSYNRITSPQVHRDAFRKLRLLRSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALA
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEPGDISKD
RGRLGKEKEEEEEEEEEEETR

シグナル配列：

アミノ酸 1-48

N-グリコシル化部位

アミノ酸 243-247, 310-314, 328-332, 439-443

カゼインキナーゼIIリン酸化部位

アミノ酸 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-ミリストイル化部位

アミノ酸 19-25, 107-113, 213-219, 217-223, 236-242, 335-341, 477-483, 498-502, 539-545, 548-554

ロイシンジッパーパターン

アミノ酸 116-138, 251-273, 258-280, 322-344, 464-486, 471-493, 535-557

【図１０９】

GGGAGGGGGCTCCGGGGCGCCGCCGCAGCAGACCTGCTCCGGCCGGCGCCTCGCCGCTGTCCTGCGGGAGCGGCAG
CAGTAGCCCGGGCGGCGAGGGCTGGGGGTTCCTCGAGACTCTCAGAGGGGCCCTCCCATCGGGGCCCACCACCC
CAACCTGTTCCTCGGCGCGCCACTGCGCTGCGCCCCAGGACCCGCTGCCCAACATGGATTTTCTCCTGGCGCTGGT
GCTGGTATCCTCGCTCTACCTGCAGGCGGCCGCCGAGTTCGACGGGAGGTGGCCCAGGCAAATAGTGTCATCGAT
TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTCAGCC
TGTGTGCCAACCACGATGCAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTCATCCTGGTTATGCTGG
AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA
CGGCAGCTACAAGTGCTACTGTCTCAACGGATATATGCTCATGCCGGATGGTTCCTGCTCAAGTGCCCTGACCTG
CTCCATGGCAAACTGTCAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCTGGCCT
GCACCTGGCTCCTGATGGCAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT
TAGGCAATGTGTCAACACTTTGGGAGCTACATCTGCAAGTGTCATAAAGGCTTCGATCTCATGTATATTGGAGG
CAAATATCAATGTCATGACATAGACGAATGCTCACTTGGTCAGTATCAGTCCAGCAGCTTTGCTCGATGTTATAA
CGTACGTGGGTCCTACAACTGCAAATGTAAACAACGGATACCAGCGTGATGCACTGACTTGTCTGTATATCCCAAA
AGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAAGGGAAATGGTACCATTTAAAGGGTGACACAGGAAA
TAATAATTGGATTCCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAA
CAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAAGCCAACACCAATTCCTACTCCACCACCACCACCACC
CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCCAGAAAGGCCAACCACCGGACTGACAACTATAGC
ACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCCTCAGAAACCCAGAGG
AGATGTGTTCAGTGTTCTGGTACACAGTTGTAATTTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGACAA
TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTCGGCAGCCAAAGCCCCAGG
GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCCTCATGCATTCAGGGGACCTGTGCCTGTCATTCAGGCA
CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAACACGGTGCCACGAGGAGCCCTGTG
GGCAAGAAATGCTGGCCATCGCTGGAGCCAAACAGACATCACCTGCGAGGGGTGACATCAAGAGCGAATCACA
AACATGATTAAAGGGTTGGAAAAAAAGATCTATGATGGAAAATTAAAGGAACTGGGATTATTGAGCCTGGAGAAG
AGAAGACTGAGGGGCAAACCATTGATCGTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTG
TTCTCCATATGCACTAAGAATAGAACAACAGGAAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTTGGCAGG
GGCCATTGTTAGAATACTTCATAAAAAAAGAAGTGTGAAAATCTCAGTATCTCTCTCTTTCTAAAAATTAGA
TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACAAATTATAGAATTTCCCAAA
AGATGTTTTGATCCTACTAGTAGTATCCAGTGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG
CATTTCCCTCTTGACCTCCTAATGGAGAGGGATTGAAAGGGGAAGAGCCCACCAAATGCTGAGCTCACTGAAATA
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAAGAAAAAAGGAAACTATTTATTCCAAATGAGAGTATGAGAC
AGATATTTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC
TTTCATTTGTTCAATGGATGATGTTTCACATTTTTTTTTTTAAGAGATCCTTCAAGGAACACAGTTCAGAGAG
ATTTTCATCGGGTGCATTCTCTTGCTTCGTGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC
ACACCGGCAGACCTTTCCTTCACCTCATCAGTATGATTCAGTTTCTCTTATCAATTGGACTCTCCCAGGTTCCAC
AGAACAGTAATATTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCATTTAACCTGGTAAAGGCAGGGCTGG
AGGGGGAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTTAATGGTTCATT
TCCTTTATGGTCATATAACTGCACAGCTCAAGATGAAAGGGGAAAATAAATGAAAATTTTACTTTCGATGCCAA
TGATACATTGCACTAAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTTTATTATTTAATGTTTT
CTAAAATAAAAAATGTTAGTGGTTTTCCAAATGCCTAATAAAAACAATTATTTGTAAATAAAAACACTGTTAGT
AAT

【図１１０】

MDFLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQP
RCKHGECIGPNKCKCHFGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD
GSCSSALTCSMANCQYGCDVVKGQIRCQCPSPGLHLAPDGRTCVDVDECATGRASCPRFRQC
VNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQG
DGLTCVYIPKVMIEPSGPIHVPKGNGTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRP
TSKPTTRPTPKPTPIPTPPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDN
RVQTDPQKPRGDVPSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGG
KAARLVLPLGRLMHSGDLCLSFRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQI
TLRGADIKSESQR

シグナル配列：

アミノ酸 1-17

N-グリコシル化部位

アミノ酸 273-277

カゼインキナーゼIIリン酸化部位

アミノ酸 166-170, 345-349

チロシンキナーゼリン酸化部位

アミノ酸 199-206

N-ミリストイル化部位

アミノ酸 109-115, 125-131, 147-153, 191-197, 221-227, 236-242, 421-427, 433-439, 462-468, 476-482

アスパラギン酸及びアスパラギンヒドロキシル化部位

アミノ酸 104-116, 186-198, 231-243

細胞付着配列

アミノ酸 382-385

EGF-様ドメインシステインパターンシグネチャー

アミノ酸 75-87

【図１１１】

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCCTTTAGATTGTGA
AATGTGGCTCAAGGTCTTCACAACTTTCCTTTCCTTTGCAACAGGTGCTTGCTCGGGGCTGA
AGGTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTCAGGCCCTCTACCTACCCGTC
CACTATGGCTTCCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA
CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAAATACC
AACACAAGTTCACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCT
GATGAAGGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACTCTATCTGCCAGTCA
GAAGATACAAGTCACGGTTGATGATCCTGTCACAAAGCCAGTGGTGCAGATTCATCCTCCCT
CTGGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAGGGGGCACTCGG
CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC
TCCCCAAAACAATACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT
GCCTGGTGAGGAACCCTGTCAGTGAAATGGAAAGTGATATCATTATGCCCATCATATATTAT
GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT
TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCCAACACCTACT
CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTT
GCATCTGAGAAAGTAGCCCAGAAGACAATGGACTATGTGTGCTGTTGCTTACAACAACATAAC
CGGCAGGCAAGATGAAACTCATTTCACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG
CACAGAAAGGAAAATCATTGTCACCTTTAGCAAGTATAACTGGAATATCACTATTTTTGATT
ATATCCATGTGTCTTCTCTTCCTATGGAAAAAATATCAACCCTACAAAGTTATAAAACAGAA
ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTCAGGCCATGAAGATG
CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG
ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT
GTATGAAGTTATTCAGCACATCCCTGCCCAGCAGCAAGACCATCCAGAGTGAACTTTCATGG
GCTAAACAGTACATTCGAGTGAAATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGT
ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA
TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT
GTGCAACAGAAAAACATGTTGGGGAAATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGG
GGAGAACGAAAGTGACAGGGGTTTCCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA
ATTAGTTCTACTCTACACTTTCACTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA
TGTGGAAACTTTACATTGTTCGATTTTTCAGCAGACTTTGTTTTATTAAATTTTTATTAGTG
TTAAGAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA
CAAAGTAATAAGGATGGTTGTCACAAAAACAAAACTATGCCTTCTCTTTTTTTTCAATCACC
AGTAGTATTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTA
TTTTTTTCAAGGAAAGATGGATTCAAATAAATTATTCTGTTTTTGCTTTTAAAAAAAAAAA
AA

【図１１２】

MWLKVFTTPLSPATGACSGLKVTVPSHTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH
TMPKYLLGSVNKSVVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQGNGTLSASQ
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFS
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTV
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMDYVCCAYNNIT
GRQDETHFTVIITSVGLEKLAQKGKSLSPLASITGISLFLIISMCLLFLWKKYQPYKVIKQK
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAPPDVSGVSRIPSRSVPASDCVSGQDLHSTV
YEVIQHIPAQQQDHPE

シグナル配列：

アミノ酸 1-18

膜貫通ドメイン：

アミノ酸 341-359

N-グリコシル化部位

アミノ酸 73-77, 92-96, 117-121, 153-157, 189-193, 204-208, 276-280, 308-312

カゼインキナーゼIIリン酸化部位

アミノ酸 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

チロシンキナーゼリン酸化部位

アミノ酸 272-280

N-ミリストイル化部位

アミノ酸 15-21, 19-25, 118-124, 163-167, 203-209, 231-237, 239-245

真核生物膜リポタンパク質脂質付着部位

アミノ酸 7-18

【図１１３】

GCAAGCGGCGAA**ATG**GCGCCCTCCGGGAGTCTTGCAGTTCCCCTGGCAGTCCTGGTGCTGTT
GCTTTGGGGTGCTCCCTGGACGCACGGGCGGCGGAGCAACGTTCGCGTCATCACGGACGAGA
ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCCTGCT
TGTCAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAA
TATTGCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTC
TTCCTACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG
AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTCATCATG
GTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATGTGGA
TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT
GTTTTTGCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC
AGATTGCCTTTGTCCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCCTTCAAAAAAAT
TATTATCAGAATCTGCACAACC**TTT**GAAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAA
GATGTTTCAGAAGAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGACTTTCCACAGAATGC
CATAAGACAACGCTCTCTGGGTCCATCATTGGCCACAGATAAATCC**TAG**TTAAATTTTATAG
TTATCTTAATATTATGATTTTGATAAAAACAGAAGATTGATCATTTTGTTTGGTTTGAAGTG
AACTGTGACTTTTTTGAATATTGCAGGGTTCAGTCTAGATTGTCATTAAATTGAAGAGTCTA
CATTCAGAACATAAAAGCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATG
GTTTAAATAGTTCTCTAATTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGT
TTAATAATAACCTATTTCAAGTCTGAGTTTTGAAAATTTACATTTCCCAAGTATTGCATTAT
TGAGGTATTTAAGAAGATTATTTTAGAGAAAAATATTTCTCATTTGATATAATTTTTCTCTG
TTTCACTGTGTGAAAAAAAGAAGATATTTCCCATAAATGGGAAGTTTGCCCATTGTCTCAAG
AAATGTGTATTTCAGTGACAATTTCGTGGTCTTTTAGAGGTATATTCCAAAATTTCCTTGT
ATTTTTAGGTTATGCAACTAATAAAAACTACCTTACATTAATTAATTACAGTTTTCTACACA
TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTTAAGTTCATGGTATTCTCTTGATTC
CAACAAAGTTTGATTTTCTCTTGTATTTTTCTTACTTACTATGGGTTACATTTTTTATTTTT
CAAATTGGATGATAATTTCTTGGAAACATTTTTTATGTTTTAGTAAACAGTATTTTTTTGTT
GTTTCAAACTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAAATT
TTGGCCACTTTTTTCAGATTTTACATCATTCTTGCTGAACTTCAACTTGAAATTGTTTTTTT
TTTCTTTTTGGATGTGAAGGTGAACATTCCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTA
TTTTACATTTTGAAAATTCAAAGAAGCTTAATATAAAAGTTTGCATTCTACTCAGGAAAAAG
CATCTTCTTGTATATGTCTTAAATGTATTTTTGTCCTCATATACAGAAAGTTCTTAATTGAT
TTTACAGTCTGTAATGCTTGATGTTTTAAAATAATAACATTTTTATATTTTTTAAAAGACAA
ACTTCATATTATCCTGTGTTCTTTCCTGACTGGTAATATTGTGTGGGATTTCACAGGTAAAA
GTCAGTAGGATGGAACATTTTAGTGTATTTTTACTCCTTAAAGAGCTAGAATACATAGTTTT
CACCTTAAAAGAAGGGGGAAAATCATAAATACAATGAATCAACTGACCATTACGTAGTAGAC
AATTTCTGTAATGTCCCCTTCTTTCTAGGCTCTGTTGCTGTGTGAATCCATTAGATTTACAG
TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAAATATTGTACCATT
AAAGAGTTTGGATGTGTAACTTGTGATGCCTTAGAAAAATATCCTAAGCACAAAATAAACCT
TTCTAACCACTTCATTAAAGCTGAAAAAAAAAAAAAAAAAA

【図１１４】

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL
QPEWESFAEWGEDLEVNIAKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI
NFISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL
ATLFSGLLLGLCMIFVADCLCPSKRRRPQPYPYPSKKLLSESAQPLKKVEEEQEADEEDVSE
EEAESKEGTNKDFFQNAIRQRSLGPSLATDKS

シグナル配列：

アミノ酸 1-26

膜貫通ドメイン：

アミノ酸 182-201

カゼインキナーゼIIリン酸化部位

アミノ酸 68-72, 119-123, 128-132, 247-251, 257-261

チロシンキナーゼリン酸化部位

アミノ酸 107-115

N-ミリストイル化部位

アミノ酸 20-26, 192-198

アミド化部位

アミノ酸 25-29

【図１１５】

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTTAACTAATTCAACAAACGGGACCCTT
CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAGAGCGGG
AAGGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACC**ATG**AAACTCTGGGTGTCTGC
ATTGCTGATGGCCTGGTTTGGTGTCCTGAGCTGTGTGCAGGCCGAATTCTTCACCTCTATTG
GGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC
CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC
TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCCTGTGAATGCCTACAAACTGG
TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCCTGCAGGACTCAGCTGCA
GGTTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGG
AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCA
GAGGGGAACTTCCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG
GGCCGCTCGGCCTACAATGAAGGGGACTATTATCATACGGTGTTGTGGATGGAGCAGGTGCT
AAAGCAGCTTGATGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCA
GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCCTGCTC
TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGA
GGAAGAGAGAGAAAAAACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCCAGAAGGCA
TCTATGAGAGGCCTGTGGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCGTGGG
GAGGGTGTCAAACTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAA
CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACA
TCGTCAGGTACTACGATGTCATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA
CCTAAACTTGCACGAGCCACCGTTCGTGATCCCAAGACAGGAGTCCTCACTGTCGCCAGCTA
CCGGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATC
GTCGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT
TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG
CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG
GTGGTGCCACCGTCTTCCCTGATCTGGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTG
TTCTGGTACAACCTCTTGCGGAGCCGGGAAGGTGACTACCGAACAAGACATGCTGCCTGCCC
TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT
TGAGACCTTGTGGATCAACAGAAGTTGAC**TGA**CATCCTTTTCTGTCCTTCCCCTTCCTGGTC
CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCCTTTGTATGTTCCTATCAGGCT
GATTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT
GTGACTGAAGTCCCAGCCCTTCCATTCAGCCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA
AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTTGTACCTCAGGTGTT
TTAGGTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACCTTGTTTACATGTTTGTTTTTAT
GGCATTTCTATCTATTGTGGCTTTACCAAAAAATAAAATGTCCCTACCAGAAAAAAAAAA

【図１１６】

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA
NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALEDLVLQDSAAGFIANLSVQRQFFP
TDEDEIGAAKALMRLQDTYRLDPGTISRGELPGTKYQAMLSVDDCFGMGRSAYNEGDYYMTV
LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLLSLDPSHERAGGNLR
YFEQLLEEEREKTLTNQTEAELATPEGIYERPVDYLPERDVYESLCRGEGVKLTPRRQKRLF
CRYHHGNRAPQLLIAPFKEEDEWDSPHIVRYYDVMSDEEIERIKEIAKPKLARATVRDPKTG
VLTVASYRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS
RRPFDSGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGEGDYR
TRHAACPVLVGCKWVSNKWFHERGQEFLRPCGSTEVD
シグナル配列：

アミノ酸 1-17

N-グリコシル化部位

アミノ酸 115-119, 264-268

グリコサミノグリカン付着部位

アミノ酸 490-494

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 477-481

カゼインキナーゼIIリン酸化部位

アミノ酸 43-47, 72-76, 125-129, 151-155, 165-169, 266-270, 346-350, 365-369, 385-389, 457-461, 530-534

チロシンキナーゼリン酸化部位

アミノ酸 71-80, 489-496

N-ミリストイル化部位

アミノ酸 14-20, 131-137, 171-177, 446-452

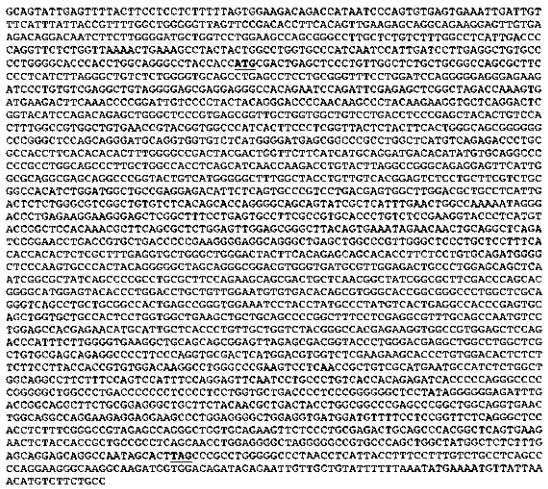
真核生物膜リポタンパク質脂質付着部位

アミノ酸 8-19

ロイシンジッパーパターン

アミノ酸 213-235

【図１１７】



【図１１８】

MRLSSLLALLRPALPLILGLSLGCSLSLLLRVSWIQGEGEDPCVEAVGERGGPQNPDSRARLD
QSDEDFKPRIVPYYRDPNKPYKKVLRTRYIQTELGSRERLLVAVLTSRATLSTLAVAVNRTV
AHHFPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETLRHLHTHFGADYDWFFIMQDDTY
VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGPGYLLSRSLLLRLRPHLDGCRG
DILSARPDEWLGRCLIDSLGVGCVSQHQGQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE
GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPEGEAGLSWPVGLPAPFTPHSRFEV
LGWDYPTEQHTFSCADGAPKCPLQGASRADVGDALETALEQLNRRYQPRLRFQKQRLLNGYR
RFDPARGMEYTLDLLLECVTQRGHRRALARRVSLLRPLSRVEILPMPYVTEATRVQLVLPLL
VAEAAAAPAFLEAFAANVLEPREHALLTLLLVYGPREGGRGAPDPFLGVKAAAAELERRYPG
TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLTTVWTRPGPEVLNRCRMNAISGWQAFFP
VHFQEPNPALSPQRSPPGPPGAGPDPPSPPGADPSRGAPIGGRFDRQASAEGCFYNADYLAA
RARLAGELAGQEEEEALEGLEVMDVFLRFSGLHLFRAVEPGLVQKFSLRDCSPRLSEELYHR
CRLSNLEGLGGRAQLAMALFEQEQANST

シグナル配列：

アミノ酸 1-15

膜貫通ドメイン：

アミノ酸 489-507

N-グリコシル化部位

アミノ酸 121-125, 342-346

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 319-323, 464-468

カゼインキナーゼIIリン酸化部位

アミノ酸 64-68, 150-154, 322-326, 331-337, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

チロシンキナーゼリン酸化部位

アミノ酸 736-743

N-ミリストイル化部位

アミノ酸 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-678

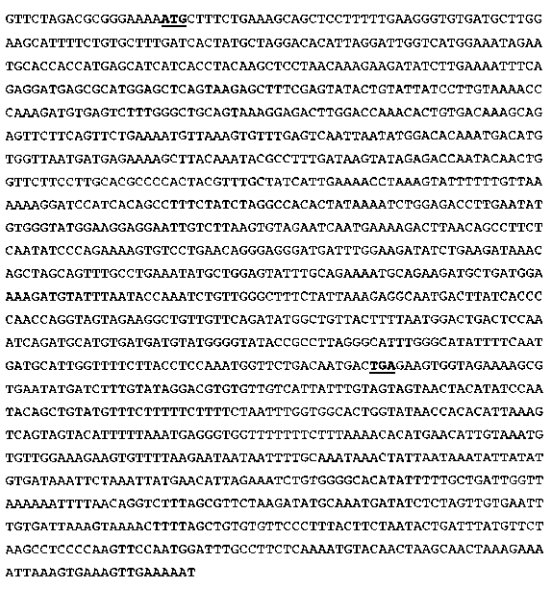
真核生物膜リポタンパク質脂質付着部位

アミノ酸 14-25

細胞付着配列

アミノ酸 247-250

【図１１９】



【図１２０】

MLSESSSFLKGVMLGSIPCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME
LSKSFRVYCIILVKPKDVSLWAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRK
AYKYAFDKYRDQYNWFFLARPTTFAIIENLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMEGG
IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNT
KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIPNDALVFL
PPNGSDND

シグナル配列：

アミノ酸 1-33

N-グリコシル化部位

アミノ酸 121-125, 342-346

cAMP-及びcGMP-依存性プロテインキナーゼリン酸化部位

アミノ酸 319-323, 464-468

カゼインキナーゼIIリン酸化部位

アミノ酸 64-132, 150-154, 322-326, 331-335, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

チロシンキナーゼリン酸化部位

アミノ酸 736-743

N-ミリストイル化部位

アミノ酸 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-672

真核生物膜リポタンパク質脂質付着部位

アミノ酸 14-25

細胞付着配列

アミノ酸 247-250

【図１２１】

CCCACGCGTCCGATCTTACCAACAAAACACTCCTGAGGAGAAAGAAAGAGAGGGAGGGAGAG
AAAAAGAGAGAGAGAGAAACAAAAAACCAAAGAGAGAGAAAAA**ATG**AATTCATCTAAATCAT
CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTTCCCAAATGTTCTTATGGACTGTTGCT
GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTCGCAT
CTTTCAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTCACAGAGCTCTCCTGCT
ACAATTATGGATCAGGTTCAGTCAAGAATTGTTGTCCATTGAACTGGGAATATTTTCAATCC
AGCTGCTACTTCTTTTCTACTGACACCATTTCCTGGGCGTTAAGTTTAAAGAACTGCTCAGC
CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCCTACAAGA
AACCTAAAATGAGAGAGTTTTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAGTGGCAA
TGGGTGGACGGCACACCTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA
CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA
ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG
AACAAAGGAAAATCTCTT**TAA**GAACAGAAGGCACAACTCAAATGTGTAAAGAAGGAAGAGCA
AGAACATGGCCACACCCACCGCCCCACACGAGAAATTTGTGCGCTG**AA**CTTCAAAGGACTTC
ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAA

【図１２２】

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN
FTELSCYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSAMGAHLVVINSQEEQ
EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCATMRDSS
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

シグナル配列：

アミノ酸 1-42

N-グリコシル化部位

アミノ酸 2-6, 62-66, 107-111

カゼインキナーゼIIリン酸化部位

アミノ酸 51-55, 120-124, 163-167, 175-179, 181-185

N-ミリストイル化部位

アミノ酸 15-21, 74-80, 155-161

真核生物膜リポタンパク質脂質付着部位

アミノ酸 27-38

【国際公開パンフレット】

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/12, C07K 14/47, C12N 15/62, C07K 16/18 — A2

(11) International Publication Number: **WO 00/15796**

(43) International Publication Date: 23 March 2000 (23.03.00)

(21) International Application Number: PCT/US99/21090

(22) International Filing Date: 15 September 1999 (15.09.99)

(30) Priority Data:
PCT/US98/19330 16 September 1998 (16.09.98) US

(71) Applicant *(for all designated States except US)*: GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: CHEN, Jian [CN/US]; 121 York Drive, Princeton, NJ 08540 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). GURNEY, Austin, L. [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). HILLAN, Kenneth [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). PENNICA, Diane [US/US]; 2417 Hale Drive, Burlingame, CA 94010 (US). WOOD, William, I. [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US). YUAN, Jean [CN/US]; 176 West 37th Avenue, San Mateo, CA 94403 (US).

(74) Agents: KRESNAK, Mark, T. et al.; Genentech Inc., 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

(57) Abstract

The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

FOR THE PURPOSES OF INFORMATION ONLY

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

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

SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptides.

BACKGROUND OF THE INVENTION

Extracellular proteins play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiationfactors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of action in the extracellular environment.

Secreted proteins have various industrial applications, including as pharmaceuticals, diagnostics, biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents. Efforts are being undertaken by both industry and academia to identify new, native secreted proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci. 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

Membrane-bound proteins and receptors can play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesin molecules like selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and nerve growth factor receptor.

Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction.

Efforts are being undertaken by both industry and academia to identify new, native receptor or membrane-bound proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor or membrane-bound proteins.

1. PRO211 and PRO217

Epidermal growth factor (EGF) is a conventional mitogenic factor that stimulates the proliferation of various types of cells including epithelial cells and fibroblasts. EGF binds to and activates the EGF receptor (EGFR), which initiates intracellular signaling and subsequent effects. The EGFR is expressed in neurons of the cerebral cortex, cerebellum, and hippocampus in addition to other regions of the central nervous system (CNS). In addition, EGF is also expressed in various regions of the CNS. Therefore, EGF acts not only on mitotic cells, but also on postmitotic neurons. In fact, many studies have indicated that EGF has neurotrophic or neuromodulatory effects on various types of neurons in the CNS. For example, EGF acts directly on cultured cerebral cortical and cerebellar neurons, enhancing neurite outgrowth and survival. On the other hand, EGF also acts on other cell types, including septal cholinergic and mesencephalic dopaminergic neurons, indirectly through glial cells. Evidence of the effects of EGF on neurons in the CNS is accumulating, but the mechanisms of action remain essentially unknown. EGF-induced signaling in mitotic cells is better understood than in postmitotic neurons. Studies of cloned pheochromocytoma PC12 cells and cultured cerebral cortical neurons have suggested that the EGF-induced neurotrophic actions are mediated by sustained activation of the EGFR and mitogen-activated protein kinase (MAPK) in response to EGF. The sustained intracellular signaling correlates with the decreased rate of EGFR down-regulation, which might determine the response of neuronal cells to EGF. It is likely that EGF is a multi-potent growth factor that acts upon various types of cells including mitotic cells and postmitotic neurons.

EGF is produced by the salivary and Brunner's glands of the gastrointestinal system, kidney, pancreas, thyroid gland, pituitary gland, and the nervous system, and is found in body fluids such as saliva, blood, cerebrospinal fluid (CSF), urine, amniotic fluid, prostatic fluid, pancreatic juice, and breast milk, Plata-Salaman, *Peptides* 12: 653-663 (1991).

EGF is mediated by its membrane specific receptor, which contains an intrinsic tyrosine kinase. Stoscheck *et al., J. Cell Biochem.* 31: 135-152 (1986). EGF is believed to function by binding to the extracellular portion of its receptor which induces a transmembrane signal that activates the intrinsic tyrosine kinase.

Purification and sequence analysis of the EGF-like domain has revealed the presence of six conserved cysteine residues which cross-bind to create three peptide loops. Savage *et al., J. Biol. Chem.* 248: 7669-7672 (1979). It is now generally known that several other peptides can react with the EGF receptor which share the same generalized motif $X_1CX_2CX_{4-5}CX_{7-9}CXCX_3GX_2CX_n$, where X represents any non-cysteine amino acid, and

n is a variable repeat number. Non isolated peptides having this motif include TGF-α, amphiregulin, schwannoma-derived growth factor (SDGF), heparin-binding EGF-like growth factors and certain virally encoded peptides (e.g., Vaccinia virus, Reisner, *Nature* 313: 801-803 (1985), Shope fibroma virus, Chang et al., Mol Cell Biol. 7: 535-540 (1987), Molluscum contagiosum, Porter and Archard, *J. Gen. Virol.* 68: 673-682 (1987), and Myxoma virus, Upton *et al., J. Virol.* 61: 1271-1275 (1987), Prigent and Lemoine, *Prog. Growth Factor Res.* 4: 1-24 (1992).

EGF-like domains are not confined to growth factors but have been observed in a variety of cell-surface and extracellular proteins which have interesting properties in cell adhesion, protein-protein interaction and development, Laurence and Gusterson, *Tumor Biol.* 11: 229-261 (1990). These proteins include blood coagulation factors (factors VI, IX, X, XII, protein C, protein S, protein Z, tissue plasminogen activator, urokinase), extracellular matrix components (laminin, cytotactin, entactin), cell surface receptors (LDL receptor, thrombomodulin receptor) and immunity-related proteins (complement C1r, uromodulin).

Even more interesting, the general structure pattern of EGF-like precursors is preserved through lower organisms as well as in mammalian cells. A number of genes with developmental significance have been identified in invertebrates with EGF-like repeats. For example, the *notch* gene of Drosophila encodes 36 tandemly arranged 40 amino acid repeats which show homology to EGF, Wharton *et al., Cell* 43: 557-581 (1985). Hydropathy plots indicate a putative membrane spanning domain, with the EGF-related sequences being located on the extracellular side of the membrane. Other homeotic genes with EGF-like repeats include Delta, 95F and 5ZD which were identified using probes based on Notch, and the nematode gene *Lin-12* which encodes a putative receptor for a developmental signal transmitted between two specified cells.

Specifically, EGF has been shown to have potential in the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions. Konturek *et al., Eur. J. Gastroenterol Hepatol.* 7 (10), 933-37 (1995), including the treatment of necrotizing enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration gastrointestinal ulcerations and congenital microvillus atrophy, Guglietta and Sullivan, *Eur. J. Gastroenterol Hepatol.* 7(10), 945-50 (1995). Additionally, EGF has been implicated in hair follicle differentiation; du Cros, *J. Invest. Dermatol.* 101 (1 Suppl.), 106S-113S (1993), Hillier, *Clin. Endocrinol.* 33(4), 427-28 (1990); kidney function, Humm *et al., Semin. Nephrol.* 13 (1): 109-15 (1993), Harris, *Am. J. Kidney Dis.* 17(6): 627-30 (1991); tear fluid, van Setten *et al., Int. Ophthalmol* 15(6); 359-62 (1991); vitamin K mediated blood coagulation, Stenflo *et al., Blood* 78(7): 1637-51 (1991). EGF is also implicated various skin disease characterized by abnormal keratinocyte differentiation, e.g., psoriasis, epithelial cancers such as squamous cell carcinomas of the lung, epidermoid carcinoma of the vulva and gliomas. King *et al., Am. J. Med. Sci.* 296: 154-158 (1988).

Of great interest is mounting evidence that genetic alterations in growth factors signaling pathways are closely linked to developmental abnormalities and to chronic diseases including cancer. Aaronson, *Science* 254: 1146-1153 (1991). For example, c-erb-2 (also known as HER-2), a proto-oncogene with close structural similarity to EGF receptor protein, is overexpressed in human breast cancer. King *et al., Science* 229: 974-976 (1985); Gullick, *Hormones and their actions*, Cooke *et al.*, eds. Amsterdam, Elsevier, pp 349-360 (1986).

We herein describe the identification and characterization of novel polypeptides having homology to EGF, wherein those polypeptides are herein designated PRO211 and PRO217.

2. **PRO230**

Nephritis is a condition characterized by inflammation of the kidney affecting the structure and normal function of the kidney. This condition can be chronic or acute and is generally caused by infection, degenerative process or vascular disease. In all cases, early detection is desirable so that the patient with nephritis can begin treatment of the condition.

An approach to detecting nephritis is to determine the antigens associated with nephritis and antibodies thereto. In rabbit, a tubulointerstitial nephritis antigen (TIN-ag) has been reported in Nelson, T. R., et al., J. Biol. Chem., 270(27):16265-70 (July 1995) (GENBANK/U24270). This study reports that the rabbit TIN-ag is a basement membrane glycoprotein having a predicted amino acid sequence which has a carboxyl-terminal region exhibiting 30% homology with human preprocathepsin B, a member of the cystein proteinase family of proteins. It is also reported that the rabbit TIN-ag has a domain in the amino-terminal region containing an epidermal growth factor-like motif that shares homology with laminin A and S chains, alpha 1 chain of type I collagen, von Willebrand's factor and mucin, indicating structural and functional similarities. Studies have also been conducted in mice. However, it is desirable to identify tubulointerstitial nephritis antigens in humans to aid in the development of early detection methods and treatment of nephritis.

Proteins which have homology to tubulointerstitial nephritis antigens are of particular interest to the medical and industrial communities. Often, proteins having homology to each other have similar function. It is also of interest when proteins having homology do not have similar functions, indicating that certain structural motifs identify information other than function, such as locality of function. We herein describe the identification and characterization of a novel polypeptide, designated herein as PRO230, which has homology to tubulointerstitial nephritis antigens.

3. ***PRO232***

Stem cells are undifferentiated cells capable of (a) proliferation, (b) self maintenance, (c) the production of a large number of differentiated functional progeny, (d) regeneration of tissue after injury and/or (e) a flexibility in the use of these options. Stem cells often express cell surface antigens which are capable of serving as cell specific markers that can be exploited to identify stem cells, thereby providing a means for identifying and isolating specific stem cell populations.

Having possession of different stem cell populations will allow for a number of important applications. For example, possessing a specific stem cell population will allow for the identification of growth factors and other proteins which are involved in their proliferation and differentiation. In addition, there may be as yet undiscovered proteins which are associated with (1) the early steps of dedication of the stem cell to a particular lineage, (2) prevention of such dedication, and (3) negative control of stem cell proliferation, all of which may be identified if one has possession of the stem cell population. Moreover, stem cells are important and ideal targets for gene therapy where the inserted genes promote the health of the individual into whom the stem cells

are transplanted. Finally, stem cells may play important roles in transplantation of organs or tissues, for example liver regeneration and skin grafting.

Given the importance of stem cells in various different applications, efforts are currently being undertaken by both industry and academia to identify new, native stem cell antigen proteins so as to provide specific cell surface markers for identifying stem cell populations as well as for providing insight into the functional roles played by stem cell antigens in cell proliferation and differentiation. We herein describe the identification and characterization of novel polypeptides having homology to a stem cell antigen, wherein those polypeptides are herein designated as PRO232 polypeptides..

4. PRO187

Growth factors are molecular signals or mediators that enhance cell growth or proliferation, alone or in concert, by binding to specific cell surface receptors. However, there are other cellular reactions than only growth upon expression to growth factors. As a result, growth factors are better characterized as multifunctional and potent cellular regulators. Their biological effects include proliferation, chemotaxis and stimulation of extracellular matrix production. Growth factors can have both stimulatory and inhibitory effects. For example, transforming growth factor (TGF-β) is highly pleiotropic and can stimulate proliferation in some cells, especially connective tissue, while being a potent inhibitor of proliferation in others, such as lymphocytes and epithelial cells.

The physiological effect of growth stimulation or inhibition by growth factors depends upon the state of development and differentiation of the target tissue. The mechanism of local cellular regulation by classical endocrine molecules involves comprehends autocrine (same cell), juxtacrine (neighbor cell), and paracrine (adjacent cells) pathways. Peptide growth factors are elements of a complex biological language, providing the basis for intercellular communication. They permit cells to convey information between each other, mediate interaction between cells and change gene expression. The effect of these multifunctional and pluripotent factors is dependent on the presence or absence of other peptides.

FGF-8 is a member of the fibroblast growth factors (FGFs) which are a family of heparin-binding, potent mitogens for both normal diploid fibroblasts and established cell lines, Gospodarowicz *et al.* (1984), *Proc. Natl. Acad. Sci. USA* 81:6963. The FGF family comprises acidic FGF (FGF-1), basic FGF (FGF-2), INT-2 (FGF-3), K-FGF/HST (FGF-4), FGF-5, FGF-6, KGF (FGF-7), AIGF (FGF-8) among others. All FGFs have two conserved cysteine residues and share 30-50% sequence homology at the amino acid level. These factors are mitogenic for a wide variety of normal diploid mesoderm-derived and neural crest-derived cells, including granulosa cells, adrenal cortical cells, chondrocytes, myoblasts, corneal and vascular endothelial cells (bovine or human), vascular smooth muscle cells, lens, retina and prostatic epithelial cells, oligodendrocytes, astrocytes, chrondocytes, myoblasts and osteoblasts.

Fibroblast growth factors can also stimulate a large number of cell types in a non-mitogenic manner. These activities include promotion of cell migration into wound area (chemotaxis), initiation of new blood vessel formulation (angiogenesis), modulation of nerve regeneration and survival (neurotrophism), modulation of endocrine functions, and stimulation or suppression of specific cellular protein expression, extracellular matrix

production and cell survival. Baird & Bohlen, Handbook of Exp. Pharmacol. 95(1): 369-418, Springer, (1990). These properties provide a basis for using fibroblast growth factors in therapeutic approaches to accelerate wound healing, nerve repair, collateral blood vessel formation, and the like. For example, fibroblast growth factors have been suggested to minimize myocardium damage in heart disease and surgery (U.S.P. 4,378,347).

FGF-8, also known as androgen-induced growth factor (AIGF), is a 215 amino acid protein which shares 30-40% sequence homology with the other members of the FGF family. FGF-8 has been proposed to be under androgenic regulation and induction in the mouse mammary carcinoma cell line SC3. Tanaka *et al., Proc. Natl. Acad. Sci. USA* 89: 8928-8932 (1992); Sato *et al., J. Steroid Biochem. Molec. Biol.* 47: 91-98 (1993). As a result, FGF-8 may have a local role in the prostate, which is known to be an androgen-responsive organ. FGF-8 can also be oncogenic, as it displays transforming activity when transfected into NIH-3T3 fibroblasts. Kouhara *et al., Oncogene* 9 455-462 (1994). While FGF-8 has been detected in heart, brain, lung, kidney, testis, prostate and ovary, expression was also detected in the absence of exogenous androgens. Schmitt *et al., J. Steroid Biochem. Mol. Biol.* 57 (3-4): 173-78 (1996).

FGF-8 shares the property with several other FGFs of being expressed at a variety of stages of murine embryogenesis, which supports the theory that the various FGFs have multiple and perhaps coordinated roles in differentiation and embryogenesis. Moreover, FGF-8 has also been identified as a protooncogene that cooperates with Wnt-1 in the process of mammary tumorigenesis (Shackleford *et al., Proc. Natl. Acad. Sci. USA* 90, 740-744 (1993); Heikinheimo *et al., Mech. Dev.* 48: 129-138 (1994)).

In contrast to the other FGFs, FGF-8 exists as three protein isoforms, as a result of alternative splicing of the primary transcript. Tanaka et al., *supra*. Normal adult expression of FGF-8 is weak and confined to gonadal tissue, however northern blot analysis has indicated that FGF-8 mRNA is present from day 10 through day 12 or murine gestation, which suggests that FGF-8 is important to normal development. Heikinheimo *et al., Mech Dev.* 48(2): 129-38 (1994). Further *in situ* hybridization assays between day 8 and 16 of gestation indicated initial expression in the surface ectoderm of the first bronchial arches, the frontonasal process, the forebrain and the midbrain-hindbrain junction. At days 10-12, FGF-8 was expressed in the surface ectoderm of the forelimb and hindlimb buds, the nasal its and nasopharynx, the infundibulum and in the telencephalon, diencephalon and metencephalon. Expression continues in the developing hindlimbs through day 13 of gestation, but is undetectable thereafter. The results suggest that FGF-8 has a unique temporal and spatial pattern in embryogenesis and suggests a role for this growth factor in multiple regions of ectodermal differentiation in the post-gastrulation embryo.

We herein describe the identification of novel polypeptides having homology to FGF-8, wherein those polypeptides are herein designated PRO187 polypeptides.

5. **PRO265**

Protein-protein interactions include receptor and antigen complexes and signalling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical

community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglubular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Nouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chtemerson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factor-β involvement for treatment for cancer, wound healing and scarring). Also of particular interest is fibromodulin and its use to prevent or reduce dermal scarring. A study of fibromodulin is found in U.S. Patent No. 5,654,270 to Ruoslahti, et al.

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as fibromodulin, the SLIT protein and platelet glycoprotein V. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. We herein describe the identification and characterization of novel polypeptides having homology to fibromodulin, herein designated as PRO265 polypeptides.

6. PRO219

Human matrilin-2 polypeptide is a member of the von Willebrand factor type A-like module superfamily. von Willebrand factor is a protein which plays an important role in the maintenence of hemostasis. More specifically, von Willebrand factor is a protein which is known to participate in platelet-vessel wall

interactions at the site of vascular injury via its ability to interact and form a complex with Factor VIII. The absence of von Willebrand factor in the blood causes an abnormality with the blood platelets that prevents platelet adhesion to the vascular wall at the site of the vascular injury. The result is the propensity for bruising, nose bleeds, intestinal bleeding, and the like comprising von Willebrand's disease.

Given the physiological importance of the blood clotting factors, efforts are currently being undertaken by both industry and academia to identify new, native proteins which may be involved in the coagulation process. We herein describe the identification of a novel full-length polypeptide which possesses homology to the human matrilin-2 precursor polypeptide.

7. PRO246

The cell surface protein HCAR is a membrane-bound protein that acts as a receptor for subgroup C of the adenoviruses and subgroup B of the coxsackieviruses. Thus, HCAR may provide a means for mediating viral infection of cells in that the presence of the HCAR receptor on the cellular surface provides a binding site for viral particles, thereby facilitating viral infection.

In light of the physiological importance of membrane-bound proteins and specficially those which serve a cell surface receptor for viruses, efforts are currently being undertaken by both industry and academia to identify new, native membrane-bound receptor proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor proteins. We herein describe a novel membrane-bound polypeptide (designated herein as PRO246) having homology to the cell surface protein HCAR and to various tumor antigens including A33 and carcinoembryonic antigen, wherein this polypeptide may be a novel cell surface virus receptor or tumor antigen.

8. PRO228

There are a number of known seven transmembrane proteins and within this family is a group which includes CD97 and EMR1. CD97 is a seven-span transmembrane receptor which has a cellular ligand, CD55, DAF. Hamann, et al., J. Exp. Med. (U.S.), 184(3):1189 (1996). Additionally, CD97 has been reported as being a dedifferentiation marker in human thyroid carcinomas and as associated with inflammation. Aust, et al., Cancer Res. (U.S.), 57(9):1798 (1997); Gray, et al., J. Immunol. (U.S.), 157(12):5438 (1996). CD97 has also been reported as being related to the secretin receptor superfamily, but unlike known members of that family, CD97 and EMR1 have extended extracellular regions that possess several EGF domains at the N-terminus. Hamann, et al., Genomics, 32(1):144 (1996); Harmann, et al., J. Immunol., 155(4):1942 (1995). EMR1 is further described in Lin, et al., Genomics, 41(3):301 (1997) and Baud, et al., Genomics, 26(2):334 (1995). While CD97 and EMR1 appear to be related to the secretin receptors, a known member of the secretin family of G protein-coupled receptors includes the alpha-latroxin receptor, latrophilin, which has been described as calcium independent and abundant among neuronal tissues. Lelianova, et al., J. Biol. Chem., 272(34), 21504 (1997); Davletov, et al., J. Biol. Chem. (U.S.), 271(38):23239 (1996). Both members of the secretin receptor superfamily and non-members which are related to the secretin receptor superfamily, or CRF and calcitonin receptors are of interest. In particular, new members of these families, identified by their homology to known

proteins, are of interest.

Efforts are being undertaken by both industry and academia to identify new membrane-bound receptor proteins, particularly transmembrane proteins with EGF repeats and large N-terminuses which may belong to the family of seven-transmembrane proteins of which CD97 and EMR1 are members. We herein describe the identification and charactization of novel polypeptides having homology to CD97 and EMR1, designated herein as PRO228 polypeptides.

9. PRO533

Growth factors are molecular signals or mediators that enhance cell growth or proliferation, alone or in concert, by binding to specific cell surface receptors. however, there are other cellular reactions than only growth upon expression to growth factors. As a result, growth factors are better characterized as multifunctional and potent cellular regulators. Their biological effects include proliferation, chemotaxis and stimulation of extracellular matrix production. Growth factors can have both stimulatory and inhibitory effects. For example, transforming growth factors (TGF-β) is highly pleiotropic and can stimulate proliferation in some cells, especially connective tissues, while being a potent inhibitor of proliferation in others, such as lymphocytes and epithelial cells.

The physiological effect of growth stimulation or inhibition by growth factors depends upon the state of development and differentiation of the target tissue. The mechanism of local cellular regulation by classical endocrine molecules comprehends autocrine (same cell), juxtacrine (neighbor cell), and paracrine (adjacent cell) pathways. Peptide growth factors are elements of a complex biological language, providing the basis for intercellular communication. They permit cells to convey information between each other, mediate interaction between cells and change gene expression. the effect of these multifunctional and pluripotent factors is dependent on the presence or absence of other peptides.

Fibroblast growth factors (FGFs) are a family of heparin-binding, potent mitogens for both normal diploid fibroblasts and established cell lines. Gospodarowicz, D. et al. (1984), Proc. Natl. Acad. Sci. USA 81: 6983. the FGF family comprises acidic FGF (FGF-1), basic FGF (FGF-2), INT-2 (FGF-3), K-FGF/HST (FGF-4), FGF-5, FGF-6, KGF (FGF-7), AIGF (FGF-8) among others. All FGFs have two conserved cysteine residues and share 30-50% sequence homology at the amino acid level. These factors are mitogenic for a wide variety of normal diploid mesoderm-derived and neural crest-derived cells, inducing granulosa cells, adrenal cortical cells, chrondocytes, myoblasts, corneal and vascular endothelial cells (bovine or human), vascular smooth muscle cells, lens, retina and prostatic epithelial cells, oligodendrocytes, astrocytes, chrondocytes, myoblasts and osteoblasts.

Fibroblast growth factors can also stimulate a large number of cell types in a non-mitogenic manner. These activities include promotion of cell migration into a wound area (chemotaxis), initiation of new blood vessel formulation (angiogenesis), modulation of nerve regeneration and survival (neurotrophism), modulation of endocrine functions, and stimulation or suppression of specific cellular protein expression, extracellular matrix production and cell survival. Baird, A. & Bohlen, P., *Handbook of Exp. Pharmacol.* 95(1): 369-418 (1990). These properties provide a basis for using fibroblast growth factors in therapeutic approaches to accelerate

wound healing, nerve repair, collateral blood vessel formation, and the like. For example, fibroblast growth factors, have been suggested to minimize myocardium damage in heart disease and surgery (U.S.P. 4,378,437).

We herein describe the identification and characterization of novel polypeptides having homology to FGF, herein designated PRO533 polypeptides.

10. PRO245

Some of the most important proteins involved in the above described regulation and modulation of cellular processes are the enzymes which regulate levels of protein phosphorylation in the cell. For example, it is known that the transduction of signals that regulate cell growth and differentiation is regulated at least in part by phosphorylation and dephosphorylation of various cellular proteins. The enzymes that catalyze these processes include the protein kinases, which function to phosphorylate various cellular proteins, and the protein phosphatases, which function to remove phosphate residues from various cellular proteins. The balance of the level of protein phosphorylation in the cell is thus mediated by the relative activities of these two types of enzymes.

Although many protein kinase enzymes have been identified, the physiological role played by many of these catalytic proteins has yet to be elucidated. It is well known, however, that a number of the known protein kinases function to phosphorylate tyrosine residues in proteins, thereby leading to a variety of different effects. Perhaps most importantly, there has been a great deal of interest in the protein tyrosine kinases since the discovery that many oncogene products and growth factors possess intrinsic protein tyrosine kinase activity. There is, therefore, a desire to identify new members of the protein tyrosine kinase family.

Given the physiological importance of the protein kinases, efforts are being undertaken by both industry and academia to identify new, native kinase proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel kinase proteins. We herein describe the identification and characterization of novel polypeptides having homology to tyrosine kinase proteins, designated herein as PRO245 polypeptides.

11. PRO220, PRO221 and PRO227

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglobular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats.

See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vonv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factorβ involvement for treatment for cancer, wound healing and scarring).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as the SLIT protein and platelet glycoprotein V.

12. PRO258

Immunoglobulins are antibody molecules, the proteins that function both as receptors for antigen on the B-cell membrane and as the secreted products of the plasma cell. Like all antibody molecules, immunoglobulins perform two major functions: they bind specifically to an antigen and they participate in a limited number of biological effector functions. Therefore, new members of the Ig superfamily are always of interest. Molecules which act as receptors by various viruses and those which act to regulate immune function are of particular interest. Also of particular interest are those molecules which have homology to known Ig family members which act as virus receptors or regulate immune function. Thus, molecules having homology to poliovirus receptors, CRTAM and CD166 (a ligand for lymphocyte antigen CD6) are of particular interest.

Extracellular and membrane-bound proteins play important roles in the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiationfactors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted

polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of action in the extracellular environment, usually at a membrane-bound receptor protein.

We herein describe the identification and characterization of novel polypeptides having homology to CRTAM, designated herein as PRO258 polypeptides.

13. PRO266

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglobular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Nouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chiemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Rueslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factorβ involvement for treatment for cancer, wound healing and scarring).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions, neuronal development and adhesin molecules. Of particular interest are those proteins having leucine rich repeats and homology to known proteins

having leucine rich repeats such as the SLIT protein. We herein describe novel polypeptides having homology to SLIT, designated herein as PRO266 polypeptides.

14. **PRO269**

Thrombomodulin binds to and regulates the activity of thrombin. It is important in the control of blood coagulation. Thrombomodulin functions as a natural anticoagulant by accelerating the activation of protein C by thrombin. Soluble thrombomodulin may have therapeutic use as an antithrombotic agent with reduced risk for hemorrhage as compared with heparin. Thrombomodulin is a cell surface trans-membrane glycoprotein, present on endothelial cells and platelets. A smaller, functionally active form of thrombomodulin circulates in the plasma and is also found in urine. (In Haeberli, A., Human Protein Data, VCH Pub., N.Y., 1992). Peptides having homology to thrombomodulin are particularly desirable.

We herein describe the identification and characterization of novel polypeptides having homology to thrombomodulin, designated herein as PRO269 polypeptides.

15. **PRO287**

Procollagen C-proteinase enhancer protein binds to and enhances the activity of bone morphogenic protein "BMP1"/procollagen C-proteinase (PCP). It plays a role in extracellular matrix deposition. BMP1 proteins may be used to induce bone and/or cartilage formation and in wound healing and tissue repair. Therefore, procollagen C-proteinase enhancer protein, BMP1 and proteins having homology thereto, are of interest to the scientific and medical communities.

We herein describe the identification and characterization of novel polypeptides having homology to procollagen C-proteinase enhancer protein precursor and procollagen C-proteinase enhancer protein, designated herein as PRO287 polypeptides.

16. **PRO214**

Growth factors are molecular signals or mediators that enhances cell growth or proliferation, alone or in concert, by binding to specific cell surface receptors. However, there are other cellular reactions than only growth upon expression to growth factors. As a result, growth factors are better characterized as multifunctional and potent cellular regulators. Their biological effects include proliferation, chemotaxis and stimulation of extracellular matrix production. Growth factors can have both stimulatory and inhibitory effects. For example, transforming growth factor β (TGF-β) is highly pleiotropic and can stimulate proliferation in some cells, especially connective tissue, while being a potent inhibitor of proliferation in others, such as lymphocytes and epithelial cells.

The physiological effect of growth stimulation or inhibition by growth factors depends upon the state of development and differentiation of the target tissue. The mechanism of local cellular regulation by classical endocrine molecules involves comprehends autocrine (same cell), juxtacrine (neighbor cell), and paracrine (adjacent cells) pathways. Peptide growth factors are elements of a complex biological language, providing the basis for intercellular communication. They permit cells to convey information between each other, mediate

interaction between cells and change gene expression. The effect of these multifunctional and pluripotent factors is dependent on the presence or absence of other peptides.

Epidermal growth factor (EGF) is a conventional mitogenic factor that stimulates the proliferation of various types of cells including epithelial cells and fibroblasts. EGF binds to and activates the EGF receptor (EGFR), which initiates intracellular signaling and subsequent effects. The EGFR is expressed in neurons of the cerebral cortex, cerebellum, and hippocampus in addition to other regions of the central nervous system (CNS). In addition, EGF is also expressed in various regions of the CNS. Therefore, EGF acts not only on mitotic cells, but also on postmitotic neurons. In fact, many studies have indicated that EGF has neurotrophic or neuromodulatory effects on various types of neurons in the CNS. For example, EGF acts directly on cultured cerebral cortical and cerebellar neurons, enhancing neurite outgrowth and survival. On the other hand, EGF also acts on other cell types, including septal cholinergic and mesencephalic dopaminergic neurons, indirectly through glial cells. Evidence of the effects of EGF on neurons in the CNS is accumulating, but the mechanisms of action remain essentially unknown. EGF-induced signaling in mitotic cells is better understood than in postmitotic neurons. Studies of cloned pheochromocytoma PC12 cells and cultured cerebral cortical neurons have suggested that the EGF-induced neurotrophic actions are mediated by sustained activation of the EGFR and mitogen-activated protein kinase (MAPK) in response to EGF. The sustained intracellular signaling correlates with the decreased rate of EGFR down-regulation, which might determine the response of neuronal cells to EGF. It is likely that EGF is a multi-potent growth factor that acts upon various types of cells including mitotic cells and postmitotic neurons.

EGF is produced by the salivary and Brunner's glands of the gastrointestinal system, kidney, pancreas, thyroid gland, pituitary gland, and the nervous system, and is found in body fluids such as saliva, blood, cerebrospinal fluid (CSF), urine, amniotic fluid, prostatic fluid, pancreatic juice, and breast milk, Plata-Salaman, CR *Peptides* 12: 653-663 (1991).

EGF is mediated by its membrane specific receptor, which contains an intrinsic tyrosine kinase. Stoscheck CM *et al., J. Cell Biochem.* 31: 135-152 (1986). EGF is believed to function by binding to the extracellular portion of its receptor which induces a transmembrane signal that activates the intrinsic tyrosine kinase.

Purification and sequence analysis of the EGF-like domain has revealed the presence of six conserved cysteine residues which cross-bind to create three peptide loops, Savage *CR et al., J. Biol. Chem.* 248: 7669-7672 (1979). It is now generally known that several other peptides can react with the EGF receptor which share the same generalized motif $X_nCX_3CX_{4-5}CX_{10}CXCX_3GX_2CX_n$, where X represents any non-cysteine amino acid, and n is a variable repeat number. Non isolated peptides having this motif include TGF-α, amphiregulin, schwannoma-derived growth factor (SDGF), heparin-binding EGF-like growth factors and certain virally encoded peptides (e.g., Vaccinia virus, Reisner AH, *Nature* 313: 801-803 (1985), Shope fibroma virus, Chang W., et al., Mol Cell Biol. 7: 535-540 (1987), Molluscum contagiosum, Porter CD & Archard LC, *J. Gen. Virol.* 68: 673-682 (1987), and Myxoma virus, Upton C *et al., J. Virol.* 61: 1271-1275 (1987). Prigent SA & Lemoine N.R., *Prog. Growth Factor Rev.* 4: 1-24 (1992).

EGF-like domains are not confined to growth factors but have been observed in a variety of cell-surface and extracellular proteins which have interesting properties in cell adhesion, protein-protein interaction and development. Laurence DJR & Gusterson BA, *Tumor Biol.* 11: 229-261 (1990). These proteins include blood coagulation factors (factors VI, IX, X, XII, protein C, protein S, protein Z, tissue plasminogen activator, urokinase), extracellular matrix components (laminin, cytotactin, entactin), cell surface receptors (LDL receptor, thrombomodulin receptor) and immunity-related proteins (complement C1r, uromodulin).

Even more interesting, the general structure pattern of EGF-like precursors is preserved through lower organisms as well as in mammalian cells. A number of genes with developmental significance have been identified in invertebrates with EGF-like repeats. For example, the *notch* gene of Drosophila encodes 36 tandemly arranged 40 amino acid repeats which show homology to EGF, Wharton W *et al.*, *Cell* 43: 557-581 (1985). Hydropathy plots indicate a putative membrane spanning domain, with the EGF-related sequences being located on the extracellular side of the membrane. Other homeotic genes with EGF-like repeats include Delta, 95F and 5ZD which were identified using probes based on Notch, and the nematode gene *Lin-12* which encodes a putative receptor for a developmental signal transmitted between two specified cells.

Specifically, EGF has been shown to have potential in the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions, Konturek, *PC et al., Eur. J. Gastroenterol Hepatol.* 7 (10), 933-37 (1995), including the treatment of necrotizing enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration gastrointestinal ulcerations and congenital microvillus atrophy, A. Guglietta & PB Sullivan, *Eur. J. Gastroenterol Hepatol.* 7(10), 945-50 (1995). Additionally, EGF has been implicated in hair follicle differentiation; C.J.. du Cros, *J. Invest. Dermatol.* 101 (1 Suppl.), 106S-113S (1993), SG Hillier, *Clin. Endocrinol.* 33(4), 427-28 (1990); kidney function, L.L. Hamm *et al.*, *Semin. Nephrol.* 13 (1): 109-15 (1993), RC Harris, *Am. J. Kidney Dis.* 17(6): 627-30 (1991); tear fluid, GB van Setten *et al.*, *Int. Ophthalmol* 15(6); 359-62 (1991); vitamin K mediated blood coagulation, J. Stenflo *et al.*, *Blood* 78(7): 1637-51 (1991). EGF is also implicated various skin disease characterized by abnormal keratinocyte differentiation, e.g., psoriasis, epithelial cancers such as squamous cell carcinomas of the lung, epidermoid carcinoma of the vulva and gliomas. King, LE *et al.*, *Am. J. Med. Sci.* 296: 154-158 (1988).

Of great interest is mounting evidence that genetic alterations in growth factors signaling pathways are closely linked to developmental abnormalities and to chronic diseases including cancer. Aaronson SA, *Science* 254: 1146-1153 (1991). For example, c-erb-2 (also known as HER-2), a proto-oncogene with close structural similarity to EGF receptor protein, is overexpressed in human breast cancer. King *et al.*, *Science* 229: 974-976 (1985); Gullick, WJ, *Hormones and their actions*, Cooke BA *et al.*, eds, Amsterdam, Elsevier, pp 349-360 (1986).

17. **PRO317**

The TGF-β supergene family, or simply TGF-β superfamily, a group of secreted proteins, includes a large number of related growth and differentiation factors expressed in virtually all phyla. Superfamily members bind to specific cell surface receptors that activate signal transduction mechanisms to elicit their multifunctional cytokine effects. Kolodziejczyk and Hall, Biochem. Cell. Biol., 74: 299-314 (1996); Attisano

and Wrana, Cytokine Growth Factor Rev., 7: 327-339 (1996); and Hill, Cellular Signaling, 8: 533-544 (1996).

Members of this family include five distinct forms of TGF-β (Sporn and Roberts, in Peptide Growth Factors and Their Receptors, Sporn and Roberts, eds. (Springer-Verlag: Berlin, 1990) pp. 419-472), as well as the differentiation factors vg1 (Weeks and Melton, Cell, 51: 861-867 (1987)) and DPP-C polypeptide (Padgett *et al.*, Nature, 325: 81-84 (1987)), the hormones activin and inhibin (Mason *et al.*, Nature, 318: 659-663 (1985); Mason *et al.*, Growth Factors, 1: 77-88 (1987)), the Mullerian-inhibiting substance (MIS) (Cate *et al.*, Cell, 45: 685-698 (1986)), the bone morphogenetic proteins (BMPs) (Wozney *et al.*, Science, 242: 1528-1534 (1988); PCT WO 88/00205 published January 14, 1988; U.S. 4,877,864 issued October 31, 1989), the developmentally regulated proteins Vgr-1 (Lyons *et al.*, Proc. Natl. Acad. Sci. USA, 86: 4554-4558 (1989)) and Vgr-2 (Jones *et al.*, Molec. Endocrinol., 6: 1961-1968 (1992)), the mouse growth differentiation factor (GDF), such as GDF-3 and GDF-9 (Kingsley, Genes Dev., 8: 133-146 (1994); McPherron and Lee, J. Biol. Chem., 268: 3444-3449 (1993)), the mouse lefty/Stra1 (Meno *et al.*, Nature, 381: 151-155 (1996); Bouillet *et al.*, Dev. Biol., 170: 420-433 (1995)), glial cell line-derived neurotrophic factor (GDNF) (Lin *et al.*, Science, 260: 1130-1132 (1993), neurturin (Kotzbauer *et al.*, Nature, 384: 467-470 (1996)), and endometrial bleeding-associated factor (EBAF) (Kothapalli *et al.*, J. Clin. Invest., 99: 2342-2350 (1997)). The subset BMP-2A and BMP-2B is approximately 75% homologous in sequence to DPP-C and may represent the mammalian equivalent of that protein.

The proteins of the TGF-β superfamily are disulfide-linked homo- or heterodimers encoded by larger precursor polypeptide chains containing a hydrophobic signal sequence, a long and relatively poorly conserved N-terminal pro region of several hundred amino acids, a cleavage site (usually polybasic), and a shorter and more highly conserved C-terminal region. This C-terminal region corresponds to the processed mature protein and contains approximately 100 amino acids with a characteristic cysteine motif, *i.e.*, the conservation of seven of the nine cysteine residues of TGF-β among all known family members. Although the position of the cleavage site between the mature and pro regions varies among the family members, the C-terminus of all of the proteins is in the identical position, ending in the sequence Cys-X-Cys-X, but differing in every case from the TGF-β consensus C-terminus of Cys-Lys-Cys-Ser. Sporn and Roberts, 1990, *supra*.

There are at least five forms of TGF-β currently identified, TGF-β1, TGF-β2, TGF-β3, TGF-β4, and TGF-β5. The activated form of TGF-β1 is a homodimer formed by dimerization of the carboxy-terminal 112 amino acids of a 390 amino acid precursor. Recombinant TGF-β1 has been cloned (Derynck *et al.*, Nature, 316:701-705 (1985)) and expressed in Chinese hamster ovary cells (Gentry *et al.*, Mol. Cell. Biol., 7: 3418-3427 (1987)). Additionally, recombinant human TGF-β2 (deMartin *et al.*, EMBO J., 6: 3673 (1987)), as well as human and porcine TGF-β3 (Derynck *et al.*, EMBO J., 7: 3737-3743 (1988); ten Dijke *et al.*, Proc. Natl. Acad. Sci. USA, 85: 4715 (1988)) have been cloned. TGF-β2 has a precursor form of 414 amino acids and is also processed to a homodimer from the carboxy-terminal 112 amino acids that shares approximately 70% homology with the active form of TGF-β1 (Marquardt *et al.*, J. Biol. Chem., 262: 12127 (1987)). See also EP 200,341; 169,016; 268,561; and 267,463; U.S. Pat. No. 4,774,322; Cheifetz *et al.*, Cell, 48: 409-415 (1987); Jakowlew *et al.*, Molecular Endocrin., 2: 747-755 (1988); Derynck *et al.*, J. Biol. Chem., 261: 4377-4379 (1986); Sharples *et al.*, DNA, 6: 239-244 (1987); Derynck *et al.*, Nucl. Acids Res., 15: 3188-3189 (1987); Derynck *et al.*, Nucl. Acids. Res., 15: 3187 (1987); Seyedin *et al.*, J. Biol. Chem., 261: 5693-5695 (1986); Madisen *et*

al., DNA, 7: 1-8 (1988); and Hanks *et al.*, Proc. Natl. Acad. Sci. (U.S.A.), 85: 79-82 (1988).

TGF-β4 and TGF-β5 were cloned from a chicken chondrocyte cDNA library (Jakowlew *et al.*, Molec. Endocrinol., 2: 1186-1195 (1988)) and from a frog oocyte cDNA library, respectively.

The pro region of TGF-β associates non-covalently with the mature TGF-β dimer (Wakefield *et al.*, J. Biol. Chem., 263: 7646-7654 (1988); Wakefield *et al.*, Growth Factors, 1: 203-218 (1989)), and the pro regions are found to be necessary for proper folding and secretion of the active mature dimers of both TGF-β and activin (Gray and Mason, Science, 247: 1328-1330 (1990)). The association between the mature and pro regions of TGF-β masks the biological activity of the mature dimer, resulting in formation of an inactive latent form. Latency is not a constant of the TGF-β superfamily, since the presence of the pro region has no effect on activin or inhibin biological activity.

A unifying feature of the biology of the proteins from the TGF-β superfamily is their ability to regulate developmental processes. TGF-β has been shown to have numerous regulatory actions on a wide variety of both normal and neoplastic cells. TGF-β is multifunctional, as it can either stimulate or inhibit cell proliferation, differentiation, and other critical processes in cell function (Sporn and Roberts, *supra*).

One member of the TGF-β superfamily, EBAF, is expressed in endometrium only in the late secretory phase and during abnormal endometrial bleeding. Kothapalli *et al.*, J. Clin. Invest., 99: 2342-2350 (1997). Human endometrium is unique in that it is the only tissue in the body that bleeds at regular intervals. In addition, abnormal endometrial bleeding is one of the most common manifestations of gynecological diseases, and is a prime indication for hysterectomy. *In situ* hybridization showed that the mRNA of EBAF was expressed in the stroma without any significant mRNA expression in the endometrial glands or endothelial cells.

The predicted protein sequence of EBAF showed a strong homology to the protein encoded by mouse *lefty/stra3* of the TGF-β superfamily. A motif search revealed that the predicted EBAF protein contains most of the cysteine residues which are conserved among the TGF-β-related proteins and which are necessary for the formation of the cysteine knot structure. The EBAF sequence contains an additional cysteine residue, 12 amino acids upstream from the first conserved cysteine residue. The only other family members known to contain an additional cysteine residue are TGF-βs, inhibins, and GDF-3. EBAF, similar to LEFTY, GDF-3/Vgr2, and GDF-9, lacks the cysteine residue that is known to form the intermolecular disulfide bond. Therefore, EBAF appears to be an additional member of the TGF-β superfamily with an unpaired cysteine residue that may not exist as a dimer. However, hydrophobic contacts between the two monomer subunits may promote dimer formation. Fluorescence *in situ* hybridization showed that the *ebaf* gene is located on human chromosome 1 at band q42.1.

Additional members of the TGF-β superfamily, such as those related to EBAF, are being searched for by industry and academics. We herein describe the identification and characterization of novel polypeptides having homology to EBAF, designated herein as PRO317 polypeptides.

18. PRO301

The widespread occurrence of cancer has prompted the devotion of considerable resources and discovering new treatments of treatment. One particular method involves the creation of tumor or cancer specific

monoclonal antibodies (mAbs) which are specific to tumor antigens. Such mAbs, which can distinguish between normal and cancerous cells are useful in the diagnosis, prognosis and treatment of the disease. Particular antigens are known to be associated with neoplastic diseases, such as colorectal cancer.

One particular antigen, the A33 antigen is expressed in more than 90% of primary or metastatic colon cancers as well as normal colon epithelium. Since colon cancer is a widespread disease, early diagnosis and treatment is an important medical goal. Diagnosis and treatment of colon cancer can be implemented using monoclonal antibodies (mAbs) specific therefore having fluorescent, nuclear magnetic or radioactive tags. Radioactive gene, toxins and/or drug tagged mAbs can be used for treatment *in situ* with minimal patient description. mAbs can also be used to diagnose during the diagnosis and treatment of colon cancers. For example, when the serum levels of the A33 antigen are elevated in a patient, a drop of the levels after surgery would indicate the tumor resection was successful. On the other hand, a subsequent rise in serum A33 antigen levels after surgery would indicate that metastases of the original tumor may have formed or that new primary tumors may have appeared. Such monoclonal antibodies can be used in lieu of, or in conjunction with surgery and/or other chemotherapies. For example, U.S.P. 4,579,827 and U.S.S.N. 424,991 (E.P. 199,141) are directed to therapeutic administration of monoclonal antibodies, the latter of which relates to the application of anti-A33 mAb.

Many cancers of epithelial origin have adenovirus receptors. In fact, adenovirus-derived vectors have been proposed as a means of inserting antisense nucleic acids into tumors (U.S.P. 5,518,885). Thus, the association of viral receptors with neoplastic tumors is not unexpected.

We herein describe the identification and characterization of novel polypeptides having homology to certain cancer-associated antigens, designated herein as PRO301 polypeptides.

19. **PRO224**

Cholesterol uptake can have serious implications on one's health. Cholesterol uptake provides cells with most of the cholesterol they require for membrane synthesis. If this uptake is blocked, cholesterol accumulates in the blood and can contribute to the formation of atherosclerotic plaques in blood vessel walls. Most cholesterol is transported in the blood bound to protein in the form of complexes known as low-density lipoproteins (LDLs). LDLs are endocytosed into cells via LDL receptor proteins. Therefore, LDL receptor proteins, and proteins having homology thereto, are of interest to the scientific and medical communities.

Membrane-bound proteins and receptors can play an important role in the formation, differentiation and maintenance of multicellular organisms. The LDL receptors are an example of membrane-bound proteins which are involved in the synthesis and formation of cell membranes, wherein the health of an individual is affected directly and indirectly by its function. Many membrane-bound proteins act as receptors such as the LDL receptor. These receptors can function to endocytose substrates or they can function as a receptor for a channel. Other membrane-bound proteins function as signals or antigens.

Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule regulators of the relevant receptor/ligand interaction. In the case of the LDL

receptor, it is desirable to find molecules which enhance endocytosis so as to lower blood cholesterol levels and plaque formation. It is also desirable to identify molecules which inhibit endocytosis so that these molecules can be avoided or regulated by individuals having high blood cholesterol. Polypeptides which are homologous to lipoprotein receptors but which do not function as lipoprotein receptors are also of interest in the determination of the function of the fragments which show homology.

The following studies report on previously known low density lipoprotein receptors and related proteins including apolipoproteins: Sawamura, et al., Nippon Chemiphar Co. Japan patent application J09098787; Novak, S., et al., J. Biol. Chem., 271:(20)11732-6 (1996); Blaas, D., J. Virol., 69(11)7244-7 (Nov. 1995); Scott, J., J. Inherit. Metab. Dis. (UK), 9/Supp. 1 (3-16) (1986); Yamamoto, et al., Cell, 39:27-38 (1984); Rebece, et al., Neurobiol. Aging, 15:5117 (1994); Novak, S., et al., J. Biol. Chemistry, 271:11732-11736 (1996); and Sestavel and Fruchart, Cell Mol. Biol., 40(4):461-81 (June 1994). These publications and others published prior to the filing of this application provide further background to peptides already known in the art.

Efforts are being undertaken by both industry and academia to identify new, native membrane-bound receptor proteins, particularly those having homology to lipoprotein receptors. We herein describe the identification and characterization of novel polypeptides having homology to lipoprotein receptors, designated herein as PRO224 polypeptides.

20. PRO222

Complement is a group of proteins found in the blood that are important in humoral immunity and inflammation. Complement proteins are sequentially activated by antigen-antibody complexes or by proteolytic enzymes. When activated, complement proteins kill bacteria and other microorganisms, affect vascular permeability, release histamine and attract white blood cells. Complement also enhances phagocytosis when bound to target cells. In order to prevent harm to autologous cells, the complement activation pathway is tightly regulated.

Deficiencies in the regulation of complement activation or in the complement proteins themselves may lead to immune-complex diseases, such as systemic lupus erythematosus, and may result in increased susceptibility to bacterial infection. In all cases, early detection of complement deficiency is desirable so that the patient can begin treatment. Thus, research efforts are currently directed toward identification of soluble and membrane proteins that regulate complement activation.

Proteins known to be important in regulating complement activation in humans include Factor H and Complement receptor type 1 (CR1). Factor H is a 150 kD soluble serum protein that interacts with complement protein C3b to accelerate the decay of C3 convertase and acts as a cofactor for Factor I-mediated cleavage of complement protein C4b. Complement receptor type 1 is a 190-280 kD membrane bound protein found in mast cells and most blood cells. CR1 interacts with complement proteins C3b, C4b, and iC3b to accelerate dissociation of C3 convertases, acts as a cofactor for Factor I-mediated cleavage of C3b and C4b, and binds immune complexes and promotes their dissolution and phagocytosis.

Proteins which have homology to complement proteins are of particular interest to the medical and industrial communities. Often, proteins having homology to each other have similar function. It is also of

interest when proteins having homology do not have similar functions, indicating that certain structural motifs identify information other than function, such as locality of function.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound proteins, particularly those having homology to known proteins involved in the complement pathway. Proteins involved in the complement pathway were reviewed in Birmingham DJ (1995), Critical Reviews in Immunology, 15(2):133-154 and in Abbas AK, et al. (1994) Cellular and Molecular Immunology, 2nd Ed. W.B. Saunders Company, Philadelphia, pp 295-315.

We herein describe the identification and characterization of novel polypeptides having homology to complement receptors, designated herein as PRO222 polypeptides.

21. PRO234

The successful function of many systems within multicellular organisms is dependent on cell-cell interactions. Such interactions are affected by the alignment of particular ligands with particular receptors in a manner which allows for ligand-receptor binding and thus a cell-cell adhesion. While protein-protein interactions in cell recognition have been recognized for some time, only recently has the role of carbohydrates in physiologically relevant recognition been widely considered (see B.K. Brandley *et al., J. Leuk. Biol.* 40: 97 (1986) and N. Sharon *et al., Science* 246: 227 (1989). Oligosaccharides are well positioned to act as recognition novel lectins due to their cell surface location and structural diversity. Many oligosaccharide structures can be created through the differential activities of a smaller number of glycosyltransferases. The diverse structures of oligosaccharides can be generated by transcription of relatively few gene products, which suggests that the oligosaccharides are a plausible mechanism by which is directed a wide range of cell-cell interactions. Examples of differential expression of cell surface carbohydrates and putative carbohydrate binding proteins (lectins) on interacting cells have been described (J. Dodd & T.M. Jessel, *J. Neurosci.* 5: 3278 (1985); L.J. Regan *et al., Proc. Natl. Acad. Sci. USA* 83: 2248 (1986); M. Constantine-Paton *et al., Nature* 324: 459 (1986); and M. Tiemeyer *et al., J. Biol. Chem.* 263: 1671 (1989). One interesting member of the lectin family are selectins.

The migration of leukocytes to sites of acute or chronic inflammation involves adhesive interactions between these cells and the endothelium. This specific adhesion is the initial event in the cascade that is initiated by inflammatory insults, and it is, therefore, of paramount importance to the regulated defense of the organism.

The types of cell adhesion molecules that are involved in the interaction between leukocytes and the endothelium during an inflammatory response currently stands at four: (1) selectins; (2) (carbohydrate and glycoprotein) ligands for selectins; (3) integrins; and (4) integrin ligands, which are members of the immunoglobulin gene superfamily.

The selectins are cell adhesion molecules that are unified both structurally and functionally. Structurally, selectins are characterized by the inclusion of a domain with homology to a calcium-dependent lectin (C-lectins), an epidermal growth factor (egf)-like domain and several complement binding-like domains. Bevilacqua, M.P. *et al., Science* 243: 1160-1165 (1989); Johnston *et al., Cell* 56: 1033-1044 (1989); Lasky *et al, Cell* 56: 1045-1055 (1989); Siegalman, M. *et al., Science* 243: 1165-1172 (1989); Stoolman, L.M., *Cell* 56: 907-910 (1989). Functionally, selectins share the common property of their ability to mediate cell binding

through interactions between their lectin domains and cell surface carbohydrate ligands (Brandley, B, *et al.*, *Cell* 63, 861-863 (1990); Springer, T. and Lasky, L.A., *Nature* 349, 19-197 (1991); Bevilacqua, M.P. and Nelson, R.M., *J. Clin. Invest.* 91 379-387 (1993) and Tedder *et al.*, *J. Exp. Med.* 170: 123-133 (1989).

There are three members identified so far in the selectin family of cell adhesion molecules: L-selectin (also called peripheral lymph node homing receptor (pnHR), LEC-CAM-1, LAM-1, gp90 ^{MEL}, gp100^{MEL}, gp110^{MEL}, MEL-14 antigen, Leu-8antigen, TQ-1 antigen, DREG antigen), E-selectin (LEC-CAM-2, LECAM-2, ELAM-1) and P-selectin (LEC-CAM-3, LECAM-3, GMP-140, PADGEM).

The identification of the C-lectin domain has led to an intense effort to define carbohydrate binding ligands for proteins containing such domains. E-selectin is believed to recognize the carbohydrate sequence NeuNAcα2-3Galβ1-4(Fucα1-3)GlcNAc (sialyl-Lewis x, or sLe^x) and related oligosaccharides, Berg *et al.*, *J. Biol. Chem.* 265: 14869-14872 (1991); Lowe *et al.*, *Cell* 63: 475-484 (1990); Phillips *et al.*, *Science* 250: 1130-1132 (1990); *Tiemeyer et al., Proc. Natl. Acad. Sci. USA* 88: 1138-1142 (1991).

L-selectin, which comprises a lectin domain, performs its adhesive function by recognizing carbohydrate-containing ligands on endothelial cells. L-selectin is expressed on the surface of leukocytes, such as lymphocytes, neutrophils, monocytes and eosinophils, and is involved with the trafficking of lymphocytes to peripheral lymphoid tissues (*Gallatin et al., Nature* 303: 30-34 (1983)) and with acute neutrophil-mediated inflammatory responses (Watson, S.R., *Nature* 349: 164-167 (1991)). The amino acid sequence of L-selectin and the encoding nucleic acid sequence are, for example, disclosed in U.S. patent No. 5,098,833 issued 24 March 1992.

L-selectin (LECAM-1) is particularly interesting because of its ability to block neutrophil influx (Watson *et al., Nature* 349: 164-167 (1991). It is expressed in chronic lymphocytic leukemia cells which bind to HEV (Spertini *et al., Nature* 349: 691-694 (1991). It is also believed that HEV structures at sites of chronic inflammation are associated with the symptoms of diseases such as rheumatoid arthritis, psoriasis and multiple sclerosis.

E-selectin (ELAM-1), is particularly interesting because of its transient expression on endothelial cells in response to IL-1 or TNF. Bevilacqua et al., *Science* 243: 1160 (1989). The time course of this induced expression (2-8 h) suggests a role for this receptor in initial neutrophil induced extravasation in response to infection and injury. It has further been reported that anti-ELAM-1 antibody blocks the influx of neutrophils in a primate asthma model and thus is beneficial for preventing airway obstruction resulting from the inflammatory response. Gundel *et al., J. Clin. Invest.* 88: 1407 (1991).

The adhesion of circulating neutrophils to stimulated vascular endothelium is a primary event of the inflammatory response. P-selectin has been reported to recognize the Lewis x structure (Galβ1-4(Fucα1-3) GlcNAc), Larsen *et al., Cell* 63: 467-474(1990). Others report that an additional terminal linked sialic acid is required for high affinity binding, Moore *et al., J. Cell. Biol.* 112: 491-499 (1991). P-selectin has been shown to be significant in acute lung injury. Anti-P-selectin antibody has been shown to have strong protective effects in a rodent lung injury model. M.S. Mulligan *et al., J. Clin. Invest.* 90: 1600 (1991).

We herein describe the identification and characterization of novel polypeptides having homology to lectin proteins, herein designated as PRO234 polypeptides.

22. PRO231

Some of the most important proteins involved in the above described regulation and modulation of cellular processes are the enzymes which regulate levels of protein phosphorylation in the cell. For example, it is known that the transduction of signals that regulate cell growth and differentiation is regulated at least in part by phosphorylation and dephosphorylation of various cellular proteins. The enzymes that catalyze these processes include the protein kinases, which function to phosphorylate various cellular proteins, and the protein phosphatases, which function to remove phosphate residues from various cellular proteins. The balance of the level of protein phosphorylation in the cell is thus mediated by the relative activities of these two types of enzymes.

Protein phosphatases represent a growing family of enzymes that are found in many diverse forms, including both membrane-bound and soluble forms. While many protein phosphatases have been described, the functions of only a very few are beginning to be understood (Tonks, *Semin. Cell Biol.* 4:373-453 (1993) and Dixon, *Recent Prog. Horm. Res.* 51:405-414 (1996)). However, in general, it appears that many of the protein phosphatases function to modulate the positive or negative signals induced by various protein kinases. Therefore, it is likely that protein phosphatases play critical roles in numerous and diverse cellular processes.

Given the physiological importance of the protein phosphatases, efforts are being undertaken by both industry and academia to identify new, native phosphatase proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel phosphatase proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to acid phosphatases, designated herein as PRO231 polypeptides.

23. PRO229

Scavenger receptors are known to protect IgG molecules from catabolic degradation. Riechmann and Hollinger, Nature Biotechnology, 15:617 (1997). In particular, studies of the CH2 and CH3 domains have shown that specific sequences of these domains are important in determining the half-lives of antibodies. Ellerson, et al., J. Immunol., 116: 510 (1976); Yasmeen, et al., J. Immunol. 116: 518 (1976; Pollock, et al., Eur. J. Immunol., 20: 2021 (1990). Scavenger receptor proteins and antibodies thereto are further reported in U.S. Patent No. 5,510,466 to Krieger, et al. Due to the ability of scavenger receptors to increase the half-life of polypeptides and their involvement in immune function, molecules having homology to scavenger receptors are of importance to the scientific and medical community.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly those having homology to scavenger receptors. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to scavenger receptors, designated herein as PRO229 polypeptides.

24. PRO238

Oxygen free radicals and antioxidants appear to play an important role in the central nervous system after cerebral ischemia and reperfusion. Moreover, cardiac injury, related to ischaemia and reperfusion has been reported to be caused by the action of free radicals. Additionally, studies have reported that the redox state of the cell is a pivotal determinant of the fate of the cells. Furthermore, reactive oxygen species have been reported to be cytotoxic, causing inflammatory disease, including tissue necrosis, organ failure, atherosclerosis, infertility, birth defects, premature aging, mutations and malignancy. Thus, the control of oxidation and reduction is important for a number of reasons including for control and prevention of strokes, heart attacks, oxidative stress and hypertension. In this regard, reductases, and particularly, oxidoreductases, are of interest. Publications further describing this subject matter include Kelsey, et al., Br. J. Cancer, 76(7):852-4 (1997); Friedrich and Weiss, J. Theor. Biol., 187(4):529-40 (1997) and Pieulle, et al., J. Bacteriol., 179(18)::5684-92 (1997).

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly secreted proteins which have homology to reductase. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to reductase, designated herein as PRO238 polypeptides.

25. PRO233

Studies have reported that the redox state of the cell is an important determinant of the fate of the cell. Furthermore, reactive oxygen species have been reported to be cytotoxic, causing inflammatory disease, including tissue necrosis, organ failure, atherosclerosis, infertility, birth defects, premature aging, mutations and malignancy. Thus, the control of oxidation and reduction is important for a number of reasons, including the control and prevention of strokes, heart attacks, oxidative stress and hypertension. Oxygen free radicals and antioxidants appear to play an important role in the central nervous system after cerebral ischemia and reperfusion. Moreover, cardiac injury, related to ischaemia and reperfusion has been reported to be caused by the action of free radicals. In this regard, reductases, and particularly, oxidoreductases, are of interest. In addition, the transcription factors, NF-kappa B and AP-1, are known to be regulated by redox state and to affect the expression of a large variety of genes thought to be involved in the pathogenesis of AIDS, cancer, atherosclerosis and diabetic complications. Publications further describing this subject matter include Kelsey, et al., Br. J. Cancer, 76(7):852-4 (1997); Friedrich and Weiss, J. Theor. Biol., 187(4):529-40 (1997) and Pieulle, et al., J. Bacteriol., 179(18):5684-92 (1997). Given the physiological importance of redox reactions *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in redox

reactions. We describe herein the identification of novel polypeptides which have homology to reductase, designated herein as PRO233 polypeptides.

26. PRO223

The carboxypeptidase family of exopeptidases constitutes a diverse group of enzymes that hydrolyze carboxyl-terminal amide bonds in polypeptides, wherein a large number of mammalian tissues produce these enzymes. Many of the carboxypeptidase enzymes that have been identified to date exhibit rather strong cleavage specificities for certain amino acids in polypeptides. For example, carboxypeptidase enzymes have been identified which prefer lysine, arginine, serine or amino acids with either aromatic or branched aliphatic side chains as substrates at the carboxyl terminus of the polypeptide.

With regard to the serine carboxypeptidases, such amino acid specific enzymes have been identified from a variety of different mammalian and non-mammalian organisms. The mammalian serine carboxypeptidase enzymes play important roles in many different biological processes including, for example, protein digestion, activation, inactivation, or modulation of peptide hormone activity, and alteration of the physical properties of proteins and enzymes.

In light of the physiological importance of the serine carboxypeptidases, efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins and specifically novel carboxypeptidases. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. We describe herein novel polypeptides having homology to one or more serine carboxypeptidase polypeptides, designated herein as PRO223 polypeptides.

27. PRO235

Plexin was first identified in Xenopus tadpole nervous system as a membrane glycoprotein which was shown to mediate cell adhesion via a homophilic binding mechanism in the presence of calcium ions. Strong evolutionary conservation between Xenopus, mouse and human homologs of plexin has been observed. [Kaneyama et al., Biochem. And Biophys. Res. Comm. 226: 524-529 (1996)]. Given the physiological importance of cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in cell adhesion. We describe herein the identification of a novel polypeptide which has homology to plexin, designated herein as PRO235.

28. PRO236 and PRO262

β-galactosidase is a well known enzymatic protein which functionsto hydrolyze β-galactoside molecules. β-galactosidase has been employed for a variety of different applications, both *in vitro* and *in vivo* and has proven to be an extremely useful research tool. As such, there is an interest in obtaining novel polypeptides which exhibit homology to the β-galactosidase polypeptide.

Given the strong interest in obtaining novel polypeptides having homology to β-galactosidase, efforts are currently being undertaken by both industry and academia to identify new, native β-galactosidase homolog

proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel β-galactosidase-like proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)]. We herein describe novel poylpeptides having siginificant homology to the β-galactosidase enzyme, designated herein as PRO236 and PRO262 polypeptides.

29. **PRO239**

Densin is a glycoprotein which has been isolated from the brain which has all the hallmarks of an adhesion molecule. It is highly concentrated at synaptic sites in the brain and is expressed prominently in dendritic processes in developing neurons. Densin has been characterized as a member of the O-linked sialoglycoproteins. Densin has relevance to medically important processes such as regeneration. Given the physiological importance of synaptic processes and cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in synaptic machinery and cell adhesion. We describe herein the identification of novel polypeptides which have homology to densin, designated herein as PRO239 polypeptides.

30. **PRO257**

Ebnerin is a cell surface protein associated with von Ebner glands in mammals. Efforts are being undertaken by both industry and academia to identify new, native cell surface receptor proteins and specifically those which possess sequence homology to cell surface proteins such as ebnerin. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor proteins. We herein describe the identification of novel polypeptides having significant homology to the von Ebner's gland-associated protein ebnerin, designated herein as PRO257 polypeptides.

31. **PRO260**

Fucosidases are enzymes that remove fucose residues from fucose containing proteoglycans. In some pathological conditions, such as cancer, rheumatoid arthritis, and diabetes, there is an abnormal fucosylation of serum proteins. Therefore, fucosidases, and proteins having homology to fucosidase, are of importance to the study and abrogation of these conditions. In particular, proteins having homology to the alpha-l-fucosidase precursor are of interest. Fucosidases and fucosidase inhibitors are further described in U.S. Patent Nos. 5,637,490, 5,382,709, 5,240,707, 5,153,325, 5,100,797, 5,096,909 and 5,017,704. Studies are also reported in Valk, et al., J. Virol., 71(9):6796 (1997), Aktogu, et al., Monaldi, Arch. Chest Dis. (Italy), 52(2):118 (1997) and Focarelli, et al., Biochem. Biophys. Res. Commun. (U.S.), 234(1):54 (1997).

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins. Of particular interest are proteins having homology to the alpha-l-fucosidase precursor. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-

7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to fucosidases, designated herein as PRO260 polypeptides.

32. **PRO263**

CD44 is a cell surface adhesion molecule involved in cell-cell and cell-matrix interactions. Hyaluronic acid, a component of the extracellular matrix is a major ligand. Other ligands include collagen, fibronectin, laminin, chrondroitin sulfate, mucosal addressin, serglycin and osteoponin. CD44 is also important in regulating cell traffic, lymph node homing, transmission of growth signals, and presentation of chemokines and growth factors to traveling cells. CD44 surface proteins are associated with metastatic tumors and CD44 has been used as a marker for HIV infection. Certain splice variants are associated with metastasis and poor prognosis of cancer patients. Therefore, molecules having homology with CD44 are of particular interest, as their homology indicates that they may have functions related to those functions of CD44. CD44 is further described in U.S. Patent Nos. 5,506,119, 5,504,194 and 5,108,904; Gerberick, et al., Toxicol. Appl. Pharmacol., 146(1):1 (1997); Wittig, et al., Immunol. Letters (Netherlands), 57(1-3):217 (1997); and Oliveira and Odell, Oral Oncol. (England), 33(4):260 (1997).

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly transmembrane proteins with homology to CD44 antigen. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to CD44 antigen, designated herein as PRO263 polypeptides.

33. **PRO270**

Thioredoxins effect reduction-oxidation (redox) state. Many diseases are potentially related to redox state and reactive oxygen species may play a role in many important biological processes. The transcription factors, NF-kappa B and AP-1, are regulated by redox state and are known to affect the expression of a large variety of genes thought to be involved in the pathogenesis of AIDS, cancer, atherosclerosis and diabetic complications. Such proteins may also play a role in cellular antioxidant defense, and in pathological conditions involving oxidative stress such as stroke and inflammation in addition to having a role in apoptosis. Therefore, thioredoxins, and proteins having homology thereto, are of interest to the scientific and medical communities.

We herein describe the identification and characterization of novel polypeptides having homology to thioredoxin, designated herein as PRO270 polypeptides.

34. **PRO271**

The proteoglycan link protein is a protein which is intimately associated with various extracellular

matrix proteins and more specifically with proteins such as collagen. For example, one primary component of collagen is a large proteoglycan called aggrecan. This molecule is retained by binding to the glycosaminoglycan hyaluronan through the amino terminal G1 globular domain of the core protein. This binding is stabilized by the proteoglycan link protein which is a protein that is also associated with other tissues containing hyaluronan binding proteoglycans such as versican.

Link protein has been identified as a potential target for autoimmune antibodies in individuals who suffer from juvenile rheumatoid arthritis (see Guerassimov et al., *J. Rheumatology* 24(5):959-964 (1997)). As such, there is strong interest in identifying novel proteins having homology to link protein. We herein describe the identification and characterization of novel polypeptides having such homology, designated herein as PRO271 polypeptides.

35. **PRO272**

Reticulocalbin is an endoplasmic reticular protein which may be involved in protein transport and luminal protein processing. Reticulocalbin resides in the lumen of the endoplasmic reticulum, is known to bind calcium, and may be involved in a luminal retention mechanism of the endoplasmic reticulum. It contains six domains of the EF-hand motif associated with high affinity calcium binding. We describe herein the identification and characterization of a novel polypeptide which has homology to the reticulocalbin protein, designated herein as PRO272.

36. **PRO294**

Collagen, a naturally occurring protein, finds wide application in industry. Chemically hydrolyzed natural collagen can be denatured and renatured by heating and cooling to produce gelatin, which is used in photographic and medical, among other applications. Collagen has important properties such as the ability to form interchain aggregates having a conformation designated as a triple helix. We herein describe the identification and characterization of a novel polypeptide which has homology to portions of the collagen molecule, designated herein as PRO294.

37. **PRO295**

The integrins comprise a supergene family of cell-surface glycoprotein receptors that promote cellular adhesion. Each cell has numerous receptors that define its cell adhesive capabilities. Integrins are involved in a wide variety of interaction between cells and other cells or matrix components. The integrins are of particular importance in regulating movement and function of immune system cells The platelet IIb/IIIA integrin complex is of particular importance in regulating platelet aggregation. A member of the integrin family, integrin β-6, is expressed on epithelial cells and modulates epithelial inflammation. Another integrin, leucocyte-associated antigen-1 (LFA-1) is important in the adhesion of lymphocytes during an immune response. The integrins are expressed as heterodimers of non-covalently associated alpha and beta subunits. Given the physiological importance of cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in cell adhesion. We describe herein the identification and characterization of a

novel polypeptide which has homology to integrin, designated herein as PRO295.

38. PRO293

Protein-proteininteractions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglubular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factorβ involvement for treatment for cancer, wound healing and scarring).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known neuronal leucine rich repeat proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113

(1996); U.S. Patent No. 5,536,637)].

We describe herein the identification and characterization of a novel polypeptide which has homology to leucine rich repeat proteins, designated herein as PRO293.

39. **PRO247**

Protein-proteininteractions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglubular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during cotogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factor β involvement for treatment for cancer, wound healing and scarring).

Densin is a glycoprotein which has been isolated from the brain which has all the hallmarks of an adhesion molecule. It is highly concentrated at synaptic sites in the brain and is expressed prominently in dendritic processes in developing neurons. Densin has been characterized as a member of the O-linked sialoglycoproteins. Densin has relevance to medically important processes such as regeneration. Given the

physiological importance of synaptic processes and cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in synaptic machinery and cell adhesion. Densin is further described in Kennedy, M.B, Trends Neurosci. (England). 20(6):264 (1997) and Apperson, et al., J. Neurosci., 16(21):6839 (1996).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as KIAA0231 and densin. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We describe herein the identification and characterization of a novel polypeptide which has homology to leucine rich repeat proteins, designated herein as PRO247.

40. **PRO302, PRO303, PRO304, PRO307 and PRO343**

Proteases are enzymatic proteins which are involved in a large number of very important biological processes in mammalian and non-mammalian organisms. Numerous different protease enzymes from a variety of different mammalian and non-mammalian organisms have been both identified and characterized. The mammalian protease enzymes play important roles in many different biological processes including, for example, protein digestion, activation, inactivation, or modulation of peptide hormone activity, and alteration of the physical properties of proteins and enzymes.

In light of the important physiological roles played by protease enzymes, efforts are currently being undertaken by both industry and academia to identify new, native protease homologs. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)]. We herein describe the identification of novel polypeptides having homology to various protease enzymes, designated herein as PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides.

41. **PRO328**

The GLIP protein family has been characterized as comprising zinc-finger proteins which play important roles in embryogenesis. These proteins may function as transcriptional regulatory proteins and are known to be amplified in a subset of human tumors. Glioma pathogenesis protein is structurally related to a group of plant pathogenesis-related proteins. It is highly expressed in glioblastoma. See US Pat. Nos. 5,582,981 (issued Dec. 10, 1996) and 5,322,801 (issued June 21, 1996), Ellington, A.D. et al., Nature, 346:818 (1990), Grindley, J.C. et al., Dev. Biol., 188(2):337 (1997), Marine, J.C. et al., Mech. Dev., 63(2):211 (1997), The CRISP or cysteine rich secretory protein family are a group of proteins which are also structurally related to a group of plant pathogenesis proteins. [Schwidetzky, U., Biochem. J., 321:325 (1997), Pfisterer, P., Mol. Cell Biol.,

16(11):6160 (1996), Kratzschmar, J., Eur. J. Biochem., 236(3):827 (1996)]. We describe herein the identification of a novel polypeptide which has homology to GLIP and CRISP, designated herein as PRO328 polypeptides.

42. PRO335, PRO331 and PRO326

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglubular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome, Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation reporting that decorin binding to transforming growth factorβ has involvement in a treatment for cancer, wound healing and scarring. Related by function to this group of proteins is the insulin like growth factor (IGF), in that it is useful in wound-healing and associated therapies concerned with re-growth of tissue, such as connective tissue, skin and bone; in promoting body growth in humans and animals; and in stimulating other growth-related processes. The acid labile subunit of IGF (ALS) is also of interest in that it increases the half-life of IGF and is part of the IGF complex in vivo.

Another protein which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Of particular interest is LIG-1, a membrane glycoprotein that is expressed specifically in glial cells in the mouse brain, and has leucine rich repeats and immunoglobulin-like domains. Suzuki, et al., J. Biol. Chem. (U.S.), 271(37):22522 (1996). Other studies reporting on the biological functions of proteins having leucine rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789

(July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as LIG-1, ALS and decorin. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We describe herein the identification and characterization of novel polypeptides which have homology to proteins of the leucine rich repeat superfamily, designated herein as PRO335, PRO331 and PRO326 polypeptides.

43. PRO332

Secreted proteins comprising a repeat characterized by an arrangement of conserved leucine residues (leucine-rich repeat motif) have diverse biological roles. Certain proteoglycans, such as biglycan, fibromodulin and decorin, are, for example, characterized by the presence of a leucine-rich repeat of about 24 amino acids [Ruoslahti, Ann. Rev. Cell. Biol. 4 229-255 (1988); Oldberg *et al.*, EMBO J. 8, 2601-2604 (1989)]. In general, proteoglycans are believed to play a role in regulating extracellular matrix, cartilage or bone function. The proteoglycan decorin binds to collagen type I and II and affects the rate of fibril formation. Fibromodulin also binds collagen and delays fibril formation. Both fibromodulin and decorin inhibit the activity of transforming growth factor beta (TGF-β) (U.S. Patent No. 5,583,103 issued December 10, 1996). TGF-β is known to play a key role in the induction of extracellular matrix and has been implicated in the development of fibrotic diseases, such as cancer and glomerulonephritis. Accordingly, proteoglycans have been proposed for the treatment of fibrotic cancer, based upon their ability to inhibit TGF-β's growth stimulating activity on the cancer cell. Proteoglycans have also been described as potentially useful in the treatment of other proliferative pathologies, including rheumatoid arthritis, arteriosclerosis, adult respiratory distress syndrome, cirrhosis of the liver, fibrosis of the lungs, post-myocardial infarction, cardiac fibrosis, post-angioplasty restenosis, renal interstitial fibrosis and certain dermal fibrotic conditions, such as keloids and scarring, which might result from burn injuries, other invasive skin injuries, or cosmetic or reconstructive surgery (U.S. Patent No. 5,654,270, issued August 5, 1997).

We describe herein the identification and characterization of novel polypeptides which have homology to proteins of the leucine rich repeat superfamily, designated herein as PRO332 polypeptides.

44. PRO334

Microfibril bundles and proteins found in association with these bundles, particularly attachment molecules, are of interest in the field of dermatology, particularly in the study of skin which has been damaged from aging, injuries or the sun. Fibrillin microfibrils define the continuous elastic network of skin, and are

present in dermis as microfibril bundles devoid of measurable elastin extending from the dermal-epithelial junction and as components of the thick elastic fibres present in the deep reticular dermis. Moreover, Marfan syndrome has been linked to mutations which interfere with multimerization of fibrillin monomers or other connective tissue elements.

Fibulin-1 is a modular glycoprotein with amino-terminal anaphlatoxin-like modules followed by nine epidermal growth factor (EGF)-like modules and, depending on alternative splicing, four possible carboxyl termini. Fibulin-2 is a novel extracellular matrix protein frequently found in close association with microfibrils containing either fibronectin or fibrillin. Thus, fibrillin, fibulin, and molecules related thereto are of interest, particularly for the use of preventing skin from being damaged from aging, injuries or the sun, or for restoring skin damaged from same. Moreover, these molecules are generally of interest in the study of connective tissue and attachment molecules and related mechanisms. Fibrillin, fibulin and related molecules are further described in Adams, et al., J. Mol. Biol., 272(2):226-36 (1997); Kielty and Shuttleworth, Microsc. Res. Tech., 38(4):413-27 (1997); and Child, J. Card. Surg., 12(2Supp.):131-5 (1997).

Currently, efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly secreted proteins which have homology to fibulin and fibrillin. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to fibulin and fibrillin, designated herein as PRO334 polypeptides.

45. **PRO346**

The widespread occurrence of cancer has prompted the devotion of considerable resources and discovering new treatments of treatment. One particular method involves the creation of tumor or cancer specific monoclonal antibodies (mAbs) which are specific to tumor antigens. Such mAbs, which can distinguish between normal and cancerous cells are useful in the diagnosis, prognosis and treatment of the disease. Particular antigens are known to be associated with neoplastic diseases, such as colorectal and breast cancer. Since colon cancer is a widespread disease, early diagnosis and treatment is an important medical goal. Diagnosis and treatment of cancer can be implemented using monoclonal antibodies (mAbs) specific therefore having fluorescent, nuclear magnetic or radioactive tags. Radioactive genes, toxins and/or drug tagged mAbs can be used for treatment *in situ* with minimal patient description.

Carcinoembryonic antigen (CEA) is a glycoprotein found in human colon cancer and the digestive organs of a 2-6 month human embryos. CEA is a known human tumor marker and is widely used in the diagnosis of neoplastic diseases, such as colon cancer. For example, when the serum levels of CEA are elevated in a patient, a drop of CEA levels after surgery would indicate the tumor resection was successful. On the other hand, a subsequent rise in serum CEA levels after surgery would indicate that metastases of the original tumor may have formed or that new primary tumors may have appeared. CEA may also be a target for mAb, antisense

nucleotides

46. **PRO268**

Protein disulfide isomerase is an enzymatic protein which is involved in the promotion of correct refolding of proteins through the establishment of correct disulfide bond formation. Protein disulfide isomerase was initially identified based upon its ability to catalyze the renaturation of reduced denatured RNAse (Goldberger et al., *J. Biol. Chem.* 239:1406-1410 (1964) and Epstein et al., *Cold Spring Harbor Symp. Quant. Biol.* 28:439-449 (1963)). Protein disulfide isomerase has been shown to be a resident enzyme of the endoplasmic reticulum which is retained in the endoplasmic reticulum via a -KDEL or -HDEL amino acid sequence at its C-terminus.

Given the importance of disulfide bond-forming enzymes and their potential uses in a number of different applications, for example in increasing the yield of correct refolding of recombinantly produced proteins, efforts are currently being undertaken by both industry and academia to identify new, native proteins having homology to protein disulfide isomerase. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel protein disulfide isomerase homologs. We herein describe a novel polypeptide having homology to protein disulfide isomerase, designated herein as PRO268.

47. **PRO330**

Prolyl 4-hydroxylase is an enzyme which functions to post-translationally hydroxylate proline residues at the Y position of the amino acid sequence Gly-X-Y, which is a repeating three amino acid sequence found in both collagen and procollagen. Hydroxylation of proline residues at the Y position of the Gly-X-Y amino acid triplet to form 4-hydroxyproline residues at those positions is required before newly synthesized collagen polypeptide chains may fold into their proper three-dimensional triple-helical conformation. If hydroxylation does not occur, synthesized collagen polypeptides remain non-helical, are poorly secreted by cells and cannot assemble into stable functional collagen fibrils. Vuorio et al., *Proc. Natl. Acad. Sci. USA* 89:7467-7470 (1992). Prolyl 4-hydroxylase is comprised of at least two different polypeptide subunits, alpha and beta.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)]. Based upon these efforts, Applicants have herein identified and describe a novel polypeptide having homology to the alpha subunit of prolyl 4-hydroxylase, designated herein as PRO330.

48. **PRO339 and PRO310**

Fringe is a protein which specifically blocks serrate-mediated activation of notch in the dorsal compartment of the Drosophila wing imaginal disc. Fleming, et al., Development, 124(15):2973-81 (1997).

Therefore, fringe is of interest for both its role in development as well as its ability to regulate serrate, particularly serrate's signaling abilities. Also of interest are novel polypeptides which may have a role in development and/or the regulation of serrate-like molecules. Of particular interest are novel polypeptides having homology to fringe as identified and described herein, designated herein as PRO339 and PRO310 polypeptides.

49. **PRO244**

Lectins are a class of proteins comprising a region that binds carbohydrates specifically and non-covalently. Numerous lectins have been identified in higher animals, both membrane-bound and soluble, and have been implicated in a variety of cell-recognition phenomena and tumor metastasis.

Most lectins can be classified as either C-type (calcium-dependent) or S-type (thiol-dependent).

Lectins are thought to play a role in regulating cellular events that are initiated at the level of the plasma membrane. For example, plasma membrane associated molecules are involved in the activation of various subsets of lymphoid cells, e.g. T-lymphocytes, and it is known that cell surface molecules are responsible for activation of these cells and consequently their response during an immune reaction.

A particular group of cell adhesion molecules, selectins, belong in the superfamily of C-type lectins. This group includes L-selectin (peripheral lymph node homing receptor (pnHR), LEC-CAM-1, LAM-1, $gp90^{MEL}$, $gp100^{MEL}$, $gp110^{MEL}$, MEL-14 antigen, Leu-8 antigen, TQ-1 antigen, DREG antigen), E-selectin (LEC-CAM-2, LECAM-2, ELAM-1), and P-selectin (LEC-CAM-3, LECAM-3, GMP-140, PADGEM). The structure of selectins consists of a C-type lectin (carbohydrate binding) domain, an epidermal growth factor-like (EGF-like) motif, and variable numbers of complement regulatory (CR) motifs. Selectins are associated with leukocyte adhesion, e.g. the attachment of neutrophils to vesular endothelial cells adjacent to inflammation (E-selectin), or with the trafficking of lymphocytes from blood to secondary lymphoid organs, e.g. lymph nodes and Peyer's patches (L-selectin).

Another exemplary lectin is the cell-associated macrophage antigen, Mac-2 that is believed to be involved in cell adhesion and immune responses. Macrophages also express a lectin that recognizes Tn Ag, a human carcinoma-associated epitope.

Another C-type lectin is CD95 (Fas antigen/APO-1) that is an important mediator of immunologically relevant regulated or programmed cell death (apoptosis). "Apoptosis" is a non-necrotic cell death that takes place in metazoan animal cells following activation of an intrinsic cell suicide program. The cloning of Fas antigen is described in PCT publication WO 91/10448, and European patent application EP510691. The mature Fas molecule consists of 319 amino acids of which 157 are extracellular, 17 constitute the transmembrane domain, and 145 are intracellular. Increased levels of Fas expression at T cell surface have been associated with tumor cells and HIV-infected cells. Ligation of CD95 triggers apoptosis in the presence of interleukin-1 (IL-2).

C-type lectins also include receptors for oxidized low-density lipoprotein (LDL). This suggests a possible role in the pathogenesis of atherosclerosis.

We herein describe the identification and characterization of novel polypeptides having homology to C-type lectins, designated herein as PRO244 polypeptides.

SUMMARY OF THE INVENTION

1. **PRO211 and PRO217**

Applicants have identified cDNA clones that encode novel polypeptides having homology to EGF, designated in the present application as "PRO211" and "PRO217" polypeptides.

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO211 or PRO217 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding EGF-like homologue PRO211 and PRO217 polypeptides of Fig. 2 (SEQ ID NO:2) and/or 4 (SEQ ID NO:4) indicated in Fig. 1 (SEQ ID NO1) and/or Fig. 3 (SEQ ID NO:3), respectively, or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO211 and PRO217 EGF-like homologue PRO211 and PRO217 polypeptides. In particular, the invention provides isolated native sequence PRO211 and PRO217 EGF-like homologue polypeptides, which in one embodiment, includes an amino acid sequence comprising residues: 1 to 353 of Fig. 2 (SEQ ID NO:2) or (2) 1 to 379 of Fig. 4 (SEQ ID NO: 4).

2. **PRO230**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO230".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO230 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO230 polypeptide having amino acid residues 1 through 467 of Figure 6 (SEQ ID NO: 12), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO230 polypeptide. In particular, the invention provides isolated native sequence PRO230 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 467 of Figure 6 (SEQ ID NO:12).

In another embodiment, the invention provides an expressed sequence tag (EST) comprising the nucleotide sequence of SEQ ID NO:13 (Figure 7) which is herein designated as DNA20088.

3. **PRO232**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO232".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO232 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO232 polypeptide having amino acid residues 1 to 114 of Figure 9 (SEQ ID NO:18), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO232 polypeptide. In particular, the invention provides isolated native sequence PRO232 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 114 of Figure 9 (SEQ ID NO:18).

4. PRO187

Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO187".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO187 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO187 polypeptide of Figure 11 (SEQ ID NO:23), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid comprising the coding sequence of Figure 10 (SEQ ID NO:22) or its complement. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA27864-1155, deposited with the ATCC under accession number ATCC 209375, alternatively the coding sequence of clone DNA27864-1155, deposited under accession number ATCC 209375.

In yet another embodiment, the invention provides isolated PRO187 polypeptide. In particular, the invention provides isolated native sequence PRO187 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 205 of Figure 11 (SEQ ID NO:23). Alternatively, the invention provides a polypeptide encoded by the nucleic acid deposited under accession number ATCC 209375.

5. PRO265

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO265".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO265 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO265 polypeptide having amino acid residues 1 to 660 of Figure 13 (SEQ ID NO:28), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO265 polypeptide. In particular, the invention provides isolated native sequence PRO265 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 660 of Figure 13 (SEQ ID NO:28). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO265 polypeptide.

6. PRO219

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO219".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO219 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO219

polypeptide having amino acid residues 1 to 915 of Figure 15 (SEQ ID NO:34), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO219 polypeptide. In particular, the invention provides isolated native sequence PRO219 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 915 of Figure 15 (SEQ ID NO:34).

7. **PRO246**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO246".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO246 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO246 polypeptide having amino acid residues 1 to 390 of Figure 17 (SEQ ID NO:39), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO246 polypeptide. In particular, the invention provides isolated native sequence PRO246 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 390 of Figure 17 (SEQ ID NO:39). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO246 polypeptide.

8. **PRO228**

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to CD97, EMR1 and latrophilin, wherein the polypeptide is designated in the present application as "PRO228".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO228 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO228 polypeptide having amino acid residues 1 to 690 of Figure 19 (SEQ ID NO:49), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO228 polypeptide. In particular, the invention provides isolated native sequence PRO228 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 690 of Figure 19 (SEQ ID NO:49). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO228 polypeptide.

In another embodiment, the invention provides an expressed sequence tag (EST) comprising the nucleotide sequence of SEQ ID NO:50, designated herein as DNA21951.

9. **PRO533**

Applicants have identified a cDNA clone (DNA49435-1219) that encodes a novel polypeptide, designated in the present application as PRO533.

In one embodiment, the invention provides an isolated nucleic acid molecule having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO533 polypeptide comprising the sequence of amino acids 23 to 216 of Figure 22 (SEQ ID NO:59), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 23 to 216 of Figure 22 (SEQ ID NO:59). Preferably, the highest degree of sequence identity occurs within the secreted portion (amino acids 23 to 216 of Figure 22, SEQ ID NO:59). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO533 polypeptide having amino acid residues 1 to 216 of Figure 22 (SEQ ID NO:59), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA49435-1219, deposited with the ATCC under accession number ATCC 209480.

In yet another embodiment, the invention provides isolated PRO533 polypeptide. In particular, the invention provides isolated native sequence PRO533 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 23 to 216 of Figure 22 (SEQ ID NO:59). Native PRO533 polypeptides with or without the native signal sequence (amino acids 1 to 22 in Figure 22 (SEQ ID NO:59)), and with or without the initiating methionine are specifically included. Alternatively, the invention provides a PRO533 polypeptide encoded by the nucleic acid deposited under accession number ATCC 209480.

10. PRO245

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO245".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO245 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO245 polypeptide having amino acid residues 1 to 312 of Fig. 24 (SEQ ID NO:64), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO245 polypeptide. In particular, the invention provides isolated native sequence PRO245 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 312 of Figure 24 (SEQ ID NO:64).

11. PRO220, PRO221 and PRO227

Applicants have identified cDNA clones that each encode novel polypeptides, all having leucine rich repeats. These polypeptides are designated in the present application as PRO220, PRO221 and PRO227.

In one embodiment, the inventionprovides isolated nucleic acid molecules comprising DNA respectively encodiog PRO220, PRO221 and PRO227, respectively. In one aspect, provided herein is an isolated nucleic acid comprises DNA encoding the PRO220 polypeptide having amino acid residues 1 through 708 of Figure 26

(SEQ ID NO:69), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Also provided herein is an isolated nucleic acid comprises DNA encoding the PRO221 polypeptide having amino acid residues 1 through 259 of Figure 28 (SEQ ID NO:71), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Moreover, also provided herein is an isolated nucleic acid comprises DNA encoding the PRO227 polypeptide having amino acid residues 1 through 620 of Figure 30 (SEQ ID NO:73), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO220, PRO221 and PRO227 polypeptides. In particular, provided herein is the isolated native sequence for the PRO220 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 708 of Figure 26 (SEQ ID NO:69). Additionally provided herein is the isolated native sequence for the PRO221 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 259 of Figure 28 (SEQ ID NO:71). Moreover, provided herein is the isolated native sequence for the PRO227 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 620 of Figure 30 (SEQ ID NO:73).

12. PRO258

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to CRTAM and poliovirus receptor precursors, wherein the polypeptide is designated in the present application as "PRO258".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO258 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO258 polypeptide having amino acid residues 1 to 398 of Figure 32 (SEQ ID NO:84), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO258 polypeptide. In particular, the invention provides isolated native sequence PRO258 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 398 of Figure 32 (SEQ ID NO:84). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO258 polypeptide.

13. PRO266

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO266".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO266 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO266 polypeptide having amino acid residues 1 to 696 of Figure 34 (SEQ ID NO:91), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO266 polypeptide. In particular, the invention provides isolated native sequence PRO266 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 696 of Figure 34 (SEQ ID NO:91).

14. PRO269

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as PRO269.

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO269 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO269 polypeptide having amino acid residues 1 to 490 of Fig. 36 (SEQ ID NO:96), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO269 polypeptide. In particular, the invention provides isolated native sequence PRO269 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 490 of Figure 36 (SEQ ID NO:96). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO269 polypeptide.

15. PRO287

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO287".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO287 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO287 polypeptide having amino acid residues 1 to 415 of Fig. 38 (SEQ ID NO:104), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO287 polypeptide. In particular, the invention provides isolated native sequence PRO287 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 415 of Figure 38 (SEQ ID NO:104).

16. PRO214

Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO214".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO214 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO214 polypeptide of Fig. 40 (SEQ ID NO:109), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid comprising the coding sequence of Fig. 39 (SEQ ID NO:108) or its complement. In another aspect, the invention provides a nucleic acid of the full length protein of clone

DNA32286-1191, deposited with ATCC under accession number ATCC 209385.

In yet another embodiment, the invention provides isolated PRO214 polypeptide. In particular, the invention provides isolated native sequence PRO214 polypeptide, which in one embodiment, includes an amino acid sequence comprising the residues of Figure 40 (SEQ ID NO:109). Alternatively, the invention provides a polypeptide encoded by the nucleic acid deposited under accession number ATCC 209385.

17. **PRO317**

Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO317".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding PRO317 polypeptide. In one aspect, the isolated nucleic acid comprises DNA (SEQ ID NO:113) encoding PRO317 polypeptide having amino acid residues 1 to 366 of Fig. 42, or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO317 polypeptide. In particular, the invention provides isolated native-sequence PRO317 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 366 of Figure 42 (SEQ ID NO:114).

In yet another embodiment, the invention supplies a method of detecting the presence of PRO317 in a sample, the method comprising:

a) contacting a detectable anti-PRO317 antibody with a sample suspected of containing PRO317; and

b) detecting binding of the antibody to the sample; wherein the sample is selected from the group consisting of a body fluid, a tissue sample, a cell extract, and a cell culture medium.

In a still further embodiment a method is provided for determining the presence of PRO317 mRNA in a sample, the method comprising:

a) contacting a sample suspected of containing PRO317 mRNA with a detectable nucleic acid probe that hybridizes under moderate to stringent conditions to PRO317 mRNA; and

b) detecting hybridization of the probe to the sample.

Preferably, in this method the sample is a tissue sample and the detecting step is by *in situ* hybridization, or the sample is a cell extract and detection is by Northern analysis.

Further, the invention provides a method for treating a PRO317-associated disorder comprising administering to a mammal an effective amount of the PRO317 polypeptide or a composition thereof containing a carrier, or with an effective amount of a PRO317 agonist or PRO317 antagonist, such as an antibody which binds specifically to PRO317.

18. **PRO301**

Applicants have identified a cDNA clone (DNA40628-1216) that encodes a novel polypeptide, designated in the present application as "PRO301".

In one embodiment, the invention provides an isolated nucleic acid molecule having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO301 polypeptide comprising the sequence of amino acids 28 to 258 of Fig. 44 (SEQ ID NO:119), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 28 to 258 of Fig. 44 (SEQ ID NO:119). Preferably, the highest degree of sequence identity occurs within the extracellular domains (amino acids 28 to 258 of Fig. 44, SEQ ID NO:119). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO301 polypeptide having amino acid residues 28 to 299 of Fig. 44 (SEQ ID NO:119), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA40628-1216, deposited with the ATCC under accession number ATCC 209432, alternatively the coding sequence of clone DNA40628-1216, deposited under accession number ATCC 209432.

In yet another embodiment, the invention provides isolated PRO301 polypeptide. In particular, the invention provides isolated native sequence PRO301 polypeptide, which in one embodiment, includes an amino acid sequence comprising the extracellular domain residues 28 to 258 of Figure 44 (SEQ ID NO:119). Native PRO301 polypeptides with or without the native signal sequence (amino acids 1 to 27 in Figure 44 (SEQ ID NO:119), and with or without the initiating methionine are specifically included. Additionally, the sequences of the invention may also comprise the transmembrane domain (residues 236 to about 258 in Figure 44; SEQ ID NO:119) and/or the intracellular domain (about residue 259 to 299 in Figure 44; SEQ ID NO:119). Alternatively, the invention provides a PRO301 polypeptide encoded by the nucleic acid deposited under accession number ATCC 209432.

19. **PRO224**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO224".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO224 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO224 polypeptide having amino acid residues 1 to 282 of Figure 46 (SEQ ID NO:127), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO224 polypeptide. In particular, the invention provides isolated native sequence PRO224 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 282 of Figure 46 (SEQ ID NO:127).

20. **PRO222**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is

designated in the present application as "PRO222".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO222 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO222 polypeptide having amino acid residues 1 to 490 of Fig. 48 (SEQ ID NO:132), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO222 polypeptide. In particular, the invention provides isolated native sequence PRO222 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 490 of Figure 48 (SEQ ID NO:132).

21. PRO234

Applicants have identified a cDNA clone that encodes a novel lectin polypeptide molecule, designated in the present application as "PRO234".

In one embodiment, the invention provides an isolated nucleic acid encoding a novel lectin comprising DNA encoding a PRO234 polypeptide. In one aspect, the isolated nucleic acid comprises the DNA encoding PRO234 polypeptides having amino acid residues 1 to 382 of Fig. 50 (SEQ ID NO:137), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides an isolated nucleic acid molecule comprising the nucleotide sequence of Fig. 49 (SEQ ID NO:136).

In another embodiment, the invention provides isolated novel PRO234 polypeptides. In particular, the invention provides isolated native sequence PRO234 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 382 of Figure 50 (SEQ ID NO:137).

In yet another embodiment, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences.

22. PRO231

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to a putative acid phosphatase, wherein the polypeptide is designated in the present application as "PRO231".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO231 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO231 polypeptide having amino acid residues 1 to 428 of Fig. 52 (SEQ ID NO:142), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO231 polypeptide. In particular, the invention provides isolated native sequence PRO231 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 428 of Figure 52 (SEQ ID NO:142).

23. PRO229

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to scavenger receptors wherein the polypeptide is designated in the present application as "PRO229".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO229 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO229 polypeptide having amino acid residues 1 to 347 of Figure 54 (SEQ ID NO:148), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO229 polypeptide. In particular, the invention provides isolated native sequence PRO229 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 347 of Figure 54 (SEQ ID NO:148).

24. PRO238

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to reductase, wherein the polypeptide is designated in the present application as "PRO238".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO238 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO238 polypeptide having amino acid residues 1 to 310 of Figure 56 (SEQ ID NO:153), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO238 polypeptide. In particular, the invention provides isolated native sequence PRO238 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 310 of Figure 56 (SEQ ID NO:153).

25. PRO233

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO233".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO233 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO233 polypeptide having amino acid residues 1 to 300 of Figure 58 (SEQ ID NO:159), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO233 polypeptide. In particular, the invention provides isolated native sequence PRO233 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 300 of Figure 58 (SEQ ID NO:159).

26. PRO223

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to serine

carboxypeptidase polypeptides, wherein the polypeptide is designated in the present application as "PRO223".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO223 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO223 polypeptide having amino acid residues 1 to 476 of Figure 60 (SEQ ID NO:164), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO223 polypeptide. In particular, the invention provides isolated native sequence PRO223 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 476 of Figure 60 (SEQ ID NO:164).

27. **PRO235**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO235".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO235 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO235 polypeptide having amino acid residues 1 to 552 of Figure 62 (SEQ ID NO:170), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO235 polypeptide. In particular, the invention provides isolated native sequence PRO235 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 552 of Figure 62 (SEQ ID NO:170).

28. **PRO236 and PRO262**

Applicants have identified cDNA clones that encode novel polypeptides having homology to β-galactosidase, wherein those polypeptides are designated in the present application as "PRO236" and "PRO262".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO236 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO236 polypeptide having amino acid residues 1 to 636 of Figure 64 (SEQ ID NO:175), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO262 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO262 polypeptide having amino acid residues 1 to 654 of Figure 66 (SEQ ID NO:177), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO236 polypeptide. In particular, the invention provides isolated native sequence PRO236 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 636 of Figure 64 (SEQ ID NO:175).

In another embodiment, the invention provides isolated PRO262 polypeptide. In particular, the invention provides isolated native sequence PRO262 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 654 of Figure 66 (SEQ ID NO:177).

29. **PRO239**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO239".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO239 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO239 polypeptide having amino acid residues 1 to 501 of Figure 68 (SEQ ID NO:185), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO239 polypeptide. In particular, the invention provides isolated native sequence PRO239 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 501 of Figure 68 (SEQ ID NO:185).

30. **PRO257**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO257".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO257 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO257 polypeptide having amino acid residues 1 to 607 of Figure 70 (SEQ ID NO:190), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO257 polypeptide. In particular, the invention provides isolated native sequence PRO257 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 607 of Figure 70 (SEQ ID NO:190). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO257 polypeptide.

31. **PRO260**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO260".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO260 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO260 polypeptide having amino acid residues 1 to 467 of Figure 72 (SEQ ID NO:195), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO260 polypeptide. In particular, the invention provides isolated native sequence PRO260 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 467 of Figure 72 (SEQ ID NO:195).

32. PRO263

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to CD44 antigen, wherein the polypeptide is designated in the present application as "PRO263".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO263 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO263 polypeptide having amino acid residues 1 to 322 of Figure 74 (SEQ ID NO:201), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO263 polypeptide. In particular, the invention provides isolated native sequence PRO263 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 322 of Figure 74 (SEQ ID NO:201). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO263 polypeptide.

33. PRO270

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO270".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO270 polypeptide. In one aspect, the isolated nucleic acid comprises DNA whivch includes the sequence encoding the PRO270 polypeptide having amino acid residues 1 to 296 of Fig. 76 (SEQ ID NO:207), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO270 polypeptide. In particular, the invention provides isolated native sequence PRO270 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 296 of Figure 76 (SEQ ID NO:207).

34. PRO271

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to the proteoglycan link protein, wherein the polypeptide is designated in the present application as "PRO271".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO271 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO271 polypeptide having amino acid residues 1 to 360 of Figure 78 (SEQ ID NO:213), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO271 polypeptide. In particular, the invention provides isolated native sequence PRO271 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 360 of Figure 78 (SEQ ID NO:213).

35. **PRO272**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO272".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO272 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO272 polypeptide having amino acid residues 1 to 328 of Figure 80 (SEQ ID NO:221), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO272 polypeptide. In particular, the invention provides isolated native sequence PRO272 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 328 of Figure 80 (SEQ ID NO:211).

36. **PRO294**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO294".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO294 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO294 polypeptide having amino acid residues 1 to 550 of Figure 82 (SEQ ID NO:227), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO294 polypeptide. In particular, the invention provides isolated native sequence PRO294 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 550 of Figure 82 (SEQ ID NO:227).

37. **PRO295**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO295".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO295 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO295 polypeptide having amino acid residues 1 to 350 of Figure 84 (SEQ ID NO:236), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO295 polypeptide. In particular, the invention provides isolated native sequence PRO295 polypeptide, which in one embodiment, includes an amino

acid sequence comprising residues 1 to 350 of Figure 84 (SEQ ID NO:236).

38. PRO293

Applicants have identified a cDNA clone that encodes a novel human neuronal leucine rich repeat polypeptide, wherein the polypeptide is designated in the present application as "PRO293".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO293 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO293 polypeptide having amino acid residues 1 to 713 of Figure 86 (SEQ ID NO:245), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO293 polypeptide. In particular, the invention provides isolated native sequence PRO293 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 713 of Figure 86 (SEQ ID NO:245). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO293 polypeptide.

39. PRO247

Applicants have identified a cDNA clone that encodes a novel polypeptide having leucine rich repeats wherein the polypeptide is designated in the present application as "PRO247".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO247 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO247 polypeptide having amino acid residues 1 to 546 of Figure 88 (SEQ ID NO:250), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO247 polypeptide. In particular, the invention provides isolated native sequence PRO247 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 546 of Figure 88 (SEQ ID NO:250). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO247 polypeptide.

40. PRO302, PRO303, PRO304, PRO307 and PRO343

Applicants have identified cDNA clones that encode novel polypeptides having homology to various proteases, wherein those polypeptide are designated in the present application as "PRO302", "PRO303", "PRO304", "PRO307" and "PRO343" polypeptides.

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO302 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO302 polypeptide having amino acid residues 1 to 452 of Figure 90 (SEQ ID NO:255), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO303 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO303 polypeptide having amino acid residues 1 to 314 of Figure 92 (SEQ ID NO:257), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In yet another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO304 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO304 polypeptide having amino acid residues 1 to 556 of Figure 94 (SEQ ID NO:259), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO307 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO307 polypeptide having amino acid residues 1 to 383 of Figure 96 (SEQ ID NO:261), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO343 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO343 polypeptide having amino acid residues 1 to 317 of Figure 98 (SEQ ID NO:263), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO302 polypeptide. In particular, the invention provides isolated native sequence PRO302 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 452 of Figure 90 (SEQ ID NO:255).

In another embodiment, the invention provides isolated PRO303 polypeptide. In particular, the invention provides isolated native sequence PRO303 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 314 of Figure 92 (SEQ ID NO:257).

In another embodiment, the invention provides isolated PRO304 polypeptide. In particular, the invention provides isolated native sequence PRO304 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 556 of Figure 94 (SEQ ID NO:259).

In another embodiment, the invention provides isolated PRO307 polypeptide. In particular, the invention provides isolated native sequence PRO307 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 383 of Figure 96 (SEQ ID NO:261).

In another embodiment, the invention provides isolated PRO343 polypeptide. In particular, the invention provides isolated native sequence PRO343 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 317 of Figure 98 (SEQ ID NO:263).

41. **PRO328**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO328".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO328 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO328 polypeptide having amino acid residues 1 to 463 of Figure 100 (SEQ ID NO:285), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO328 polypeptide. In particular, the invention provides isolated native sequence PRO328 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 463 of Figure 100 (SEQ ID NO:285). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO306 polypeptide.

42. **PRO335, PRO331 and PRO326**

Applicants have identified three cDNA clones that respectively encode three novel polypeptides, each having leucine rich repeats and homology to LIG-1 and ALS. These polypeptides are designated in the present application as PRO335, PRO331 and PRO326, respectively.

In one embodiment, the invention provides three isolated nucleic acid molecules comprising DNA respectively encoding PRO335, PRO331 and PRO326, respectively. In one aspect, herein is provided an isolated nucleic acid comprising DNA encoding the PRO335 polypeptide having amino acid residues 1 through 1059 of Figure 102 (SEQ ID NO:290), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Also provided herein is an isolated nucleic acid comprises DNA encoding the PRO331 polypeptide having amino acid residues 1 through 640 of Figure 104 (SEQ ID NO:292), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Additionally provided herein is an isolated nucleic acid comprises DNA encoding the PRO326 polypeptide having amino acid residues 1 through 1119 of Figure 106 (SEQ ID NO:294), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO335, PRO331 and PRO326 polypeptides or extracellular domains thereof. In particular, the invention provides isolated native sequence for the PRO335 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 1059 of Figure 102 (SEQ ID NO:290). Also provided herein is the isolated native sequence for the PRO331 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 640 of Figure 104 (SEQ ID NO:292). Also provided herein is the isolated native sequence for the PRO326 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 1119 of Figure 106 (SEQ ID NO:294).

43. **PRO332**

Applicants have identified a cDNA clone (DNA40982-1235) that encodes a novel polypeptide, designated in the present application as "PRO332."

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO358 polypeptide comprising the sequence of amino acids 49 to 642 of Fig. 108 (SEQ ID NO:310), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 1 to 642 of Fig. 108 (SEQ ID NO:310). Preferably, the highest degree of sequence identity occurs within the leucine-rich repeat domains (amino acids 116 to 624 of Fig. 108, SEQ ID NO:310). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO332 polypeptide having amino acid residues 49 to 642 of Fig. 108 (SEQ ID NO:310), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO332 polypeptides. In particular, the invention provides isolated native sequence PRO332 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 49 to 624 of Figure 108 (SEQ ID NO:310). Native PRO332 polypeptides with or without the native signal sequence (amino acids 1 to 48 in Figure 108, SEQ ID NO:310), and with or without the initiating methionine are specifically included.

44. **PRO334**

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to fibulin and fibrillin, wherein the polypeptide is designated in the present application as "PRO334".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO334 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO334 polypeptide having amino acid residues 1 to 509 of Figure 110 (SEQ ID NO:315), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO334 polypeptide. In particular, the invention provides isolated native sequence PRO334 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 509 of Figure 110 (SEQ ID NO:315).

45. **PRO346**

Applicants have identified a cDNA clone (DNA44167-1243) that encodes a novel polypeptide, designated in the present application as "PRO346."

In one embodiment, the invention provides an isolated nucleic acid molecule having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO346 polypeptide comprising the sequence of amino

acids 19 to 339 of Fig. 112 (SEQ ID NO: 320), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 19 to 339 of Fig. 112 (SEQ ID NO:320). Preferably, the highest degree of sequence identity occurs within the extracellular domains (amino acids 19 to 339 of Fig. 112, SEQ ID NO:320). In alternative embodiments, the polypeptide by which the homology is measured comprises the residues 1-339, 19-360 or 19-450 of Fig. 112, SEQ ID NO:320). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO346 polypeptide having amino acid residues 19 to 339 of Fig. 112 (SEQ ID NO:320), alternatively residues 1-339, 19-360 or 19-450 of Fig. 112 (SEQ ID NO:320) or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA44167-1243, deposited with the ATCC under accession number ATCC 209434, alternatively the coding sequence of clone DNA44167-1243, deposited under accession number ATCC 209434.

In yet another embodiment, the invention provides isolated PRO346 polypeptide. In particular, the invention provides isolated native sequence PRO346 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 19 to 339 of Figure 112 (SEQ ID NO:320). Native PRO346 polypeptides with or without the native signal sequence (residues 1 to 18 in Figure 112 (SEQ ID NO:320), with or without the initiating methionine, with or without the transmembrane domain (residues 340 to 360) and with or without the intracellular domain (residues 361 to 450) are specifically included. Alternatively, the invention provides a PRO346 polypeptide encoded by the nucleic acid deposited under accession number ATCC 209434.

46. **PRO268**

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to protein disulfide isomerase, wherein the polypeptide is designated in the present application as "PRO268".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO268 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO268 polypeptide having amino acid residues 1 to 280 of Figure 114 (SEQ ID NO:325), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO268 polypeptide. In particular, the invention provides isolated native sequence PRO268 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 280 of Figure 114 (SEQ ID NO:325). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO268 polypeptide.

47. **PRO330**

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to the alpha subunit of prolyl 4-hydroxylase, wherein the polypeptide is designated in the present application as "PRO330".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO330 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO330 polypeptide having amino acid residues 1 to 533 of Figure 116 (SEQ ID NO:332), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO330 polypeptide. In particular, the invention provides isolated native sequence PRO330 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 533 of Figure 116 (SEQ ID NO:332).

48. **PRO339 and PRO310**

Applicants have identified two cDNA clones wherein each clone encodes a novel polypeptide having homology to fringe, wherein the polypeptides are designated in the present application as "PRO339" and "PRO310".

In one embodiment, the invention provides isolated nucleic acid molecules comprising DNA encoding a PRO339 and/or a PRO310 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO339 polypeptide having amino acid residues 1 to 772 of Figure 118 (SEQ ID NO:339), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the isolated nucleic acid comprises DNA encoding the PRO310 polypeptide having amino acid residues 1 to 318 of Figure 120 (SEQ ID NO:341), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO339 as well as isolated PRO310 polypeptides. In particular, the invention provides isolated native sequence PRO339 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 772 of Figure 118 (SEQ ID NO:339). The invention further provides isolated native sequence PRO310 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 318 of Figure 120 (SEQ ID NO:341).

49. **PRO244**

Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO244".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding PRO244 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding PRO244 polypeptide having amino acid residues 1 to 219 of Fig. 122 (SEQ ID NO:377), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO244 polypeptide. In particular, the invention provides isolated native sequence PRO244 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 219 of Figure 122 (SEQ ID NO:377).

50. Additional Embodiments

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells. *E. coli*, or yeast. A process for producing any of the herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody which specifically binds to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

In yet other embodiments, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences, wherein those probes may be derived from any of the above or below described nucleotide sequences.

In other embodiments, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein or an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein, or (b) the complement of the DNA molecule of (a).

In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at

least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein or the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein, or (b) the complement of the DNA molecule of (a).

In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes or for encoding fragments of a PRO polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody. Such nucleic acid fragments are usually at least about 20 nucleotides in length, preferably at least about 30 nucleotides in length, more preferably at least about 40 nucleotides in length, yet more preferably at least about 50 nucleotides in length, yet more preferably at least about 60 nucleotides in length, yet more preferably at least about 70 nucleotides in length, yet more preferably at least about 80 nucleotides in length, yet more preferably at least about 90 nucleotides in length, yet more preferably at least about 100 nucleotides in length, yet more preferably at least about 110 nucleotides in length, yet more preferably at least about 120 nucleotides in length, yet more

preferably at least about 130 nucleotides in length, yet more preferably at least about 140 nucleotides in length, yet more preferably at least about 150 nucleotides in length, yet more preferably at least about 160 nucleotides in length, yet more preferably at least about 170 nucleotides in length, yet more preferably at least about 180 nucleotides in length, yet more preferably at least about 190 nucleotides in length, yet more preferably at least about 200 nucleotides in length, yet more preferably at least about 250 nucleotides in length, yet more preferably at least about 300 nucleotides in length, yet more preferably at least about 350 nucleotides in length, yet more preferably at least about 400 nucleotides in length, yet more preferably at least about 450 nucleotides in length, yet more preferably at least about 500 nucleotides in length, yet more preferably at least about 600 nucleotides in length, yet more preferably at least about 700 nucleotides in length, yet more preferably at least about 800 nucleotides in length, yet more preferably at least about 900 nucleotides in length and yet more preferably at least about 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypepride-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

In another embodiment, the invention provides isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein or an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity,

yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence scoring at least about 80% positives, preferably at least about 81% positives, more preferably at least about 82% positives, yet more preferably at least about 83% positives, yet more preferably at least about 84% positives, yet more preferably at least about 85% positives, yet more preferably at least about 86% positives, yet more preferably at least about 87% positives, yet more preferably at least about 88% positives, yet more preferably at least about 89% positives, yet more preferably at least about 90% positives, yet more preferably at least about 91% positives, yet more preferably at least about 92% positives, yet more preferably at least about 93% positives, yet more preferably at least about 94% positives, yet more preferably at least about 95% positives, yet more preferably at least about 96% positives, yet more preferably at least about 97% positives, yet more preferably at least about 98% positives and yet more preferably at least about 99% positives when compared with the amino acid sequence of a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein or an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

In yet another embodiment, the invention concerns agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a

biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO211 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA32292-1131".

Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO217 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA33094-1131".

Figure 4 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

Figure 5 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO230 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA33223-1136".

Figure 6 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 5.

Figure 7 shows a nucleotide sequence designated herein as DNA20088 (SEQ ID NO:13).

Figure 8 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO232 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA34435-1140".

Figure 9 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in Figure 8.

Figure 10 shows a nucleotide sequence (SEQ ID NO:22) of a native sequence PRO187 cDNA, wherein SEQ ID NO:22 is a clone designated herein as "DNA27864-1155".

Figure 11 shows the amino acid sequence (SEQ ID NO:23) derived from the coding sequence of SEQ ID NO:22 shown in Figure 10.

Figure 12 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO265 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA36350-1158".

Figure 13 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 12.

Figure 14 shows a nucleotide sequence (SEQ ID NO:33) of a native sequence PRO219 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "DNA32290-1164".

Figure 15 shows the amino acid sequence (SEQ ID NO:34) derived from the coding sequence of SEQ ID NO:33 shown in Figure 14.

Figure 16 shows a nucleotide sequence (SEQ ID NO:38) of a native sequence PRO246 cDNA, wherein SEQ ID NO:38 is a clone designated herein as "DNA35639-1172".

Figure 17 shows the amino acid sequence (SEQ ID NO:39) derived from the coding sequence of SEQ ID NO:38 shown in Figure 16.

Figure 18 shows a nucleotide sequence (SEQ ID NO:48) of a native sequence PRO228 cDNA, wherein SEQ ID NO:48 is a clone designated herein as "DNA33092-1202".

Figure 19 shows the amino acid sequence (SEQ ID NO:49) derived from the coding sequence of SEQ ID NO:48 shown in Figure 18.

Figure 20 shows a nucleotide sequence designated herein as DNA21951 (SEQ ID NO:50).

Figure 21 shows a nucleotide sequence (SEQ ID NO:58) of a native sequence PRO533 cDNA, wherein SEQ ID NO:58 is a clone designated herein as "DNA49435-1219".

Figure 22 shows the amino acid sequence (SEQ ID NO:59) derived from the coding sequence of SEQ ID NO:58 shown in Figure 21.

Figure 23 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO245 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA35638-1141".

Figure 24 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in Figure 23.

Figure 25 shows a nucleotide sequence (SEQ ID NO:68) of a native sequence PRO220 cDNA, wherein SEQ ID NO:68 is a clone designated herein as "DNA32298-1132".

Figure 26 shows the amino acid sequence (SEQ ID NO:69) derived from the coding sequence of SEQ ID NO:68 shown in Figure 25.

Figure 27 shows a nucleotide sequence (SEQ ID NO:70) of a native sequence PRO221 cDNA, wherein SEQ ID NO:70 is a clone designated herein as "DNA33089-1132".

Figure 28 shows the amino acid sequence (SEQ ID NO:71) derived from the coding sequence of SEQ ID NO:70 shown in Figure 27.

Figure 29 shows a nucleotide sequence (SEQ ID NO:72) of a native sequence PRO227 cDNA, wherein SEQ ID NO:72 is a clone designated herein as "DNA33786-1132".

Figure 30 shows the amino acid sequence (SEQ ID NO:73) derived from the coding sequence of SEQ ID NO:72 shown in Figure 29.

Figure 31 shows a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO258 cDNA, wherein SEQ ID NO:83 is a clone designated herein as "DNA35918-1174".

Figure 32 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:83 shown in Figure 31.

Figure 33 shows a nucleotide sequence (SEQ ID NO:90) of a native sequence PRO266 cDNA, wherein SEQ ID NO:90 is a clone designated herein as "DNA37150-1178".

Figure 34 shows the amino acid sequence (SEQ ID NO:91) derived from the coding sequence of SEQ ID NO:90 shown in Figure 33.

Figure 35 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO269 cDNA, wherein SEQ ID NO:95 is a clone designated herein as "DNA38260-1180".

Figure 36 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in Figure 35.

Figure 37 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO287 cDNA, wherein SEQ ID NO:103 is a clone designated herein as "DNA39969-1185".

Figure 38 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 37.

Figure 39 shows a nucleotide sequence (SEQ ID NO:108) of a native sequence PRO214 cDNA, wherein SEQ ID NO:108 is a clone designated herein as "DNA32286-1191".

Figure 40 shows the amino acid sequence (SEQ ID NO:109) derived from the coding sequence of SEQ ID NO:108 shown in Figure 39.

Figure 41 shows a nucleotide sequence (SEQ ID NO:113) of a native sequence PRO317 cDNA, wherein SEQ ID NO:113 is a clone designated herein as "DNA33461-1199".

Figure 42 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:113 shown in Figure 41.

Figure 43 shows a nucleotide sequence (SEQ ID NO:118) of a native sequence PRO301 cDNA, wherein SEQ ID NO:118 is a clone designated herein as "DNA40628-1216".

Figure 44 shows the amino acid sequence (SEQ ID NO:119) derived from the coding sequence of SEQ ID NO:118 shown in Figure 43.

Figure 45 shows a nucleotide sequence (SEQ ID NO:126) of a native sequence PRO224 cDNA, wherein SEQ ID NO:126 is a clone designated herein as "DNA33221-1133".

Figure 46 shows the amino acid sequence (SEQ ID NO:127) derived from the coding sequence of SEQ ID NO:126 shown in Figure 45.

Figure 47 shows a nucleotide sequence (SEQ ID NO:131) of a native sequence PRO222 cDNA, wherein SEQ ID NO:131 is a clone designated herein as "DNA33107-1135".

Figure 48 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:131 shown in Figure 47.

Figure 49 shows a nucleotide sequence (SEQ ID NO:136) of a native sequence PRO234 cDNA, wherein SEQ ID NO:136 is a clone designated herein as "DNA35557-1137".

Figure 50 shows the amino acid sequence (SEQ ID NO:137) derived from the coding sequence of SEQ ID NO:136 shown in Figure 49.

Figure 51 shows a nucleotide sequence (SEQ ID NO:141) of a native sequence PRO231 cDNA, wherein SEQ ID NO:141 is a clone designated herein as "DNA34434-1139".

Figure 52 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:141 shown in Figure 51.

Figure 53 shows a nucleotide sequence (SEQ ID NO:147) of a native sequence PRO229 cDNA, wherein SEQ ID NO:147 is a clone designated herein as "DNA33100-1159".

Figure 54 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:147 shown in Figure 53.

Figure 55 shows a nucleotide sequence (SEQ ID NO:152) of a native sequence PRO238 cDNA, wherein SEQ ID NO:152 is a clone designated herein as "DNA35600-1162".

Figure 56 shows the amino acid sequence (SEQ ID NO:153) derived from the coding sequence of SEQ ID NO:152 shown in Figure 55.

Figure 57 shows a nucleotide sequence (SEQ ID NO:158) of a native sequence PRO233 cDNA, wherein SEQ ID NO:158 is a clone designated herein as "DNA34436-1238".

Figure 58 shows the amino acid sequence (SEQ ID NO:159) derived from the coding sequence of SEQ ID NO:158 shown in Figure 57.

Figure 59 shows a nucleotide sequence (SEQ ID NO:163) of a native sequence PRO223 cDNA, wherein SEQ ID NO:163 is a clone designated herein as "DNA33206-1165".

Figure 60 shows the amino acid sequence (SEQ ID NO:164) derived from the coding sequence of SEQ ID NO:163 shown in Figure 59.

Figure 61 shows a nucleotide sequence (SEQ ID NO:169) of a native sequence PRO235 cDNA, wherein SEQ ID NO:169 is a clone designated herein as "DNA35558-1167".

Figure 62 shows the amino acid sequence (SEQ ID NO:170) derived from the coding sequence of SEQ ID NO:169 shown in Figure 61.

Figure 63 shows a nucleotide sequence (SEQ ID NO:174) of a native sequence PRO236 cDNA, wherein SEQ ID NO:174 is a clone designated herein as "DNA35599-1168".

Figure 64 shows the amino acid sequence (SEQ ID NO:175) derived from the coding sequence of SEQ ID NO:174 shown in Figure 63.

Figure 65 shows a nucleotide sequence (SEQ ID NO:176) of a native sequence PRO262 cDNA, wherein SEQ ID NO:176 is a clone designated herein as "DNA36992-1168".

Figure 66 shows the amino acid sequence (SEQ ID NO:177) derived from the coding sequence of SEQ ID NO:176 shown in Figure 65.

Figure 67 shows a nucleotide sequence (SEQ ID NO:184) of a native sequence PRO239 cDNA, wherein SEQ ID NO:184 is a clone designated herein as "DNA34407-1169".

Figure 68 shows the amino acid sequence (SEQ ID NO:185) derived from the coding sequence of SEQ ID NO:184 shown in Figure 67.

Figure 69 shows a nucleotide sequence (SEQ ID NO:189) of a native sequence PRO257 cDNA, wherein SEQ ID NO:189 is a clone designated herein as "DNA35841-1173".

Figure 70 shows the amino acid sequence (SEQ ID NO:190) derived from the coding sequence of SEQ ID NO:189 shown in Figure 69.

Figure 71 shows a nucleotide sequence (SEQ ID NO:194) of a native sequence PRO260 cDNA, wherein SEQ ID NO:194 is a clone designated herein as "DNA33470-1175".

Figure 72 shows the amino acid sequence (SEQ ID NO:195) derived from the coding sequence of SEQ ID NO:194 shown in Figure 71.

Figure 73 shows a nucleotide sequence (SEQ ID NO:200) of a native sequence PRO263 cDNA, wherein SEQ ID NO:200 is a clone designated herein as "DNA34431-1177".

Figure 74 shows the amino acid sequence (SEQ ID NO:201) derived from the coding sequence of SEQ ID NO:200 shown in Figure 73.

Figure 75 shows a nucleotide sequence (SEQ ID NO:206) of a native sequence PRO270 cDNA, wherein SEQ ID NO:206 is a clone designated herein as "DNA39510-1181".

Figure 76 shows the amino acid sequence (SEQ ID NO:207) derived from the coding sequence of SEQ ID NO:206 shown in Figure 75.

Figure 77 shows a nucleotide sequence (SEQ ID NO:212) of a native sequence PRO271 cDNA, wherein SEQ ID NO:212 is a clone designated herein as "DNA39423-1182".

Figure 78 shows the amino acid sequence (SEQ ID NO:213) derived from the coding sequence of SEQ ID NO:212 shown in Figure 77.

Figure 79 shows a nucleotide sequence (SEQ ID NO:220) of a native sequence PRO272 cDNA, wherein SEQ ID NO:220 is a clone designated herein as "DNA40620-1183".

Figure 80 shows the amino acid sequence (SEQ ID NO:221) derived from the coding sequence of SEQ ID NO:220 shown in Figure 79.

Figure 81 shows a nucleotide sequence (SEQ ID NO:226) of a native sequence PRO294 cDNA, wherein SEQ ID NO:226 is a clone designated herein as "DNA40604-1187".

Figure 82 shows the amino acid sequence (SEQ ID NO:227) derived from the coding sequence of SEQ ID NO:226 shown in Figure 81.

Figure 83 shows a nucleotide sequence (SEQ ID NO:235) of a native sequence PRO295 cDNA, wherein SEQ ID NO:235 is a clone designated herein as "DNA38268-1188".

Figure 84 shows the amino acid sequence (SEQ ID NO:236) derived from the coding sequence of SEQ ID NO:235 shown in Figure 83.

Figure 85 shows a nucleotide sequence (SEQ ID NO:244) of a native sequence PRO293 cDNA, wherein SEQ ID NO:244 is a clone designated herein as "DNA37151-1193".

Figure 86 shows the amino acid sequence (SEQ ID NO:245) derived from the coding sequence of SEQ ID NO:244 shown in Figure 85.

Figure 87 shows a nucleotide sequence (SEQ ID NO:249) of a native sequence PRO247 cDNA, wherein SEQ ID NO:249 is a clone designated herein as "DNA35673-1201".

Figure 88 shows the amino acid sequence (SEQ ID NO:250) derived from the coding sequence of SEQ ID NO:249 shown in Figure 87.

Figure 89 shows a nucleotide sequence (SEQ ID NO:254) of a native sequence PRO302 cDNA, wherein SEQ ID NO:254 is a clone designated herein as "DNA40370-1217".

Figure 90 shows the amino acid sequence (SEQ ID NO:255) derived from the coding sequence of SEQ ID NO:254 shown in Figure 89.

Figure 91 shows a nucleotide sequence (SEQ ID NO:256) of a native sequence PRO303 cDNA, wherein SEQ ID NO:256 is a clone designated herein as "DNA42551-1217".

Figure 92 shows the amino acid sequence (SEQ ID NO:257) derived from the coding sequence of SEQ ID NO:256 shown in Figure 91.

Figure 93 shows a nucleotide sequence (SEQ ID NO:258) of a native sequence PRO304 cDNA, wherein SEQ ID NO:258 is a clone designated herein as "DNA39520-1217".

Figure 94 shows the amino acid sequence (SEQ ID NO:259) derived from the coding sequence of SEQ ID NO:258 shown in Figure 93.

Figure 95 shows a nucleotide sequence (SEQ ID NO:260) of a native sequence PRO307 cDNA, wherein SEQ ID NO:260 is a clone designated herein as "DNA41225-1217".

Figure 96 shows the amino acid sequence (SEQ ID NO:261) derived from the coding sequence of SEQ ID NO:260 shown in Figure 95.

Figure 97 shows a nucleotide sequence (SEQ ID NO:262) of a native sequence PRO343 cDNA, wherein SEQ ID NO:262 is a clone designated herein as "DNA43318-1217".

Figure 98 shows the amino acid sequence (SEQ ID NO:263) derived from the coding sequence of SEQ ID NO:262 shown in Figure 97.

Figure 99 shows a nucleotide sequence (SEQ ID NO:284) of a native sequence PRO328 cDNA, wherein SEQ ID NO:284 is a clone designated herein as "DNA40587-1231".

Figure 100 shows the amino acid sequence (SEQ ID NO:285) derived from the coding sequence of SEQ ID NO:284 shown in Figure 99.

Figure 101 shows a nucleotide sequence (SEQ ID NO:289) of a native sequence PRO335 cDNA, wherein SEQ ID NO:289 is a clone designated herein as "DNA41388-1234".

Figure 102 shows the amino acid sequence (SEQ ID NO:290) derived from the coding sequence of SEQ ID NO:289 shown in Figure 101.

Figure 103 shows a nucleotide sequence (SEQ ID NO:291) of a native sequence PRO331 cDNA, wherein SEQ ID NO:291 is a clone designated herein as "DNA40981-1234".

Figure 104 shows the amino acid sequence (SEQ ID NO:292) derived from the coding sequence of SEQ ID NO:291 shown in Figure 103.

Figure 105 shows a nucleotide sequence (SEQ ID NO:293) of a native sequence PRO326 cDNA, wherein SEQ ID NO:293 is a clone designated herein as "DNA37140-1234".

Figure 106 shows the amino acid sequence (SEQ ID NO:294) derived from the coding sequence of SEQ ID NO:293 shown in Figure 105.

Figure 107 shows a nucleotide sequence (SEQ ID NO:309) of a native sequence PRO332 cDNA, wherein SEQ ID NO:309 is a clone designated herein as "DNA40982-1235".

Figure 108 shows the amino acid sequence (SEQ ID NO:310) derived from the coding sequence of SEQ ID NO:309 shown in Figure 107.

Figure 109 shows a nucleotide sequence (SEQ ID NO:314) of a native sequence PRO334 cDNA, wherein SEQ ID NO:314 is a clone designated herein as "DNA41379-1236".

Figure 110 shows the amino acid sequence (SEQ ID NO:315) derived from the coding sequence of SEQ ID NO:314 shown in Figure 109.

Figure 111 shows a nucleotide sequence (SEQ ID NO:319) of a native sequence PRO346 cDNA, wherein SEQ ID NO:319 is a clone designated herein as "DNA44167-1243".

Figure 112 shows the amino acid sequence (SEQ ID NO:320) derived from the coding sequence of SEQ ID NO:319 shown in Figure 111.

Figure 113 shows a nucleotide sequence (SEQ ID NO:324) of a native sequence PRO268 cDNA, wherein SEQ ID NO:324 is a clone designated herein as "DNA39427-1179".

Figure 114 shows the amino acid sequence (SEQ ID NO:325) derived from the coding sequence of SEQ ID NO:324 shown in Figure 113.

Figure 115 shows a nucleotide sequence (SEQ ID NO:331) of a native sequence PRO330 cDNA, wherein SEQ ID NO:331 is a clone designated herein as "DNA40603-1232".

Figure 116 shows the amino acid sequence (SEQ ID NO:332) derived from the coding sequence of SEQ ID NO:331 shown in Figure 115.

Figure 117 shows a nucleotide sequence (SEQ ID NO:338) of a native sequence PRO339 cDNA, wherein SEQ ID NO:338 is a clone designated herein as "DNA43466-1225".

Figure 118 shows the amino acid sequence (SEQ ID NO:339) derived from the coding sequence of SEQ ID NO:338 shown in Figure 117.

Figure 119 shows a nucleotide sequence (SEQ ID NO:340) of a native sequence PRO310 cDNA, wherein SEQ ID NO:340 is a clone designated herein as "DNA43046-1225".

Figure 120 shows the amino acid sequence (SEQ ID NO:341) derived from the coding sequence of SEQ ID NO:340 shown in Figure 119.

Figure 121 shows a nucleotide sequence (SEQ ID NO:376) of a native sequence PRO244 cDNA, wherein SEQ ID NO:376 is a clone designated herein as "DNA35668-1171".

Figure 122 shows the amino acid sequence (SEQ ID NO:377) derived from the coding sequence of SEQ ID NO:376 shown in Figure 121.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

1. Definitions

The terms "PRO polypeptide" and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (i.e., PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and "PRO/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the PRO polypeptides.

The PRO polypeptide "extracellular domain" or "ECD" refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

The approximate location of the "signal peptides" of the various PRO polypeptides disclosed herein are shown in the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (e.g.,

Nielsen et al., Prot. Eng. 10:1-6 (1997) and von Heinje et al., Nucl. Acids. Res. 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

"PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for instance, PRO polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80% amino acid sequence identity, preferably at least about 81% amino acid sequence identity, more preferably at least about 82% amino acid sequence identity, more preferably at least about 83% amino acid sequence identity, more preferably at least about 84% amino acid sequence identity, more preferably at least about 85% amino acid sequence identity, more preferably at least about 86% amino acid sequence identity, more preferably at least about 87% amino acid sequence identity, more preferably at least about 88% amino acid sequence identity, more preferably at least about 89% amino acid sequence identity, more preferably at least about 90% amino acid sequence identity, more preferably at least about 91% amino acid sequence identity, more preferably at least about 92% amino acid sequence identity, more preferably at least about 93% amino acid sequence identity, more preferably at least about 94% amino acid sequence identity, more preferably at least about 95% amino acid sequence identity, more preferably at least about 96% amino acid sequence identity, more preferably at least about 97% amino acid sequence identity, more preferably at least about 98% amino acid sequence identity and most preferably at least about 99% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10 amino acids in length, often at least about 20 amino acids in length, more often at least about 30 amino acids in length, more often at least about 40 amino acids in length, more often at least about 50 amino acids in length, more often at least about 60 amino acids in length, more often at least about 70 amino acids in length, more often at least about 80 amino acids in length, more often at least about 90 amino acids in length, more often at least about 100 amino acids in length, more often at least about 150 amino acids in length, more often at least about 200 amino acids in length, more often at least about 300 amino acids in length, or more.

"Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific PRO polypeptide sequence, after aligning the sequences and introducing gaps,

if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 10 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 10 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 10 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction X/Y

where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this method, Tables 11 and 12 below demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO".

Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % amino acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acid residues between the amino acid sequence of the PRO polypeptide of interest having a sequence derived from the native PRO polypeptide and the

comparison amino acid sequence of interest (i.e., the sequence against which the PRO polypeptide of interest is being compared which may be a PRO variant polypeptide) as determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest. For example, in the statement "a polypeptide comprising an the amino acid sequence A which has or having at least 80% amino acid sequence identity to the amino acid sequence B", the amino acid sequence A is the comparison amino acid sequence of interest and the amino acid sequence B is the amino acid sequence of the PRO polypeptide of interest.

Percent amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from http://www.ncbi.nlm.nih.gov. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction X/Y

where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

"PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, more preferably at least about 81% nucleic acid sequence identity, more preferably at least about 82% nucleic acid sequence identity, more preferably at least about 83% nucleic acid sequence identity, more preferably at least about 84% nucleic acid sequence identity, more preferably at least about 85% nucleic acid sequence identity, more preferably at least about 86% nucleic acid sequence identity, more preferably at least about 87% nucleic acid sequence identity, more preferably at least about 88% nucleic acid sequence identity, more preferably at least about 89% nucleic acid sequence identity, more preferably at least about 90% nucleic acid sequence

identity, more preferably at least about 91% nucleic acid sequence identity, more preferably at least about 92% nucleic acid sequence identity, more preferably at least about 93% nucleic acid sequence identity, more preferably at least about 94% nucleic acid sequence identity, more preferably at least about 95% nucleic acid sequence identity, more preferably at least about 96% nucleic acid sequence identity, more preferably at least about 97% nucleic acid sequence identity, more preferably at least about 98% nucleic acid sequence identity and yet more preferably at least about 99% nucleic acid sequence identity with the nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, often at least about 60 nucleotides in length, more often at least about 90 nucleotides in length, more often at least about 120 nucleotides in length, more often at least about 150 nucleotides in length, more often at least about 180 nucleotides in length, more often at least about 210 nucleotides in length, more often at least about 240 nucleotides in length, more often at least about 270 nucleotides in length, more often at least about 300 nucleotides in length, more often at least about 450 nucleotides in length, more often at least about 600 nucleotides in length, more often at least about 900 nucleotides in length, or more.

"Percent (%) nucleic acid sequence identity" with respect to PRO-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 10 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 10 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 10 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction W/Z

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 13 and 14 below demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA".

Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % nucleic acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % nucleic acid sequence identity value is determined by dividing (a) the number of matching identical nucleotides between the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest. For example, in the statement "an isolated nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the nucleic acid sequence B", the nucleic acid sequence A is the comparison nucleic acid molecule of interest and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest.

Percent nucleic acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from http://www.ncbi.nlm.nih.gov. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction W/Z

where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C.

In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length PRO polypeptide as disclosed herein. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

The term "positives", in the context of sequence comparison performed as described above, includes residues in the sequences compared that are not identical but have similar properties (e.g. as a result of conservative substitutions, see Table 1 below). For purposes herein, the % value of positives is determined by dividing (a) the number of amino acid residues scoring a positive value between the PRO polypeptide amino acid sequence of interest having a sequence derived from the native PRO polypeptide sequence and the comparison amino acid sequence of interest (i.e., the amino acid sequence against which the PRO polypeptide sequence is being compared) as determined in the BLOSUM62 matrix of WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest.

Unless specifically stated otherwise, the % value of positives is calculated as described in the immediately preceding paragraph. However, in the context of the amino acid sequence identity comparisons performed as described for ALIGN-2 and NCBI-BLAST2 above, includes amino acid residues in the sequences compared that are not only identical, but also those that have similar properties. Amino acid residues that score a positive value to an amino acid residue of interest are those that are either identical to the amino acid residue of interest or are a preferred substitution (as defined in Table 1 below) of the amino acid residue of interest.

For amino acid sequence comparisons using ALIGN-2 or NCBI-BLAST2, the % value of positives of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % positives to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction X/Y

where X is the number of amino acid residues scoring a positive value as defined above by the sequence alignment program ALIGN-2 or NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % positives of A to B will not equal the % positives of B to A.

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic

uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" PRO polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRO polypeptide nucleic acid. An isolated PRO polypeptide nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated PRO polypeptide nucleic acid molecules therefore are distinguished from the specific PRO polypeptide nucleic acid molecule as it exists in natural cells. However, an isolated PRO polypeptide nucleic acid molecule includes PRO polypeptide nucleic acid molecules contained in cells that ordinarily express the PRO polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-PRO monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-PRO antibody compositions with polyepitopic specificity, single chain anti-PRO antibodies, and fragments of anti-PRO antibodies (see below). The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when

complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent that those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any

immunoglobulin, such as IgG-1, IgG 2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

"Active" or "activity" for the purposes herein refers to form(s) of a PRO polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring PRO, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native PRO polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for identifying agonists or antagonists of a PRO polypeptide may comprise contacting a PRO polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the PRO polypeptide.

"Treatment" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

"Mammal" for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-

forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', $F(ab')_2$, and Fv fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an $F(ab')_2$ fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the VH-VL dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. $F(ab')_2$ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

"Single-chain Fv" or "sFv" antibody fragments comprise the VH and VL domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the VH and VL domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenburg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (VH) connected to a light-chain variable domain (VL) in the same polypeptide chain (VH - VL). By using a linker that is too short to allow pairing between the two domains

on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

An "isolated" antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody in situ within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

By "solid phase" is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and siliconcs. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a PRO polypeptide or antibody thereto) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

A "small molecule" is defined herein to have a molecular weight below about 500 Daltons.

"PRO317-associated disorder" refers to a pathological condition or disease wherein PRO317 is over- or underexpressed. Such disorders include diseases of the female genital tract or of the endometrium of a mammal, including hyperplasia, endometritis, endometriosis, wherein the patient is at risk for infertility due to endometrial factor, endometrioma, and endometrial cancer, especially those diseases involving abnormal bleeding such as a gynecological disease. They also include diseases involving angiogenesis, wherein the angiogenesis results in a pathological condition, such as cancer involving solid tumors (the therapy for the disorder would result in decreased vascularization and a decline in growth and metastasis of a variety of tumors). Alternatively, the angiogenesis may be beneficial, such as for ischemia, especially coronary ischemia. Hence, these disorders include those found in patients whose hearts are functioning but who have a blocked blood supply due to atherosclerotic coronary artery disease, and those with a functioning but underperfused heart, including patients

with coronary arterial disease who are not optimal candidates for angioplasty and coronary artery by-pass surgery. The disorders also include diseases involving the kidney or originating from the kidney tissue, such as polycystic kidney disease and chronic and acute renal failure.

II. Compositions and Methods of the Invention

A. Full-Length PRO Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding various PRO polypeptides have been identified and isolated, as disclosed in further detail in the Examples below. It is noted that proteins produced in separate expression rounds may be given different PRO numbers but the UNQ number is unique for any given DNA and the encoded protein, and will not be changed. However, for sake of simplicity, in the present specification the protein encoded by the full length native nucleic acid molecules disclosed herein as well as all further native homologues and variants included in the foregoing definition of PRO, will be referred to as "PRO/number", regardless of their origin or mode of preparation.

As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The actual nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the PRO polypeptides and encoding nucleic acids described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

1. **Full-length PRO211 and PRO217 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO211 and PRO217. In particular, Applicants have identified and isolated cDNA encoding PRO211 and PRO217 polypeptides, as disclosed in further detail in the Examples below. Using BLAST (FastA format) sequence alignment computer programs, Applicants found that cDNA sequences encoding full-length native sequence PRO211 and PRO217 have homologies to known proteins having EGF-like domains. Specifically, the cDNA sequence DNA32292-1131 (Figure 1, SEQ ID NO:1) has certain identity and a Blast score of 209 with PAC6_RAT and certain identity and a Blast score of 206 with Fibulin-1, isoform c precursor. The cDNA sequence DNA33094-1131 (Figure 3, SEQ ID NO:3) has 36% identity and a Blast score of 336 with eastern newt tenascin, and 37% identity and a Blast score of 331 with human tenascin-X precursor. Accordingly, it is presently believed that PRO211 and PRO217 polypeptides disclosed in the present application are newly identified members of the EGF-like family and possesses properties typical of the EGF-like protein family.

2. **Full-length PRO230 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO230. In particular, Applicants have identified and isolated cDNA

encoding a PRO230 polypeptide, as disclosed in further detail in the Examples below. Using known programs such as BLAST and FastA sequence alignment computer programs, Applicants found that a cDNA sequence encoding full-length native sequence PRO230 has 48% amino acid identity with the rabbit tubulointerstitial nephritis antigen precursor. Accordingly, it is presently believed that PRO230 polypeptide disclosed in the present application is a newly identified member of the tubulointerstitial nephritis antigen family and possesses the ability to be recognized by human autoantibodies in certain forms of tubulointerstitial nephritis.

3. **Full-length PRO232 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO232. In particular, Applicants have identified and isolated cDNA encoding a PRO232 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a portion of the full-length native sequence PRO232 (shown in Figure 9 and SEQ ID NO:18) has 35% sequence identity with a stem cell surface antigen from Gallus gallus. Accordingly, it is presently believed that the PRO232 polypeptide disclosed in the present application may be a newly identified stem cell antigen.

4. **Full-length PRO187 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO187. In particular, Applicants have identified and isolated cDNA encoding a PRO187 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO187 (shown in Figure 15) has 74% amino acid sequence identity and BLAST score of 310 with various androgen-induced growth factors and FGF-8. Accordingly, it is presently believed that PRO187 polypeptide disclosed in the present application is a newly identified member of the FGF-8 protein family and may possess identify activity or property typical of the FGF-8-like protein family.

5. **Full-length PRO265 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO265. In particular, Applicants have identified and isolated cDNA encoding a PRO265 polypeptide, as disclosed in further detail in the Examples below. Using programs such as BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO265 polypeptide have significant homology with the fibromodulin protein and fibromodulin precursor protein. Applicants have also found that the DNA encoding the PRO265 polypeptide has significant homology with platelet glycoprotein V, a member of the leucine rich related protein family involved in skin and wound repair. Accordingly, it is presently believed that PRO265 polypeptide disclosed in the present application is a newly identified member of the leucine rich repeat family and possesses protein protein binding capabilities, as well as be involved in skin and wound repair as typical of this family.

6. Full-length PRO219 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO219. In particular, Applicants have identified and isolated cDNA encoding a PRO219 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO219 polypeptide have significant homology with the mouse and human matrilin-2 precursor polypeptides. Accordingly, it is presently believed that PRO219 polypeptide disclosed in the present application is related to the matrilin-2 precursor polypeptide.

7. Full-length PRO246 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO246. In particular, Applicants have identified and isolated cDNA encoding a PRO246 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a portion of the PRO246 polypeptide has significant homology with the human cell surface protein HCAR. Accordingly, it is presently believed that PRO246 polypeptide disclosed in the present application may be a newly identified membrane-bound virus receptor or tumor cell-specific antigen.

8. Full-length PRO228 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO228. In particular, Applicants have identified and isolated cDNA encoding a PRO228 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO228 polypeptide have significant homology with the EMR1 protein. Applicants have also found that the DNA encoding the PRO228 polypeptide has significant homology with latrophilin, macrophage restricted cell surface glycoprotein, B0457.1 and leucocyte antigen CD97 precursor. Accordingly, it is presently believed that PRO228 polypeptide disclosed in the present application is a newly identified member of the seven transmembrane superfamily and possesses characteristics and functional properties typical of this family. In particular, it is believed that PRO228 is a new member of the subgroup within this family to which CD97 and EMR1 belong.

9. Full-length PRO533 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO533. In particular, Applicants have identified and isolated cDNA encoding a PRO533 polypeptide, as disclosed in further detail in the Examples below. Using BLAST-2 and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO533 (shown in Figure 22 and SEQ ID NO:59) has a Blast score of 509 and 53% amino acid sequence identity with fibroblast growth factor (FGF). Accordingly, it is presently believed that PRO533 disclosed in the present application is a newly identified member of the fibroblast growth factor family and may possess activity typical

of such polypeptides.

10. **Full-length PRO245 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO245. In particular, Applicants have identified and isolated cDNA encoding a PRO245 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a portion of the amino acid sequence of the PRO245 polypeptide has 60% amino acid identity with the human c-myb protein. Accordingly, it is presently believed that the PRO245 polypeptide disclosed in the present application may be a newly identified member of the transmembrane protein tyrosine kinase family.

11. **Full-length PRO220, PRO221 and PRO227 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO220, PRO221 and PRO227. In particular, Applicants have identified and isolated cDNAs encoding a PRO220, PRO221 and PRO227 polypeptide, respectively, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, PRO220 has amino acid identity with the amino acid sequence of a leucine rich protein wherein the identity is 87%. PRO220 additionally has amino acid identity with the neuronal leucine rich protein wherein the identity is 55%. The neuronal leucine rich protein is further described in Taguchi, *et al.*, Mol. Brain Res., 35:31-40 (1996).

PRO221 has amino acid identity with the SLIT protein precursor, wherein different portions of these two proteins have the respective percent identities of 39%, 38%, 34%, 31%, and 30%.

PRO227 has amino acid identity with the amino acid sequence of platelet glycoprotein V precursor. The same results were obtained for human glycoprotein V. Different portions of these two proteins show the following percent identities of 30%, 28%, 28%, 31%, 35%, 39% and 27%.

Accordingly, it is presently believed that PRO220, PRO221 and PRO227 polypeptides disclosed in the present application are newly identified members of the leucine rich repeat protein superfamily and that each possesses protein-protein binding capabilities typical of the leucine rich repeat protein superfamily. It is also believed that they have capabilities similar to those of SLIT, the leucine rich repeat protein and human glycoprotein V.

12. **Full-length PRO258 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO258. In particular, Applicants have identified and isolated cDNA encoding a PRO258 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO258 polypeptide have significant homology with the CRTAM and poliovirus receptors. Accordingly, it is presently believed that PRO258 polypeptide disclosed in the present application is a newly identified member of the Ig superfamily and possesses virus receptor capabilities or regulates immune function as typical of this family.

13. Full-length PRO266 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO266. In particular, Applicants have identified and isolated cDNA encoding a PRO266 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO266 polypeptide have significant homology with the SLIT protein from Drosophilia. Accordingly, it is presently believed that PRO266 polypeptide disclosed in the present application is a newly identified member of the leucine rich repeat family and possesses ligand-ligand binding activity and neuronal development typical of this family. SLIT has been shown to be useful in the study and treatment of Alzheimer's disease, *supra*, and thus, PRO266 may have involvement in the study and cure of this disease.

14. Full-length PRO269 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO269. In particular, Applicants have identified and isolated cDNA encoding a PRO269 polypeptide, as disclosed in further detail in the Examples below. Using BLAST, FastA and sequence alignment computer programs, Applicants found that the amino acid sequence encoded by nucleotides 314 to 1783 of the full-length native sequence PRO269 (shown in Figure 35 and SEQ ID NO:95) has significant homology to human urinary thrombomodulin and various thrombomodulin analogues respectively, to which it was aligned. Accordingly, it is presently believed that PRO269 polypeptide disclosed in the present application is a newly identified member of the thrombomodulin family.

15. Full-length PRO287 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO287. In particular, Applicants have identified and isolated cDNA encoding a PRO287 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO287 polypeptide have significant homology with the type I procollagen C-proteinase enhancer protein precursor and type I procollagen C-proteinase enhancer protein. Accordingly, it is presently believed that PRO287 polypeptide disclosed in the present application is a newly identified member of the C-proteinase enhancer protein family.

16. Full-length PRO214 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO214. In particular, Applicants have identified and isolated cDNA encoding a PRO214 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO214 polypeptide (shown in Figure 40 and SEQ ID NO:109) has 49% amino acid sequence identity with HT protein, a known member of the EGF-family. The comparison resulted in a BLAST score of 920, with 150 matching nucleotides. Accordingly, it is presently believed that the PRO214 polypeptide disclosed in the present application is a newly

identified member of the family comprising EGF domains and may possess activities or properties typical of the EGF-domain containing family.

17. Full-length PRO317 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO317. In particular, cDNA encoding a PRO317 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below. Using BLAST ™ and FastA™ sequence alignment computer programs, it was found that a full-length native-sequence PRO317 (shown in Figure 42 and SEQ ID NO:114) has 92% amino acid sequence identity with EBAF-1. Further, it is closely aligned with many other members of the TGF- superfamily.

Accordingly, it is presently believed that PRO317 disclosed in the present application is a newly identified member of the TGF- superfamily and may possess properties that are therapeutically useful in conditions of uterine bleeding, etc. Hence, PRO317 may be useful in diagnosing or treating abnormal bleeding involved in gynecological diseases, for example, to avoid or lessen the need for a hysterectomy. PRO317 may also be useful as an agent that affects angiogenesis in general, so PRO317 may be useful in anti-tumor indications, or conversely, in treating coronary ischemic conditions.

Library sources reveal that ESTs used to obtain the consensus DNA for generating PRO317 primers and probes were found in normal tissues (uterus, prostate, colon, and pancreas), in several tumors (colon, brain (twice), pancreas, and mullerian cell), and in a heart with ischemia. PRO317 has shown up in several tissues as well, but it does look to have a greater concentration in uterus. Hence, PRO317 may have a broader use by the body than EBAF-1. It is contemplated that, at least for some indications, PRO317 may have opposite effects from EBAF-1.

18. Full-length PRO301 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO301. In particular, Applicants have identified and isolated cDNA encoding a PRO301 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO301 (shown in Figure 44 and SEQ ID NO:119) has a Blast score of 246 corresponding to 30% amino acid sequence identity with human A33 antigen precursor. Accordingly, it is presently believed that PRO301 disclosed in the present application is a newly identified member of the A33 antigen protein family and may be expressed in human neoplastic diseases such as colorectal cancer.

19. Full-length PRO224 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO224. In particular, Applicants have identified and isolated cDNA encoding a PRO224 polypeptide, as disclosed in further detail in the Examples below. Using known programs such as BLAST and FastA sequence alignment computer programs, Applicants found that full-length native

PRO224 (Figure 46, SEQ ID NO:127) has amino acid identity with apolipoprotein E receptor 2906 from homo sapiens. The alignments of different portions of these two polypeptides show amino acid identities of 37%, 36%, 30%, 44%, 44% and 28% respectively. Full-length native PRO224 (Figure 46, SEQ ID NO:127) also has amino acid identity with very low-density lipoprotein receptor precursor from gall. The alignments of different portions of these two polypeptides show amino acid identities of 38%, 37%, 42%, 33%, and 37% respectively. Additionally, full-length native PRO224 (Figure 46, SEQ ID NO:127) has amino acid identity with the chicken oocyte receptor P95 from Gallus gallus. The alignments of different portions of these two polypeptides show amino acid identities of 38%, 37%, 42%, 33%, and 37% respectively. Moreover, full-length native PRO224 (Figure 46, SEQ ID NO:127) has amino acid identity with very low density lipoprotein receptor short form precursor from humans. The alignments of different portions of these two polypeptides show amino acid identities of 32%, 38%, 34%, 45%, and 31%, respectively. Accordingly, it is presently believed that PRO224 polypeptide disclosed in the present application is a newly identified member of the low density lipoprotein receptor family and possesses the structural characteristics required to have the functional ability to recognize and endocytose low density lipoproteins typical of the low density lipoprotein receptor family. (The alignments described above used the following scoring parameters: T=7, S+65, S2=36, Matrix: BLOSUM62.)

20. **Full-length PRO222 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO222. In particular, Applicants have identified and isolated cDNA encoding a PRO222 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a sequence encoding full-length native sequence PRO222 (shown in Figure 48 and SEQ ID NO:132) has 25-26% amino acid identity with mouse complement factor h precursor, has 27-29% amino acid identity with complement receptor, has 25-47% amino acid identity with mouse complement C3b receptor type 2 long form precursor, has 40% amino acid identity with human hypothetical protein kiaa0247. Accordingly, it is presently believed that PRO222 polypeptide disclosed in the present application is a newly identified member of the complement receptor family and possesses activity typical of the complement receptor family.

21. **Full-length PRO234 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO234. In particular, Applicants have identified and isolated cDNA encoding a PRO234 polypeptide, as disclosed in further detail in the Examples below. Using BLAST (FastA-format) sequence alignment computer programs, Applicants found that a cDNA sequence encoding full-length native sequence PRO234 has 31% identity and Blast score of 134 with E-selectin precursor. Accordingly, it is presently believed that the PRO234 polypeptides disclosed in the present application are newly identified members of the lectin/selectin family and possess activity typical of the lectin/selectin family.

22. Full-length PRO231 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO231. In particular, Applicants have identified and isolated cDNA encoding a PRO231 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the full-length native sequence PRO231 polypeptide (shown in Figure 52 and SEQ ID NO:142) has 30 % and 31 % amino acid identity with human and rat prostatic acid phosphatase precursor proteins, respectively. Accordingly, it is presently believed that the PRO231 polypeptide disclosed in the present application may be a newly identified member of the acid phosphatase protein family.

23. Full-length PRO229 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO229. In particular, Applicants have identified and isolated cDNA encoding a PRO229 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO229 polypeptide have significant homology with antigen wc1.1, M130 antigen, T cell surface glycoprotein CD6 and CD6. It also is related to Sp-alpha. Accordingly, it is presently believed that PRO229 polypeptide disclosed in the present application is a newly identified member of the family containing scavenger receptor homology, a sequence motif found in a number of proteins involved in immune function and thus possesses immune function and /or segments which resist degradation, typical of this family.

24. Full-length PRO238 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO238. In particular, Applicants have identified and isolated cDNA encoding a PRO238 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO238 polypeptide have significant homology with reductases, including oxidoreductase and fatty acyl-CoA reductase. Accordingly, it is presently believed that PRO238 polypeptide disclosed in the present application is a newly identified member of the reductase family and possesses reducing activity typical of the reductase family.

25. Full-length PRO233 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO233. In particular, Applicants have identified and isolated cDNA encoding a PRO233 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO233 polypeptide have significant homology with the reductase protein. Applicants have also found that the DNA encoding the PRO233 polypeptide has significant homology with proteins from *Caenorhabditis elegans*. Accordingly, it is presently believed that PRO233 polypeptide disclosed in the present application is a newly identified member of the

reductase family and possesses the ability to effect the redox state of the cell typical of the reductase family.

26. **Full-length PRO223 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO223. In particular, Applicants have identified and isolated cDNA encoding a PRO223 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO223 polypeptide has significant homology with various serine carboxypeptidase polypeptides. Accordingly, it is presently believed that PRO223 polypeptide disclosed in the present application is a newly identified serine carboxypeptidase.

27. **Full-length PRO235 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO235. In particular, Applicants have identified and isolated cDNA encoding a PRO235 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO235 polypeptide have significant homology with the various plexin proteins. Accordingly, it is presently believed that PRO235 polypeptide disclosed in the present application is a newly identified member of the plexin family and possesses cell adhesion properties typical of the plexin family.

28. **Full-length PRO236 and PRO262 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO236 and PRO262. In particular, Applicants have identified and isolated cDNA encoding PRO236 and PRO262 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO236 and PRO262 polypeptides have significant homology with various β-galactosidase and β-galactosidase precursor polypeptides. Accordingly, it is presently believed that the PRO236 and PRO262 polypeptides disclosed in the present application are newly identified β-galactosidase homologs.

29. **Full-length PRO239 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO239. In particular, Applicants have identified and isolated cDNA encoding a PRO239 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO239 polypeptide have significant homology with densin proteins. Accordingly, it is presently believed that PRO239 polypeptide disclosed in the present application is a newly identified member of the densin family and possesses cell adhesion and the ability to effect synaptic processes as is typical of the densin family.

30. **Full-length PRO257 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO257. In particular, Applicants have identified and isolated cDNA encoding a PRO257 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO257 polypeptide have significant homology with the ebnerin precursor and ebnerin protein. Accordingly, it is presently believed that PRO257 polypeptide disclosed in the present application is a newly identified protein member which is related to the ebnerin protein.

31. **Full-length PRO260 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO260. In particular, Applicants have identified and isolated cDNA encoding a PRO260 polypeptide, as disclosed in further detail in the Examples below. Using programs such as BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO260 polypeptide have significant homology with the alpha-l-fucosidase precursor. Accordingly, it is presently believed that PRO260 polypeptide disclosed in the present application is a newly identified member of the fucosidase family and possesses enzymatic activity related to fucose residues typical of the fucosidase family.

32. **Full-length PRO263 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO263. In particular, Applicants have identified and isolated cDNA encoding a PRO263 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO263 polypeptide have significant homology with the CD44 antigen and related proteins. Accordingly, it is presently believed that PRO263 polypeptide disclosed in the present application is a newly identified member of the CD44 antigen family and possesses at least one of the properties associated with these antigens, i.e., cancer and HIV marker, cell-cell or cell-matrix interactions, regulating cell traffic, lymph node homing, transmission of growth signals, and presentation of chemokines and growth factors to traveling cells.

33. **Full-length PRO270 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO270. In particular, Applicants have identified and isolated cDNA encoding a PRO270 polypeptide, as disclosed in further detail in the Examples below. Using BLAST, FastA and sequence alignment computer programs, Applicants found that that various portions of the PRO270 polypeptide have significant homology with various thioredoxin proteins. Accordingly, it is presently believed that PRO270 polypeptide disclosed in the present application is a newly identified member of the thioredoxin family and possesses the ability to effect reduction-oxidation (redox) state typical of the thioredoxin family.

34. Full-length PRO271 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO271. In particular, Applicants have identified and isolated cDNA encoding a PRO271 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO271 polypeptide has significant homology with various link proteins and precursors thereof. Accordingly, it is presently believed that PRO271 polypeptide disclosed in the present application is a newly identified link protein homolog.

35. Full-length PRO272 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO272. In particular, Applicants have identified and isolated cDNA encoding a PRO272 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO272 polypeptide have significant homology with the human reticulocalbin protein and its precursors. Applicants have also found that the DNA encoding the PRO272 polypeptide has significant homology with the mouse reticulocalbin precursor protein. Accordingly, it is presently believed that PRO272 polypeptide disclosed in the present application is a newly identified member of the reticulocalbin family and possesses the ability to bind calcium typical of the reticulocalbin family.

36. Full-length PRO294 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO294. In particular, Applicants have identified and isolated cDNA encoding a PRO294 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO294 polypeptide have significant homology with the various portions of a number of collagen proteins. Accordingly, it is presently believed that PRO294 polypeptide disclosed in the present application is a newly identified member of the collagen family.

37. Full-length PRO295 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO295. In particular, Applicants have identified and isolated cDNA encoding a PRO295 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO295 polypeptide have significant homology with integrin proteins. Accordingly, it is presently believed that PRO295 polypeptide disclosed in the present application is a newly identified member of the integrin family and possesses cell adhesion typical of the integrin family.

38. Full-length PRO293 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO293. In particular, Applicants have identified and isolated cDNA encoding a PRO293 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that portions of the PRO293 polypeptide have significant homology with the neuronal leucine rich repeat proteins 1 and 2, (NLRR-1 and NLRR-2), particularly NLRR-2. Accordingly, it is presently believed that PRO293 polypeptide disclosed in the present application is a newly identified member of the neuronal leucine rich repeat protein family and possesses ligand-ligand binding activity typical of the NRLL protein family.

39. Full-length PRO247 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO247. In particular, Applicants have identified and isolated cDNA encoding a PRO247 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO247 polypeptide have significant homology with densin. Applicants have also found that the DNA encoding the PRO247 polypeptide has significant homology with a number of other proteins, including KIAA0231. Accordingly, it is presently believed that PRO247 polypeptide disclosed in the present application is a newly identified member of the leucine rich repeat family and possesses ligand binding abilities typical of this family.

40. Full-length PRO302, PRO303, PRO304, PRO307 and PRO343 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO302, PRO303, PRO304, PRO307 and PRO343. In particular, Applicants have identified and isolated cDNA encoding PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides have significant homology with various protease proteins. Accordingly, it is presently believed that the PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides disclosed in the present application are newly identified protease proteins.

41. Full-length PRO328 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO328. In particular, Applicants have identified and isolated cDNA encoding a PRO328 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO328 polypeptide have significant homology with the human glioblastoma protein ("GLIP"). Further, Applicants found that various portions of the PRO328 polypeptide have significant homology with the cysteine rich secretory protein ("CRISP") as identified by BLAST homology [ECCRISP3_1, S68683, and CRS3_HUMAN]. Accordingly, it

is presently believed that PRO328 polypeptide disclosed in the present application is a newly identified member of the GLIP or CRISP families and possesses transcriptional regulatory activity typical of the GLIP or CRISP families.

42. Full-length PRO335, PRO331 and PRO326 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO335, PRO331 or PRO326. In particular, Applicants have identified and isolated cDNA encoding a PRO335, PRO331 or PRO326 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO335, PRO331 or PRO326 polypeptide have significant homology with LIG-1, ALS and in the case of PRO331, additionally, decorin. Accordingly, it is presently believed that the PRO335, PRO331 and PRO326 polypeptides disclosed in the present application are newly identified members of the leucine rich repeat superfamily, and particularly, are related to LIG-1 and possess the biological functions of this family as discussed and referenced herein.

43. Full-length PRO332 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO332. In particular, Applicants have identified and isolated cDNA encoding PRO332 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO332 (shown in Figure 108 and SEQ ID NO:310) has about 30-40% amino acid sequence identity with a series of known proteoglycan sequences, including, for example, fibromodulin and fibromodulin precursor sequences of various species (FMOD_BOVIN, FMOD_CHICK, FMOD_RAT, FMOD_MOUSE, FMOD_HUMAN, P_R36773), osteomodulin sequences (AB000114_1, AB007848_1), decorin sequences (CFUR3141_1, OCU03394_1, P R42266, P_R42267, P_R42260, P_R89439), keratan sulfate proteoglycans (BTU48360_1, AF022890_1), corneal proteoglycan (AF022256_1), and bone/cartilage proteoglycans and proteoglycane precursors (PGS1_BOVIN, PGS2_MOUSE, PGS2_HUMAN). Accordingly, it is presently believed that PRO332 disclosed in the present application is a new proteoglycan-type molecule, and may play a role in regulating extracellular matrix, cartilage, and/or bone function.

44. Full-length PRO334 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO334. In particular, Applicants have identified and isolated cDNA encoding a PRO334 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO334 polypeptide have significant homology with fibulin and fibrillin. Accordingly, it is presently believed that PRO334 polypeptide disclosed in the present application is a newly identified member of the epidermal growth factor family and possesses properties and activities typical of this family.

45. **Full-length PRO346 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO346. In particular, Applicants have identified and isolated cDNA encoding a PRO346 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO346 (shown in Figure 112 and SEQ ID NO:320) has 28 % amino acid sequence identity with carcinoembryonic antigen. Accordingly, it is presently believed that PRO346 disclosed in the present application is a newly identified member of the carcinoembryonic protein family and may be expressed in association with neoplastic tissue disorders.

46. **Full-length PRO268 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO268. In particular, Applicants have identified and isolated cDNA encoding a PRO268 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that portions of the PRO268 polypeptide have significant homology with the various protein disulfide isomerase proteins. Accordingly, it is presently believed that PRO268 polypeptide disclosed in the present application is a homolog of the protein disulfide isomerase p5 protein.

47. **Full-length PRO330 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO330. In particular, Applicants have identified and isolated cDNA encoding a PRO330 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO330 polypeptide have significant homology with the murine prolyl 4-hydroxylase alpha-II subunit protein. Accordingly, it is presently believed that PRO330 polypeptide disclosed in the present application is a novel prolyl 4-hydroxylase subunit polypeptide.

48. **Full-length PRO339 and PRO310 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO339 and PRO310. In particular, Applicants have identified and isolated cDNA encoding a PRO339 polypeptide, as disclosed in further detail in the Examples below. Applicants have also identified and isolated cDNA encoding a PRO310 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO339 and PRO310 polypeptides have significant homology with small secreted proteins from *C. elegans* and are distantly related to fringe. PRO339 also shows homology to collagen-like polymers. Sequences which were used to identify PRO310, designated herein as DNA40533 and DNA42267, also show homology to proteins from *C. elegans*. Accordingly, it is presently believed that the PRO339 and

PRO310 polypeptides disclosed in the present application are newly identified member of the family of proteins involved in development, and which may have regulatory abilities similar to the capability of fringe to regulate serrate.

49. Full Length PRO244 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding C-type lectins referred to in the present application as PRO244. In particular, applicants have identified and isolated cDNA encoding PRO244 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO244 (shown in Figure 122 and SEQ ID NO:377) has 43% amino acid sequence identity with the hepatic lectin gallus gallus (LECH-CHICK), and 42% amino acid sequence identity with an HIV gp120 binding C-type lectin (A46274). Accordingly, it is presently believed that PRO244 disclosed in the present application is a newly identified member of the C-lectin superfamily and may play a role in immune function, apoptosis, or in the pathogenesis of atherosclerosis. In addition, PRO244 may be useful in identifying tumor-associated epitopes.

B. PRO Polypeptide Variants

In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the native full-length sequence PRO or in various domains of the PRO described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO that results in a change in the amino acid sequence of the PRO as compared with the native sequence PRO. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO

polypeptide.

PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide disclosed herein.

In particular embodiments, conservative substitutions of interest are shown in Table 1 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 1, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 1

Original Residue	Exemplary Substitutions	Preferred Substitutions
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucine	leu
Leu (L)	norleucine; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala; tyr	leu
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucine	leu

Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

(1) hydrophobic: norleucine, met, ala, val, leu, ile;

(2) neutral hydrophilic: cys, ser, thr;

(3) acidic: asp, glu;

(4) basic: asn, gln, his, lys, arg;

(5) residues that influence chain orientation: gly, pro; and

(6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant

DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

C. Modifications of PRO

Covalent modifications of PRO are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the PRO. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α-amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence PRO. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO (for O-linked glycosylation sites). The PRO amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The PRO of the present invention may also be modified in a way to form a chimeric molecule comprising PRO fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the PRO with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO. The presence of such epitope-tagged forms of the PRO can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α-tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

D. Preparation of PRO

The description below relates primarily to production of PRO by culturing cells transformed or transfected with a vector containing PRO nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO. For instance, the PRO sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO.

1. Isolation of DNA Encoding PRO

DNA encoding PRO may be obtained from a cDNA library prepared from tissue believed to possess the PRO mRNA and to express it at a detectable level. Accordingly, human PRO DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

Libraries can be screened with probes (such as antibodies to the PRO or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ^{32}P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, $CaCl_2$, $CaPO_4$, liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA* ; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan*r; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7*

ilvG kan^r; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., Bio/Technology, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 737 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickeramii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilarum* (ATCC 36,906; Van den Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus; yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida; Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora, Penicillium, Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn et al., Gene, 26:205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylotropic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula, Candida, Kloeckera, Pichia, Saccharomyces, Torulopsis,* and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

Suitable host cells for the expression of glycosylated PRO are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as Drosophila S2 and Spodoptera Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally

include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The PRO may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α-factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp*1 gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp*1 gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β-lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid

promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphateisomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding the PRO by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of PRO may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO produced.

E. Uses for PRO

Nucleotide sequences (or their complement) encoding PRO have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO nucleic acid will also be useful for the preparation of PRO polypeptides by the recombinant techniques described herein.

The full-length native sequence PRO gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of PRO or PRO from other species) which have a desired sequence identity to the native PRO sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO. By way of example, a screening method will comprise isolating the coding region of the PRO gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

Any EST sequences disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

Other useful fragments of the PRO nucleic acids include antisense or sense oligonucleotides comprising a singe-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target PRO mRNA (sense) or PRO DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of PRO DNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. The antisense oligonucleotides thus may be used to block expression of PRO proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid

sequence by any gene transfer method, including, for example, $CaPO_4$-mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retrovirus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 90/13641).

Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

Antisense RNA or DNA molecules are generally at least about 5 bases in length, about 10 bases in length, about 15 bases in length, about 20 bases in length, about 25 bases in length, about 30 bases in length, about 35 bases in length, about 40 bases in length, about 45 bases in length, about 50 bases in length, about 55 bases in length, about 60 bases in length, about 65 bases in length, about 70 bases in length, about 75 bases in length, about 80 bases in length, about 85 bases in length, about 90 bases in length, about 95 bases in length, about 100 bases in length, or more.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO coding sequences.

Nucleotide sequences encoding a PRO can also be used to construct hybridization probes for mapping the gene which encodes that PRO and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO encode a protein which binds to another protein (example, where the PRO is a receptor), the PRO can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO or a receptor for PRO. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a

variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO can be used to construct a PRO "knock out" animal which has a defective or altered gene encoding PRO as a result of homologous recombination between the endogenous gene encoding PRO and altered genomic DNA encoding PRO introduced into an embryonic stem cell of the animal. For example, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques. A portion of the genomic DNA encoding PRO can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, Cell, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., Cell, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO polypeptide.

Nucleic acid encoding the PRO polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve *in vivo* synthesis of a therapeutically effective genetic product, for example for replacement of a defective gene. "Gene therapy" includes both conventional gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA. Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes *in vivo*. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane. (Zamecnik *et al.*, Proc. Natl. Acad. Sci. USA 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by uncharged groups.

There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. The currently preferred *in vivo* gene transfer techniques include transfection with viral (typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau et al., Trends in Biotechnology 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellularlocalization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu et al., J. Biol. Chem. 262, 4429-4432 (1987); and Wagner et al., Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). For review of gene marking and gene therapy protocols see Anderson et al., Science 256, 808-813 (1992).

The PRO polypeptides described herein may also be employed as molecular weight markers for protein electrophoresis purposes and the isolated nucleic acid sequences may be used for recombinantly expressing those markers.

The nucleic acid molecules encoding the PRO polypeptides or fragments thereof described herein are useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each PRO nucleic acid molecule of the present invention can be used as a chromosome marker.

The PRO polypeptides and nucleic acid molecules of the present invention may also be used for tissue typing, wherein the PRO polypeptides of the present invention may be differentially expressed in one tissue as compared to another. PRO nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

The PRO polypeptides described herein may also be employed as therapeutic agents. The PRO polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the PRO product hereof is combined in admixture with a pharmaceutically acceptable carrier vehicle. Therapeutic formulations are prepared for storage by mixing the active ingredient having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, PLURONICS™ or PEG.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

Therapeutic compositions herein generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

The route of administration is in accord with known methods, e.g. injection or infusion by intravenous, intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial or intralesional routes, topical administration, or by sustained release systems.

Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician. Animal experiments provide reliable guidance for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordenti, J. and Chappell, W. "The use of interspecies scaling in toxicokinetics" In Toxicokinetics and New Drug Development, Yacobi et al., Eds., Pergamon Press, New York 1989, pp. 42-96.

When *in vivo* administration of a PRO polypeptide or agonist or antagonist thereof is employed, normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 μg/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

Where sustained-release administration of a PRO polypeptide is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of the PRO polypeptide, microencapsulation of the PRO polypeptide is contemplated. Microencapsulation of recombinant

proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon-(rhIFN-), interleukin-2, and MN rgp120. Johnson et al., Nat. Med., 2:795-799 (1996); Yasuda, Biomed. Ther., 27:1221-1223 (1993); Hora et al., Bio/Technology, 8:755-758 (1990); Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit and Adjuvant Approach, Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S. Pat. No. 5,654,010.

The sustained-release formulations of these proteins were developed using poly-lactic-coglycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation products of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: M. Chasin and R. Langer (Eds.), Biodegradable Polymers as Drug Delivery Systems (Marcel Dekker: New York, 1990), pp. 1-41.

This invention encompasses methods of screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the PRO polypeptides encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

All assays for antagonists are common in that they call for contacting the drug candidate with a PRO polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, the PRO polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the PRO polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the PRO polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

If the candidate compound interacts with but does not bind to a particular PRO polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers (Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β-galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a PRO polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

To assay for antagonists, the PRO polypeptide may be added to a cell along with the compound to be screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the PRO polypeptide indicates that the compound is an antagonist to the PRO polypeptide. Alternatively, antagonists may be detected by combining the PRO polypeptide and a potential antagonist with membrane-bound PRO polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO polypeptide can be labeled, such as by radioactivity, such that the number of PRO polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991).

Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the PRO polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the PRO polypeptide. Transfected cells that are grown on glass slides are exposed to labeled PRO polypeptide. The PRO polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single clone that encodes the putative receptor.

As an alternative approach for receptor identification, labeled PRO polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro- sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with labeled PRO polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with PRO polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the PRO polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the PRO polypeptide.

Another potential PRO polypeptide antagonist is an antisense RNA or DNA construct prepared using antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature PRO polypeptides herein, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al., Science, 241: 456 (1988); Dervan et al., Science, 251:1360 (1991)), thereby preventing transcription and the production of the PRO polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the PRO polypeptide (antisense - Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the PRO

polypeptide. When antisense DNA is used, oligodeoxyribonucleotidesderived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the PRO polypeptide, thereby blocking the normal biological activity of the PRO polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, Current Biology, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

With regard to the PRO211 and PRO217 polypeptide, therapeutic indications include disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g., enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g., psoriasis, epithelial cancers such as lung squamous cell carcinoma, epidermoid carcinoma of the vulva and gliomas.

Since the PRO232 polypeptide and nucleic acid encoding it possess sequence homology to a cell surface stem cell antigen and its encoding nucleic acid, probes based upon the PRO232 nucleotide sequence may be employed to identify other novel stem cell surface antigen proteins. Soluble forms of the PRO232 polypeptide may be employed as antagonists of membrane bound PRO232 activity both *in vitro* and *in vivo*. PRO232 polypeptides may be employed in screening assays designed to identify agonists or antagonists of the native PRO232 polypeptide, wherein such assays may take the form of any conventional cell-type or biochemical binding assay. Moreover, the PRO232 polypeptide may serve as a molecular marker for the tissues in which the polypeptide is specifically expressed.

With regard to the PRO187 polypeptides disclosed herein, FGF-8 has been implicated in cellular differentiation and embryogenesis, including the patterning which appears during limb formation. FGF-8 and the PRO187 molecules of the invention therefore are likely to have potent effects on cell growth and development. Diseases which relate to cellular growth and differentiation are therefore suitable targets for therapeutics based on functionality similar to FGF-8. For example, diseases related to growth or survival of nerve cells including Parkinson's disease, Alzheimer's disease, ALS, neuropathies. Additionally, disease related

to uncontrolled cell growth, e.g., cancer, would also be expected therapeutic targets.

With regard to the PRO265 polypeptides disclosed herein, other methods for use with PRO265 are described in U.S. Patent 5,654,270 to Ruoslahti et al. In particular, PRO265 can be used in comparison with the fibromodulin disclosed thereio to compare its effects on reducing dermal scarring and other properties of the fibromodulin described therein including where it is located and with what it binds and does not.

The PRO219 polypeptides of the present invention which play a regulatory role in the blood coagulation cascade may be employed *in vivo* for therapeutic purposes as well as for *in vitro* purposes. Those of ordinary skill in the art will well know how to employ PRO219 polypeptides for such uses.

The PRO246 polypeptides of the present invention which serve as cell surface receptors for one or more viruses will find other uses. For example, extracellular domains derived from these PRO246 polypeptides may be employed therapeutically *in vivo* for lessening the effects of viral infection. Those PRO246 polypeptides which serves as tumor specific antigens may be exploited as therapeutic targets for anti-tumor drugs, and the like. Those of ordinary skill in the art will well know how to employ PRO246 polypeptides for such uses.

Assays in which connective growth factor and other growth factors are usually used should be performed with PRO261. An assay to determine whether TGF beta induces PRO261, indicating a role in cancer is performed as known in the art. Wound repair and tissue growth assays are also performed with PRO261. The results are applied accordingly.

PRO228 polypeptides should be used in assays in which EMR1, CD97 and latrophilin would be used in to determine their relative activities. The results can be applied accordingly. For example, a competitive binding assay with PRO228 and CD97 can be performed with the ligand for CD97, CD55.

Native PRO533 is a 216 amino acid polypeptide of which residues 1-22 are the signal sequence. Residues 3 to 216 have a Blast score of 509, corresponding to 53% homology to fibroblast growth factor. At the nucleotide level, DNA47412, the EST from which PCR oligos were generated to isolate the full length DNA49435-1219, has been observed to map to 11p15. Sequence homology to the 11p15 locus would indicate that PRO533 may have utility in the treatment of Usher Syndrome or Atrophia areata.

As mentioned previously, fibroblast growth factors can act upon cells in both a mitogenic and non-mitogenic manner. These factors are mitogenic for a wide variety of normal diploid mesoderm-derived and neural crest-derived cells, inducing granulosa cells, adrenal cortical cells, chrondrocytes, myoblasts, corneal and vascular endothelial cells (bovine or human), vascular smooth muscle cells, lens, retina and prostatic epithelial cells, oligodendrocytes, astrocytes, chrondocytes, myoblasts and osteoblasts.

Non-mitogenic actions of fibroblast growth factors include promotion of cell migration into a wound area (chemotaxis), initiation of new blood vessel formulation (angiogenesis), modulation of nerve regeneration and survival (neurotrophism), modulation of endocrine functions, and stimulation or suppression of specific cellular protein expression, extracellular matrix production and cell survival. Baird, A. & Bohlen, P., *Handbook of Exp. Pharmacol.* 95(1): 369-418 (1990). These properties provide a basis for using fibroblast growth factors in therapeutic approaches to accelerate wound healing, nerve repair, collateral blood vessel formation, and the like. For example, fibroblast growth factors, have been suggested to minimize myocardium damage in heart disease and surgery (U.S.P. 4,378,437).

Since the PRO245 polypeptide and nucleic acid encoding it possess sequence homology to a transmembrane protein tyrosine kinase protein and its encoding nucleic acid, probes based upon the PRO245 nucleotide sequence may be employed to identify other novel transmembrane tyrosine kinase proteins. Soluble forms of the PRO245 polypeptide may be employed as antagonists of membrane bound PRO245 activity both *in vitro* and *in vivo*. PRO245 polypeptides may be employed in screening assays designed to identify agonists or antagonists of the native PRO245 polypeptide, wherein such assays may take the form of any conventional cell-type or biochemical binding assay. Moreover, the PRO245 polypeptide may serve as a molecular marker for the tissues in which the polypeptide is specifically expressed.

PRO220, PRO221 and PRO227 all have leucine rich repeats. Additionally, PRO220 and PRO221 have homology to SLIT and leucine rich repeat protein. Therefore, these proteins are useful in assays described in the literature, supra, wherein the SLIT and leucine rich repeat protein are used. Regarding the SLIT protein, PRO227 can be used in an assay to determine the affect of PRO227 on neurodegenerative disease. Additionally, PRO227 has homology to human glycoprotein V. In the case of PRO227, this polypeptide is used in an assay to determine its affect on bleeding, clotting, tissue repair and scarring.

The PRO266 polypeptide can be used in assays to determine if it has a role in neurodegenerative diseases or their reversal.

PRO269 polypeptides and portions thereof which effect the activity of thrombin may also be useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO269 polypeptides and portions thereof may have therapeutic use as an antithrombotic agent with reduced risk for hemorrhage as compared with heparin. Peptides having homology to thrombomodulin are particularly desirable.

PRO287 polypeptides and portions thereof which effect the activity of bone morphogenic protein "BMP1"/procollagen C-proteinase (PCP) may also be useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO287 polypeptides and portions thereof may have therapeutic applications in wound healing and tissue repair. Peptides having homology to procollagen C-proteinase enhancer protein and its precursor may also be used to induce bone and/or cartilage formation and are therefore of particular interest to the scientific and medical communities.

Therapeutic indications for PRO214 polypeptides include disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g., enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g., psoriasis, epithelial cancers such as lung squamous cell carcinoma, epidermoid carcinoma of the vulva and gliomas.

Studies on the generation and analysis of mice deficient in members of the TGF- superfamily are reported in Matzuk, Trends in Endocrinol. and Metabol., 6: 120-127 (1995).

The PRO317 polypeptide, as well as PRO317-specificantibodies, inhibitors, agonists, receptors, or their analogs, herein are useful in treating PRO317-associated disorders. Hence, for example, they may be employed in modulating endometrial bleeding angiogenesis, and may also have an effect on kidney tissue. Endometrial bleeding can occur in gynecological diseases such as endometrial cancer as abnormal bleeding. Thus, the compositions herein may find use in diagnosing and treating abnormal bleeding conditions in the endometrium,

as by reducing or eliminating the need for a hysterectomy. The molecules herein may also find use in angiogenesis applications such as anti-tumor indications for which the antibody against vascular endothelial growth factor is used, or, conversely, ischemic indications for which vascular endothelial growth factor is employed.

Bioactive compositions comprising PRO317 or agonists or antagonists thereof may be administered in a suitable therapeutic dose determined by any of several methodologies including clinical studies on mammalian species to determine maximal tolerable dose and on normal human subjects to determine safe dose. Additionally, the bioactive agent may be complexed with a variety of well established compounds or compositions which enhance stability or pharmacological properties such as half-life. It is contemplated that the therapeutic, bioactive composition may be delivered by intravenous infusion into the bloodstream or any other effective means which could be used for treating problems of the kidney, uterus, endometrium, blood vessels, or related tissue, *e.g.*, in the heart or genital tract.

Dosages and administration of PRO317, PRO317 agonist, or PRO317 antagonist in a pharmaceutical composition may be determined by one of ordinary skill in the art of clinical pharmacology or pharmacokinetics. See, for example, Mordenti and Rescigno, Pharmaceutical Research, 9:17-25 (1992); Morenti *et al.*, Pharmaceutical Research, 8:1351-1359 (1991); and Mordenti and Chappell, "The use of interspecies scaling in toxicokinetics" in Toxicokinetics and New Drug Development, Yacobi *et al.* (eds) (Pergamon Press: NY, 1989), pp. 42-96. An effective amount of PRO317, PRO317 agonist, or PRO317 antagonist to be employed therapeutically will depend, for example, upon the therapeutic objectives, the route of administration, and the condition of the mammal. Accordingly, it will be necessary for the therapist to titer the dosage and modify the route of administration as required to obtain the optimal therapeutic effect. A typical daily dosage might range from about 10 ng/kg to up to 100 mg/kg of the mammal's body weight or more per day, preferably about 1 μg/kg/day to 10 mg/kg/day. Typically, the clinician will administer PRO317, PRO317 agonist, or PRO317 antagonist, until a dosage is reached that achieves the desired effect for treatment of the above mentioned disorders.

PRO317 or an PRO317 agonist or PRO317 antagonist may be administered alone or in combination with another to achieve the desired pharmacological effect. PRO317 itself, or agonists or antagonists of PRO317 can provide different effects when administered therapeutically. Such compounds for treatment will be formulated in a nontoxic, inert, pharmaceutically acceptable aqueous carrier medium preferably at a pH of about 5 to 8, more preferably 6 to 8, although the pH may vary according to the characteristics of the PRO317, agonist, or antagonist being formulated and the condition to be treated. Characteristics of the treatment compounds include solubility of the molecule, half-life, and antigenicity/immunogenicity; these and other characteristics may aid in defining an effective carrier.

PRO317 or PRO317 agonists or PRO317 antagonists may be delivered by known routes of administration including but not limited to topical creams and gels; transmucosal spray and aerosol, transdermal patch and bandage; injectable, intravenous, and lavage formulations; and orally administered liquids and pills, particularly formulated to resist stomach acid and enzymes. The particular formulation, exact dosage, and route of administration will be determined by the attending physician and will vary according to each specific situation.

Such determinations of administration are made by considering multiple variables such as the condition to be treated, the type of mammal to be treated, the compound to be administered, and the pharmacokinetic profile of the particular treatment compound. Additional factors which may be taken into account include disease state (*e.g.* severity) of the patient, age, weight, gender, diet, time of administration, drug combination, reaction sensitivities, and tolerance/response to therapy. Long-acting treatment compound formulations (such as liposomally encapsulated PRO317 or PEGylated PRO317 or PRO317 polymeric microspheres, such as polylactic acid-based microspheres) might be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular treatment compound.

Normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 μg/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting the uterus, for example, may necessitate delivery in a manner different from that to another organ or tissue, such as cardiac tissue.

Where sustained-release administration of PRO317 is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of PRO317, microencapsulation of PRO317 is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon- (rhIFN-), interleukin-2, and MN rgp120. Johnson *et al.*, Nat. Med., 2: 795-799 (1996); Yasuda, Biomed. Ther., 27: 1221-1223 (1993); Hora *et al.*, Bio/Technology, 8: 755-758 (1990); Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit and Adjuvant Approach , Powell and Newman, eds. (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S Pat. No. 5,654,010.

It is contemplated that conditions or diseases of the uterus, endometrial tissue, or other genital tissues or cardiac tissues may precipitate damage that is treatable with PRO317 or PRO317 agonist where PRO317 expression is reduced in the diseased state; or with antibodies to PRO317 or other PRO317 antagonists where the expression of PRO317 is increased in the diseased state. These conditions or diseases may be specifically diagnosed by the probing tests discussed above for physiologic and pathologic problems which affect the function of the organ.

The PRO317, PRO317 agonist, or PRO317 antagonist may be administered to a mammal with another biologically active agent, either separately or in the same formulation to treat a common indication for which they are appropriate. For example, it is contemplated that PRO317 can be administered together with EBAF-1 for those indications on which they demonstrate the same qualitative biological effects. Alternatively, where they have opposite effects, EBAF-1 may be administered together with an antagonist to PRO317, such as an anti-PRO317 antibody. Further, PRO317 may be administered together with VEGF for coronary ischemia where such indication is warranted, or with an anti-VEGF for cancer as warranted, or, conversely, an antagonist to PRO317 may be administered with VEGF for coronary ischemia or with anti-VEGF to treat cancer as warranted. These administrations would be in effective amounts for treating such disorders.

Native PRO301 (SEQ ID NO:119) has a Blast score of 246 and 30% homology at residues 24 to 282 of Figure 44 with A33_HUMAN, an A33 antigen precursor. A33 antigen precursor, as explained in the Background is a tumor-specific antigen, and as such, is a recognized marker and therapeutic target for the diagnosis and treatment of colon cancer. The expression of tumor-specific antigens is often associated with the progression of neoplastic tissue disorders. Native PRO301 (SEQ ID NO:119) and A33_HUMAN also show a Blast score of 245 and 30% homology at residues 21 to 282 of Fig. 44 with A33_HUMAN, the variation dependent upon how spaces are inserted into the compared sequences. Native PRO301 (SEQ ID NO:119) also has a Blast score of 165 and 29% homology at residues 60 to 255 of Fig. 44 with HS46KDA_1, a human coxsackie and adenovirus receptor protein, also known as cell surface protein HCAR. This region of PRO301 also shows a similar Blast score and homology with HSU90716_1. Expression of such proteins is usually associated with viral infection and therapeutics for the prevention of such infection may be accordingly conceived. As mentioned in the Background, the expression of viral receptors is often associated with neoplastic tumors.

Therapeutic uses for the PRO234 polypeptides of the invention includes treatments associated with leukocyte homing or the interaction between leukocytes and the endothelium during an inflammatory response. Examples include asthma, rheumatoid arthritis, psoriasis and multiple sclerosis.

Since the PRO231 polypeptide and nucleic acid encoding it possess sequence homology to a putative acid phosphatase and its encoding nucleic acid, probes based upon the PRO231 nucleotide sequence may be employed to identify other novel phosphatase proteins. Soluble forms of the PRO231 polypeptide may be employed as antagonists of membrane bound PRO231 activity both *in vitro* and *in vivo*. PRO231 polypeptides may be employed in screening assays designed to identify agonists or antagonists of the native PRO231 polypeptide, wherein such assays may take the form of any conventional cell-type or biochemical binding assay. Moreover, the PRO231 polypeptide may serve as a molecular marker for the tissues in which the polypeptide is specifically expressed.

PRO229 polypeptides can be fused with peptides of interest to determine whether the fusion peptide has an increased half-life over the peptide of interest. The PRO229 polypeptides can be used accordingly to increase the half-life of polypeptides of interest. Portions of PRO229 which cause the increase in half-life are an embodiment of the invention herein.

PRO238 can be used in assays which measure its ability to reduce substrates, including oxygen and Acetyl-CoA, and particularly, measure PRO238's ability to produce oxygen free radicals. This is done by using assays which have been previously described. PRO238 can further be used to assay for candidates which block, reduce or reverse its reducing abilities. This is done by performing side by side assays where candidates are added in one assay having PRO238 and a substrate to reduce, and not added in another assay, being the same but for the lack of the presence of the candidate.

PRO233 polypeptides and portions thereof which have homology to reductase may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel reductase proteins and related molecules may be relevant to a number of human disorders such as inflammatory disease, organ failure, atherosclerosis, cardiac injury, infertility, birth defects, premature aging, AIDS, cancer, diabetic

complications and mutations in general. Given that oxygen free radicals and antioxidants appear to play important roles in a number of disease processes, the identification of new reductase proteins and reductase-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research, as well as various industrial applications. As a result, there is particular scientific and medical interest in new molecules, such as PRO233.

The PRO223 polypeptides of the present invention which exhibit serine carboxypeptidase activity may be employed *in vivo* for therapeutic purposes as well as for *in vitro* purposes. Those of ordinary skill in the art will well know how to employ PRO223 polypeptides for such uses.

PRO235 polypeptides and portions thereof which may be involved in cell adhesion are also useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO235 polypeptides and portions thereof may have therapeutic applications in disease states which involve cell adhesion. Given the physiological importance of cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in cell adhesion. Therefore, peptides having homology to plexin are of particular interest to the scientific and medical communities.

Because the PRO236 and PRO262 polypeptides disclosed herein are homologous to various known β-galactosidase proteins, the PRO236 and PRO262 polypeptides disclosed herein will find use in conjugates of monoclonal antibodies and the polypeptide for specific killing of tumor cells by generation of active drug from a galactosylated prodrug (e.g., the generation of 5-fluorouridine from the prodrug β-D-galactosyl-5-fluorouridine). The PRO236 and PRO262 polypeptides disclosed herein may also find various uses both *in vivo* and *in vitro*, wherein those uses will be similar or identical to uses for which β-galactosidase proteins are now employed. Those of ordinary skill in the art will well know how to employ PRO236 and PRO262 polypeptides for such uses.

PRO239 polypeptides and portions thereof which have homology to densin may also be useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO239 polypeptides and portions thereof may have therapeutic applications in disease states which involve synaptic mechanisms, regeneration or cell adhesion. Given the physiological importance of synaptic processes, regeneration and cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in synaptic machinery and cell adhesion. Therefore, peptides having homology to densin are of particular interest to the scientific and medical communities.

The PRO260 polypeptides described herein can be used in assays to determine their relation to fucosidase. In particular, the PRO260 polypeptides can be used in assays in determining their ability to remove fucose or other sugar residues from proteoglycans. The PRO260 polypeptides can be assayed to determine if they have any functional or locational similarities as fucosidase. The PRO260 polypeptides can then be used to regulate the systems in which they are integral.

PRO263 can be used in assays wherein CD44 antigen is generally used to determine PRO263 activity relative to that of CD44. The results can be used accordingly.

PRO270 polypeptides and portions thereof which effect reduction-oxidation (redox) state may also be

useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. More specifically, PRO270 polypeptides may affect the expression of a large variety of genes thought to be involved in the pathogenesis of AIDS, cancer, atherosclerosis, diabetic complications and in pathological conditions involving oxidative stress such as stroke and inflammation. In addition, PRO270 polypeptides and portions thereof may affect the expression of a genes which have a role in apoptosis. Therefore, peptides having homology to thioredoxin are particularly desirable to the scientific and medical communities.

PRO272 polypeptides and portions thereof which possess the ability to bind calcium may also have numerous *in vivo* therapeutic uses, as well as various *in vitro* applications. Therefore, peptides having homology to reticulocalbin are particularly desirable. Those with ordinary skill in the art will know how to employ PRO272 polypeptides and portions thereof for such purposes.

PRO294 polypeptides and portions thereof which have homology to collagen may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel collagens and collage-like molecules may have relevance to a number of human disorders. Thus, the identification of new collagens and collage-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research as well as various industrial applications. Given the large number of uses for collagen, there is substantial interest in polypeptides with homology to the collagen molecule.

PRO295 polypeptides and portions thereof which have homology to integrin may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel integrins and integrin-like molecules may have relevance to a number of human disorders such as modulating the binding or activity of cells of the immune system. Thus, the identification of new integrins and integrin-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research as well as various industrial applications. As a result, there is particular scientific and medical interest in new molecules, such as PRO295.

As the PRO293 polypeptide is clearly a leucine rich repeat polypeptide homologue, the peptide can be used in all applications that the known NLRR-1 and NLRR-2 polypeptides are used. The activity can be compared between these peptides and thus applied accordingly.

The PRO247 polypeptides described herein can be used in assays in which densin is used to determine the activity of PRO247 relative to densin or these other proteins. The results can be used accordingly in diagnostics and/or therapeutic applications with PRO247.

PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides of the present invention which possess protease activity may be employed both *in vivo* for therapeutic purposes and *in vitro*. Those of ordinary skill in the art will well know how to employ the PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides of the present invention for such purposes.

PRO328 polypeptides and portions thereof which have homology to GLIP and CRISP may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel GLIP and CRISP-like molecules may have relevance to a number of human disorders which involve transcriptional

regulation or are over expressed in human tumors. Thus, the identification of new GLIP and CRISP-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research as well as in various industrial applications. As a result, there is particular scientific and medical interest in new molecules, such as PRO328.

Uses for PRO335, PRO331 or PRO326 including uses in competitive assays with LIG-1, ALS and decorin to determine their relative activities. The results can be used accordingly. PRO335, PRO331 or PRO326 can also be used in assays where LIG-1 would be used to determine if the same effects are incurred.

PRO332 contains GAG repeat (GKEK) at amino acid positions 625-628 in Fig. 108 (SEQ ID NO:310). Slippage in such repeats can be associated with human disease. Accordingly, PRO332 can use useful for the treatment of such disease conditions by gene therapy, i.e. by introduction of a gene containing the correct GKEK sequence motif.

Other uses of PRO334 include use in assays in which fibrillin or fibulin would be used to determine the relative activity of PRO334 to fibrillin or fibulin. In particular, PRO334 can be used in assays which require the mechanisms imparted by epidermal growth factor repeats.

Native PRO346 (SEQ ID NO:320) has a Blast score of 230, corresponding to 27% homology between amino acid residues 21 to 343 with residues 35 to 1040 CGM6_HUMAN, a carcinoembryonic antigen cgm6 precursor. This homology region includes nearly all but 2 N-terminal extracellular domain residues, including an immunoglobulin superfamily homology at residues 148 to 339 of PRO346 in addition to several transmembrane residues (340-343). Carcinoembryonic antigen precursor, as explained in the Background is a tumor-specific antigen, and as such, is a recognized marker and therapeutic target for the diagnosis and treatment of colon cancer. The expression of tumor-specific antigens is often associated with the progression of neoplastic tissue disorders. Native PRO346 (SEQ ID NO:320) and P_W06874, a human carcinoembryonic antigen CEA-d have a Blast score of 224 and homology of 28% between residues 2 to 343 and 67 to 342, respectively. This homology includes the entire extracellular domain residues of native PRO346, minus the initiator methionine (residues 2 to 18) as well as several transmembrane residues (340-343).

PRO268 polypeptides which have protein disulfide isomerase activity will be useful for many applications where protein disulfide isomerase activity is desirable including, for example, for use in promoting proper disulfide bond formation in recombinantly produced proteins so as to increase the yield of correctly folded protein. Those of ordinary skill in the art will readily know how to employ such PRO268 polypeptides for such purposes.

PRO330 polypeptides of the present invention which possess biological activity related to that of the prolyl 4-hydroxylase alpha subunit protein may be employed both *in vivo* for therapeutic purposes and *in vitro*. Those of ordinary skill in the art will well know how to employ the PRO330 polypeptides of the present invention for such purposes.

F. Anti-PRO Antibodies

The present invention further provides anti-PRO antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

The anti-PRO antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

2. Monoclonal Antibodies

The anti-PRO antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas,

Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against PRO. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., supra] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Human and Humanized Antibodies

The anti-PRO antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')$_2$ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks *et al.*,

Bio/Technology 10, 779-783 (1992); Lonberg *et al.*, Nature 368 856-859 (1994); Morrison, Nature 368, 812-13 (1994); Fishwild *et al.*, Nature Biotechnology 14, 845-51 (1996); Neuberger, Nature Biotechnology 14, 826 (1996); Lonberg and Huszar, Intern. Rev. Immunol. 13 65-93 (1995).

4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Millstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')$_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared can be prepared using chemical linkage. Brennan *et al.*, Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')$_2$ fragments. These fragments are reduced in the presence of the dithiol complexing

agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby *et al.*, J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various technique for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny *et al.*, J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger *et al.*, Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber *et al.*, J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt *et al.*, J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies may bind to two different epitopes on a given PRO polypeptide herein. Alternatively, an anti-PRO polypeptide arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular PRO polypeptide. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express a particular PRO polypeptide. These antibodies possess a PRO-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO polypeptide and further binds tissue factor (TF).

5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate

antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

6. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, J Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, Anti-Cancer Drug Design, 3: 219-230 (1989).

7. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-

isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (*e.g.*, avidin) that is conjugated to a cytotoxic agent (*e.g.*, a radionucleotide).

8. Immunoliposomes

The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013.556.

Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine(PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin *et al* ., J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon *et al.*, J. National Cancer Inst., 81(19): 1484 (1989).

9. Pharmaceutical Compositions of Antibodies

Antibodies specifically binding a PRO polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders in the form of pharmaceutical compositions.

If the PRO polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, *e.g.*, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993). The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation

techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, *supra*.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT ™ (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

G. Uses for anti-PRO Antibodies

The anti-PRO antibodies of the invention have various utilities. For example, anti-PRO antibodies may be used in diagnostic assays for PRO, *e.g.*, detecting its expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^{3}H, ^{14}C, ^{32}P, ^{35}S, or ^{125}I, a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-PRO antibodies also are useful for the affinity purification of PRO from recombinant cell culture or natural sources. In this process, the antibodies against PRO are immobilized on a suitable support, such a

Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO from the antibody.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Rockville, Maryland.

EXAMPLE 1: Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA Encoding Therefor

The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public databases (e.g., Dayhoff, GenBank), and proprietary databases (e.g. LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul, and Gish, Methods in Enzymology 266: 460-80 (1996); http://blast.wustl/edu/blast/README.html) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a Blast score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

Using this extracellular domain homology screen, consensus DNA sequences were assembled relative to the other identified EST sequences. In addition, the consensus DNA sequences obtained were often (but not always) extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward (.f) and reverse (.r) PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe (.p) sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification,

as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

EXAMPLE 2: Isolation of cDNA Clones Encoding PRO211 and PRO217

Consensus DNA sequences were assembled as described in Example 1 above and were designated as DNA28730 and DNA28760, respectively. Based on these consensus sequences, oligonucleotides were synthesized and used to identify by PCR a cDNA library that contained the sequences of interest and for use as probes to isolate a clone of the full-length coding sequence for the PRO211 and PRO217 polypeptides. The libraries used to isolate DNA32292-1131 and DNA33094-1131 were fetal lung libraries.

cDNA clones were sequenced in their entirety. The entire nucleotide sequences of PRO211 (DNA32292-1131) and PRO217 (UNQ191) are shown in Figure 1 (SEQ ID NO: 1) and Figure 3 (SEQ ID NO:3), respectively. The predicted polypeptides are 353 and 379 amino acid in length, respectively, with respective molecular weights of approximately 38,190 and 41,520 daltons.

The oligonucleotide sequences used in the above procedures were the following:

28730.p (OLI 516) (SEQ ID NO:5)
5'-AGGGAGCACGGACAGTGTGCAGATGTGGACGAGTGCTCACTAGCA-3'
28730.f (OLI 517) (SEQ ID NO:6)
5'-AGAGTGTATCTCTGGCTACGC-3'
28730.r (OLI 518) (SEQ ID NO:7)
5'-TAAGTCCGGCACATTACAGGTC-3'
28760.p (OLI 617) (SEQ ID NO:8)
5'-CCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATC-3'
28760.f (OLI 618) (SEQ ID NO:9)
5'-AAAGACGCATCTGCGAGTGTCC-3'
28760.r (OLI 619) (SEQ ID NO:10)
5'-TGCTGATTTCACACTGCTCTCCC-3'

EXAMPLE 3: Isolation of cDNA Clones Encoding Human PRO230

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA30857. An EST proprietary to Genentech was employed in the consensus assembly. The EST is designated as DNA20088 and has the

nucleotide sequence shown in Figure 7 (SEQ ID NO:13).

Based on the DNA30857 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO230.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TTCGAGGCCTCTGAGAAGTGGCCC-3' (SEQ ID NO:14)

reverse PCR primer 5'-GGCGGTATCTCTCTGGCCTCCC-3' (SEQ ID NO:15)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30857 sequence which had the following nucleotide sequence

hybridization probe

5'-TTCTCCACAGCAGCTGTGGCATCCGATCGTGTCTCAATCCATTCTCTGGG-3' (SEQ ID NO:16)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO230 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO230 (herein designated as DNA33223-1136 and the derived protein sequence for PRO230.

The entire nucleotide sequence of DNA33223-1136 is shown in Figure 5 (SEQ ID NO:11). Clone DNA33223-1136 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 100-103 and ending at the stop codon at nucleotide positions 1501-1503 (Figure 5; SEQ ID NO:11). The predicted polypeptide precursor is 467 amino acids long (Figure 6).

EXAMPLE 4: Isolation of cDNA Clones Encoding Human PRO232

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA30935. Based on the DNA30935 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO232.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGCTGTGCTACTCCTGCAAAGCCC-3' (SEQ ID NO:19)

reverse PCR primer 5'-TGCACAAGTCGGTGTCACAGCACG-3' (SEQ ID NO:20)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30935 sequence which had the following nucleotide sequence

hybridization probe

5'-AGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGG-3' (SEQ ID NO:21)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO232 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO232 [herein designated as DNA34435-1140] and the derived protein sequence for PRO232.

The entire nucleotide sequence of DNA34435-1140 is shown in Figure 8 (SEQ ID NO:17). Clone DNA34435-1140 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 17-19 and ending at the stop codon at nucleotide positions 359-361 (Fig. 8; SEQ ID NO:17). The predicted polypeptide precursor is 114 amino acids long (Fig. 9). Clone DNA34435-1140 has been deposited with ATCC on September 16, 1997 and is assigned ATCC deposit no. ATCC 209250.

Analysis of the amino acid sequence of the full-length PRO232 suggests that it possesses 35% sequence identity with a stem cell surface antigen from Gallus gallus.

EXAMPLE 5: Isolation of cDNA Clones Encoding PRO187

A proprietary expressed sequence tag (EST) DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST (#843193) was identified which showed homology to fibroblast growth factor (FGF-8) also known as androgen-induced growth factor. mRNA was isolated from human fetal lung tissue using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents (e.g., Invitrogen, San Diego, CA, Life Technologies, Gaithersburg, MD). The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into the cloning vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). The double-stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linkered cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

Several libraries from various tissue sources were screened by PCR amplification with the following oligonucleotide probes:

IN843193.f (OLI315) (SEQ ID NO:24)

5'-CAGTACGTGAGGGACCAGGGCGCCATGA-3'

IN843193.r (OLI 317) (SEQ ID NO:25)

5'-CCGGTGACCTGCACGTGCTTGCCA-3'

A positive library was then used to isolate clones encoding the PRO187 gene using one of the above oligonucleotides and the following oligonucleotide probe:

IN843193.p (OLI 316) (SEQ ID NO:26)

5'-GCGGATCTGCCGCCTGCTCANCTGGTCGGTCATGGCGCCCT-3'

A cDNA clone was sequenced in entirety. The entire nucleotide sequence of PRO187 (DNA27864-1155) is shown in Figure 10 (SEQ ID NO:22). Clone DNA27864-1155 contains a single open reading frame with an apparent translational initiation site at nucleotide position 1 (Figure 10; SEQ ID NO:22). The predicted polypeptide precursor is 205 amino acids long. Clone DNA27864-1155 has been deposited with the ATCC

(designation: DNA27864-1155) and is assigned ATCC deposit no. ATCC 209375.

Based on a BLAST and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, the PRO187 polypeptide shows 74% amino acid sequence identity (Blast score 310) to human fibroblast growth factor-8 (androgen-induced growth factor).

EXAMPLE 6: Isolation of cDNA Clones Encoding PRO265

A consensus DNA sequence was assembled relative to other EST sequences as described in Example 1 above using phrap. This consensus sequence is herein designated DNA33679. Based on the DNA33679 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO265.

PCR primers (two forward and one reverse) were synthesized:

forward PCR primer A: 5'-CGGTCTACCTGTATGGCAACC-3' (SEQ ID NO:29);

forward PCR primer B: 5'-GCAGGACAACCAGATAAACCAC-3' (SEQ ID NO:30);

reverse PCR primer 5'-ACGCAGATTTGAGAAGGCTGTC-3' (SEQ ID NO:31)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA33679 sequence which had the following nucleotide sequence

hybridization probe

5'-TTCACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCAC-3' (SEQ ID NO:32)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO265 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human a fetal brain library.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO265 [herein designated as DNA36350-1158] (SEQ ID NO:27) and the derived protein sequence for PRO265.

The entire nucleotide sequence of DNA36350-1158 is shown in Figure 12 (SEQ ID NO:27). Clone DNA36350-1158 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 352-354 and ending at the stop codon at positions 2332-2334 (Figure 12). The predicted polypeptide precursor is 660 amino acids long (Figure 13). Clone DNA36350-1158 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209378.

Analysis of the amino acid sequence of the full-length PRO265 polypeptide suggests that portions of it possess significant homology to the fibromodulin and the fibromodulin precursor, thereby indicating that PRO265 may be a novel member of the leucine rich repeat family, particularly related to fibromodulin.

EXAMPLE 7: Isolation of cDNA Clones Encoding Human PRO219

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28729. Based on the DNA28729

consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO219.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GTGACCCTGGTTGTGAATACTCC-3' (SEQ ID NO:35)

reverse PCR primer 5'-ACAGCCATGGTCTATAGCTTGG-3' (SEQ ID NO:36)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28729 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCTGTCAGTGTCCTGAGGGACACGTGCTCCGCAGCGATGGGAAG-3' (SEQ ID NO:37)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO219 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO219 [herein designated as DNA32290-1164] (SEQ ID NO:33) and the derived protein sequence for PRO219.

The entire nucleotide sequence of DNA32290-1164 is shown in Figure 14 (SEQ ID NO:33). Clone DNA32290-1164 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 204-206 and ending at the stop codon at nucleotide positions 2949-2951 (Figure 14). The predicted polypeptide precursor is 915 amino acids long (Figure 15). Clone DNA32290-1164 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209384.

Analysis of the amino acid sequence of the full-length PRO219 polypeptide suggests that portions of it possess significant homology to the mouse and human matrilin-2 precursor polypeptides.

EXAMPLE 8: Isolation of cDNA Clones Encoding Human PRO246

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30955. Based on the DNA30955 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO246.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGGGTCTCCAGGAGAAAGACTC-3' (SEQ ID NO:40)

reverse PCR primer 5'-ATTGTGGGCCTTGCAGACATAGAC-3' (SEQ ID NO:41)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30955 sequence which had the following nucleotide sequence

hybridization probe

5'-GGCCACAGCATCAAAACCTTAGAACTCAATGTACTGGTTCCTCCAGCTCC-3' (SEQ ID NO:42)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO246 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO246 [herein designated as DNA35639-1172] (SEQ ID NO:38) and the derived protein sequence for PRO246.

The entire nucleotide sequence of DNA35639-1172 is shown in Figure 16 (SEQ ID NO:38). Clone DNA35639-1172 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 126-128 and ending at the stop codon at nucleotide positions 1296-1298 (Figure 16). The predicted polypeptide precursor is 390 amino acids long (Figure 17). Clone DNA35639-1172 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209396.

Analysis of the amino acid sequence of the full-length PRO246 polypeptide suggests that it possess significant homology to the human cell surface protein HCAR, thereby indicating that PRO246 may be a novel cell surface virus receptor.

EXAMPLE 9: Isolation of cDNA Clones Encoding Human PRO228

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28758. An EST proprietary to Genentech was employed in the consensus assembly. This EST is shown in Figure 20 (SEQ ID NO:50) and is herein designated as DNA21951.

Based on the DNA28758 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO228.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GGTAATGAGCTCCATTACAG-3' (SEQ ID NO:51)
forward PCR primer 5'-GGAGTAGAAAGCGCATGG-3' (SEQ ID NO:52)
forward PCR primer 5'-CACCTGATACCATGAATGGCAG-3' (SEQ ID NO:53)
reverse PCR primer 5'-CGAGCTCGAATTAATTCG-3' (SEQ ID NO:54)
reverse PCR primer 5'-GGATCTCCTGAGCTCAGG-3' (SEQ ID NO:55)
reverse PCR primer 5'-CCTAGTTGAGTGATCCTTGTAAG-3' (SEQ ID NO:56)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28758 sequence which had the following nucleotide sequence

hybridization probe

5'-ATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACATTTTGCAATT-3' (SEQ ID NO:57)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO228 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO228 [herein designated as DNA33092-1202] (SEQ ID NO:48) and the derived protein sequence for PRO228.

The entire nucleotide sequence of DNA33092-1202 is shown in Figure 18 (SEQ ID NO:48). Clone DNA33092-1202 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 24-26 of SEQ ID NO:48 and ending at the stop codon after nucleotide position 2093 of SEQ ID NO:48. The predicted polypeptide precursor is 690 amino acids long (Figure 19). Clone DNA33092-1202 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209420.

Analysis of the amino acid sequence of the full-length PRO228 polypeptide suggests that portions of it possess significant homology to the secretin-related proteins CD97 and EMR1 as well as the secretin member, latrophilin, thereby indicating that PRO228 may be a new member of the secretin related proteins.

EXAMPLE 10: Isolation of cDNA Clones Encoding Human PRO533

The EST sequence accession number AF007268, a murine fibroblast growth factor (FGF-15) was used to search various public EST databases (e.g., GenBank, Dayhoff, etc.). The search was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996); http://blast.wustl/edu/blast/README.html] as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. The search resulted in a hit with GenBank EST AA220994, which has been identified as stratagene NT2 neuronal precursor 937230.

Based on the Genbank EST AA220994 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence. Forward and reverse PCR primers may range from 20 to 30 nucleotides (typically about 24), and are designed to give a PCR product of 100-1000 bp in length. The probe sequences are typically 40-55 bp (typically about 50) in length. In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification, as per *Ausubel et al., Current Protocols in Molecular Biology*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the PCR primers.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified below. A positive library was then used to isolate clones encoding the PRO533 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal retina. The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents (e.g., Invitrogen, San Diego, CA; Clontech, etc.) The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of PRO533 is shown in Figure 21 (SEQ ID NO:58). Clone DNA49435-1219 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 459-461 (Figure 21; SEQ ID NO:58). The predicted polypeptide precursor is 216 amino acids long. Clone DNA47412-1219 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209480.

Based on a BLAST-2 and FastA sequence alignment analysis of the full length sequence, PRO533 shows amino acid sequence identity to fibroblast growth factor (53%).

The oligonucleotide sequences used in the above procedure were the following:

FGF15.forward: 5'-ATCCGCCCAGATGGCTACAATGTGTA-3' (SEQ ID NO:60);

FGF15.probe: 5'-GCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGTGTA-3' (SEQ ID NO:61);

FGF15.reverse: 5'-CCAGTCCGGTGACAAGCCCAAA-3' (SEQ ID NO:62).

EXAMPLE 11: Isolation of cDNA Clones Encoding Human PRO245

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA30954.

Based on the DNA30954 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO245.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5' ATCGTTGTGAAGTTAGTGCCCC-3' (SEQ ID NO:65)

reverse PCR primer 5'-ACCTGCGATATCCAACAGAATTG-3' (SEQ ID NO:66)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30954 sequence which had the following nucleotide sequence

hybridization probe

5'-GGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTTCC-3' (SEQ ID NO:67)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO245 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO245 [herein designated as DNA35638-1141] and the derived protein sequence for PRO245.

The entire nucleotide sequence of DNA35638-1141 is shown in Figure 23 (SEQ ID NO:63). Clone DNA35638-1141 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 89-91 and ending at the stop codon at nucleotide positions 1025-1027 (Fig. 23; SEQ ID NO:63). The predicted polypeptide precursor is 312 amino acids long (Fig. 24). Clone DNA35638-1141 has been deposited with ATCC on September 16, 1997 and is assigned ATCC deposit no. ATCC 209265.

Analysis of the amino acid sequence of the full-length PRO245 suggests that a portion of it possesses 60% amino acid identity with the human c-myb protein and, therefore, may be a new member of the

transmembrane protein receptor tyrosine kinase family.

EXAMPLE 12: Isolation of cDNA Clones Encoding Human PRO220, PRO221 and PRO227

(a) PRO220

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28749. Based on the DNA28749 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO220.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TCACCTGGAGCCTTTATTGGCC-3' (SEQ ID NO:74)

reverse PCR primer 5'-ATACCAGCTATAACCAGGCTGCG-3' (SEQ ID NO:75)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28749 sequence which had the following nucleotide sequence:

hybridization probe

5'-CAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGATTCTGATGATTGGG-3'(SEQ ID NO:76).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO220 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO220 (herein designated as DNA32298-1132 and the derived protein sequence for PRO220.

The entire nucleotide sequence of DNA32298-1132 is shown in Figure 25 (SEQ ID NO:68). Clone DNA32298-1132 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 480-482 and ending at the stop codon at nucleotide positions 2604-2606 (Figure 25). The predicted polypeptide precursor is 708 amino acids long (Figure 26). Clone DNA32298-1132 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209257.

Analysis of the amino acid sequence of the full-length PRO220 shows it has homology to member of the leucine rich repeat protein superfamily, including the leucine rich repeat protein and the neuronal leucine-rich repeat protein 1.

(b) PRO221

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28756. Based on the DNA28756 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO221.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCATGTGTCTCCTCCTACAAAG-3' (SEQ ID NO:77)

reverse PCR primer 5'-GGGAATAGATGTGATCTGATTGG-3' (SEQ ID NO:78)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28756 sequence which had the following nucleotide sequence:

hybridization probe

5'-CACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTG-3' (SEQ ID NO:79)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO221 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO221 [herein designated as DNA33089-1132 and the derived protein sequence for PRO221.

The entire nucleotide sequence of DNA33089-1132 is shown in Figure 27 (SEQ ID NO:70). Clone DNA33089-1132 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 179-181 and ending at the stop codon at nucleotide positions 956-958 (Figure 27). The predicted polypeptide precursor is 259 amino acids long (Figure 28). PRO221 is believed to have a transmembrane region at amino acids 206-225. Clone DNA33089-1132 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209262.

Analysis of the amino acid sequence of the full-length PRO221 shows it has homology to member of the leucine rich repeat protein superfamily, including the SLIT protein.

(c) PRO227

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28740. Based on the DNA28740 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO227.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGCAACCGCCTGAAGCTCATCC-3' (SEQ ID NO:80)

reverse PCR primer 5'-AAGGCGCGGTGAAAGATGTAGACG-3' (SEQ ID NO:81)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28740 sequence which had the following nucleotide sequence:

hybridization probe

5'GACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCGA-3' (SEQ ID NO:82).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO227 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNA

sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO227 [herein designated as DNA33786-1132 and the derived protein sequence for PRO227.

The entire nucleotide sequence of DNA33786-1132 is shown in Figure 29 (SEQ ID NO:72). Clone DNA33786-1132 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 33-35 and ending at the stop codon at nucleotide positions 1893-1895 (Figure 29). The predicted polypeptide precursor is 620 amino acids long (Figure 30). PRO227 is believed to have a transmembrane region. Clone DNA33786-1132 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209253.

Analysis of the amino acid sequence of the full-length PRO221 shows it has homology to member of the leucine rich repeat protein superfamily, including the platelet glycoprotein V precursor and the human glycoprotein V.

EXAMPLE 13: Isolation of cDNA Clones Encoding Human PRO258

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28746.

Based on the DNA28746 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO258.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GCTAGGAATTCCACAGAAGCCC-3' (SEQ ID NO:85)

reverse PCR primer 5'-AACCTGGAATGTCACCGAGCTG-3' (SEQ ID NO:86)

reverse PCR primer 5'-CCTAGCACAGTGACGAGGGACTTGGC-3' (SEQ ID NO:87)

Additionally, synthetic oligonucleotide hybridization probes were constructed from the consensus DNA28740 sequence which had the following nucleotide sequence:

hybridization probe

5'-AAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAGCCTGCAGCC-3' (SEQ ID NO:88)

5'-GCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT-3' (SEQ ID NO:89)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO258 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO258 [herein designated as DNA35918-1174] (SEQ ID NO:83) and the derived protein sequence for PRO258.

The entire nucleotide sequence of DNA35918-1174 is shown in Figure 31 (SEQ ID NO:83). Clone DNA35918-1174 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 147-149 of SEQ ID NO:83 and ending at the stop codon after nucleotide position 1340 of SEQ ID NO:83 (Figure 31). The predicted polypeptide precursor is 398 amino acids long (Figure 32). Clone DNA35918-1174 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209402.

Analysis of the amino acid sequence of the full-length PRO258 polypeptide suggests that portions of it possess significant homology to the CRTAM and the poliovirus receptor and have an Ig domain, thereby indicating that PRO258 is a new member of the Ig superfamily.

EXAMPLE 14: Isolation of cDNA Clones Encoding Human PRO266

An expressed sequence tag database was searched for ESTs having homology to SLIT, resulting in the identification of a single EST sequence designated herein as T73996. Based on the T73996 EST sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO266.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GTTGGATCTGGGCAACAATAAC-3' (SEQ ID NO:92)

reverse PCR primer 5'-ATTGTTGTGCAGGCTGAGTTTAAG-3' (SEQ ID NO:93)

Additionally, a synthetic oligonucleotide hybridization probe was constructed which had the following nucleotide sequence

hybridization probe

5'-GGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGG-3' (SEQ ID NO:94)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO266 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO266 [herein designated as DNA37150-1178] (SEQ ID NO:90) and the derived protein sequence for PRO266.

The entire nucleotide sequence of DNA37150-1178 is shown in Figure 33 (SEQ ID NO:90). Clone DNA37150-1178 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 167-169 and ending at the stop codon after nucleotide position 2254 of SEQ ID NO:90. The predicted polypeptide precursor is 696 amino acids long (Figure 34). Clone DNA37150-1178 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209401.

Analysis of the amino acid sequence of the full-length PRO266 polypeptide suggests that portions of it possess significant homology to the SLIT protein, thereby indicating that PRO266 may be a novel leucine rich repeat protein.

EXAMPLE 15: Isolation of cDNA Clones Encoding Human PRO269

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35705. Based on the DNA35705 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO269.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-TGGAAGGAGATGCGATGCCACCTG -3'
(SEQ ID NO:97)
forward PCR primer (.f2) 5'-TGACCAGTGGGGAAGGACAG-3' (SEQ ID NO:98)
forward PCR primer (.f3) 5'-ACAGAGCAGAGGGTGCCTTG-3' (SEQ ID NO:99)
reverse PCR primer (.r1) 5'-TCAGGGACAAGTGGTGTCTCTCCC-3'
(SEQ ID NO:100)
reverse PCR primer (.r2) 5'-TCAGGGAAGGAGTGTGCAGTTCTG-3'
(SEQ ID NO:101)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35705 sequence which had the following nucleotide sequence:

hybridization probe
5'-ACAGCTCCCGATCTCAGTTACTTGCATCGCGGACGAAATCGGCGCTCGCT-3' (SEQ ID NO:102)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO269 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO269 [herein designated as DNA38260-1180] (SEQ ID NO:95) and the derived protein sequence for PRO269.

The entire nucleotide sequence of DNA38260-1180 is shown in Figure 35 (SEQ ID NO:95). Clone DNA38260-1180 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 314-316 and ending at the stop codon at nucleotide positions 1784-1786 (Fig. 35; SEQ ID NO:95). The predicted polypeptide precursor is 490 amino acids long (Fig. 36). Clone DNA38260-1180 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209397.

Analysis of the amino acid sequence of the full-length PRO269 suggests that portions of it possess significant homology to the human thrombomodulin proteins, thereby indicating that PRO269 may possess one or more thrombomodulin-like domains.

EXAMPLE 16: Isolation of cDNA Clones Encoding Human PRO287

A consensus DNA sequence encoding PRO287 was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28728. Based on the DNA28728 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO287.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCGATTCATAGACCTCGAGAGT-3' (SEQ ID NO:105)
reverse PCR primer 5'-GTCAAGGAGTCCTCCACAATAC-3' (SEQ ID NO:106)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28728

sequence which had the following nucleotide sequence

hybridization probe

5'-GTGTACAATGGCCATGCCAATGGCCAGCGCATTGGCCGCTTCTGT-3'

(SEQ ID NO:107)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO287 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO287 [herein designated as DNA39969-1185, SEQ ID NO:103] and the derived protein sequence for PRO287.

The entire nucleotide sequence of DNA39969-1185 is shown in Figure 37 (SEQ ID NO:103). Clone DNA39969-1185 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 307-309 and ending at the stop codon at nucleotide positions 1552-1554 (Fig. 37; SEQ ID NO:103). The predicted polypeptide precursor is 415 amino acids long (Fig. 38). Clone DNA39969-1185 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209400.

Analysis of the amino acid sequence of the full-length PRO287 suggests that it may possess one or more procollagen C-proteinase enhancer protein precursor or procollagen C-proteinase enhancer protein-like domains. Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO287 shows nucleic acid sequence identity to procollagen C-proteinase enhancer protein precursor and procollagen C-proteinase enhancer protein (47 and 54%, respectively).

EXAMPLE 17: Isolation of cDNA Clones Encoding Human PRO214

A consensus DNA sequence was assembled using phrap as described in Example 1 above. This consensus DNA sequence is designated herein as DNA28744. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified below. A positive library was then used to isolate clones encoding the PRO214 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of DNA32286-1191 is shown in Figure 39 (SEQ ID NO:108). DNA32286-1191 contains a single open reading frame with an apparent translational initiation site at nucleotide position 103 (Fig. 39; SEQ ID NO:108). The predicted polypeptide precursor is 420 amino acids long (SEQ ID NO:109).

Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO214 polypeptide shows amino acid sequence identity to HT protein and/or Fibulin (49% and 38%, respectively).

The oligonucleotide sequences used in the above procedure were the following:

28744.p (OLI555)
5'-CCTGGCTATCAGCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGA-3' (SEQ ID NO:110)
28744.f (OLI556)
5'-ATTCTGCGTGAACACTGAGGGC-3' (SEQ ID NO:111)
28744.r (OLI557)
5'-ATCTGCTTGTAGCCCTCGGCAC-3' (SEQ ID NO:112)

EXAMPLE 18: Isolation of cDNA Clones Encoding Human PRO317

A consensus DNA sequence was assembled using phrap as described in Example 1 above, wherein the consensus sequence is herein designated as DNA28722. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence. The forward and reverse PCR primers, respectively, synthesized for this purpose were:

5'-AGGACTGCCATAACTTGCCTG (OLI489) (SEQ ID NO:115) and

5'-ATAGGAGTTGAAGCAGCGCTGC (OLI490) (SEQ ID NO:116).

The probe synthesized for this purpose was:

5'-TGTGTGGACATAGACGAGTGCCGCTACCGCTACTGCCAGCACCGC (OLI488) (SEQ ID NO:117)

mRNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel *et al.*, *Current Protocols in Molecular Biology* (1989), with the PCR primer pair identified above. A positive library was then used to isolate clones containing the PRO317 gene using the probe oligonucleotide identified above and one of the PCR primers.

A cDNA clone was sequenced in its entirety. The entire nucleotide sequence of DNA33461-1199 (encoding PRO317) is shown in Figure 41 (SEQ ID NO:113). Clone DNA33461-1199 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 68-70 (Fig. 41; SEQ ID NO:113). The predicted polypeptide precursor is 366 amino acids long. The predicted signal sequence is amino acids 1-18 of Figure 42 (SEQ ID NO:114). There is one predicted N-linked glycosylation site at amino acid residue 160. Clone DNA33461-1199 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209367.

Based on BLAST™ and FastA™ sequence alignment analysis (using the ALIGN™ computer program) of the full-length PRO317sequence, PRO317 shows the most amino acid sequence identity to EBAF-1 (92%). The results also demonstrate a significant homology between human PRO317 and mouse LEFTY protein. The C-terminal end of the PRO317 protein contains many conserved sequences consistent with the pattern expected of a member of the TGF- superfamily.

In situ expression analysis in human tissues performed as described below evidences that there is distinctly strong expression of the PRO317 polypeptide in pancreatic tissue.

EXAMPLE 19: Isolation of cDNA clones Encoding Human PRO301

A consensus DNA sequence designated herein as DNA35936 was assembled using phrap as described in Example 1 above. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified below. A positive library was then used to isolate clones encoding the PRO301 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney.

A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of native sequence PRO301 is shown in Figure 43 (SEQ ID NO:118). Clone DNA40628-1216 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 52-54 (Fig. 43; SEQ ID NO:118). The predicted polypeptide precursor is 299 amino acids long with a predicted molecular weight of 32,583 daltons and pI of 8.29. Clone DNA40628-1216 has been deposited with ATCC and is assigned ATCC deposit No. ATCC 209432.

Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO301 shows amino acid sequence identity to A33 antigen precursor (30%) and coxsackie and adenovirus receptor protein (29%).

The oligonucleotide sequences used in the above procedure were the following:

OLI2162 (35936.f1) 5'-TCGCGGAGCTGTGTTCTGTTTCCC-3' (SEQ ID NO:120)
OLI2163 (35936.p1)
5'-TGATCGCGATGGGGACAAAGGCGCAAGCTCGAGAGGAAACTGTTGTGCCT-3' (SEQ ID NO:121)
OLI2164 (35936.f2)
5'-ACACCTGGTTCAAAGATGGG-3' (SEQ ID NO:122)
OLI2165 (35936.r1)
5'-TAGGAAGAGTTGCTGAAGGCACGG-3' (SEQ ID NO:123)
OLI2166 (35936.f3)
5'-TTGCCTTACTCAGGTGCTAC-3' (SEQ ID NO:124)
OLI2167 (35936.r2)
5'-ACTCAGCAGTGGTAGGAAAG-3' (SEQ ID NO:125)

EXAMPLE 20: Isolation of cDNA Clones Encoding Human PRO224

A consensus DNA sequence assembled relative to the other identified EST sequences as described in Example 1, wherein the consensus sequence is designated herein as DNA30845. Based on the DNA30845 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO224.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AAGTTCCAGTGCCGCACCAGTGGC-3' (SEQ ID NO:128)
reverse PCR primer 5'-TTGGTTCCACAGCCGAGCTCGTCG-3' (SEQ ID NO:129)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30845 sequence which had the following nucleotide sequence

hybridization probe

5'-GAGGAGGAGTGCAGGATTGAGCCATGTACCCAGAAAGGGCAATGCCCACC-3' (SEQ ID NO:130)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO224 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO224 [herein designated as DNA33221-1133] and the derived protein sequence for PRO224.

The entire nucleotide sequence of DNA33221-1133 is shown in Figure 45 (SEQ ID NO:126). Clone DNA33221-1133 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 33-35 and ending at the stop codon at nucleotide positions 879-899 (Figure 45; SEQ ID NO:126). The start of a transmembrane region begins at nucleotide position 777. The predicted polypeptide precursor is 282 amino acids long (Figure 46). Clone DNA33221-1133 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209263.

Analysis of the amino acid sequence of the full-length PRO224 suggests that it has homology to very low-density lipoprotein receptors, apolipoprotein E receptor and chicken oocyte receptors P95. Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO224 has amino acid identity to portions of these proteins in the range from 28% to 45%, and overall identity with these proteins in the range from 33% to 39%.

EXAMPLE 21: Isolation of cDNA Clones Encoding Human PRO222

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28771. Based on the DNA28771 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO222.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-ATCTCCTATCGCTGCTTTCCCGG-3' (SEQ ID NO:133)

reverse PCR primer 5'-AGCCAGGATCGCAGTAAAACTCC-3' (SEQ ID NO:134)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28771 sequence which had the following nucleotide sequence:

hybridization probe

5'-ATTTAAACTTGATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCT-3' (SEQ ID NO:135)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO222 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO222 [herein designated as DNA33107-1135] and the derived protein sequence for PRO222.

The entire nucleotide sequence of DNA33107-1135 is shown in Figure 47 (SEQ ID NO:131). Clone DNA33107-1135 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 159-161 and ending at the stop codon at nucleotide positions 1629-1631 (Fig. 47; SEQ ID NO:131). The predicted polypeptide precursor is 490 amino acids long (Fig. 48). Clone DNA33107-1135 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209251.

Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO222 shows amino acid sequence identity to mouse complement factor h precursor (25-26%), complement receptor (27-29%), mouse complement C3b receptor type 2 long form precursor (25-47%) and human hypothetical protein ksia0247 (40%).

EXAMPLE 22: Isolation of cDNA clones Encoding PRO234

A consensus DNA sequence was assembled (DNA30926) using phrap as described in Example 1 above. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence.

RNA for the construction of the cDNA libraries was isolated using standard isolation protocols, e.g., Ausubel *et al.*, *Current Protocols in Molecular Biology*, from tissue or cell line sources or it was purchased from commercial sources (e.g., Clontech). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods (e.g., Ausubel *et al.*) using commercially available reagents (e.g., Invitrogen). This library was derived from 22 week old fetal brain tissue.

A cDNA clone was sequenced in its entirety. The entire nucleotide sequence of PRO234 is shown in Figure 49 (SEQ ID NO:136). The predicted polypeptide precursor is 382 amino acids long and has a calculated molecular weight of approximately 43.1 kDa.

The oligonucleotide sequences used in the above procedure were the following:

30926.p (OLI826) (SEQ ID NO:138): 5'-GTTCATTGAAAACCTCTTGCCATCT GATGGTGACTTCTGGATTGGGCTCA-3'

30926.f (OLI827) (SEQ ID NO:139): 5'-AAGCCAAAGAAGCCTGCAGGAGGG-3'

30926.r (OLI828) (SEQ ID NO:140): 5'-CAGTCCAAGCATAAAGGTCCTGGC-3'

EXAMPLE 23: Isolation of cDNA Clones Encoding Human PRO231

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence was designated herein as DNA30933. Based on the DNA30933 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO231.

Three PCR primers (two forward and one reverse) were synthesized:

forward PCR primer 1 5'-CCAACTACCAAAGCTGCTGGAGCC-3' (SEQ ID NO:143)

forward PCR primer 2 5'-GCAGCTCTATTACCACGGGAAGGA-3' (SEQ ID NO:144)

reverse PCR primer 5'-TCCTTCCCGTGGTAATAGAGCTGC-3' (SEQ ID NO:145)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30933 sequence which had the following nucleotide sequence

hybridization probe

5'-GGCAGAGAACCAGAGGCCGGAGGAGACTGCCTCTTTACAGCCAGG-3' (SEQ ID NO:146)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO231 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO231 [herein designated as DNA34434-1139] and the derived protein sequence for PRO231.

The entire nucleotide sequence of DNA34434-1139 is shown in Figure 51 (SEQ ID NO:141). Clone DNA34434-1139 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 173-175 and ending at the stop codon at nucleotide positions 1457-1459 (Fig. 51; SEQ ID NO:141). The predicted polypeptide precursor is 428 amino acids long (Fig. 52). Clone DNA34434-1139 has been deposited with ATCC on September 16, 1997 and is assigned ATCC deposit no. ATCC 209252.

Analysis of the amino acid sequence of the full-length PRO231 suggests that it possesses 30% and 31% amino acid identity with the human and rat prostatic acid phosphatase precursor proteins, respectively.

EXAMPLE 24: Isolation of cDNA Clones Encoding Human PRO229

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28762. Based on the DNA28762 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO229.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TTCAGCTCATCACCTTCACCTGCC-3' (SEQ ID NO:149)

reverse PCR primer 5'-GGCTCATACAAAATACCACTAGGG-3' (SEQ ID NO:150)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28762 sequence which had the following nucleotide sequence

hybridization probe

5'-GGGCCTCCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGT-3' (SEQ ID NO:151)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO229 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO229 [herein designated as DNA33100-1159] (SEQ ID NO:147) and the derived protein sequence for PRO229.

The entire nucleotide sequence of DNA33100-1159 is shown in Figure 53 (SEQ ID NO:147). Clone DNA33100-1159 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 98-100 and ending at the stop codon at nucleotide positions 1139-1141 (Figure 53). The predicted polypeptide precursor is 347 amino acids long (Figure 54). Clone DNA33100-1159 has been deposited with ATCC and is assigned ATCC deposit no.ATCC 209377

Analysis of the amino acid sequence of the full-length PRO229 polypeptide suggests that portions of it possess significant homology to antigen wc1.1, M130 antigen and CD6.

EXAMPLE 25: Isolation of cDNA Clones Encoding Human PRO238

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described above in Example 1. This consensus sequence is herein designated DNA30908. Based on the DNA30908 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO238.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 1 5'-GGTGCTAAACTGGTGCTCTGTGGC-3' (SEQ ID NO:154)
forward PCR primer 2 5'-CAGGGCAAGATGAGCATTCC-3' (SEQ ID NO:155)
reverse PCR primer 5'-TCATACTGTTCCATCTCGGCACGC-3' (SEQ ID NO:156)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30908 sequence which had the following nucleotide sequence

hybridization probe
5'-AATGGTGGGGCCCTAGAAGAGCTCATCAGAGAACTCACCGCTTCTCATGC-3' (SEQ ID NO:157)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO238 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO238 and the derived protein sequence for PRO238.

The entire nucleotide sequence of DNA35600-1162 is shown in Figure 55 (SEQ ID NO:152). Clone DNA35600-1162 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 134-136 and ending prior to the stop codon at nucleotide positions 1064-1066 (Figure 55). The predicted polypeptide precursor is 310 amino acids long (Figure 56). Clone DNA35600-1162 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209370.

Analysis of the amino acid sequence of the full-length PRO238 polypeptide suggests that portions of it possess significant homology to reductase, particularly oxidoreductase, thereby indicating that PRO238 may be a novel reductase.

EXAMPLE 26: Isolation of cDNA Clones Encoding Human PRO233

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; http://bozeman.mbt.washington.edu/phrap.docs/phrap.html).

An expressed sequence tag (EST) was identified by the EST database search and a consensus DNA sequence was assembled relative to other EST sequences using phrap. This consensus sequence is herein designated DNA30945. Based on the DNA30945 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO233.

Forward and reverse PCR primers were synthesized:

forward PCR primer 5'-GGTGAAGGCAGAAATTGGAGATG-3' (SEQ ID NO:160)

reverse PCR primer 5'-ATCCCATGCATCAGCCTGTTTACC-3' (SEQ ID NO:161)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30945 sequence which had the following nucleotide sequence

hybridization probe

5'-GCTGGTGTAGTCTATACATCAGATTTGTTTGCTACACAAGATCCTCAG-3'

(SEQ ID NO:162)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO233 gene using the probe oligonucleotide.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO233 [herein designated as DNA34436-1238] (SEQ ID NO:158) and the derived protein sequence for PRO233.

The entire nucleotide sequence of DNA34436-1238 is shown in Figure 57 (SEQ ID NO:158). Clone DNA34436-1238 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 101-103 and ending at the stop codon at nucleotide positions 1001-1003 (Figure 57). The predicted polypeptide precursor is 300 amino acids long (Figure 58). The full-length PRO233 protein shown in Figure 58 has an estimated molecular weight of about 32,964 daltons and a pI of about 9.52. Clone DNA34436-1238 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209523.

Analysis of the amino acid sequence of the full-length PRO233 polypeptide suggests that portions of it possess significant homology to reductase proteins, thereby indicating that PRO233 may be a novel reductase.

EXAMPLE 27: Isolation of cDNA Clones Encoding Human PRO223

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30836. Based on the DNA30836 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO223.

PCR primer pairs (one forward and two reverse) were synthesized:

forward PCR primer 5'-TTCCATGCCACCTAAGGGAGACTC-3' (SEQ ID NO:165)

reverse PCR primer 1 5'-TGGATGAGGTGTGCAATGGCTGGC-3' (SEQ ID NO:166)

reverse PCR primer 2 5'-AGCTCTCAGAGGCTGGTCATAGGG-3' (SEQ ID NO:167)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30836 sequence which had the following nucleotide sequence

hybridization probe

5'-GTCGGCCCTTTCCCAGGACTGAACATGAAGAGTTATGCCGGCTTCCTCAC-3' (SEQ ID NO:168)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO223 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO223 [herein designated as DNA33206-1165] (SEQ ID NO:163) and the derived protein sequence for PRO223.

The entire nucleotide sequence of DNA33206-1165 is shown in Figure 59 (SEQ ID NO:163). Clone DNA33206-1165 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 97-99 and ending at the stop codon at nucleotide positions 1525-1527 (Figure 59). The predicted polypeptide precursor is 476 amino acids long (Figure 60). Clone DNA33206-1165 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209372.

Analysis of the amino acid sequence of the full-length PRO223 polypeptide suggests that it possesses significant homology to various serine carboxypeptidase proteins, thereby indicating that PRO223 may be a novel serine carboxypeptidase.

EXAMPLE 28: Isolation of cDNA Clones Encoding Human PRO235

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated "DNA30927". Based on the DNA30927 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO235.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGAATACCGCCTCCTGCAG-3' (SEQ ID NO:171)

reverse PCR primer 5'-CTTCTGCCCTTTGGAGAAGATGGC-3' (SEQ ID NO:172)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30927 sequence which had the following nucleotide sequence

hybridization probe

5'-GGACTCACTGGCCCAGGCCTTCAATATCACCAGCCAGGACGAT-3' (SEQ ID NO:173)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO235 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO235 [herein designated as DNA35558-1167] (SEQ ID NO:169) and the derived protein sequence for PRO235.

The entire nucleotide sequence of DNA35558-1167 is shown in Figure 61 (SEQ ID NO:169). Clone DNA35558-1167 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 667-669 and ending at the stop codon at nucleotide positions 2323-2325 (Figure 61). The predicted polypeptide precursor is 552 amino acids long (Figure 62). Clone DNA35558-1167 has been deposited with ATCC and is assigned ATCC deposit no. 209374.

Analysis of the amino acid sequence of the full-length PRO235 polypeptide suggests that portions of it possess significant homology to the human, mouse and *Xenopus* plexin protein, thereby indicating that PRO235 may be a novel plexin protein.

EXAMPLE 29: Isolation of cDNA Clones Encoding Human PRO236 and Human PRO262

Consensus DNA sequences were assembled relative to other EST sequences using phrap as described in Example 1 above. These consensus sequences are herein designated DNA30901 and DNA30847. Based on the DNA30901 and DNA30847 consensus sequences, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO236 and PRO262, respectively.

Based upon the DNA30901 consensus sequence, a pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGCTACTCCAAGACCCTGGCATG-3' (SEQ ID NO:178)

reverse PCR primer 5'-TGGACAAATCCCCTTGCTCAGCCC-3' (SEQ ID NO:179)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30901 sequence which had the following nucleotide sequence

hybridization probe

5'-GGGCTTCACCGAAGCAGTGGACCTTTATTTTGACCACCTGATGTCCAGGG-3' (SEQ ID NO:180)

Based upon the DNA30847 consensus sequence, a pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCAGCTATGACTATGATGCACC-3' (SEQ ID NO:181)

reverse PCR primer 5'-TGGCACCCAGAATGGTGTTGGCTC-3' (SEQ ID NO:182)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30847 sequence which had the following nucleotide sequence

hybridization probe

5'-CGAGATGTCATCAGCAAGTTCCAGGAAGTTCCTTTGGGACCTTTACCTCC-3' (SEQ ID NO:183)

In order to screen several libraries for a source of full-length clones, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. Positive libraries were then used to isolate clones encoding the PRO236 and PRO262 genes using the probe oligonucleotides and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human feral lung tissue for PRO236 and human fetal liver tissue for PRO262.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO236 [herein designated as DNA35599-1168] (SEQ ID NO:174), the derived protein sequence for PRO236, the full-length DNA sequence for PRO262 [herein designated as DNA36992-1168] (SEQ ID NO:176) and the derived protein sequence for PRO262.

The entire nucleotide sequence of DNA35599-1168 is shown in Figure 63 (SEQ ID NO:174). Clone DNA35599-1168 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 69-71 and ending at the stop codon at nucleotide positions 1977-1979 (Figure 63). The predicted polypeptide precursor is 636 amino acids long (Figure 64). Clone DNA35599-1168 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209373.

The entire nucleotide sequence of DNA36992-1168 is shown in Figure 65 (SEQ ID NO:176). Clone DNA36992-1168 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 240-242 and ending at the stop codon at nucleotide positions 2202-2204 (Figure 65). The predicted polypeptide precursor is 654 amino acids long (Figure 66). Clone DNA36992-1168 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209382.

Analysis of the amino acid sequence of the full-length PRO236 and PRO262 polypeptides suggests that portions of those polypeptides possess significant homology to β-galactosidase proteins derived from various sources, thereby indicating that PRO236 and PRO262 may be novel β-galactosidase homologs.

EXAMPLE 30: Isolation of cDNA Clones Encoding Human PRO239

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30909. Based on the DNA30909 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO239.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCTCCCTCTATTACCCATGTC-3' (SEQ ID NO:186)

reverse PCR primer 5'-GACCAACTTTCTCTGGGAGTGAGG-3' (SEQ ID NO:187)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30909 sequence which had the following nucleotide sequence

hybridization probe

5'-GTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAG-3'

(SEQ ID NO:188)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO239 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO239 [herein designated as DNA34407-1169] (SEQ ID NO:184) and the derived protein sequence for PRO239.

The entire nucleotide sequence of DNA34407-1169 is shown in Figure 67 (SEQ ID NO:184). Clone DNA34407-1169 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 72-74 and ending at the stop codon at nucleotide positions 1575-1577 (Figure 67). The predicted polypeptide precursor is 501 amino acids long (Figure 68). Clone DNA34407-1169 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209383.

Analysis of the amino acid sequence of the full-length PRO239 polypeptide suggests that portions of it possess significant homology to the densin protein, thereby indicating that PRO239 may be a novel molecule in the densin family.

EXAMPLE 31: Isolation of cDNA Clones Encoding Human PRO257

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28731. Based on the DNA28731 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO257.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TCTCTATTCCAAACTGTGGCG-3' (SEQ ID NO:191)

reverse PCR primer 5'-TTTGATGACGATTCGAAGGTGG-3' (SEQ ID NO:192)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28731 sequence which had the following nucleotide sequence

hybridization probe

5'-GGAAGGATCCTTCACCAGCCCCAATTACCCAAAGCCGCATCCTGAGC-3' (SEQ ID NO:193)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO257 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO257 [herein designated as DNA35841-1173 (SEQ ID NO:189] and the derived protein sequence for PRO257.

The entire nucleotide sequence of DNA35841-1173 is shown in Figure 69 (SEQ ID NO:189). Clone DNA35841-1173 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 964-966 and ending at the stop codon at nucleotide positions 2785-2787 (Figure 69). The predicted polypeptide precursor is 607 amino acids long (Figure 70). Clone DNA35841-1173 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209403.

Analysis of the amino acid sequence of the full-length PRO257 polypeptide suggests that portions of it possess significant homology to the ebnerin protein, thereby indicating that PRO257 may be a novel protein member related to the ebnerin protein.

EXAMPLE 32: Isolation of cDNA Clones Encoding Human PRO260

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30834. Based on the DNA30834 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO260.

PCR primers (forward and two reverse) were synthesized:

forward PCR primer: 5'-TGGTTTGACCAGGCCAAGTTCGG-3' (SEQ ID NO:196);

reverse PCR primer A: 5'-GGATTCATCCTCAAGGAAGAGCGG-3' (SEQ ID NO:197); and

reverse PCR primer B: 5'AACTTGCAGCATCAGCCACTCTGC-3' (SEQ ID NO:198)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30834 sequence which had the following nucleotide sequence:

hybridization probe:

5'-TTCCGTGCCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCA-3' (SEQ ID NO:199)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO260 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO260 [herein designated as DNA33470-1175] (SEQ ID NO:194) and the derived protein sequence for PRO260.

The entire nucleotide sequence of DNA33470-1175 is shown in Figure 71 (SEQ ID NO:194). Clone DNA33470-1175 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 67-69 and ending at the stop codon 1468-1470 (see Figure 71). The predicted polypeptide precursor is 467 amino acids long (Figure 72). Clone DNA33470-1175 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209398.

Analysis of the amino acid sequence of the full-length PRO260 polypeptide suggests that portions of it possess significant homology to the alpha-l-fucosidase precursor, thereby indicating that PRO260 may be a novel fucosidase.

EXAMPLE 33: Isolation of cDNA Clones Encoding Human PRO263

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30914. Based on the DNA30914 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO263.

PCR primers (tow forward and one reverse) were synthesized:

forward PCR primer 1: 5'-GAGCTTTCCATCCAGGTGTCATGC-3' (SEQ ID NO:202);

forward PCR primer 2: 5'-GTCAGTGACAGTACCTACTCGG-3' (SEQ ID NO:203); reverse PCR primer: 5'-TGGAGCAGGAGGAGTAGTAGTAGG-3' (SEQ ID NO:204)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30914 sequence which had the following nucleotide sequence:

hybridization probe:

5'-AGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTT-3' (SEQ ID NO:205)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO263 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO263 [herein designated as DNA34431-1177] (SEQ ID NO:200) and the derived protein sequence for PRO263.

The entire nucleotide sequence of DNA34431-1177 is shown in Figure 73 (SEQ ID NO:200). Clone DNA34431-1177 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 160-162 of SEQ ID NO:200 and ending at the stop codon after the nucleotide at position 1126-1128 of SEQ ID NO:200 (Figure 73). The predicted polypeptide precursor is 322 amino acids long (Figure 74). Clone DNA34431-1177 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209399.

Analysis of the amino acid sequence of the full-length PRO263 polypeptide suggests that portions of it possess significant homology to CD44 antigen, thereby indicating that PRO263 may be a novel cell surface adhesion molecule.

EXAMPLE 34: Isolation of cDNA Clones Encoding Human PRO270

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence was designated herein as DNA35712. Based on the DNA35712 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that

contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO270. Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-GCTTGGATATTCGCATGGGCCTAC-3' (SEQ ID NO:208)
forward PCR primer (.f2) 5'-TGGAGACAATATCCCTGAGG-3' (SEQ ID NO:209)
reverse PCR primer (.r1) 5'-AACAGTTGGCCACAGCATGGCAGG-3' (SEQ ID NO:210)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35712 sequence which had the following nucleotide sequence

hybridization probe
5'-CCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAG-3'
(SEQ ID NO:211)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO270 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO270 [herein designated as DNA39510-1181] (SEQ ID NO:206) and the derived protein sequence for PRO270.

The entire nucleotide sequence of DNA39510-1181 is shown in Figure 75 (SEQ ID NO:206). Clone DNA39510-1181 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 3-5 and ending at the stop codon at nucleotide positions 891-893 (Fig. 75; SEQ ID NO:206). The predicted polypeptide precursor is 296 amino acids long (Fig. 76). Clone DNA39510-1181 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209392.

Analysis of the amino acid sequence of the full-length PRO270 suggests that portions of it possess significant homology to the thioredoxin-protein, thereby indicating that the PRO270 protein may be a novel member of the thioredoxin family.

EXAMPLE 35: Isolation of cDNA Clones Encoding Human PRO271

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35737. Based on the DNA35737 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO271.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-TGCTTCGCTACTGCCCTC-3' (SEQ ID NO:214)
forward PCR primer 2 5'-TTCCCTTGTGGGTTGGAG-3' (SEQ ID NO:215)
forward PCR primer 3 5'-AGGGCTGGAAGCCAGTTC-3' (SEQ ID NO:216)
reverse PCR primer 1 5'-AGCCAGTGAGGAAATGCG-3' (SEQ ID NO:217)
reverse PCR primer 2 5'-TGTCCAAAGTACACACACCTGAGG-3' (SEQ ID NO:218)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35737 sequence which had the following nucleotide sequence

hybridization probe

5'-GATGCCACGATCGCCAAGGTGGGACAGCTCTTTGCCGCCTGGAAG-3' (SEQ ID NO:219)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO271 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO271 [herein designated as DNA39423-1182] (SEQ ID NO:212) and the derived protein sequence for PRO271.

The entire nucleotide sequence of DNA39423-1182 is shown in Figure 77 (SEQ ID NO:212). Clone DNA39423-1182 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 101-103 and ending at the stop codon at nucleotide positions 1181-1183 (Figure 77). The predicted polypeptide precursor is 360 amino acids long (Figure 78). Clone DNA39423-1182 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209387.

Analysis of the amino acid sequence of the full-length PRO271 polypeptide suggests that it possess significant homology to the proteoglycan link protein, thereby indicating that PRO271 may be a link protein homolog.

EXAMPLE 36: Isolation of cDNA Clones Encoding Human PRO272

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA36460. Based on the DNA36460 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO272.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1)	5'-CGCAGGCCCTCATGGCCAGG-3'	(SEQ ID NO:222)
forward PCR primer (.f2)	5'-GAAATCCTGGGTAATTGG-3'	(SEQ ID NO:223)
reverse PCR primer	5'-GTGCGCGGTGCTCACAGCTCATC-3'	(SEQ ID NO:224)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA36460 sequence which had the following nucleotide sequence

hybridization probe

5'-CCCCCCTGAGCGACGCTCCCCCATGATGACGCCCACGGGAACTTC-3' (SEQ ID NO:225)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO272 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO272 [herein designated as DNA40620-1183] (SEQ ID NO:220) and the derived protein sequence for PRO272.

The entire nucleotide sequence of DNA40620-1183 is shown in Figure 79 (SEQ ID NO:220). Clone DNA40620-1183 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 35-37 and ending at the stop codon at nucleotide positions 1019-1021 (Figure 79). The predicted polypeptide precursor is 328 amino acids long (Figure 80). Clone DNA40620-1183 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209388.

Analysis of the amino acid sequence of the full-length PRO272 polypeptide suggests that portions of it possess significant homology to the human and mouse reticulocalbin proteins, respectively, thereby indicating that PRO272 may be a novel reticulocalbin protein.

EXAMPLE 37: Isolation of cDNA Clones Encoding Human PRO294

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35731. Based on the DNA35731 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO294.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-TGGTCTCGCACACCGATC-3' (SEQ ID NO:228)
forward PCR primer (.f2) 5'-CTGCTGTCCACAGGGGAG-3' (SEQ ID NO:229)
forward PCR primer (.f3) 5'-CCTTGAAGCATACTGCTC-3' (SEQ ID NO:230)
forward PCR primer (.f4) 5'-GAGATAGCAATTTCCGCC-3' (SEQ ID NO:231)
reverse PCR primer (.r1) 5'-TTCCTCAAGAGGGCAGCC-3' (SEQ ID NO:232)
reverse PCR primer (.r2) 5'-CTTGGCACCAATGTCCGAGATTTC-3' (SEQ ID NO:233)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35731 sequence which had the following nucleotide sequence

hybridization probe
5'-GCTCTGAGGAAGGTGACGCGCGGGGCCTCCGAACCCTTGGCCTTG-3' (SEQ ID NO:234)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO294 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO294 [herein designated as DNA40604-1187] (SEQ ID NO:226) and the derived protein sequence for PRO294.

The entire nucleotide sequence of DNA40604-1187 is shown in Figure 81 (SEQ ID NO:226). Clone DNA40604-1187 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 396-398 and ending at the stop codon at nucleotide positions 2046-2048 (Figure 81). The predicted polypeptide precursor is 550 amino acids long (Figure 82). Clone DNA40604-1187 has been deposited with ATCC and is assigned ATCC deposit no. 209394.

Analysis of the amino acid sequence of the full-length PRO294 polypeptide suggests that portions of it possess significant homology to portions of various collagen proteins, thereby indicating that PRO294 may be collagen-like molecule.

EXAMPLE 38: Isolation of cDNA Clones Encoding Human PRO295

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35814. Based on the DNA35814 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO295.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-GCAGAGCGGAGATGCAGCGGCTTG-3' (SEQ ID NO:238)

forward PCR primer (.f2) 5'-CCCAGCATGTACTGCCAG-3' (SEQ ID NO:239)

forward PCR primer (.f3) 5'-TTGGCAGCTTCATGGAGG-3' (SEQ ID NO:240)

forward PCR primer (.f4) 5'-CCTGGGCAAAAATGCAAC-3' (SEQ ID NO:241)

reverse PCR primer (.r1) 5'-CTCCAGCTCCTGGCGCACCTCCTC-3' (SEQ ID NO:242)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35814 sequence which had the following nucleotide sequence

hybridization probe

5'-GGCTCTCAGCTACCGCGCAGGAGCGAGGCCACCCTCAATGAGATG-3' (SEQ ID NO:243)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO295 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO295 [herein designated as DNA38268-1188] (SEQ ID NO:235) and the derived protein sequence for PRO295.

The entire nucleotide sequence of DNA38268-1188 is shown in Figure 83 (SEQ ID NO:235). Clone DNA38268-1188 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 153-155 and ending at the stop codon at nucleotide positions 1202-1204 (Figure 83). The predicted polypeptide precursor is 350 amino acids long (Figure 84). Clone DNA38268-1188 has been deposited with

ATCC and is assigned ATCC deposit no. 209421.

Analysis of the amino acid sequence of the full-length PRO295 polypeptide suggests that portions of it possess significant homology to the integrin proteins, thereby indicating that PRO295 may be a novel integrin.

EXAMPLE 39: Isolation of cDNA Clones Encoding Human PRO293

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; http://bozeman.mbt.washington.edu/phrap.docs/phrap.html).

Based on an expression tag sequence designated herein as T08294 identified in the above analysis, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO293.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AACAAGGTAAGATGCCATCCTG-3' (SEQ ID NO:246)

reverse PCR primer 5'-AAACTTGTCGATGGAGACCAGCTC-3' (SEQ ID NO:247)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the expression sequence tag which had the following nucleotide sequence

hybridization probe

5'-AGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTCTATGACAACCAGC-3' (SEQ ID NO:248)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO293 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO293 [herein designated as DNA37151-1193] (SEQ ID NO:244) and the derived protein sequence for PRO293.

The entire nucleotide sequence of DNA37151-1193 is shown in Figure 85 (SEQ ID NO:244). Clone DNA37151-1193 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 881-883 and ending at the stop codon after nucleotide position 3019 of SEQ ID NO:244, Figure 85). The predicted polypeptide precursor is 713 amino acids long (Figure 86). Clone DNA37151-1193 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209393.

Analysis of the amino acid sequence of the full-length PRO293 polypeptide suggests that portions of it possess significant homology to the NLRR proteins, thereby indicating that PRO293 may be a novel NLRR

protein.

EXAMPLE 40: Isolation of cDNA Clones Encoding Human PRO247

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA33480. Based on the DNA33480 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO247.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CAACAATGAGGGCACCAAGC-3' (SEQ ID NO:251)

reverse PCR primer 5'-GATGGCTAGGTTCTGGAGGTTCTG-3' (SEQ ID NO:252)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA33480 expression sequence tag which had the following nucleotide sequence

hybridization probe

5'-CAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCG-3' (SEQ ID NO:253)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO247 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO247 (herein designated as DNA35673-1201) (SEQ ID NO:249) and the derived protein sequence for PRO247.

The entire nucleotide sequence of DNA35673-1201 is shown in Figure 89 (SEQ ID NO:249). Clone DNA35673-1201 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 80-82 of SEQ ID NO:249 and ending at the stop codon after nucleotide position 1717 of SEQ ID NO:249 (Figure 89). The predicted polypeptide precursor is 546 amino acids long (Figure 88). Clone DNA35673-1201 has been deposited with ATCC and is assigned ATCC deposit no. 209418.

Analysis of the amino acid sequence of the full-length PRO247 polypeptide suggests that portions of it possess significant homology to the densin molecule and KIAA0231, thereby indicating that PRO247 may be a novel leucine rich repeat protein.

EXAMPLE 41: Isolation of cDNA Clones Encoding Human PRO302, PRO303, PRO304, PRO307 and PRO343

Consensus DNA sequences were assembled relative to other EST sequences using phrap as described in Example 1 above. These consensus sequences are herein designated DNA35953, DNA35955, DNA35958, DNA37160 and DNA30895. Based on the DNA35953 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO302.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 1 5'-GTCCGCAAGGATGCCTACATGTTC-3' (SEQ ID NO:264)

forward PCR primer 2 5'-GCAGAGGTGTCTAAGGTTG-3' (SEQ ID NO:265)

reverse PCR primer 5'-AGCTCTAGACCAATGCCAGCTTCC-3' (SEQ ID NO:266)

Also, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35953 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCTGGTCATG-3' (SEQ ID NO:267)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO302 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB228).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO302 [herein designated as DNA40370-1217] (SEQ ID NO:254) and the derived protein sequence for PRO302.

The entire nucleotide sequence of DNA40370-1217 is shown in Figure 89 (SEQ ID NO:254). Clone DNA40370-1217 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 34-36 and ending at the stop codon at nucleotide positions 1390-1392 (Figure 89). The predicted polypeptide precursor is 452 amino acids long (Figure 90). Various unique aspects of the PRO302 protein are shown in Figure 90. Clone DNA40370-1217 has been deposited with the ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209485.

Based on the DNA35955 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO303.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GGGGAATTCACCCTATGACATTGCC-3' (SEQ ID NO:268)

reverse PCR primer 5'-GAATGCCCTGCAAGCATCAACTGG-3' (SEQ ID NO:269)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35955 sequence which had the following nucleotide sequence:

hybridization probe

5'-GCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCCTC-3' (SEQ ID NO:270)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO303 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO303 [herein designated as DNA42551-1217] (SEQ ID NO:256) and the derived protein sequence for PRO303.

The entire nucleotide sequence of DNA42551-1217 is shown in Figure 91 (SEQ ID NO:256). Clone DNA42551-1217 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 20-22 and ending at the stop codon at nucleotide positions 962-964 (Figure 91). The predicted polypeptide precursor is 314 amino acids long (Figure 92). Various unique aspects of the PRO303 protein are shown in Figure 92. Clone DNA42551-1217 has been deposited on November 21, 1997 with the ATCC and is assigned ATCC deposit no. ATCC 209483.

Based on the DNA35958 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO304.

Pairs of PCR primers (forward and reverse) were synthesized:

forward PCR primer 1 5'-GCGGAAGGGCAGAATGGGACTCCAAG-3' (SEQ ID NO:271)

forward PCR primer 2 5'-CAGCCCTGCCACATGTGC-3' (SEQ ID NO:272)

forward PCR primer 3 5'-TACTGGGTGGTCAGCAAC-3' (SEQ ID NO:273)

reverse PCR primer 5'-GGCGAAGAGCAGGGTGAGACCCCG-3' (SEQ ID NO:274)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35958 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGAC-3' (SEQ ID NO:275)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO304 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from 22 week human fetal brain tissue (LIB153).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO304 [herein designated as DNA39520-1217] (SEQ ID NO:258) and the derived protein sequence for PRO304.

The entire nucleotide sequence of DNA39520-1217 is shown in Figure 93 (SEQ ID NO:258). Clone DNA39520-1217 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 34-36 and ending at the stop codon at nucleotide positions 1702-1704 (Figure 93). The predicted polypeptide precursor is 556 amino acids long (Figure 94). Various unique aspects of the PRO304 protein are shown in Figure 94. Clone DNA39520-1217 has been deposited with ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209482.

Based on the DNA37160 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO307.

Pairs of PCR primers (forward and reverse) were synthesized:

forward PCR primer 1 5'-GGGCAGGGATTCCAGGGCTCC-3' (SEQ ID NO:276)

forward PCR primer 2 5'-GGCTATGACAGCAGGTTC-3' (SEQ ID NO:277)

forward PCR primer 3 5'-TGACAATGACCGACCAGG-3' (SEQ ID NO:278)
reverse PCR primer 5'-GCATCGCATTGCTGGTAGAGCAAG-3' (SEQ ID NO:279)
Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA37160 sequence which had the following nucleotide sequence
hybridization probe
5'-TTACAGTGCCCCCTGGAAACCCACTTGGCCTGCATACCGCCTCCC-3' (SEQ ID NO:280)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO307 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue (LIB229).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO307 [herein designated as DNA41225-1217] (SEQ ID NO:260) and the derived protein sequence for PRO307.

The entire nucleotide sequence of DNA41225-1217 is shown in Figure 95 (SEQ ID NO:260). Clone DNA41225-1217 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 92-94 and ending at the stop codon at nucleotide positions 1241-1243 (Figure 95). The predicted polypeptide precursor is 383 amino acids long (Figure 96). Various unique aspects of the PRO307 protein are shown in Figure 96. Clone DNA41225-1217 has been deposited with ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209491.

Based on the DNA30895 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO343.

A pair of PCR primers (forward and reverse) were synthesized:
forward PCR primer 5'-CGTCTCGAGCGCTCCATACAGTTCCCTTGCCCCA-3' (SEQ ID NO:281)
reverse PCR primer
5'-TGGAGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCC
CCCTCATGTGCCAGGTGGA-3' (SEQ ID NO:282)
Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30895 sequence which had the following nucleotide sequence
hybridization probe
5'-CCCTCAGACCCTGCAGAAGCTGAAGGTTCCTATCATCGAC
TCGGAAGTCTGCAGCCATCTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGC
TGTGTGCCGGCTACT-3' (SEQ ID NO:283)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO343 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB26).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for

PRO343 [herein designated as DNA43318-1217] (SEQ ID NO:262) and the derived protein sequence for PRO343.

The entire nucleotide sequence of DNA43318-1217 is shown in Figure 97 (SEQ ID NO:262). Clone DNA43318-1217 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 53-55 and ending at the stop codon at nucleotide positions 1004-1006 (Figure 97). The predicted polypeptide precursor is 317 amino acids long (Figure 98). Various unique aspects of the PRO343 protein are shown in Figure 98. Clone DNA43318-1217 has been deposited with ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209481.

EXAMPLE 42: Isolation of cDNA Clones Encoding Human PRO328

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35615. Based on the DNA35615 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO328.

Forward and reverse PCR primers were synthesized:

forward PCR primer	5'-TCCTGCAGTTTCCTGATGC-3'	(SEQ ID NO:286)
reverse PCR primer	5'-CTCATATTGCACACCAGTAATTCG-3'	(SEQ ID NO:287)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35615 sequence which had the following nucleotide sequence

hybridization probe

5'-ATGAGGAGAAACGTTTGATGGTGGAGCTGCACAACCTCTACCGGG-3'

(SEQ ID NO:288)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO328 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO328 [herein designated as DNA40587-1231] (SEQ ID NO:284) and the derived protein sequence for PRO328.

The entire nucleotide sequence of DNA40587-1231 is shown in Figure 99 (SEQ ID NO:284). Clone DNA40587-1231 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 15-17 and ending at the stop codon at nucleotide positions 1404-1406 (Figure 99). The predicted polypeptide precursor is 463 amino acids long (Figure 100). Clone DNA40587-1231 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209438.

Analysis of the amino acid sequence of the full-length PRO328 polypeptide suggests that portions of it possess significant homology to the human glioblastoma protein and to the cysteine rich secretory protein thereby indicating that PRO328 may be a novel glioblastoma protein or cysteine rich secretory protein.

EXAMPLE 43: Isolation of cDNA Clones Encoding Human PRO335, PRO331 or PRO326

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA36685. Based on the DNA36685 consensus sequence, and Incyte EST sequence no. 2228990, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO335, PRO331 or PRO326.

Forward and reverse PCR primers were synthesized for the determination of PRO335:

forward PCR primer 5'-GGAACCGAATCTCAGCTA-3' (SEQ ID NO:295)
forward PCR primer 5'-CCTAAACTGAACTGGACCA-3' (SEQ ID NO:296)
forward PCR primer 5'-GGCTGGAGACACTGAACCT-3' (SEQ ID NO:297)
forward PCR primer 5'-ACAGCTGCACAGCTCAGAACAGTG-3' (SEQ ID NO:298)
reverse PCR primer 5'-CATTCCCAGTATAAAAATTTTC-3' (SEQ ID NO:299)
reverse PCR primer 5'-GGGTCTTGGTGAATGAGG-3' (SEQ ID NO:300)
reverse PCR primer 5'-GTGCCTCTCGGTTACCACCAATGG-3' (SEQ ID NO:301)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO335 which had the following nucleotide sequence

hybridization probe
5'-GCGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGCCGTCCTAC-3'
(SEQ ID NO:302)

Forward and reverse PCR primers were synthesized for the determination of PRO331:

forward PCR primer 5'-GCCTTTGACAACCTTCAGTCACTAGTGG-3' (SEQ ID NO:303)
reverse PCR primer 5'-CCCCATGTGTCCATGACTGTTCCC-3' (SEQ ID NO:304)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO331 which had the following nucleotide sequence

hybridization probe
5'-TACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTTAGAGCGG-3'
(SEQ ID NO:305)

Forward and reverse PCR primers were synthesized for the determination of PRO326:

forward PCR primer 5'-ACTCCAAGGAAATCGGATCCGTTC-3' (SEQ ID NO:306)
reverse PCR primer 5'-TTAGCAGCTGAGGATGGGCACAAC-3' (SEQ ID NO:307)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO331 which had the following nucleotide sequence

hybridization probe
5'-GCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGTGACAACGC-3'
(SEQ ID NO:308)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO335, PRO331 or PRO326 gene using the probe oligonucleotide and one of

the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (PRO335 and PRO326) and human fetal brain (PRO331).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO335, PRO331 or PRO326 [herein designated as SEQ ID NOS:289, 291 and 293, respectively; see Figures 103, 105 and 107, respectively], and the derived protein sequence for PRO335, PRO331 or PRO326 (see Figures 104, 106 and 108, respectively; SEQ ID NOS:290, 292 and 294, respectively).

The entire nucleotide sequences are shown in Figures 103, 105 and 107, deposited with the ATCC on June 2, 1998, November 7, 1997 and November 21, 1997, respectively.

Analysis of the amino acid sequence of the full-length PRO335, PRO331 or PRO326 polypeptide suggests that portions of it possess significant homology to the LIG-1 protein, thereby indicating that PRO335, PRO331 and PRO326 may be a novel LIG-1-related protein.

EXAMPLE 44: Isolation of cDNA clones Encoding Human PRO332

Based upon an ECD homology search performed as described in Example 1 above, a consensus DNA sequence designated herein as DNA36688 was assembled. Based on the DNA36688 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO332.

A pair of PCR primers (forward and reverse) were synthesized:

5'-GCATTGGCCGCGAGACTTTGCC-3' (SEQ ID NO:311)

5'-GCGGCCACGGTCCTTGGAAATG-3' (SEQ ID NO:312)

A probe was also synthesized:

5'-TGGAGGAGCTCAACCTCAGCTACAACCGCATCACCAGCCCACAGG-3'

(SEQ ID NO:313)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO332 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from a human fetal liver library (LIB229).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for DNA40982-1235 and the derived protein sequence for PRO332.

The entire nucleotide sequence of DNA40982-1235 is shown in Figure 109 (SEQ ID NO:309). Clone DNA40982-1235 contains a single open reading frame (with an apparent translational initiation site at nucleotide positions 342-344, as indicated in Figure 109). The predicted polypeptide precursor is 642 amino acids long, and has a calculated molecular weight of 72,067 (pI: 6.60). Clone DNA40982-1235 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209433.

Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO332 shows about 30-40% amino acid sequence identity with a series of known proteoglycan sequences, including, for example, fibromodulin and fibromodulin precursor sequences of various species (FMOD_BOVIN, FMOD

CHICK, FMOD_RAT, FMOD_MOUSE, FMOD_HUMAN, P_R36773), osteomodulin sequences (AB000114 1, AB007848_1), decorin sequences (CFU83141_1, OCU03394_1, P_R42266, P_R42267, P_R42260, P_R89439), keratan sulfate proteoglycans (BTU48360_1, AF022890_1), corneal proteoglycan (AF022256_1), and bone/cartilage proteoglycans and proteoglycane precursors (PGS1_BOVIN, PGS2_MOUSE, PGS2_HUMAN).

EXAMPLE 45: Isolation of cDNA clones Encoding Human PRO334

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. Based on the consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO334.

Forward and reverse PCR primers were synthesized for the determination of PRO334:

forward PCR primer 5'-GATGGTTCCTGCTCAAGTGCCCTG-3' (SEQ ID NO:316)

reverse PCR primer 5'-TTGCACTTGTAGGACCCACGTACG-3' (SEQ ID NO:317)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO334 which had the following nucleotide sequence

hybridization probe

5'-CTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCC-3'

(SEQ ID NO:318)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO334 gene using the probe oligonucleotide and one of the PCR primers.

Human feral kidney cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO334 [herein designated as DNA41379-1236] (SEQ ID NO:314) and the derived protein sequence for PRO334.

The entire nucleotide sequence of DNA41379-1236 (also referred to as UNQ295) is shown in Figure 109 (SEQ ID NO:314). Clone DNA41379-1236 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 203-205 and ending at the stop codon at nucleotide positions 1730-1732 (Figure 109). The predicted polypeptide precursor is 509 amino acids long (Figure 110). Clone DNA41379-1236 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209488.

Analysis of the amino acid sequence of the full-length PRO334 polypeptide suggests that portions of it possess significant homology to the fibulin and fibrillin proteins, thereby indicating that PRO334 may be a novel member of the EGF protein family.

EXAMPLE 46: Isolation of cDNA Clones Encoding Human PRO346

A consensus DNA sequence was identified using phrap as described in Example 1 above. Specifically, this consensus sequence is herein designated DNA38240. Based on the DNA38240 consensus sequence,

oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length PRO346 coding sequence.

RNA for construction of the cDNA libraries was isolated from human fetal liver. The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents (e.g., Invitrogen, San Diego, CA; Clontech, etc.) The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site: see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A cDNA clone was sequenced in entirety. The entire nucleotide sequence of DNA44167-1243 is shown in Figure 111 (SEQ ID NO:319). Clone DNA44167-1243 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 64-66 (Fig. 113; SEQ ID NO:319). The predicted polypeptide precursor is 450 amino acids long. Clone DNA44167-1243 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209434 (designation DNA44167-1243).

Based on a BLAST, BLAST-2 and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO346 shows amino acid sequence identity to carcinoembryonic antigen (28%).

The oligonucleotide sequences used in the above procedure were the following:

OLI2691 (38240.f1)

5'-GATCCTGTCACAAAGCCAGTGGTGC-3' (SEQ ID NO:321)

OLI2693 (38240.c1)

5'-CACTGACAGGGTTCCTCACCCAGG-3' (SEQ ID NO:322)

OLI2692 (38240.p1)

5'-CTCCCTCTGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATG-3' (SEQ ID NO:323)

EXAMPLE 47: Isolation of cDNA Clones Encoding Human PRO268

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35698. Based on the DNA35698 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO268.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-TGAGGTGGGCAAGCGGCGAAATG-3' (SEQ ID NO:326)

forward PCR primer 2 5'-TATGTGGATCAGGACGTGCC-3' (SEQ ID NO:327)

forward PCR primer 3 5'-TGCAGGGTTCAGTCTAGATTG-3' (SEQ ID NO:328)

reverse PCR primer 5'-TTGAAGGACAAAGGCAATCTGCCAC-3' (SEQ ID NO:329)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35698 sequence which had the following nucleotide sequence

hybridization probe

5'-GGAGTCTTGCAGTTCCCCTGGCAGTCCTGGTGCTGTTGCTTTGGG-3' (SEQ ID NO:330)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO268 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO268 [herein designated as DNA39427-1179] (SEQ ID NO:324) and the derived protein sequence for PRO268.

The entire nucleotide sequence of DNA39427-1179 is shown in Figure 113 (SEQ ID NO:324). Clone DNA39427-1179 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 13-15 and ending at the stop codon at nucleotide positions 853-855 (Figure 113). The predicted polypeptide precursor is 280 amino acids long (Figure 114). Clone DNA39427-1179 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209395.

Analysis of the amino acid sequence of the full-length PRO268 polypeptide suggests that it possess significant homology to protein disulfide isomerase, thereby indicating that PRO268 may be a novel protein disulfide isomerase.

EXAMPLE 48: Isolation of cDNA Clones Encoding Human PRO330

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35730. Based on the DNA35730 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO330.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-CCAGGCACAATTTCCAGA-3' (SEQ ID NO:333)

forward PCR primer 2 5'-GGACCCTTCTGTGTGCCAG-3' (SEQ ID NO:334)

reverse PCR primer 1 5'-GGTCTCAAGAACTCCTGTC-3' (SEQ ID NO:335)

reverse PCR primer 2 5'-ACACTCAGCATTGCCTGGTACTTG-3' (SEQ ID NO:336)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which had the following nucleotide sequence

hybridization probe

5'-GGGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAG-3' (SEQ ID NO:337)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO330 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for

PRO330 [herein designated as DNA40603-1232] (SEQ ID NO:331) and the derived protein sequence for PRO330.

The entire nucleotide sequence of DNA40603-1232 is shown in Figure 115 (SEQ ID NO:331). Clone DNA40603-1232 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 167-169 and ending at the stop codon at nucleotide positions 1766-1768 (Figure 115). The predicted polypeptide precursor is 533 amino acids long (Figure 116). Clone DNA40603-1232 has been deposited with ATCC and is assigned ATCC deposit no.ATCC 209486 on November 21, 1997.

Analysis of the amino acid sequence of the full-length PRO330 polypeptide suggests that portions of it possess significant homology to the mouse prolyl 4-hydroxylase alpha subunit protein, thereby indicating that PRO330 may be a novel prolyl 4-hydroxylase alpha subunit polypeptide.

EXAMPLE 49: Isolation of cDNA Clones Encoding Human PRO310

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA40553. Based on the DNA40553 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO310.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-TCCCCAAGCCGTTCTAGACGCGG-3' (SEQ ID NO:342)

forward PCR primer 2 5'-CTGGTTCTTCCTTGCACG-3' (SEQ ID NO:343)

reverse PCR primer 5'-GCCCAAATGCCCTAAGGCGGTATACCCC-3' (SEQ ID NO:344)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which had the following nucleotide sequence

hybridization probe

5'-GGGTGTGATGCTTGGAAGCATTTTCTGTGCTTTGATCACTATGCTAGGAC-3' (SEQ ID NO:345)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO310 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO310 [herein designated as DNA43046-1225 (SEQ ID NO:340) and the derived protein sequence for PRO310 (SEQ ID NO:341).

The entire nucleotide sequence of DNA43046-1225 is shown in Figure 119 (SEQ ID NO:340). Clone DNA43046-1225 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 81-83 and ending at the stop codon at nucleotide positions 1035-1037 (Figure 119). The predicted polypeptide precursor is 318 amino acids long (Figure 120) and has a calculated molecular weight of approximately 36,382 daltons. Clone DNA43046-1225 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209484.

Analysis of the amino acid sequence of the full-length PRO310 polypeptide suggests that portions of it possess homology to *C. elegans* proteins and to fringe, thereby indicating that PRO310 may be involved in development.

EXAMPLE 50: Isolation of cDNA clones Encoding Human PRO339

An expressed sequence tag (EST) DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) was searched and ESTs were identified. An assembly of Incyte clones and a consensus sequence was formed using phrap as described in Example 1 above.

Forward and reverse PCR primers were synthesized based upon the assembly-created consensus sequence:

forward PCR primer 1 5'-GGGATGCAGGTGGTGTCTCATGGGG-3' (SEQ ID NO:346)
forward PCR primer 2 5'-CCCTCATGTACCGGCTCC-3' (SEQ ID NO:347)
forward PCR primer 3 5'-GTGTGACACAGCGTGGGC-3' (SEQ ID NO:43)
forward PCR primer 4 5'-GACCGGCAGGCTTCTGCG-3' (SEQ ID NO:44)
reverse PCR primer 1 5'-CAGCAGCTTCAGCCACCAGGAGTGG-3' (SEQ ID NO:45)
reverse PCR primer 2 5'-CTGAGCCGTGGGCTGCAGTCTCGC-3' (SEQ ID NO:46)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which had the following nucleotide sequence

hybridization probe

5'-CCGACTACGACTGGTTCTTCATCATGCAGGATGACACATATGTGC-3' (SEQ ID NO:47)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO339 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

A cDNA clone was sequenced in entirety. The entire nucleotide sequence of DNA43466-1225 is shown in Figure 117 (SEQ ID NO:338). Clone DNA43466-1225 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 333-335 and ending at the stop codon found at nucleotide positions 2649-2651 (Figure 117; SEQ ID NO:338). The predicted polypeptide precursor is 772 amino acids long and has a calculated molecular weight of approximately 86,226 daltons. Clone DNA43466-1225 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209490.

Based on a BLAST and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO339 has homology to *C. elegans* proteins and collagen-like polymer sequences as well as to fringe, thereby indicating that PRO339 may be involved in development or tissue growth.

EXAMPLE 51: Isolation of cDNA Clones Encoding Human PRO244

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. Based on this consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length

coding sequence for PRO244.

A pair of PCR primers (forward and reverse) were synthesized:

5'-TTCAGCTTCTGGGATGTAGGG-3' (30923.f1) (SEQ ID NO:378)

5'-TATTCCTACCATTTCACAAATCCG-3' (30923.r1) (SEQ ID NO:379)

A probe was also synthesized:

5'-GGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGG-3' (30923.p1) (SEQ IDNO:380)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO244 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from a human fetal kidney library. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence and the derived protein sequence for PRO244.

The entire nucleotide sequence of PRO244 is shown in Figure 121 (SEQ ID NO:376). Clone DNA35668-1171 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 106-108 (Fig. 121). The predicted polypeptide precursor is 219 amino acids long. Clone DNA35668-1171 has been deposited with ATCC (designated as DNA35663-1171) and is assigned ATCC deposit no. ATCC209371. The protein has a cytoplasmic domain (aa 1-20), a transmembrane domain (aa 21-46), and an extracellular domain (aa 47-219), with a C-lectin domain at aa 55-206.

Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO244 shows notable amino acid sequence identity to hepatic lectin gallus gallus (43%), HIC hp120-binding C-type lectin (42%), macrophage lectin 2 (HUMHML2-1, 41%), and sequence PR32188 (44%).

EXAMPLE 52: Use of PRO Polypeptide-Encoding Nucleic Acid as Hybridization Probes

The following method describes use of a nucleotide sequence encoding a PRO polypeptide as a hybridization probe.

DNA comprising the coding sequence of of a PRO polypeptide of interest as disclosed herein may be employed as a probe or used as a basis from which to prepare probes to screen for homologous DNAs (such as those encoding naturally-occurring variants of the PRO polypeptide) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO polypeptide-encoding nucleic acid-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO polypeptide can then be identified using standard techniques known in the art.

EXAMPLE 53: Expression of PRO Polypeptides in *E. coli*

This example illustrates preparation of an unglycosylated form of a desired PRO polypeptide by recombinant expression in *E. coli*.

The DNA sequence encoding the desired PRO polypeptide is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the specific PRO polypeptide coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO polypeptide can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

PRO187, PRO317, PRO301, PRO224 and PRO238 were successfully expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO187, PRO317, PRO301, PRO224 or PRO238 was initially amplified using selected PCR primers. The primers contained restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences were then ligated into an expression vector, which was used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants were first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 was reached. Cultures were then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g $(NH_4)_2SO_4$, 0.71 g sodium citrate·2H2O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM $MgSO_4$) and grown for approximately 20-30 hours at 30°C with shaking. Samples were removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets were frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) was resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution was stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution was centrifuged at 40,000 rpm in a Beckman Ultracentifuge for 30 min. The supernatant was diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. Depending the clarified extract was loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column was washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein was eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein were pooled and stored at 4°C. Protein concentration was estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins were refolded by diluting sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes were chosen so that the final protein concentration was between 50 to 100 micrograms/ml. The refolding solution was stirred gently at 4°C for 12-36 hours. The refolding reaction was quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution was filtered through a 0.22 micron filter and acetonitrile was added to 2-10% final concentration. The refolded protein was chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance were analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein were pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded PRO187, PRO317, PRO301, PRO224 and PRO238 proteins, respectively, were pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins were formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

EXAMPLE 54: Expression of PRO Polypeptides in Mammalian Cells

This example illustrates preparation of a glycosylated form of a desired PRO polypeptide by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO polypeptide-encoding DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO polypeptide DNA using ligation methods such as described in Sambrook et al., supra. The resulting vector is called pRK5-PRO polypeptide.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μg pRK5-PRO polypeptide DNA is mixed with about 1 μg DNA encoding the VA RNA gene [Thimmappaya et al., Cell, 31:543 (1982)] and dissolved in 500 μl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M $CaCl_2$. To this mixture is added, dropwise, 500 μl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM $NaPO_4$, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μCi/ml ^{35}S-cysteine and 200 μCi/ml ^{35}S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO polypeptide may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μg pRK5-PRO polypeptide DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μg/ml bovine insulin and 0.1 μg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO polypeptide can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO polypeptides can be expressed in CHO cells. The pRK5-PRO polypeptide can be transfected into CHO cells using known reagents such as $CaPO_4$ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO polypeptide can then be concentrated and purified by any selected method.

Epitope-tagged PRO polypeptide may also be expressed in host CHO cells. The PRO polypeptide may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO polypeptide insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO polypeptide can then be concentrated and purified by any selected method, such

as by Ni^{2+}-chelate affinity chromatography.

PRO211, PRO217, PRO230, PRO219, PRO245, PRO221, PRO258, PRO301, PRO224, PRO222, PRO234, PRO229, PRO223, PRO328 and PRO332 were successfully expressed in CHO cells by both a transient and a stable expression procedure. In addition, PRO232, PRO265, PRO246, PRO228, PRO227, PRO220, PRO266, PRO269, PRO287, PRO214, PRO231, PRO233, PRO238, PRO244, PRO235, PRO236, PRO262, PRO239, PRO257, PRO260, PRO263, PRO270, PRO271, PRO272, PRO294, PRO295, PRO293, PRO247, PRO303 and PRO268 were successfully transiently expressed in CHO cells.

Stable expression in CHO cells was performed using the following procedure. The proteins were expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins were fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs were subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., *Current Protocols of Molecular Biology*, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas *et al.*, *Nucl. Acids Res.* 24: 9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA were introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect® (Quiagen), Dosper® or Fugene® (Boehringer Mannheim). The cells were grown and described in Lucas *et al.*, supra. Approximately 3×10^{-7} cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA were thawed by placement into water bath and mixed by vortexing. The contents were pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant was aspirated and the cells were resuspended in 10 mL of selective media (0.2 μm filtered PS20 with 5% 0.2 μm diafiltered fetal bovine serum). The cells were then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells were transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, a 250 mL, 500 mL and 2000 mL spinners were seeded with 3×10^5 cells/mL. The cell media was exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in US Patent No. 5,122,469, issued June 16, 1992 was actually used. 3L production spinner is seeded at 1.2×10^6 cells/mL. On day 0, the cell number pH were determined. On day 1, the spinner was sampled and sparging with filtered air was commenced. On day 2, the spinner was sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion). Throughout the production, pH was adjusted as necessary to keep at around 7.2. After 10 days, or until viability dropped below 70%, the cell culture was harvested by centrifugtion and filtering through a 0.22 μm filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins were purified using a Ni-NTA column (Qiagen). Before purification, imidazole was added to the conditioned media to a concentration of 5 mM. The conditioned media was pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column was washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein was subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc containing) constructs of were purified from the conditioned media as follows. The conditioned medium was pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column was washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein was immediately neutralized by collecting 1 ml fractions into tubes containing 275 μL of 1 M Tris buffer, pH 9. The highly purified protein was subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity was assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

PRO211, PRO217, PRO230, PRO232, PRO187, PRO265, PRO219, PRO246, PRO228, PRO533, PRO245, PRO221, PRO227, PRO220, PRO258, PRO266, PRO269, PRO287, PRO214, PRO317, PRO301, PRO224, PRO222, PRO234, PRO231, PRO229, PRO233, PRO238, PRO223, PRO235, PRO236, PRO262, PRO239, PRO257, PRO260, PRO263, PRO270, PRO271, PRO272, PRO294, PRO295, PRO293, PRO247, PRO304, PRO302, PRO307, PRO303, PRO343, PRO328, PRO326, PRO331, PRO332, PRO334, PRO346, PRO268, PRO330, PRO310 and PRO339 were also successfully transiently expressed in COS cells.

EXAMPLE 55: Expression of PRO Polypeptides in Yeast

The following method describes recombinant expression of a desired PRO polypeptide in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO polypeptides from the ADH2/GAPDH promoter. DNA encoding a desired PRO polypeptide, a selected signal peptide and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of the PRO polypeptide. For secretion, DNA encoding the PRO polypeptide can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, the yeast alpha-factor secretory signal/leader sequence, and linker sequences (if needed) for expression of the PRO polypeptide.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO polypeptide can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing the PRO polypeptide may further be purified using selected column chromatography resins.

EXAMPLE 56: Expression of PRO Polypeptides in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO polypeptides in Baculovirus-infected insect cells.

The desired PRO polypeptide is fused upstream of an epitope tag contained with a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the PRO polypeptide or the desired portion of the PRO polypeptide (such as the sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4-5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression is performed as described by O'Reilley et al., Baculovirus expression vectors: A laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO polypeptide can then be purified, for example, by Ni^{2+}-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert *et al.*, *Nature*, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM $MgCl_2$; 0.1 mM EDTA; 10% Glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% Glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni^{2+}-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A_{280} with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% Glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or western blot with Ni^{2+}-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His_{10}-tagged PRO polypeptide are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO polypeptide can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

PRO211, PRO217, PRO230, PRO187, PRO265, PRO246, PRO228, PRO533, PRO245, PRO221, PRO220, PRO258, PRO266, PRO269, PRO287, PRO214, PRO301, PRO224, PRO222, PRO234, PRO231, PRO229, PRO235, PRO239, PRO257, PRO272, PRO294, PRO295, PRO328, PRO326, PRO331, PRO334, PRO346 and PRO310 were successfully expressed in baculovirus infected Sf9 or high5 insect cells. While the expression was actually performed in a 0.5-2 L scale, it can be readily scaled up for larger (e.g. 8 L)

preparations. The proteins were expressed as an IgG construct (immunoadhesin), in which the protein extracellular region was fused to an IgG1 constant region sequence containing the hinge, CH2 and CH3 domains and/or in poly-His tagged forms.

Following PCR amplification, the respective coding sequences were subcloned into a baculovirus expression vector (pb.PH.IgG for IgG fusions and pb.PH.His.c for poly-His tagged proteins), and the vector and Baculogold® baculovirus DNA (Pharmingen) were co-transfected into 105 *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711), using Lipofectin (Gibco BRL). pb.PH.IgG and pb.PH.His are modifications of the commercially available baculovirus expression vector pVL1393 (Pharmingen), with modified polylinker regions to include the His or Fc tag sequences. The cells were grown in Hink's TNM-FH medium supplemented with 10% FBS (Hyclone). Cells were incubated for 5 days at 28°C. The supernatant was harvested and subsequently used for the first viral amplification by infecting Sf9 cells in Hink's TNM-FH medium supplemented with 10% FBS at an approximate multiplicity of infection (MOI) of 10. Cells were incubated for 3 days at 28°C. The supernatant was harvested and the expression of the constructs in the baculovirus expression vector was determined by batch binding of 1 ml of supernatant to 25 mL of Ni-NTA beads (QIAGEN) for histidine tagged proteins or Protein-A Sepharose CL-4B beads (Pharmacia) for IgG tagged proteins followed by SDS-PAGE analysis comparing to a known concentration of protein standard by Coomassie blue staining.

The first viral amplification supernatant was used to infect a spinner culture (500 ml) of Sf9 cells grown in ESF-921 medium (Expression Systems LLC) at an approximate MOI of 0.1. Cells were incubated for 3 days at 28°C. The supernatant was harvested and filtered. Batch binding and SDS-PAGE analysis was repeated, as necessary, until expression of the spinner culture was confirmed.

The conditioned medium from the transfected cells (0.5 to 3 L) was harvested by centrifugation to remove the cells and filtered through 0.22 micron filters. For the poly-His tagged constructs, the protein construct were purified using a Ni-NTA column (Qiagen). Before purification, imidazole was added to the conditioned media to a concentration of 5 mM. The conditioned media were pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column was washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein was subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc containing) constructs of proteins were purified from the conditioned media as follows. The conditioned media were pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column was washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein was immediately neutralized by collecting 1 ml fractions into tubes containing 275 mL of 1 M Tris buffer, pH 9. The highly purified protein was subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity of the proteins was verified by SDS polyacrylamide gel (PEG) electrophoresis and N-terminal amino acid sequencing by Edman degradation.

EXAMPLE 57: Preparation of Antibodies that Bind to PRO Polypeptides

This example illustrates preparation of monoclonal antibodies which can specifically bind to a PRO polypeptide.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified PRO polypeptide, fusion proteins containing the PRO polypeptide, and cells expressing recombinant PRO polypeptide on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO polypeptide immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO polypeptide antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO polypeptide. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against the PRO polypeptide. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against the PRO polypeptide is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO polypeptide monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

EXAMPLE 58: Chimeric PRO Polypeptides

PRO polypeptides may be expressed as chimeric proteins with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS™ extension/affinity purification system (Immunex Corp., Seattle Wash.). The inclusion of a cleavable linker sequence such as Factor XA or enterokinase (Invitrogen, San Diego Calif.) between the

purification domain and the PRO polypeptide sequence may be useful to facilitate expression of DNA encoding the PRO polypeptide.

EXAMPLE 59: Purification of PRO Polypeptides Using Specific Antibodies

Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

EXAMPLE 60: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (I) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

EXAMPLE 61: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of an PRO polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above,

and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

EXAMPLE 62: Diagnostic Test Using PRO317 Polypeptide-Specific Antibodies

Particular anti-PRO317 polypeptide antibodies are useful for the diagnosis of prepathologic conditions, and chronic or acute diseases such as gynecological diseases or ischemic diseases which are characterized by differences in the amount or distribution of PRO317. PRO317 has been found to be expressed in human kidney and is thus likely to be associated with abnormalities or pathologies which affect this organ. Further, since it is so closely related to EBAF-1, it is likely to affect the endometrium and other genital tissues. Further, due to library sources of certain ESTs, it appears that PRO317 may be involved as well in forming blood vessels and hence to be a modulator of angiogenesis.

Diagnostic tests for PRO317 include methods utilizing the antibody and a label to detect PRO317 in human body fluids, tissues, or extracts of such tissues. The polypeptide and antibodies of the present invention may be used with or without modification. Frequently, the polypeptide and antibodies will be labeled by joining them, either covalently or noncovalently, with a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and have been reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent agents, chemiluminescent agents, magnetic particles, and the like. Patents teaching the use of such labels include U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced as shown in U.S. Pat. No. 4,816,567.

A variety of protocols for measuring soluble or membrane-bound PRo317, using either polyclonal or monoclonal antibodies specific for that PRO317, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), radioreceptor assay (RRA), and fluorescent activated cell sorting (FACS). A two-site monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on PRO317 is preferred, but a competitive binding assay may be employed. These assays are described, among other places, in Maddox *et al.* J Exp. Med., 158:1211 (1983).

EXAMPLE 63: Identification of PRO317 Receptors

Purified PRO317 is useful for characterization and purification of specific cell surface receptors and other binding molecules. Cells which respond to PRO317 by metabolic changes or other specific responses are

likely to express a receptor for PRO317. Such receptors include, but are not limited to, receptors associated with and activated by tyrosine and serine/threonine kinases. See Kolodziejczyk and Hall, *supra*, for a review on known receptors for the TGF- superfamily. Candidate receptors for this superfamily fall into two primary groups, termed type I and type II receptors. Both types are serine/threonine kinases. Upon activation by the appropriate ligand, type I and type II receptors physically interact to form hetero-oligomers and subsequently activate intracellular signaling cascades, ultimately regulating gene transcription and expression. In addition, TGF- binds to a third receptor class, type III, a membrane-anchored proteoglycan lacking the kinase activity typical of signal transducing molecules.

PRO317 receptors or other PRO317-binding molecules may be identified by interaction with radiolabeled PRO317. Radioactive labels may be incorporated into PRO317 by various methods known in the art. A preferred embodiment is the labeling of primary amino groups in PRO317 with ^{125}I Bolton-Hunter reagent (Bolton and Hunter, Biochem. J., 133:529 (1973)), which has been used to label other polypeptides without concomitant loss of biological activity (Hebert *et al.*, J. Biol. Chem., 266:18989 (1991); McColl *et al.*, J. Immunol., 150:4550-4555 (1993)). Receptor-bearing cells are incubated with labeled PRO317. The cells are then washed to removed unbound PRO317, and receptor-bound PRO317 is quantified. The data obtained using different concentrations of PRO317 are used to calculate values for the number and affinity of receptors.

Labeled PRO317 is useful as a reagent for purification of its specific receptor. In one embodiment of affinity purification, PRO317 is covalently coupled to a chromatography column. Receptor-bearing cells are extracted, and the extract is passed over the column. The receptor binds to the column by virtue of its biological affinity for PRO317. The receptor is recovered from the column and subjected to N-terminal protein sequencing. This amino acid sequence is then used to design degenerate oligonucleotide probes for cloning the receptor gene.

In an alternative method, mRNA is obtained from receptor-bearing cells and made into a cDNA library. The library is transfected into a population of cells, and those cells expressing the receptor are selected using fluorescently labeled PRO317. The receptor is identified by recovering and sequencing recombinant DNA from highly labeled cells.

In another alternative method, antibodies are raised against the surface of receptor bearing cells, specifically monoclonal antibodies. The monoclonal antibodies are screened to identify those which inhibit the binding of labeled PRO317. These monoclonal antibodies are then used in affinity purification or expression cloning of the receptor.

Soluble receptors or other soluble binding molecules are identified in a similar manner. Labeled PRO317 is incubated with extracts or other appropriate materials derived from the uterus. After incubation, PRO317 complexes larger than the size of purified PRO317 are identified by a sizing technique such as size-exclusion chromatography or density gradient centrifugation and are purified by methods known in the art. The soluble receptors or binding protein(s) are subjected to N-terminal sequencing to obtain information sufficient for database identification, if the soluble protein is known, or for cloning, if the soluble protein is unknown.

EXAMPLE 64: Determination of PRO317-Induced Cellular Response

The biological activity of PRO317 is measured, for example, by binding of an PRO317 of the invention

to an PRO317 receptor. A test compound is screened as an antagonist for its ability to block binding of PRO317 to the receptor. A test compound is screened as an agonist of the PRO317 for its ability to bind an PRO317 receptor and influence the same physiological events as PRO317 using, for example, the KIRA-ELISA assay described by Sadick *et al.*, Analytical Biochemistry, 235:207-214 (1996) in which activation of a receptor tyrosine kinase is monitored by immuno-capture of the activated receptor and quantitation of the level of ligand-induced phosphorylation. The assay may be adapted to monitor PRO317-induced receptor activation through the use of an PRO317 receptor-specific antibody to capture the activated receptor. These techniques are also applicable to other PRO polypeptides described herein.

EXAMPLE 65: Use of PRO224 for Screening Compounds

PRO224 is expressed in a cell stripped of membrane proteins and capable of expressing PRO224. Low density lipoproteins having a detectable label are added to the cells and incubated for a sufficient time for endocytosis. The cells are washed. The cells are then analysed for label bound to the membrane and within the cell after cell lysis. Detection of the low density lipoproteins within the cell determines that PRO224 is within the family of low density lipoprotein receptor proteins. Members found within this family are then used for screening compounds which affect these receptors, and particularly the uptake of cholesterol via these receptors.

EXAMPLE 66: Ability of PRO Polypeptides to Inhibit Vascular Endothelial Growth Factor (VEGF) Stimulated Proliferation of Endothelial Cell Growth

The ability of various PRO polypeptides to inhibit VEGF stimulated proliferation of endothelial cells was tested. Specifically, bovine adrenal cortical capillary endothelial (ACE) cells (from primary culture, maximum 12-14 passages) were plated on 96-well microtiter plates (Amersham Life Science) at a density of 500 cells/well per 100 μL in low glucose DMEM, 10% calf serum, 2 mM glutamine, 1x pen/strept and fungizone, supplemented with 3 ng/mL VEGF. Controls were plated the same way but some did not include VEGF. A test sample of the PRO polypeptide of interest was added in a 100 μl volume for a 200 μl final volume. Cells were incubated for 6-7 days at 37°C. The media was aspirated and the cells washed 1x with PBS. An acid phosphatase reaction mixture (100 μL, 0.1M sodium acetate, pH 5.5, 0.1% Triton-100, 10 mM p-nitrophenyl phosphate) was added. After incubation for 2 hours at 37°C, the reaction was stopped by addition of 10 μl 1N NaOH. OD was measured on microtiter plate reader at 405 nm. Controls were no cells, cells alone, cells + FGF (5 ng/mL), cells + VEGF (3 ng/mL), cells + VEGF (3 ng/ml) + TGF-β (1 ng/ml), and cells + VEGF (3ng/mL) + LIF (5 ng/mL). (TGF-β at a 1 ng/ml concentration is known to block 70-90% of VEGF stimulated cell proliferation.)

The results were assessed by calculating the percentage inhibition of VEGF (3 ng/ml) stimulated cells proliferation, determined by measuring acid phosphatase activity at OD405 nm, (1) relative to cells without stimulation, and (2) relative to the reference TGF-β inhibition of VEGF stimulated activity. The results, as shown in Table 2 below, are indicative of the utility of the PRO polypeptides in cancer therapy and specifically in inhibiting tumor angiogenesis. The numerical values (relative inhibition) shown in Table 2 are determined by calculating the percent inhibition of VEGF stimulated proliferation by the PRO polypeptide relative to cells

without stimulation and then dividing that percentage into the percent inhibition obtained by TGF-β at 1 ng/ml which is known to block 70-90% of VEGF stimulated cell proliferation.

Table 2

PRO Name	PRO Concentration	Relative Inhibition
PRO211	0.01%	99.0
PRO211	0.01%	1.09
PRO211	0.1%	0.95
PRO211	0.1%	67.0
PRO211	1.0%	0.27
PRO211	1.0%	20.0
PRO217	0.01%	1.06
PRO217	0.1%	0.84
PRO217	1.0%	0.39
PRO217	2.5 μM	0.2
PRO217	25 nM	0.88
PRO217	250 nM	0.58
PRO187	0.01%	0.91
PRO187	0.1%	0.82
PRO187	1.0%	0.44
PRO219	5.7 μM	0.61
PRO219	57 nM	1.09
PRO219	570 nM	0.97
PRO246	0.01%	1.04
PRO246	0.1%	1.0
PRO246	1.0%	0.49
PRO228	0.01%	0.99
PRO228	0.1%	0.93
PRO228	1.0%	0.57
PRO228	0.01%	0.95
PRO228	0.01%	0.98
PRO228	0.1%	0.77
PRO228	0.1%	0.88
PRO228	1.0%	0.16
PRO228	1.0%	0.48
PRO245	0.01%	0.76
PRO245	0.1%	0.35
PRO245	1.0%	0.11
PRO245	0.48 nM	1.03
PRO245	4.8 nM	0.95
PRO245	48 nM	0.49
PRO221	0.01%	1.03
PRO221	0.01%	1.06
PRO221	0.1%	0.82
PRO221	0.1%	0.93
PRO221	1.0%	0.31
PRO221	1.0%	0.43
PRO258	0.01%	0.98
PRO258	0.01%	1.06
PRO258	0.1%	0.95
PRO258	0.1%	1.02
PRO258	1.0%	0.6
PRO258	1.0%	0.69

PAGE INTENTIONALLY LEFT BLANK

Table 2 cont'

PRO Name	PRO Concentration	Relative Inhibition
PRO301	7.0 μM	1.02
PRO301	70 μM	0.88
PRO301	700 μM	0.44
PRO301	0.01%	0.92
PRO301	0.1%	0.85
PRO301	1.0%	0.68
PRO224	0.01%	101.0
PRO224	0.1%	65.0
PRO224	1.0%	23.0
PRO272	0.01%	0.95
PRO272	0.1%	0.57
PRO272	1.0%	0.18
PRO328	0.01%	0.98
PRO328	0.1%	0.96
PRO328	1.0%	0.6
PRO331	0.01%	0.88
PRO331	0.1%	0.82
PRO331	1.0%	0.56

EXAMPLE 67: Retinal Neuron Survival

This example demonstrates that PRO220 polypeptides have efficacy in enhancing the survival of retinal neuron cells.

Sprague Dawley rat pups at postnatal day 7 (mixed population: glia and retinal neuronal types) are killed by decapitation following CO_2 anesthesia and the eyes are removed under sterile conditions. The neural retina is dissected away from the pigment epithelium and other ocular tissue and then dissociated into a single cell suspension using 0.25% trypsin in Ca^{2+}, Mg^{2+}-free PBS. The retinas are incubated at 37°C for 7-10 minutes after which the trypsin is inactivated by adding 1 ml soybean trypsin inhibitor. The cells are plated at 100,000 cells per well in 96 well plates in DMEM/F12 supplemented with N2 and with or without the specific test PRO polypeptide. Cells for all experiments are grown at 37°C in a water saturated atmosphere of 5% CO_2. After 2-3 days in culture, cells are stained with calcein AM then fixed using 4% paraformaldehyde and stained with DAPI for determination of total cell count. The total cells (fluorescent) are quantified at 20X objective magnification using CCD camera and NIH image software for Macintosh. Fields in the well are chosen at random.

The effect of various concentration of PRO220 polypeptides are reported in Table 3 below where percent survival is calculated by dividing the total number of calcein AM positive cells at 2-3 days in culture by the total number of DAPI-labeled cells at 2-3 days in culture. Anything above 30% survival is considered positive.

Table 3

PRO Name	PRO Concentration	Percent Survival
PRO220	0.01%	2.4%
PRO220	0.01%	4.1%
PRO220	0.1%	3.0%
PRO220	0.1%	3.1%
PRO220	1.0%	72.4%
PRO220	1.0%	42.1%

EXAMPLE 68: Rod Photoreceptor Survival

This example demonstrates that PRO220 polypeptides have efficacy in enhancing the survival of rod photoreceptor cells.

Sprague Dawley rat pups at 7 day postnatal (mixed population: glia and retinal neuronal cell types) are killed by decapitation following CO_2 anesthesis and the eyes are removed under sterile conditions. The neural retina is dissected away form the pigment epithelium and other ocular tissue and then dissociated into a single cell suspension using 0.25% trypsin in Ca^{2+}, Mg^{2+}-free PBS. The retinas are incubated at 37°C for 7-10 minutes after which the trypsin is inactivated by adding 1 ml soybean trypsin inhibitor. The cells are plated at 100,000 cells per well in 96 well plates in DMEM/F12 supplemented with N2 and with or without the specific test PRO polypeptide. Cells for all experiments are grown at 37°C in a water saturated atmosphere of 5% CO2. After 2-3 days in culture, cells are fixed using 4% paraformaldehyde, and then stained using CellTracker Green CMFDA. Rho 4D2 (ascites or IgG 1:100), a monoclonal antibody directed towards the visual pigment rhodopsin is used to detect rod photoreceptor cells by indirect immunofluorescence. The results are reported as % survival: total number of calcein/CellTracker - rhodopsin positive cells at 2-3 days in culture, divided by the total number of rhodopsin positive cells at time 2-3 days in culture. The total cells (fluorescent) are quantified at 20x objective magnification using a CCD camera and NIH image software for MacIntosh. Fields in the well are chosen at random.

The effect of various concentration of PRO220 polypeptides are reported in Table 4 below. Anything above 10% survival is considered positive..

Table 4

PRO Name	PRO Concentration	Percent Survival
PRO220	0.01%	0.0%
PRO220	0.1%	0.0%
PRO220	2.0%	0.0%
PRO220	10%	0.0%
PRO220	20%	66.9%
PRO220	1.0%	56.9%

EXAMPLE 69: Induction of Endothelial Cell Apoptosis

The ability of PRO228 polypeptides to induce apoptosis in endothelial cells was tested in human venous umbilical vein endothelial cells (HUVEC, Cell Systems), using a 96-well format, in 0% serum media supplemented with 100 ng/ml VEGF. (As HUVEC cells are easily dislodged from the plating surface, all pipetting in the wells must be done as gently as practicable.)

The media was aspired and the cells washed once with PBS. 5 ml of 1 x trypsin was added to the cells in a T-175 flask, and the cells were allowed to stand until they were released from the plate (about 5-10 minutes). Trypsinization was stopped by adding 5 ml of growth media. The cells were spun at 1000 rpm for 5 minutes at 4 °C. The media was aspirated and the cells were resuspended in 10 ml of 10% serum complemented medium (Cell Systems), 1 x penn/strep.

The cells were plated on 96-well microtiter plates (Amersham Life Science, cytostar-T scintillating microplate, RPNQ160, sterile, tissue-culture treated, individually wrapped), in 10% serum (CSG-medium, Cell Systems), at a density of 2×10^4 cells per well in a total volume of 100 μl. The PRO228 polypeptide was added in triplicate at dilutions of 1%, 0.33% and 0.11%. Wells without cells were used as a blank and wells with cells only as a negative control. As a positive control 1:3 serial dilutions of 50 μl of a 3x stock of staurosporine were used. The ability of the PRO228 polypeptide to induce apoptosis was determined using Annexin V, a member of the calcium and phospholipid binding proteins, to detect apoptosis.

0.2 ml Annexin V - Biotin stock solution (100 μg/ml) were diluted in 4.6 ml 2 x Ca^{2+} binding buffer and 2.5% BSA (1:25 dilution). 50 μls of the diluted Annexin V - Biotin solution were added to each well (except controls) to a final concentration of 1.0 μg/ml. The samples were incubated for 10-15 minutes with Annexin-Biotin prior to direct addition of ^{35}S-Streptavidin. ^{35}S-Streptavidin was diluted in 2x Ca^{2+} binding buffer, 2.5% BSA and was added to all wells at a final concentration of 3×10^6 cpm/well. The plates were then sealed, centrifuged at 1000 rpm for 15 minutes and placed on orbital shaker for 2 hours. The analysis was performed on 1450 Microbeta Trilux (Wallac). The results are shown in Table 5 below where percent above background represents the percentage amount of counts per minute above the negative controls. Percents greater than or equal to 30% above background are considered positive.

Table 5

PRO Name	PRO Concentration	Percent Above Background
PRO228	0.11%	0.7%
PRO228	0.11%	47.6%
PRO228	0.33%	92.2%
PRO228	0.33%	123.7%
PRO228	1.0%	51.4%
PRO228	1.0%	95.3%

EXAMPLE 70: PDB12 Cell Inhibition

This example demonstrates that various PRO polypeptides have efficacy in inhibiting protein production by PDB12 pancreatic ductal cells.

PDB12 pancreatic ductal cells are plated on fibronectin coated 96 well plates at 1.5×10^4 cells per well in 100 μL/180 μL of growth media. 100 μL of growth media with the PRO polypeptide test sample or negative control lacking the PRO polypeptide is then added to well, for a final volume of 200 μL. Controls contain growth medium containing a protein shown to be inactive in this assay. Cells are incubated for 4 days at 37°C. 20 μL of Alamar Blue Dye (AB) is then added to each well and the flourescent reading is measured at 4 hours post addition of AB, on a microtiter plate reader at 530 nm excitation and 590 nm emission. The standard employed is cells without Bovine Pituitary Extract (BPE) and with various concentrations of BPE. Buffer or CM controls

from unknowns are run 2 times on each 96 well plate.

The results from these assays are shown in Table 6 below wherein percent decrease in protein production is calculated by comparing the Alamar Blue Dye calculated protein concentration produced by the PRO polypeptide-treated cells with the Alamar Blue Dye calculated protein concentration produced by the negative control cells. A percent decrease in protein production of greater than or equal to 25% as compared to the negative control cells is considered positive.

Table 6

PRO Name	PRO Concentration	Percent Decrease in Protein Production
PRO211	0.1%	0.0%
PRO211	0.01%	0.6%
PRO211	1.0%	59.7%
PRO287	2.0%	22.3%
PRO287	10%	18.2%
PRO287	50%	67.5%
PRO287	2.0%	45.53%
PRO287	10%	57.3%
PRO287	50%	52.24%
PRO301	2.0%	0.0%
PRO301	10%	59.8%
PRO301	50%	65.6%
PRO293	2.0%	0.0%
PRO293	10%	40.4%
PRO293	50%	56.7%

EXAMPLE 71: Stimulation of Adult Heart Hypertrophy

This assay is designed to measure the ability of various PRO polypeptides to stimulate hypertrophy of adult heart.

Ventricular myocytes freshly isolated from adult (250g) Sprague Dawley rats are plated at 2000 cell/well in 180 μl volume. Cells are isolated and plated on day 1, the PRO polypeptide-containing test samples or growth medium only (negative control) (20 μl volume) is added on day 2 and the cells are then fixed and stained on day 5. After staining, cell size is visualized wherein cells showing no growth enhancement as compared to control cells are given a value of 0.0, cells showing small to moderate growth enhancement as compared to control cells are given a value of 1.0 and cells showing large growth enhancement as compared to control cells are given a value of 2.0. Any degree of growth enhancement as compared to the negative control cells is considered positive for the assay. The results are shown in Table 7 below.

Table 7

PRO Name	PRO Concentration	Growth Enhancement Score
PRO287	20%	1.0
PRO287	20%	1.0
PRO301	20%	1.0
PRO301	20%	1.0
PRO293	20%	1.0
PRO293	20%	1.0
PRO303	20%	1.0
PRO303	20%	1.0

EXAMPLE 72: PDB12 Cell Proliferation

This example demonstrates that various PRO polypeptides have efficacy in inducing proliferation of PDB12 pancreatic ductal cells.

PDB12 pancreatic ductal cells are plated on fibronectin coated 96 well plates at $1.5x10^3$ cells per well in 100 µL/180 µL of growth media. 100 µL of growth media with the PRO polypeptide test sample or negative control lacking the PRO polypeptide is then added to well, for a final volume of 200 µL. Controls contain growth medium containing a protein shown to be inactive in this assay. Cells are incubated for 4 days at 37°C. 20 µL of Alamar Blue Dye (AB) is then added to each well and the flouresecent reading is measured at 4 hours post addition of AB, on a microtiter plate reader at 530 nm excitation and 590 nm emission. The standard employed is cells without Bovine Pituitary Extract (BPE) and with various concentrations of BPE. Buffer or growth medium only controls from unknowns are run 2 times on each 96 well plate.

The results from these assays are shown in Table 8 below wherein percent increase in protein production is calculated by comparing the Alamar Blue Dye calculated protein concentration produced by the PRO polypeptide-treated cells with the Alamar Blue Dye calculated protein concentration produced by the negative control cells. A percent increase in protein production of greater than or equal to 25% as compared to the negative control cells is considered positive.

Table 8

PRO Name	PRO Concentration	Percent Increase in Protein Production
PRO301	2.0%	44.0%
PRO301	10%	67.4%
PRO301	50%	185.8%
PRO303	2.0%	27.9%
PRO303	10%	174.9%
PRO303	50%	193.1%

EXAMPLE 73: Enhancement of Heart Neonatal Hypertrophy Induced by PRO224

This assay is designed to measure the ability of PRO224 polypeptides to stimulate hypertrophy of neonatal heart.

Cardiac myocytes from 1-day old Harlan Sprague Dawley rats were obtained. Cells (180 µl at 7.5 x 10^5/ml, serum < 0.1%, freshly isolated) are added on day 1 to 96-well plates previously coated with DMEM/F12 + 4% FCS. Test samples containing the test PRO224 polypeptide or growth medium only (negative control) (20 µl/well) are added directly to the wells on day 1. PGF (20 µl/well) is then added on day 2 at final concentration of 10^{-6} M. The cells are then stained on day 4 and visually scored on day 5, wherein cells showing no increase in size as compared to negative controls are scored 0.0, cells showing a small to moderate increase in size as compared to negative controls are scored 1.0 and cells showing a large increase in size as compared to negative controls are scored 2.0. The results are shown in Table 9 below.

Table 9

PRO Name	PRO Concentration	Growth Enhancement Score
PRO224	0.01%	0.0
PRO224	0.1%	0.0
PRO224	1.0%	1.0

EXAMPLE 74: *In situ* Hybridization

In situ hybridization is a powerful and versatile technique for the detection and localization of nucleic acid sequences within cell or tissue preparations. It may be useful, for example, to identify sites of gene expression, analyze the tissue distribution of transcription, identify and localize viral infection, follow changes in specific mRNA synthesis and aid in chromosome mapping.

In situ hybridization was performed following an optimized version of the protocol by Lu and Gillett, Cell Vision 1:169-176 (1994), using PCR-generated ^{33}P-labeled riboprobes. Briefly, formalin-fixed, paraffin-embedded human tissues were sectioned, deparaffinized, deproteinated in proteinase K (20 g/ml) for 15 minutes at 37°C, and further processed for *in situ* hybridization as described by Lu and Gillett, *supra*. A [^{33}-P] UTP-labeled antisense riboprobe was generated from a PCR product and hybridized at 55°C overnight. The slides were dipped in Kodak NTB2 nuclear track emulsion and exposed for 4 weeks.

^{33}P-Riboprobe synthesis

6.0 μl (125 mCi) of ^{33}P-UTP (Amersham BF 1002, SA < 2000 Ci/mmol) were speed vac dried. To each tube containing dried ^{33}P-UTP, the following ingredients were added:

2.0 μl 5x transcription buffer

1.0 μl DTT (100 mM)

2.0 μl NTP mix (2.5 mM : 10 μl each of 10 mM GTP, CTP & ATP + 10 μl H_2O)

1.0 μl UTP (50 μM)

1.0 μl Rnasin

1.0 μl DNA template (1 μg)

1.0 μl H_2O

1.0 μl RNA polymerase (for PCR products T3 = AS, T7 = S, usually)

The tubes were incubated at 37°C for one hour. 1.0 μl RQ1 DNase were added, followed by incubation at 37°C for 15 minutes. 90 μl TE (10 mM Tris pH 7.6/1mM EDTA pH 8.0) were added, and the mixture was pipetted onto DE81 paper. The remaining solution was loaded in a Microcon-50 ultrafiltration unit, and spun using program 10 (6 minutes). The filtration unit was inverted over a second tube and spun using program 2 (3 minutes). After the final recovery spin, 100 μl TE were added. 1 μl of the final product was pipetted on DE81 paper and counted in 6 ml of Biofluor II.

The probe was run on a TBE/urea gel. 1-3 μl of the probe or 5 μl of RNA Mrk III were added to 3 μl of loading buffer. After heating on a 95°C heat block for three minutes, the gel was immediately placed on ice. The wells of gel were flushed, the sample loaded, and run at 180-250 volts for 45 minutes. The gel was wrapped in saran wrap and exposed to XAR film with an intensifying screen in -70°C freezer one hour to overnight.

^{33}P-Hybridization

A. Pretreatment of frozen sections

The slides were removed from the freezer, placed on aluminium trays and thawed at room temperature for 5 minutes. The trays were placed in 55°C incubator for five minutes to reduce condensation. The slides were fixed for 10 minutes in 4% paraformaldehyde on ice in the fume hood, and washed in 0.5 x SSC for 5 minutes, at room temperature (25 ml 20 x SSC + 975 ml SQ H_2O). After deproteination in 0.5 μg/ml proteinase K for 10 minutes at 37°C (12.5 μl of 10 mg/ml stock in 250 ml prewarmed RNase-free RNAse buffer), the sections were washed in 0.5 x SSC for 10 minutes at room temperature. The sections were dehydrated in 70%, 95%, 100% ethanol, 2 minutes each.

B. Pretreatment of paraffin-embedded sections

The slides were deparaffinized, placed in SQ H_2O, and rinsed twice in 2 x SSC at room temperature, for 5 minutes each time. The sections were deproteinated in 20 μg/ml proteinase K (500 μl of 10 mg/ml in 250 ml RNase-free RNase buffer; 37°C, 15 minutes) - human embryo, or 8 x proteinase K (100 μl in 250 ml Rnase buffer, 37°C, 30 minutes) - formalin tissues. Subsequent rinsing in 0.5 x SSC and dehydration were performed as described above.

C. Prehybridization

The slides were laid out in a plastic box lined with Box buffer (4 x SSC, 50% formamide) - saturated filter paper. The tissue was covered with 50 μl of hybridization buffer (3.75g Dextran Sulfate + 6 ml SQ H_2O), vortexed and heated in the microwave for 2 minutes with the cap loosened. After cooling on ice, 18.75 ml formamide, 3.75 ml 20 x SSC and 9 ml SQ H_2O were added, the tissue was vortexed well, and incubated at 42°C for 1-4 hours.

D. Hybridization

1.0 x 10^6 cpm probe and 1.0 μl tRNA (50 mg/ml stock) per slide were heated at 95°C for 3 minutes. The slides were cooled on ice, and 48 μl hybridization buffer were added per slide. After vortexing, 50 μl ^{33}P mix were added to 50 μl prehybridization on slide. The slides were incubated overnight at 55°C.

E. Washes

Washing was done 2 x 10 minutes with 2xSSC, EDTA at room temperature (400 ml 20 x SSC + 16 ml 0.25M EDTA, V_f=4L), followed by RNaseA treatment at 37°C for 30 minutes (500 μl of 10 mg/ml in 250 ml Rnase buffer = 20 μg/ml). The slides were washed 2 x 10 minutes with 2 x SSC, EDTA at room temperature. The stringency wash conditions were as follows: 2 hours at 55°C, 0.1 x SSC, EDTA (20 ml 20 x SSC + 16 ml EDTA, V_f=4L).

F. Oligonucleotides

In situ analysis was performed on a variety of DNA sequences disclosed herein. The oligonucleotides employed for these analyses are as follows.

(1) DNA33094-1131 (PRO217)

p1 5'-GGATTCTAATACGACTCACTATAGGGCTCAGAAAAGCGCAACAGAGAA-3' (SEQ ID NO:348)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGATGTCTTCCATGCCAACCTTC-3(SEQ ID NO:349)

(2) DNA33223-1136 (PRO230)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGGCGATGTCCACTGGGGCTAC-3' (SEQ ID NO:350)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACGAGGAAGATGGGCGGATGGT-3' (SEQ ID NO:351)

(3) DNA34435-1140 (PRO232)

p1 5'-GGATTCTAATACGACTCACTATAGGGCACCCACGCGTCCGGCTGCTT-3(SEQ ID NO:352)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACGGGGGACACCACGGACCAGA-3' (SEQ ID NO:353)

(4) DNA35639-1172 (PRO246)

p1 5'-GGATTCTAATACGACTCACTATAGGGCTTGCTGCGGTTTTTGTTCCTG-3(SEQ ID NO:354)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGCTGCCGATCCCACTGGTATT-3' (SEQ ID NO:355)

(5) DNA49435-1219 (PRO533)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGGATCCTGGCCGGCCTCTG-3' (SEQ ID NO:356)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGCCCGGGCATGGTCTCAGTTA-3' (SEQ ID NO:357)

(6) DNA35638-1141 (PRO245)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGGGAAGATGGCGAGGAGGAG-3(SEQ ID NO:358)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACCAAGGCCACAAACGGAAATC-3' (SEQ ID NO:359)

(7) DNA33089-1132 (PRO221)

p1 5'-GGATTCTAATACGACTCACTATAGGGCTGTGCTTTCATTCTGCCAGTA-3(SEQ ID NO:360)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGGGTACAATTAAGGGGTGGAT-3' (SEQ ID NO:361)

(8) DNA35918-1174 (PRO258)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCCGCCTCGCTCCTGCTCCTG-3(SEQ ID NO:362)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGGATTGCCGCGACCCTCACAG-3' (SEQ ID NO:363)

(9) DNA32286-1191 (PRO214)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCCCTCCTGCCTTCCCTGTCC-3'(SEQ ID NO:364)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGTGGTGGCCGCGATTATCTGC-3' (SEQ ID NO:365)

(10) DNA33221-1133 (PRO224)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGCAGCGATGGCAGCGATGAGG-3' (SEQ ID NO:366)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACAGACGGGGCAGAGGGAGTG-3'(SEQ ID NO:367)

(11) DNA35557-1137 (PRO234)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCAGGAGGCGTGAGGAGAAAC-3'(SEQ ID NO:368)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAAAGACATGTCATCGGGAGTGG-3' (SEQ ID NO:369)

(12) DNA33100-1159 (PRO229)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCGGGTGGAGGTGGAACAGAAA-3' (SEQ ID NO:370)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACACAGACAGAGCCCCATACGC-3' (SEQ ID NO:371)

(13) DNA34431-1177 (PRO263)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCAGGGAAATCCGGATGTCTC-3(SEQ ID NO:372)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGTAAGGGGATGCCACCGAGTA-3' (SEQ ID NO:373)

(14) DNA38268-1188 (PRO295)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCAGCTACCCGCAGGAGGAGG-3'(SEQ ID NO:374)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGATCCCAGGTGATGAGGTCCAGA-3' (SEQ ID NO:375)

G. Results

In situ analysis was performed on a variety of DNA sequences disclosed herein. The results from these analyses are as follows.

(1) DNA33094-1131 (PRO217)

Highly distinctive expression pattern, that does not indicate an obvious biological function. In the human embryo it was expressed in outer smooth muscle layer of the GI tract, respiratory cartilage, branching respiratory epithelium, osteoblasts, tendons, gonad, in the optic nerve head and developing dermis. In the adult expression was observed in the epidermal pegs of the chimp tongue, the basal epithelial/myoepithelial cells of the prostate and urinary bladder. Also expressed in the alveolar lining cells of the adult lung, mesenchymal cells juxtaposed

to erectile tissue in the penis and the cerebral cortex (probably glial cells). In the kidney, expression was only seen in disease, in cells surrounding thyroidized renal tubules.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, myocardium, aorta, spleen, lymph node, gall bladder, pancreas, lung, skin, eye (inc. retina), prostate, bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

(a) Chimp Tissues: Salivary gland, stomach, thyroid, parathyroid, skin, thymus, ovary, lymph node.

(b) Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum, penis.

(2) DNA33223-1136 (PRO230)

Sections show an intense signal associated with arterial and venous vessels in the fetus. In arteries the signal appeared to be confined to smooth-muscle/pericytic cells. The signal is also seen in capillary vessels and in glomeruli. It is not clear whether or not endothelial cells are expressing this mRNA. Expression is also observed in epithelial cells in the fetal lens. Strong expression was also seen in cells within placental trophoblastic villi, these cells lie between the trophoblast and the fibroblast-like cells that express HGF - uncertain histogenesis. In the adult, there was no evidence of expression and the wall of the aorta and most vessels appear to be negative. However, expression was seen over vascular channels in the normal prostate and in the epithelium lining the gallbladder. Insurers expression was seen in the vessels of the soft-tissue sarcoma and a renal cell carcinoma. In summary, this is a molecule that shows relatively specific vascular expression in the fetus as well as in some adult organs. Expression was also observed in the fetal lens and the adult gallbladder.

In a secondary screen, vascular expression was observed, similar to that observed above, seen in fetal blocks. Expression is on vascular smooth muscle, rather than endothelium. Expression also seen in smooth muscle of the developing oesophagus, so as reported previously, this molecule is not vascular specific. Expression was examined in 4 lung and 4 breast carcinomas. Substantial expression was seen in vascular smooth muscle of at least 3/4 lung cancers and 2/4 breast cancers. In addition, in one breast carcinoma, expression was observed in peritumoral stromal cells of uncertain histogenesis (possibly myofibroblasts). No endothelial cell expression was observed in this study.

(3) DNA34435-1140 (PRO232)

Strong expression in prostatic epithelium and bladder epithelium, lower level of expression in bronchial epithelium. High background / low level expression seen in a number of sites, including among others: bone, blood, chondrosarcoma, adult heart and fetal liver. It is felt that this level of signal represents background, partly because signal at this level was seen over the blood. All other tissues negative.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

Rhesus Monkey Tissues: Cerebral cortex, hippocampus

In a secondary screen, expression was observed in the epithelium of the prostate, the superficial layers of the urethelium of the urinary bladder, the urethelium lining the renal pelvis and the urethelium of the ureter (1 out of 2 experiments). The urethra of a rhesus monkey was negative; it is unclear whether this represents a true lack of expression by the urethra, or if it is the result of a failure of the probe to cross react with rhesus tissue. The findings in the prostate and bladder are similar to those previously described using an isotopic detection technique. Expression of the mRNA for this antigen is NOT prostate epithelial specific. The antigen may serve as a useful marker for urethelial derived tissues. Expression in the superficial, post-mitotic cells, of the urinary tract epithelium also suggest that it is unlikely to represent a specific stem cell marker, as this would be expected to be expressed specifically in basal epithelium.

(4) DNA35639-1172 (PRO246)

Strongly expressed in fetal vascular endothelium, including tissues of the CNS. Lower level of expression in adult vasculature, including the CNS. Not obviously expressed at higher levels in tumor vascular endothelium. Signal also seen over bone matrix and adult spleen, not obviously cell associated, probably related to non-specific background at these sites.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

Rhesus Monkey Tissues: Cerebral cortex, hippocampus

(5) DNA49435-1219 (PRO533)

Moderate expression over cortical neurones in the fetal brain. Expression over the inner aspect of the fetal retina, possible expression in the developing lens. Expression over fetal skin, cartilage, small intestine, placental villi and umbilical cord. In adult tissues there is an extremely high level of expression over the gallbladder epithelium. Moderate expression over the adult kidney, gastric and colonic epithelia. Low-level expression was observed over many cell types in many tissues, this may be related to stickiness of the probe,

these data should therefore be interpreted with a degree of caution.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum.

(6) DNA35638-1141 (PRO245)

Expression observed in the endothelium lining a subset of fetal and placental vessels. Endothelial expression was confined to these tissue blocks. Expression also observed over intermediate trophoblast cells of placenta. All other tissues negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), penis, eye, bladder, stomach, gastric carcinoma, colon, colonic carcinoma, thyroid (chimp), parathyroid (chimp) ovary (chimp) and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis

(7) DNA33089-1132 (PRO221)

Specific expression over fetal cerebral white and grey matter, as well as over neurones in the spinal cord. Probe appears to cross react with rat. Low level of expression over cerebellar neurones in adult rhesus brain. All other tissues negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), penis, eye, bladder, stomach, gastric carcinoma, colon, colonic carcinoma and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis

(8) DNA35918-1174 (PRO258)

Strong expression in the nervous system. In the rhesus monkey brain expression is observed in cortical, hippocampal and cerebellar neurones. Expression over spinal neurones in the fetal spinal cord, the developing

brain and the inner aspects of the fetal retina. Expression over developing dorsal root and autonomic ganglia as well as enteric nerves. Expression observed over ganglion cells in the adult prostate. In the rat, there is strong expression over the developing hind brain and spinal cord. Strong expression over interstitial cells in the placental villi. All other tissues were negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, renal cell carcinoma, adrenal, aorta, spleen, lymph node, pancreas, lung, myocardium, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), bladder, prostate, stomach, gastric carcinoma, colon, colonic carcinoma, thyroid (chimp), parathyroid (chimp) ovary (chimp) and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis.

(9) DNA32286-1191 (PRO214)

Fetal tissue: Low level throughout mesenchyme. Moderate expression in placental stromal cells in membranous tissues and in thyroid. Low level expression in cortical neurones. Adult tissue: all negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined include: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung and skin.

(10) DNA33221-1133 (PRO224)

Expression limited to vascular endothelium in fetal spleen, adult spleen, fetal liver, adult thyroid and adult lymph node (chimp). Additional site of expression is the developing spinal ganglia. All other tissues negative.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: Salivary gland, stomach, thyroid, parathyroid, skin, thymus, ovary, lymph node.

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum, penis.

(11) DNA35557-1137 (PRO234)

Specific expression over developing motor neurones in ventral aspect of the fetal spinal cord (will develop into ventral horns of spinal cord). All other tissues negative. Possible role in growth, differentiation and/or development of spinal motor neurons.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), penis, eye, bladder, stomach, gastric carcinoma, colon, colonic carcinoma and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis

(12) DNA33100-1159 (PRO229)

Striking expression in mononuclear phagocytes (macrophages) of fetal and adult spleen, liver, lymph node and adult thymus (in tingible body macrophages). The highest expression is in the spleen. All other tissues negative. Localisation and homology are entirely consistent with a role as a scavenger receptor for cells of the reticuloendothelial system. Expression also observed in placental mononuclear cells.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, myocardium, aorta, spleen, lymph node, gall bladder, pancreas, lung, skin, eye (inc. retina), prostate, bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: Salivary gland, stomach, thyroid, parathyroid, skin, thymus, ovary, lymph node.

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum, penis.

(13) DNA34431-1177 (PRO263)

Widepread expression in human fetal tissues and placenta over mononuclear cells, probably macrophages +/- lymphocytes. The cellular distribution follows a perivascular pattern in many tissues. Strong expression also seen in epithelial cells of the fetal adrenal cortex. All adult tissues were negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, adrenal, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), bladder, stomach, colon and colonic carcinoma. Acetominophen induced liver injury and hepatic cirrhosis.

A secondary screen evidenced expression over stromal mononuclear cells probably histiocytes.

(14) DNA38268-1188 (PRO295)

High expression over ganglion cells in human fetal spinal ganglia and over large neurones in the anterior horns of the developing spinal cord. In the adult there is expression in the chimp adrenal medulla (neural), neurones

of the rhesus monkey brain (hippocampus [+++] and cerebral cortex) and neurones in ganglia in the normal adult human prostate (the only section that contains ganglion cells, ie expression in this cell type is presumed NOT to be confined to the prostate). All other tissues negative.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, great vessels, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum.

Table 10

```
/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int     _day[26][26] = {
/*      A  B  C  D  E  F  G  H  I  J  K  L  M  N  O  P  Q  R  S  T  U  V  W  X  Y  Z */
/* A */  { 2, 0,-2, 0, 0,-4, 1,-1,-1, 0,-1,-2,-1, 0, _M, 1, 0,-2, 1, 1, 0, 0,-6, 0,-3, 0},
/* B */  { 0, 3,-4, 3, 2,-5, 0, 1,-2, 0, 0,-3,-2, 2, _M,-1, 1, 0, 0, 0, 0,-2,-5, 0,-3, 1},
/* C */  {-2,-4,15,-5,-5,-4,-3,-3,-2, 0,-5,-6,-5,-4, _M,-3,-5,-4, 0,-2, 0,-2,-8, 0, 0,-5},
/* D */  { 0, 3,-5, 4, 3,-6, 1, 1,-2, 0, 0,-4,-3, 2, _M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 2},
/* E */  { 0, 2,-5, 3, 4,-5, 0, 1,-2, 0, 0,-3,-2, 1, _M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 3},
/* F */  {-4,-5,-4,-6,-5, 9,-5,-2, 1, 0,-5, 2, 0,-4, _M,-5,-5,-4,-3,-3, 0,-1, 0, 0, 7,-5},
/* G */  { 1, 0,-3, 1, 0,-5, 5,-2,-3, 0,-2,-4,-3, 0, _M,-1,-1,-1,-3, 1, 0, 0,-1,-7, 0,-5, 0},
/* H */  {-1, 1,-3, 1, 1,-2,-2, 6,-2, 0, 0,-2,-2, 2, _M, 0, 3, 2,-1,-1, 0,-2,-3, 0, 0, 2},
/* I */  {-1,-2,-2,-2,-2, 1,-3,-2, 5, 0,-2, 2, 2,-2, _M,-2,-2,-2,-1, 0, 0, 4,-5, 0,-1,-2},
/* J */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */  {-1, 0,-5, 0, 0,-5,-2, 0,-2, 0, 5,-3, 0, 1, _M,-1, 1, 3, 0, 0, 0,-2,-3, 0,-4, 0},
/* L */  {-2,-3,-6,-4,-3, 2,-4,-2, 2, 0,-3, 6, 4, 3, _M,-3,-2,-3,-3,-1, 0, 2,-2, 0,-1,-2},
/* M */  {-1,-2,-5,-3,-2, 0,-3,-2, 2, 0, 0, 4, 6,-2, _M,-2,-1, 0,-2,-1, 0, 2,-4, 0,-2,-1},
/* N */  { 0, 2,-4, 2, 1,-4, 0, 2,-2, 0, 1,-3,-2, 2, _M,-1, 1, 0, 1, 0, 0,-2,-4, 0,-2, 1},
/* O */  {_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M, _M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M},
/* P */  { 1,-1,-3,-1,-1,-5,-1, 0,-2, 0,-1,-3,-2,-1, _M, 6, 0, 0, 1, 0, 0,-1,-6, 0,-5, 0},
/* Q */  { 0, 1,-5, 2, 2,-5,-1, 3,-2, 0, 1,-2,-1, 1, _M, 0, 4, 1,-1,-1, 0,-2,-5, 0,-4, 3},
/* R */  {-2, 0,-4,-1,-1,-4,-3, 2,-2, 0, 3,-3, 0, 0, _M, 0, 1, 6, 0,-1, 0,-2, 2, 0,-4, 0},
/* S */  { 1, 0, 0, 0, 0,-3, 1,-1,-1, 0, 0,-3,-2, 1, _M, 1,-1, 0, 2, 1, 0,-1,-2, 0,-3, 0},
/* T */  { 1, 0,-2, 0, 0,-3, 0,-1, 0, 0, 0,-1,-1, 0, _M, 0,-1,-1, 1, 3, 0, 0,-5, 0,-3, 0},
/* U */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */  { 0,-2,-2,-2,-2,-1,-1,-2, 4, 0,-2, 2, 2,-2, _M,-1,-2,-2,-1, 0, 0, 4,-6, 0,-2,-2},
/* W */  {-6,-5,-8,-7,-7, 0,-7,-3,-5, 0,-3,-2,-4,-4, _M,-6,-5, 2,-2,-5, 0,-6,17, 0, 0,-6},
/* X */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */  {-3,-3, 0,-4,-4, 7,-5, 0,-1, 0,-4,-1,-2,-2, _M,-5,-4,-4,-3,-3, 0,-2, 0, 0,10,-4},
/* Z */  { 0, 1,-5, 2, 3,-5, 0, 2,-2, 0, 0,-2,-1, 1, _M, 0, 3, 0, 0, 0, 0,-2,-6, 0,-4, 4}
};
```

Table 10 (cont')

```
/*
 */
#include <stdio.h>
#include <ctype.h>

#define  MAXJMP          16      /* max jumps in a diag */
#define  MAXGAP          24      /* don't continue to penalize gaps larger than this */
#define  JMPS            1024    /* max jmps in an path */
#define  MX              4       /* save if there's at least MX-1 bases since last jmp */

#define  DMAT            3       /* value of matching bases */
#define  DMIS            0       /* penalty for mismatched bases */
#define  DINS0           8       /* penalty for a gap */
#define  DINS1           1       /* penalty per base */
#define  PINS0           8       /* penalty for a gap */
#define  PINS1           4       /* penalty per residue */

struct jmp {
        short           n[MAXJMP];      /* size of jmp (neg for dely) */
        unsigned short  x[MAXJMP];      /* base no. of jmp in seq x */
};                                      /* limits seq to 2^16 -1 */

struct diag {
        int             score;          /* score at last jmp */
        long            offset;         /* offset of prev block */
        short           ijmp;           /* current jmp index */
        struct jmp      jp;             /* list of jmps */
};

struct path {
        int             spc;            /* number of leading spaces */
        short           n[JMPS];        /* size of jmp (gap) */
        int             x[JMPS];        /* loc of jmp (last elem before gap) */
};

char            *ofile;                 /* output file name */
char            *namex[2];              /* seq names: getseqs() */
char            *prog;                  /* prog name for err msgs */
char            *seqx[2];                       /* seqs: getseqs() */
int             dmax;                   /* best diag: nw() */
int             dmax0;                  /* final diag */
int             dna;                    /* set if dna: main() */
int             endgaps;                /* set if penalizing end gaps */
int             gapx, gapy;             /* total gaps in seqs */
int             len0, len1;             /* seq lens */
int             ngapx, ngapy;           /* total size of gaps */
int             smax;                   /* max score: nw() */
int             *xbm;                   /* bitmap for matching */
long            offset;                 /* current offset in jmp file */
struct  diag    *dx;                    /* holds diagonals */
struct  path    pp[2];                  /* holds path for seqs */

char            *calloc(), *malloc(), *index(), *strcpy();
char            *getseq(), *g_calloc();
```

Table 10 (cont')

```
/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 *  where file1 and file2 are two dna or two protein sequences.
 *  The sequences can be in upper- or lower-case an may contain ambiguity
 *  Any lines beginning with ';', '>' or '<' are ignored
 *  Max file length is 65535 (limited by unsigned short x in the jmp struct)
 *  A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 *  Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static  _dbval[26] = {
        1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static  _pbval[26] = {
        1, 2|(1<<('D'-'A')),(1<<('N'-'A')), 4, 8, 16, 32, 64,
        128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
        1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
        1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)                                                    main
        int     ac;
        char    *av[];
{
        prog = av[0];
        if (ac != 3) {
                fprintf(stderr, "usage: %s file1 file2\n", prog);
                fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
                fprintf(stderr, "The sequences can be in upper- or lower-case\n");
                fprintf(stderr, "Any lines beginning with ';' or '<' are ignored\n");
                fprintf(stderr, "Output is in the file \"align.out\"\n");
                exit(1);
        }
        namex[0] = av[1];
        namex[1] = av[2];
        seqx[0] = getseq(namex[0], &len0);
        seqx[1] = getseq(namex[1], &len1);
        xbm = (dna)? _dbval : _pbval;

        endgaps = 0;                    /* 1 to penalize endgaps */
        ofile = "align.out";            /* output file */

        nw();           /* fill in the matrix, get the possible jmps */
        readjmps();     /* get the actual jmps */
        print();        /* print stats, alignment */

        cleanup(0);     /* unlink any tmp files */
}
```

Table 10 (cont')

```
/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()                                                              nw
{
        char            *px, *py;               /* seqs and ptrs */
        int             *ndely, *dely;  /* keep track of dely */
        int             ndelx, delx;    /* keep track of delx */
        int             *tmp;           /* for swapping row0, row1 */
        int             mis;            /* score for each type */
        int             ins0, ins1;     /* insertion penalties */
        register        id;             /* diagonal index */
        register        ij;             /* jmp index */
        register        *col0, *col1;   /* score for curr, last row */
        register        xx, yy;         /* index into seqs */

        dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

        ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
        dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
        col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
        col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
        ins0 = (dna)? DINS0 : PINS0;
        ins1 = (dna)? DINS1 : PINS1;

        smax = -10000;
        if (endgaps) {
                for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
                        col0[yy] = dely[yy] = col0[yy-1] - ins1;
                        ndely[yy] = yy;
                }
                col0[0] = 0;            /* Waterman Bull Math Biol 84 */
        }
        else
                for (yy = 1; yy <= len1; yy++)
                        dely[yy] = -ins0;

        /* fill in match matrix
         */
        for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
                /* initialize first entry in col
                 */
                if (endgaps) {
                        if (xx == 1)
                                col1[0] = delx = -(ins0+ins1);
                        else
                                col1[0] = delx = col0[0] - ins1;
                        ndelx = xx;
                }
                else {
                        col1[0] = 0;
                        delx = -ins0;
                        ndelx = 0;
                }
```

Table 10 (cont')

...nw

```
for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
        mis = col0[yy-1];
        if (dna)
                mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
        else
                mis += _day[*px-'A'][*py-'A'];

        /* update penalty for del in x seq;
         * favor new del over ongong del
         * ignore MAXGAP if weighting endgaps
         */
        if (endgaps || ndely[yy] < MAXGAP) {
                if (col0[yy] - ins0 >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else {
                        dely[yy] -= ins1;
                        ndely[yy]++;
                }
        } else {
                if (col0[yy] - (ins0+ins1) >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else
                        ndely[yy]++;
        }

        /* update penalty for del in y seq;
         * favor new del over ongong del
         */
        if (endgaps || ndelx < MAXGAP) {
                if (col1[yy-1] - ins0 >= delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else {
                        delx -= ins1;
                        ndelx++;
                }
        } else {
                if (col1[yy-1] - (ins0+ins1) >= delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else
                        ndelx++;
        }

        /* pick the maximum score; we're favoring
         * mis over any del and delx over dely
         */
```

Table 10 (cont')

...nw

```
                id = xx - yy + len1 - 1;
                if (mis >= delx && mis >= dely[yy])
                        col1[yy] = mis;
                else if (delx >= dely[yy]) {
                        col1[yy] = delx;
                        ij = dx[id].ijmp;
                        if (dx[id].jp.n[0] && (!dna    | (ndelx >= MAXJMP
                        && xx > dx[id].jp.x[ij+MX] || mis > dx[id].score+DINS0)) {
                                dx[id].ijmp++;
                                if (++ij >= MAXJMP) {
                                        writejmps(id);
                                        ij = dx[id].ijmp = 0;
                                        dx[id].offset = offset;
                                        offset += sizeof(struct jmp) + sizeof(offset);
                                }
                        }
                        dx[id].jp.n[ij] = ndelx;
                        dx[id].jp.x[ij] = xx;
                        dx[id].score = delx;
                }
                else {
                        col1[yy] = dely[yy];
                        ij = dx[id].ijmp;
        if (dx[id].jp.n[0] && (!dna |  (ndely[yy] >= MAXJMP
                        && xx > dx[id].jp.x[ij+MX] || mis > dx[id].score+DINS0)) {
                                dx[id].ijmp++;
                                if (++ij >= MAXJMP) {
                                        writejmps(id);
                                        ij = dx[id].ijmp = 0;
                                        dx[id].offset = offset;
                                        offset += sizeof(struct jmp) + sizeof(offset);
                                }
                        }
                        dx[id].jp.n[ij] = -ndely[yy];
                        dx[id].jp.x[ij] = xx;
                        dx[id].score = dely[yy];
                }
                if (xx == len0 && yy < len1) {
                        /* last col
                        */
                        if (endgaps)
                                col1[yy] += ins0+ins1[*(len1-yy);
                        if (col1[yy] > smax) {
                                smax = col1[yy];
                                dmax = id;
                        }
                }
        }
        if (endgaps && xx < len0)
                col1[yy-1] += ins0+ins1[*(len0-xx);
        if (col1[yy-1] > smax) {
                smax = col1[yy-1];
                dmax = id;
        }
        tmp = col0; col0 = col1; col1 = tmp;
    }
    (void) free((char *)ndely);
    (void) free((char *)dely);
    (void) free((char *)col0);
    (void) free((char *)col1);
}
```

Table 10 (cont')

```
/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() -- put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC     3
#define P_LINE  256     /* maximum output line */
#define P_SPC   3       /* space between name or num and seq */

extern  _day[26][26];
int     olen;           /* set output line length */
FILE    *fx;            /* output file */

print()                                                    print
{
        int     lx, ly, firstgap, lastgap;      /* overlap */

        if ((fx = fopen(ofile, "w")) == 0) {
                fprintf(stderr, "%s: can't write %s\n", prog, ofile);
                cleanup(1);
        }
        fprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
        fprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
        olen = 60;
        lx = len0;
        ly = len1;
        firstgap = lastgap = 0;
        if (dmax < len1 - 1) {          /* leading gap in x */
                pp[0].spc = firstgap = len1 - dmax - 1;
                ly -= pp[0].spc;
        }
        else if (dmax >= len1 - 1) {    /* leading gap in y */
                pp[1].spc = firstgap = dmax - (len1 - 1);
                lx -= pp[1].spc;
        }
        if (dmax0 < len0 - 1) {         /* trailing gap in x */
                lastgap = len0 - dmax0 -1;
                lx -= lastgap;
        }
        else if (dmax0 >= len0 - 1) {   /* trailing gap in y */
                lastgap = dmax0 - (len0 - 1);
                ly -= lastgap;
        }
        getmat(lx, ly, firstgap, lastgap);
        pr_align();
}
```

Table 10 (cont')

getmat

```
/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)
        int     lx, ly;                 /* "core" (minus endgaps) */
        int     firstgap, lastgap;      /* leading trailing overlap */
{
        int             nm, i0, i1, siz0, siz1;
        char            outx[32];
        double          pct;
        register        n0, n1;
        register char   *p0, *p1;

        /* get total matches, score
         */
        i0 = i1 = siz0 = siz1 = 0;
        p0 = seqx[0] + pp[1].spc;
        p1 = seqx[1] + pp[0].spc;
        n0 = pp[1].spc + 1;
        n1 = pp[0].spc + 1;

        nm = 0;
        while ( *p0 && *p1 ) {
                if (siz0) {
                        p1++;
                        n1++;
                        siz0--;
                }
                else if (siz1) {
                        p0++;
                        n0++;
                        siz1--;
                }
                else {
                        if (xbm[*p0-'A']&xbm[*p1-'A'])
                                nm++;
                        if (n0++ == pp[0].x[i0])
                                siz0 = pp[0].n[i0++];
                        if (n1++ == pp[1].x[i1])
                                siz1 = pp[1].n[i1++];
                        p0++;
                        p1++;
                }
        }

        /* pct homology:
         * if penalizing endgaps, base is the shorter seq
         * else, knock off overhangs and take shorter core
         */
        if (endgaps)
                lx = (len0 < len1)? len0 : len1;
        else
                lx = (lx < ly)? lx : ly;
        pct = 100.*(double)nm/(double)lx;
        fprintf(fx, "\n");
        fprintf(fx, " < %d match%s in an overlap of %d: %.2f percent similarity\n",
                nm, (nm == 1)? "" : "es", lx, pct);
}
```

Table 10 (cont')

...getmat

```
        fprintf(fx, "< gaps in first sequence: %d", gapx);
        if (gapx) {
                (void) sprintf(outx, " (%d %s%s)",
                        ngapx, (dna)? "base":"residue", (ngapx == 1)? "":"s");
                fprintf(fx,"%s", outx);

        fprintf(fx, ", gaps in second sequence: %d", gapy);
        if (gapy) {
                (void) sprintf(outx, " (%d %s%s)",
                        ngapy, (dna)? "base":"residue", (ngapy == 1)? "":"s");
                fprintf(fx,"%s", outx);
        }
        if (dna)
                fprintf(fx,
                "\n< score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
                smax, DMAT, DMIS, DINS0, DINS1);
        else
                fprintf(fx,
                "\n< score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
                smax, PINS0, PINS1);
        if (endgaps)
                fprintf(fx,
                "< endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
                firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
                lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
        else
                fprintf(fx, "< endgaps not penalized\n");
}

static  nm;             /* matches in core -- for checking */
static  lmax;           /* lengths of stripped file names */
static  ij[2];          /* jmp index for a path */
static  nc[2];          /* number at start of current line */
static  ni[2];          /* current elem number -- for gapping */
static  siz[2];
static char     *ps[2];         /* ptr to current element */
static char     *po[2];         /* ptr to next output char slot */
static char     out[2][P_LINE]; /* output line */
static char     star[P_LINE];   /* set by stars() */

/*
* print alignment of described in struct path pp[]
*/
static
```

pr_align

```
pr_align()
{
        int             nn;     /* char count */
        int             more;
        register        i;

        for (i = 0, lmax = 0; i < 2; i++) {
                nn = stripname(namex[i]);
                if (nn > lmax)
                        lmax = nn;

                nc[i] = 1;
                ni[i] = 1;
                siz[i] = ij[i] = 0;
                ps[i] = seqx[i];
                po[i] = out[i];
        }
```

Table 10 (cont')

...pr_align

```
        for (nn = nm = 0, more = 1; more; ) {
                for (i = more = 0; i < 2; i++) {
                        /*
                         * do we have more of this sequence?
                         */
                        if (!*ps[i])
                                continue;

                        more++;

                        if (pp[i].spc) {        /* leading space */
                                *po[i]++ = ' ';
                                pp[i].spc--;
                        }
                        else if (siz[i]) {      /* in a gap */
                                *po[i]++ = '-';
                                siz[i]--;
                        }
                        else {                  /* we're putting a seq element
                                                 */
                                *po[i] = *ps[i];
                                if (islower(*ps[i]))
                                        *ps[i] = toupper(*ps[i]);
                                po[i]++;
                                ps[i]++;

                                /*
                                 * are we at next gap for this seq?
                                 */
                                if (ni[i] == pp[i].x[ij[i]]) {
                                        /*
                                         * we need to merge all gaps
                                         * at this location
                                         */
                                        siz[i] = pp[i].nj[i]++];
                                        while (ni[i] == pp[i].x[ij[i]])
                                                siz[i] += pp[i].nj[i]++];
                                }
                                ni[i]++;
                        }
                }
                if (++nn == olen || !more && nn) {
                        dumpblock();
                        for (i = 0; i < 2; i++)
                                po[i] = out[i];
                        nn = 0;
                }
        }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()                                                    dumpblock
{
        register  i;

        for (i = 0; i < 2; i++)
                *po[i]-- = '\0';
```

Table 10 (cont')

...dumpblock

```
        (void) putc('\n', fx);
        for (i = 0; i < 2; i++) {
                if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
                        if (i == 0)
                                nums(i);
                        if (i == 0 && *out[1])
                                stars();
                        putline(i);
                        if (i == 0 && *out[1])
                                fprintf(fx, star);
                        if (i == 1)
                                nums(i);
                }
        }
}

/*
 * put out a number line: dumpblock()
 */
static
nums(ix)                                                        nums
        int     ix;     /* index in out[] holding seq line */
{
        char            nline[P_LINE];
        register        i, j;
        register char   *pn, *px, *py;

        for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
                *pn = ' ';
        for (i = nc[ix], py = out[ix]; *py; py++, pn++, i--) {
                if (*py == ' ' || *py == '-')
                        *pn = ' ';
                else {
                        if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                                j = (i < 0)? -i : i;
                                for (px = pn; j /= 10, px--)
                                        *px = j%10 + '0';
                                if (i < 0)
                                        *px = '-';
                        }
                        else
                                *pn = ' ';
                        i--;
                }
        }
        *pn = '\0';
        nc[ix] = i;
        for (pn = nline; *pn; pn++)
                (void) putc(*pn, fx);
        (void) putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
putline(ix)                                                     putline
        int     ix;
{
```

Table 10 (cont')

...putline

```
        int             i;
        register char   *px;

        for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
                (void) putc(*px, fx);
        for (; i < lmax+P_SPC; i++)
                (void) putc(' ', fx);

        /* these count from 1:
         * ni[] is current element (from 1)
         * nc[] is number at start of current line
         */
        for (px = out[ix]; *px; px++)
                (void) putc(*px&0x7F, fx);
        (void) putc('\n', fx);
}

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()
{
        int             i;
        register char   *p0, *p1, cx, *px;

        if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
            !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
                return;
        px = star;
        for (i = lmax+P_SPC; i; i--)
                *px++ = ' ';

        for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
                if (isalpha(*p0) && isalpha(*p1)) {

                        if (xbm[*p0-'A']&xbm[*p1-'A']) {
                                cx = '*';
                                nn++;
                        }
                        else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                                cx = '.';
                        else
                                cx = ' ';
                }
                else
                        cx = ' ';
                *px++ = cx;
        }
        *px++ = '\n';
        *px = '\0';
}
```

stars

Table 10 (cont')

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)
        char    *pn;    /* file name (may be path) */
{
        register char   *px, *py;

        py = 0;
        for (px = pn; *px; px++)
                if (*px == '/')
                        py = px + 1;
        if (py)
                (void) strcpy(pn, py);
        return(strlen(pn));
}
```

stripname

Table 10 (cont')

```
/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 * readjmps() -- get the good jmps, from tmp file if necessary
 * writejmps() -- write a filled array of jmps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";             /* tmp file for jmps */
FILE    *fj;

int     cleanup();                              /* cleanup tmp file */
long    lseek();

/*
 * remove any tmp file if we blow
 */
cleanup(i)                                                              cleanup
        int     i;
{
        if (fj)
                (void) unlink(jname);
        exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)                                                       getseq
        char    *file;  /* file name */
        int     *len;   /* seq len */
{
        char            line[1024], *pseq;
        register char   *px, *py;
        int             natgc, tlen;
        FILE            *fp;

        if ((fp = fopen(file,"r")) == 0) {
                fprintf(stderr,"%s: can't read %s\n", prog, file);
                exit(1);
        }
        tlen = natgc = 0;
        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '\n'; px++)
                        if (isupper(*px) || islower(*px))
                                tlen++;
        }
        if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
                fprintf(stderr,"%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
                exit(1);
        }
        pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
```

Table 10 (cont')

...getseq

```
        py = pseq + 4;
        *len = tlen;
        rewind(fp);

        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '\n'; px++) {
                        if (isupper(*px))
                                *py++ = *px;
                        else if (islower(*px))
                                *py++ = toupper(*px);
                        if (index("ATGCU",*(py-1)))
                                natgc++;
                }
        }
        *py++ = '\0';
        *py = '\0';
        (void) fclose(fp);
        dna = natgc > (tlen/3);
        return(pseq+4);
}

char    *
g_calloc(msg, nx, sz)                                                   g_calloc
        char    *msg;           /* program, calling routine */
        int     nx, sz;         /* number and size of elements */
{
        char            *px, *calloc();

        if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
                if (*msg) {
                        fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
                        exit(1);
                }
        }
        return(px);
}

/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()                                                              readjmps
{
        int             fd = -1;
        int             siz, i0, i1;
        register  i, j, xx;

        if (fj) {
                (void) fclose(fj);
                if ((fd = open(jname, O_RDONLY, 0)) < 0) {
                        fprintf(stderr, "%s: can't open() %s\n", prog, jname);
                        cleanup(1);
                }
        }
        for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
                while (1) {
                        for (j = dx[dmax].jmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                                ;
```

Table 10 (cont')

...readjmps

```
                if (j < 0 && dx[dmax].offset && fj) {
                        (void) lseek(fd, dx[dmax].offset, 0);
                        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
                        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
                        dx[dmax].ijmp = MAXJMP-1;
                }
                else
                        break;
        }
        if (i >= JMPS) {
                fprintf(stderr, "%s: too many gaps in alignment\n", prog);
                cleanup(1);
        }
        if (j >= 0) {
                siz = dx[dmax].jp.n[j];
                xx = dx[dmax].jp.x[j];
                dmax1 = siz;
                if (siz < 0) {                  /* gap in second seq */
                        pp[1].n[i1] = -siz;
                        xx1 += siz;
                        /* id = xx - yy + len1 - 1
                         */
                        pp[1].x[i1] = xx - dmax1 + len1 - 1;
                        gapy++;
                        ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
                        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
                        i1++;
                }
                else if (siz > 0) {  /* gap in first seq */
                        pp[0].n[i0] = siz;
                        pp[0].x[i0] = xx;
                        gapx++;
                        ngapx += siz;
/* ignore MAXGAP when doing endgaps */
                        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
                        i0++;
                }
        }
        else
                break;
}

/* reverse the order of jmps
 */
for (j = 0, i0--; j < i0; j++, i0--) {
        i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
        i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
        i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
        i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
        (void) close(fd);
if (fj) {
        (void) unlink(jname);
        fj = 0;
        offset = 0;
}
}
```

Table 10 (cont')

writejmps

```
/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejmps(ix)
        int     ix;
{
        char    *mktemp();

        if (!fj) {
                if (mktemp(jname) < 0) {
                        fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
                        cleanup(1);
                }
                if ((fj = fopen(jname, "w")) == 0) {
                        fprintf(stderr, "%s: can't write %s\n", prog, jname);
                        exit(1);
                }
        }
        (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
        (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}
```

Table 11

PRO	XXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXYYYYYY	(Length = 12 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

5 divided by 15 = 33.3%

Table 12

PRO	XXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXYYYYYYZZYZ	(Length = 15 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

5 divided by 10 = 50%

Table 13

PRO-DNA	NNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLL	(Length = 16 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

6 divided by 14 = 42.9%

Table 14

PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLVV	(Length = 9 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

4 divided by 12 = 33.3%

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

Material	ATCC Dep. No.	Deposit Date
DNA32292-1131	ATCC 209258	September 16, 1997
DNA33094-1131	ATCC 209256	September 16, 1997
DNA33223-1136	ATCC 209264	September 16, 1997
DNA34435-1140	ATCC 209250	September 16, 1997
DNA27864-1155	ATCC 209375	October 16, 1997
DNA36350-1158	ATCC 209378	October 16, 1997
DNA32290-1164	ATCC 209384	October 16, 1997
DNA35639-1172	ATCC 209396	October 17, 1997
DNA33092-1202	ATCC 209420	October 28, 1997
DNA49435-1219	ATCC 209480	November 21, 1997
DNA35638-1141	ATCC 209265	September 16, 1997
DNA32298-1132	ATCC 209257	September 16, 1997
DNA33089-1132	ATCC 209262	September 16, 1997
DNA33786-1132	ATCC 209253	September 16, 1997
DNA35918-1174	ATCC 209402	October 17, 1997
DNA37150-1178	ATCC 209401	October 17, 1997
DNA38260-1180	ATCC 209397	October 17, 1997
DNA39969-1185	ATCC 209400	October 17, 1997
DNA32286-1191	ATCC 209385	October 16, 1997
DNA33461-1199	ATCC 209367	October 15, 1997
DNA40628-1216	ATCC 209432	November 7, 1997
DNA33221-1133	ATCC 209263	September 16, 1997
DNA33107-1135	ATCC 209251	September 16, 1997
DNA35557-1137	ATCC 209255	September 16, 1997
DNA34434-1139	ATCC 209252	September 16, 1997
DNA33100-1159	ATCC 209373	October 16, 1997
DNA35600-1162	ATCC 209370	October 16, 1997
DNA34436-1238	ATCC 209523	December 10, 1997
DNA33206-1165	ATCC 209372	October 16, 1997
DNA35558-1167	ATCC 209374	October 16, 1997
DNA35599-1168	ATCC 209371	October 16, 1997
DNA36992-1168	ATCC 209382	October 16, 1997
DNA34407-1169	ATCC 209383	October 16, 1997
DNA35841-1173	ATCC 209403	October 17, 1997
DNA33470-1175	ATCC 209398	October 17, 1997
DNA34431-1177	ATCC 209399	October 17, 1997
DNA39510-1181	ATCC 209392	October 17, 1997
DNA39423-1182	ATCC 209387	October 17, 1997
DNA40620-1183	ATCC 209388	October 17, 1997
DNA40604-1187	ATCC 209394	October 17, 1997
DNA38268-1188	ATCC 209421	October 28, 1997
DNA37151-1193	ATCC 209393	October 17, 1997
DNA35673-1201	ATCC 209418	October 28, 1997
DNA40370-1217	ATCC 209485	November 21, 1997
DNA42551-1217	ATCC 209483	November 21, 1997
DNA39520-1217	ATCC 209482	November 21, 1997
DNA41225-1217	ATCC 209491	November 21, 1997
DNA43318-1217	ATCC 209481	November 21, 1997
DNA40587-1231	ATCC 209438	November 7, 1997
DNA41338-1234	ATCC 209927	June 2, 1998
DNA40981-1234	ATCC 209439	November 7, 1997

DNA37140-1234	ATCC 209489	November 21, 1997
DNA40982-1235	ATCC 209433	November 7, 1997
DNA41379-1236	ATCC 209488	November 21, 1997
DNA44167-1243	ATCC 209434	November 7, 1997
DNA39427-1179	ATCC 209395	October 17, 1997
DNA40603-1232	ATCC 209486	November 21, 1997
DNA43466-1225	ATCC 209490	November 21, 1997
DNA43046-1225	ATCC 209484	November 21, 1997
DNA35668-1171	ATCC 209371	October 16, 1997

These deposit were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. Isolated nucleic acid having at least 80% sequence identity to a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:12), Figure 9 (SEQ ID NO:18), Figure 11 (SEQ ID NO:23), Figure 13 (SEQ ID NO:28), Figure 15 (SEQ ID NO:34), Figure 17 (SEQ ID NO:39), Figure 19 (SEQ ID NO:49), Figure 22 (SEQ ID NO:59), Figure 24 (SEQ ID NO:64), Figure 26 (SEQ ID NO:69), Figure 28 (SEQ ID NO:71), Figure 30 (SEQ ID NO:73), Figure 32 (SEQ ID NO:84), Figure 34 (SEQ ID NO:91), Figure 36 (SEQ ID NO:96), Figure 38 (SEQ ID NO:104), Figure 40 (SEQ ID NO:109), Figure 42 (SEQ ID NO:114), Figure 44 (SEQ ID NO:119), Figure 46 (SEQ ID NO:127), Figure 48 (SEQ ID NO:132), Figure 50 (SEQ ID NO:137), Figure 52 (SEQ ID NO:142), Figure 54 (SEQ ID NO:148), Figure 56 (SEQ ID NO:153), Figure 58 (SEQ ID NO:159), Figure 60 (SEQ ID NO:164), Figure 62 (SEQ ID NO:170), Figure 64 (SEQ ID NO:175), Figure 66 (SEQ ID NO:177), Figure 68 (SEQ ID NO:185), Figure 70 (SEQ ID NO:190), Figure 72 (SEQ ID NO:195), Figure 74 (SEQ ID NO:201), Figure 76 (SEQ ID NO:207), Figure 78 (SEQ ID NO:213), Figure 80 (SEQ ID NO:221), Figure 82 (SEQ ID NO:227), Figure 84 (SEQ ID NO:236), Figure 86 (SEQ ID NO:245), Figure 88 (SEQ ID NO:250), Figure 90 (SEQ ID NO:255), Figure 92 (SEQ ID NO:257), Figure 94 (SEQ ID NO:259), Figure 96 (SEQ ID NO:261), Figure 98 (SEQ ID NO:263), Figure 100 (SEQ ID NO:285), Figure 102 (SEQ ID NO:290), Figure 104 (SEQ ID NO:292), Figure 106 (SEQ ID NO:294), Figure 108 (SEQ ID NO:310), Figure 110 (SEQ ID NO:315), Figure 112 (SEQ ID NO:320), Figure 114 (SEQ ID NO:325), Figure 116 (SEQ ID NO:332), Figure 118 (SEQ ID NO:339), Figure 120 (SEQ ID NO:341) and Figure 122 (SEQ ID NO:377).

2. The nucleic acid of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:11), Figure 8 (SEQ ID NO:17), Figure 10 (SEQ ID NO:22), Figure 12 (SEQ ID NO:27), Figure 14 (SEQ ID NO:33), Figure 16 (SEQ ID NO:38), Figure 18 (SEQ ID NO:48), Figure 21 (SEQ ID NO:58), Figure 23 (SEQ ID NO:63), Figure 25 (SEQ ID NO:68), Figure 27 (SEQ ID NO:70), Figure 29 (SEQ ID NO:72), Figure 31 (SEQ ID NO:83), Figure 33 (SEQ ID NO:90), Figure 35 (SEQ ID NO:95), Figure 37 (SEQ ID NO:103), Figure 39 (SEQ ID NO:108), Figure 41 (SEQ ID NO:113), Figure 43 (SEQ ID NO:118), Figure 45 (SEQ ID NO:126), Figure 47 (SEQ ID NO:131), Figure 49 (SEQ ID NO:136), Figure 51 (SEQ ID NO:141), Figure 53 (SEQ ID NO:147), Figure 55 (SEQ ID NO:152), Figure 57 (SEQ ID NO:158), Figure 59 (SEQ ID NO:163), Figure 61 (SEQ ID NO:169), Figure 63 (SEQ ID NO:174), Figure 65 (SEQ ID NO:176), Figure 67 (SEQ ID NO:184), Figure 69 (SEQ ID NO:189), Figure 71 (SEQ ID NO:194), Figure 73 (SEQ ID NO:200), Figure 75 (SEQ ID NO:206), Figure 77 (SEQ ID NO:212), Figure 79 (SEQ ID NO:220), Figure 81 (SEQ ID NO:226), Figure 83 (SEQ ID NO:235), Figure 85 (SEQ ID NO:244), Figure 87 (SEQ ID NO:249), Figure 89 (SEQ ID NO:254), Figure 91 (SEQ ID NO:256), Figure 93 (SEQ ID NO:258), Figure 95 (SEQ ID NO:260), Figure 97 (SEQ ID NO:262), Figure 99 (SEQ ID NO:284), Figure 101 (SEQ ID NO:289), Figure 103 (SEQ ID NO:291), Figure 105 (SEQ ID NO:293), Figure 107 (SEQ ID NO:309), Figure 109 (SEQ ID NO:314), Figure 111 (SEQ ID NO:319), Figure 113 (SEQ ID NO:324), Figure 115 (SEQ ID NO:331), Figure

117 (SEQ ID NO:338), Figure 119 (SEQ ID NO:340) and Figure 121 (SEQ ID NO:376), or the complement thereof.

3. The nucleic acid of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the full-length coding sequence of the sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:11), Figure 8 (SEQ ID NO:17), Figure 10 (SEQ ID NO:22), Figure 12 (SEQ ID NO:27), Figure 14 (SEQ ID NO:33), Figure 16 (SEQ ID NO:38), Figure 18 (SEQ ID NO:48), Figure 21 (SEQ ID NO:58), Figure 23 (SEQ ID NO:63), Figure 25 (SEQ ID NO:68), Figure 27 (SEQ ID NO:70), Figure 29 (SEQ ID NO:72), Figure 31 (SEQ ID NO:83), Figure 33 (SEQ ID NO:90), Figure 35 (SEQ ID NO:95), Figure 37 (SEQ ID NO:103), Figure 39 (SEQ ID NO:108), Figure 41 (SEQ ID NO:113), Figure 43 (SEQ ID NO:118), Figure 45 (SEQ ID NO:126), Figure 47 (SEQ ID NO:131), Figure 49 (SEQ ID NO:136), Figure 51 (SEQ ID NO:141), Figure 53 (SEQ ID NO:147), Figure 55 (SEQ ID NO:152), Figure 57 (SEQ ID NO:158), Figure 59 (SEQ ID NO:163), Figure 61 (SEQ ID NO:169), Figure 63 (SEQ ID NO:174), Figure 65 (SEQ ID NO:176), Figure 67 (SEQ ID NO:184), Figure 69 (SEQ ID NO:189), Figure 71 (SEQ ID NO:194), Figure 73 (SEQ ID NO:200), Figure 75 (SEQ ID NO:206), Figure 77 (SEQ ID NO:212), Figure 79 (SEQ ID NO:220), Figure 81 (SEQ ID NO:226), Figure 83 (SEQ ID NO:235), Figure 85 (SEQ ID NO:244), Figure 87 (SEQ ID NO:249), Figure 89 (SEQ ID NO:254), Figure 91 (SEQ ID NO:256), Figure 93 (SEQ ID NO:258), Figure 95 (SEQ ID NO:260), Figure 97 (SEQ ID NO:262), Figure 99 (SEQ ID NO:284), Figure 101 (SEQ ID NO:289), Figure 103 (SEQ ID NO:291), Figure 105 (SEQ ID NO:293), Figure 107 (SEQ ID NO:309), Figure 109 (SEQ ID NO:314), Figure 111 (SEQ ID NO:319), Figure 113 (SEQ ID NO:324), Figure 115 (SEQ ID NO:331), Figure 117 (SEQ ID NO:338), Figure 119 (SEQ ID NO:340) and Figure 121 (SEQ ID NO:376), or the complement thereof.

4. Isolated nucleic acid which comprises the full-length coding sequence of the DNA deposited under accession number ATCC 209258, ATCC 209256, ATCC 209264, ATCC 209250, ATCC 209375, ATCC 209378, ATCC 209384, ATCC 209396, ATCC 209420, ATCC 209480, ATCC 209265, ATCC 209257, ATCC 209262, ATCC 209253, ATCC 209402, ATCC 209401, ATCC 209397, ATCC 209400, ATCC 209385, ATCC 209367, ATCC 209432, ATCC 209263, ATCC 209251, ATCC 209255, ATCC 209252, ATCC 209373, ATCC 209370, ATCC 209523, ATCC 209372, ATCC 209374, ATCC 209373, ATCC 209382, ATCC 209383, ATCC 209403, ATCC 209398, ATCC 209399, ATCC 209392, ATCC 209387, ATCC 209388, ATCC 209394, ATCC 209421, ATCC 209393, ATCC 209418, ATCC 209485, ATCC 209483, ATCC 209482, ATCC 209491, ATCC 209481, ATCC 209438, ATCC 209927, ATCC 209439, ATCC 209489, ATCC 209433, ATCC 209488, ATCC 209434, ATCC 209395, ATCC 209486, ATCC 209490, ATCC 209484 or ATCC 209371.

5. A vector comprising the nucleic acid of Claim 1.

6. The vector of Claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.

7. A host cell comprising the vector of Claim 5.

8. The host cell of Claim 7 wherein said cell is a CHO cell.

9. The host cell of Claim 7 wherein said cell is an *E. coli*.

10. The host cell of Claim 7 wherein said cell is a yeast cell.

11. A process for producing a PRO polypeptides comprising culturing the host cell of Claim 7 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.

12. Isolated native sequence PRO polypeptide having at least 80% sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:12), Figure 9 (SEQ ID NO:18), Figure 11 (SEQ ID NO:23), Figure 13 (SEQ ID NO:28), Figure 15 (SEQ ID NO:34), Figure 17 (SEQ ID NO:39), Figure 19 (SEQ ID NO:49), Figure 22 (SEQ ID NO:59), Figure 24 (SEQ ID NO:64), Figure 26 (SEQ ID NO:69), Figure 28 (SEQ ID NO:71), Figure 30 (SEQ ID NO:73), Figure 32 (SEQ ID NO:84), Figure 34 (SEQ ID NO:91), Figure 36 (SEQ ID NO:96), Figure 38 (SEQ ID NO:104), Figure 40 (SEQ ID NO:109), Figure 42 (SEQ ID NO:114), Figure 44 (SEQ ID NO:119), Figure 46 (SEQ ID NO:127), Figure 48 (SEQ ID NO:132), Figure 50 (SEQ ID NO:137), Figure 52 (SEQ ID NO:142), Figure 54 (SEQ ID NO:148), Figure 56 (SEQ ID NO:153), Figure 58 (SEQ ID NO:159), Figure 60 (SEQ ID NO:164), Figure 62 (SEQ ID NO:170), Figure 64 (SEQ ID NO:175), Figure 66 (SEQ ID NO:177), Figure 68 (SEQ ID NO:185), Figure 70 (SEQ ID NO:190), Figure 72 (SEQ ID NO:195), Figure 74 (SEQ ID NO:201), Figure 76 (SEQ ID NO:207), Figure 78 (SEQ ID NO:213), Figure 80 (SEQ ID NO:221), Figure 82 (SEQ ID NO:227), Figure 84 (SEQ ID NO:236), Figure 86 (SEQ ID NO:245), Figure 88 (SEQ ID NO:250), Figure 90 (SEQ ID NO:255), Figure 92 (SEQ ID NO:257), Figure 94 (SEQ ID NO:259), Figure 96 (SEQ ID NO:261), Figure 98 (SEQ ID NO:263), Figure 100 (SEQ ID NO:285), Figure 102 (SEQ ID NO:290), Figure 104 (SEQ ID NO:292), Figure 106 (SEQ ID NO:294), Figure 108 (SEQ ID NO:310), Figure 110 (SEQ ID NO:315), Figure 112 (SEQ ID NO:320), Figure 114 (SEQ ID NO:325), Figure 116 (SEQ ID NO:332), Figure 118 (SEQ ID NO:339), Figure 120 (SEQ ID NO:341) and Figure 122 (SEQ ID NO:377).

13. Isolated PRO polypeptide having at least 80% sequence identity to the amino acid sequence encoded by the nucleotide deposited under accession number ATCC 209258, ATCC 209256, ATCC 209264, ATCC 209250, ATCC 209375, ATCC 209378, ATCC 209384, ATCC 209396, ATCC 209420, ATCC 209480, ATCC 209265, ATCC 209257, ATCC 209262, ATCC 209253, ATCC 209402, ATCC 209401, ATCC 209397, ATCC 209400, ATCC 209385, ATCC 209367, ATCC 209432, ATCC 209263, ATCC 209251, ATCC 209255, ATCC 209252, ATCC 209373, ATCC 209370, ATCC 209523, ATCC 209372, ATCC 209374, ATCC 209373,

ATCC 209382, ATCC 209383, ATCC 209403, ATCC 209398, ATCC 209399, ATCC 209392, ATCC 209387, ATCC 209388, ATCC 209394, ATCC 209421, ATCC 209393, ATCC 209418, ATCC 209485, ATCC 209483, ATCC 209482, ATCC 209491, ATCC 209481, ATCC 209438, ATCC 209927, ATCC 209439, ATCC 209489, ATCC 209433, ATCC 209488, ATCC 209434, ATCC 209395, ATCC 209486, ATCC 209490, ATCC 209484 or ATCC 209371.

14. A chimeric molecule comprising a polypeptide according to Claim 12 fused to a heterologous amino acid sequence.

15. The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is an epitope tag sequence.

16. The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

17. An antibody which specifically binds to a PRO polypeptide according to Claim 12.

18. The antibody of Claim 17 wherein said antibody is a monoclonal antibody.

FIGURE 1

ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA
CCCACGCGTCCGGGCCGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCGTCTTCCCG
CAGCGCTACCCGCCATGCGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTG
CTGCTGCCGCCCGCGCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCT
GGTGGACAAGTTTAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACA
CGGCTTGGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTCGCCTGCTGGAGATC
CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA
GCACCTGGAGGCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGC
CAGGGCGGATCCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA
GGGCGACGGGTCCTGCCGGTGCCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGG
ACGGCTACTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC
TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT
GGACGAGGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTG
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTG
GGCTGCACAGGGGAAGGCCCAGGAAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA
CGGACAGTGTGCAGATGTGGACGAGTGCTCACTAGCAGAAAAAACCTGTGTGAGGAAAAACG
AAAACTGCTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCCTGACGGCTTCGAAGAAACG
GAAGATGCCTGTGTGCCGCCGGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT
GCCCTCCCGCGAAGACCTGTAATGTGCCGGACTTACCCTTTAAATTATTCAGAAGGATGTCC
CGTGGAAAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGC
CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTA
AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG
AGGAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGC
TTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA
TCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTC
ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTCGGCGCAGCACCATGGCCTGAAAT
AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG
TGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAGCATGC
ATCTCAATTAGTCAGCAACCCAGTTTT

FIGURE 2

><subunit 1 of 1, 353 aa, 0 stop
><MW: 38192, pI: 4.53, NX(S/T): 2
MRLPRRAALGLLPLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEEKTLSKYESSEIRL
LEILEGLCESSDFECNQMLEAQEEHLSAWWLQLKSEYPDLFEWFCVKTLKVCCSPGTYGPDCLACQGGSQRPCSG
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESCKTCSGLTNRDCGECEVGWVLDE
GACVDVDECAAEPPPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKT
CVRKNENCYNTPGSYVCVCPDGFEETEDACVPPAEAEATEGESPTQLPSREDL

Signal peptide:
amino acids 1-24

N-glycosylation sites.
amino acids 190-194 and 251-255

Glycosaminoglycan attachment sites.
amino acids 149-153 and 155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 26-30

Casein kinase II phosphorylation sites.
amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343 and 349-353

Tyrosine kinase phosphorylation site.
amino acids 303-310

N-myristoylation sites.
amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and 313-319

Aspartic acid and asparagine hydroxylation site.
amino acids 308-320

EGF-like domain cysteine pattern signature.
amino acids 166-178

Leucine zipper pattern.
amino acids 94-116

FIGURE 3

CAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTC
GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCGCCCAGCCGTCTAAACGGGAACA
GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCG
GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGC**ATG**GCCCGGAGGAGCGCCTTC
CCTGCCGCCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGC
CGGGCCGCCGCAGGAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA
TAGGATTTGAAGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGAT
TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCCTGTCAATATCCATTCCATGAATTTTAC
CTGGCAAGCTGCAGGGCAGGCAGAATACTTCTATGAATTCCTGTCCTTGCGCTCCCTGGATA
AAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCA
TCAGTTGTTCAAGTTGGTTTCCCATGTCTTGGAAAACAGGATGGGGTGGCAGCATTTGAAGT
GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCT
TCTTTAAAACATGTCAACAAGCTGAGTGCCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAAT
GAAAGACGCATCTGCGAGTGTCCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTG
TACCCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCCACCTG
GATTCTATGGAGTGAACTGTGACAAAGCAAACTGCTCAACCACCTGCTTTAATGGAGGGACC
TGTTTCTACCCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG
CAAATGCCCACAACCCTGTCGAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTT
CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACAT
GGAACCTGCCATGAACCCAACAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAA
TAAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGCAGGCGCCCAGCTCAGGCAGC
ACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG
TGAACTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTCA
TGTGTTGAATGTTCAAATAATGTTCATTACACTTAAGAATACTGGCCTGAATTTTATTAGCT
TCATTATAAATCACTGAGCTGATATTTACTCTTCCTTTTAAGTTTTCTAAGTACGTCTGTAG
CATGATGGTATAGATTTTCTTGTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGA
TCAGGTTAAAATTTTCAGTGTGTAGTTGGCAGATATTTTCAAAATTACAATGCATTTATGGT
GTCTGGGGGCAGGGGAACATCAGAAAGGTTAAATTGGGCAAAAATGCGTAAGTCACAAGAAT
TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGAT
GTTTGTTACATTTTTAAAAATTGCTCTTAATTTTTAAACTCTCAATACAATATATTTTGACC
TTACCATTATTCCAGAGATTCAGTATTAAAAAAAAAAAAATTACACTGTGGTAGTGGCATTT
AAACAATATAATATATTCTAAACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCAT
TGGCTTGAAGCAATATAATATATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTAT
ACTGTTTGTATGTATAAAATAAAGGTGCTGCTTTAGTTTTTTGGAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGC
CGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 4

>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094
><subunit 1 of 1, 379 aa, 0 stop
><MW: 41528, pI: 7.97, NX(S/T): 2
MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFLSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPGGC
RNGGFCNERRICECPDGFHGPHCEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEERRDP
PESNYIW

Signal peptide:
amino acids 1-28

N-glycosylation site.
amino acids 88-92, 245-249

Casein kinase II phosphorylation site.
amino acids 319-323

Tyrosine kinase phosphorylation site.
amino acids 370-378

N-myristoylation sites.
amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).
amino acids 285-293

EGF-like domain cysteine pattern signature.
amino acids 198-210, 230-242, 262-274, 294-306, 326-338

FIGURE 5

CGGACGCGTGGGCGTCCGGCGGTCGCAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGG
CCCCAGCCCACACCTTCACCAGGGCCCAGGAGCCACC**ATG**TGGCGATGTCCACTGGGGCTAC
TGCTGTTGCTGCCGCTGGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTCGTGGGCGCCGG
GAGCTAGCACCGGGTCTGCACCTGCGGGGCATCCGGGACGCGGGAGGCCGGTACTGCCAGGA
GCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT
GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTC
TGCCTCGGCGTGCCACCCCCTTTTCCCCCGATCCAAGGATGTATGCATGGAGGTCGTATCTA
TCCAGTCTTGGGAACGTACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGT
GGCATGGTGGATCCAGACATGATCAAAGCCATCAACCAGGGCAACTATGGCTGGCAGGCTGG
GAACCACAGCGCCTTCTGGGGCATGACCCTGGA**TGA**GGGCATTCGCTACCGCCTGGGCACCA
TCCGCCCATCTTCCTCGGTCATGAACATGCATGAAATTTATACAGTGCTGAACCCAGGGGAG
GTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCATGAGCCTCTTGA
CCAAGGCAACTGTGCAGGCTCCTGGGCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT
CAATCCATTCTCTGGGACACATGACGCCTGTCCTGTCGCCCCAGAACCTGCTGTCTTGTGAC
ACCCACCAGCAGCAGGGCTGCCGCGGTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCG
CCGAGGGGTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCC
CTGCGCCCCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCACTGCC
CACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCT
CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA
TGGAGGTGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGC
CTTGGGAGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG
AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCCTGGGGCC
CAGCCTGGGGCGAGAGGGGCCACTTCCGCATCGTGCGCGGCGTCAATGAGTGCGACATCGAG
AGCTTCGTGCTGGGCGTCTGGGGCCGCGTGGGCATGGAGGACATGGGTCATCACTGAGGCTG
CGGGCACCACGCGGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCCAATG
GGGCGGTGACCCCAGCCTCGCCCGACAGAGCCCGGGGCGCAGGCGGGCGCCAGGGCGCTAAT
CCCGGCGCGGGTTCCGCTGACGCAGCGCCCCGCCTGGGAGCCGCGGGCAGGCGAGACTGGCG
GAGCCCCCAGACCTCCCAGTGGGGACGGGGCAGGGCCTGGCCTGGGAAGAGCACAGCTGCAG
ATCCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGC
CCCAATACCCCACCCCAATCCCGTATTCTTTTTTTTTTTTTTTAGACAGGGTCTTGCTCCG
TTGCCCAGGTTGGAGTGCAGTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCA
AGTGACCCTCCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCACCACACCTGGC
TAATTTTTGTATTTTTTGTAAAGAGGGGGGTCTCACTGTGTTGCCCAGGCTGGTTTCGAACT
CCTGGGCTCAAGCGGTCCACCTGCCTCCGCCTCCCAAAGTGCTGGGATTGCAGGCATGAGCC
ACTGCACCCAGCCCTGTATTCTTATTCTTCAGATATTTATTTTTCTTTTCACTGTTTTAAAA
TAAAACCAAAGTATTGATAAAAAAAAA

FIGURE 6

>>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223

>>subunit 1 of 1, 164 aa, 1 stop

>>MW: 18359, pI: 7.45, NX(S/T): 1

MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQDLCCRGRADDC

ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPPPPPPIQGCMHGGRIYPVLGTYWDNCNR

CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

N-glycosylation site.

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300, 411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230, 269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 398-409

FIGURE 7

AGGCTCCTTGGCCCTTTTTCCACAGCAAGCTTNTGCNATCCCGATTCGTTGTCTCAAATCCA
ATTCTCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCGCCGAGG
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC
CCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCACTGCCCACTGC
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG
TGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG
AGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAG

FIGURE 8

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG
GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG
CAGCTTGAACTGCGTGGATGACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT
GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT
GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCCGGCCAGCTATAGGCTCTGGGGGG
CCCCGCTGCAGCCCACACTGGGTGTGGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG
GCCCAGTGGGAGCCTGTCCTGGTTCCTGAGGCACATCCTAACGCAAGTCTGACCATGTATGT
CTGCACCCCTGTCCCCCACCCTGACCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA
TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCTCCAACCCTCTCTGCTGCTGTTTC
CATGGCCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT
CCCTGCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA
GGGGACAGGCACTCAGGAGGGCCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA
GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA
GGGGCCAGGCCTCACATTCGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT
AATAAACACCTGTTGGATAAGCCAAAAAAA

FIGURE 9

MTHRTTTWARRTSRAVTPTCATPAGPMPCSRLPPSLRCSLHSACCSGDPASYRLWGAPLQPT
LGVVPQASVPLLTDLAQWEPVLVPEAHPNASLTMYVCTPVPHPDPPMALSRTPTRQISSSDT
DPPADGPSNPLCCCFHGPAFSTLNPVLRHLFPQEAFPAIPIYDLSQVWSVVSPAPSRGQALRRAQ

Signal peptide:
amino acids 1-47

N-glycosylation site.
amino acids 31-35, 74-78, 84-88

Casein kinase II phosphorylation site.
amino acids 22-26, 76-80

N-myristoylation site.
amino acids 56-60

Amidation site.
amino acids 70-74

FIGURE 10

CCCACGCGTCCGAACCTCTCCAGCGATGGGAGCCGCCCGCCTGCTGCCCAACCTCACTCTGT
GCTTACAGCTGCTGATTCTCTGCTGTCAAACTCAGTACGTGAGGGACCAGGGCGCCATGACC
GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA
CGTGCAGGTCACCGGGCGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC
TCATAGTGGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGCTGAGAGTGAGAAG
TACATCTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTG
CGTGTTCACGGAGATCGTGCTGGAGAACAACTATACGGCCTTCCAGAACGCCCGGCACGAGG
GCTGGTTCATGGCCTTCACGCGGCAGGGGCGGCCCCGCCAGGCTTCCCGCAGCCGCCAGAAC
CAGCGCGAGGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCACGC
CGAGAACCAGAAGCAGTTCGAGTTTGTGGGCTCCGCCCCCACCCGCCGGACCAAGCGCACAC
GGCGGCCCCAGCCCCTCACGTAGTCTGGGAGGCAGGGGGCAGCAGCCCCTGGGCCGCCTCCC
CACCCCTTTCCCTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCC
GAGGGAGGACCCTGAGGGCCGCGAAGCATCCGAGCCCCCAGCTGGGAAGGGCAGGCCGGTG
CCCCAGGGGCGGCTGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACT
GAGTGTCACCCTGATCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCC
CGCTGAAAGGTCAGCGACTGAAGGCCTTGCAGACAACCGTCTGGAGGTGGCTGTCCTCAAAA
TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCCAGCCCCCAAACTCCTCCTGGCTAGACTGTA
GGAAGGGACTTTTGTTTGTTTGTTTGTTTCAGGAAAAAAGAAAGGGAGAGAGAGGAAAATAG
AGGGTTGTCCACTCCTCACATTCCACGACCCAGGCCTGCACCCCACCCCCAACTCCCAGCCC
CGGAATAAAACCATTTTCCTGC

FIGURE 11

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRI
SATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLIGKPSGKSKDCVFTEIVLE
NNYTAFQNARHEGWFMAFTRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQKQFEF
VGSAPTRRTKRTRRPQPLT

Signal peptide:
amino acids 1-22

N-glycosylation site.
amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 60-64

Casein kinase II phosphorylation site.
amino acids 65-69

Tyrosine kinase phosphorylation site.
amino acids 39-48, 89-97

N-myristoylation site.
amino acids 69-75, 188-194

Amidation site.
amino acids 58-62

HBGF/FGF family signature.
amino acids 103-128

FIGURE 12

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG
GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT
CATTTTGATTTTGCTGTTTATTTTTTTTTCTTTTTCTTTTTCCCACCACATTGTATTTTAT
TTCCGTACTTCAGAA**ATG**GGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCCT
GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC
CTAGTGTGTGCCGCTGCGACAGGAACTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC
TGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC
AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC
AATATTCAGACCATTTCACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCACCT
GGATGACAACTCCATATCCACAGTGGGGGTGGAAGACCGGGCCTTCCGGGAGGCTATTAGCC
TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGAC
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCATATCCGACATGGCCTTCCAGAA
TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG
AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC
CACCCTCCTCCCGATCTCCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT
AAACCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA
ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTGATAATCTCTCCAACCTGAAGCAGCTC
ACTGCTCGGAATAACCCTTGGTTTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAA
ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTCCCAAGGTCCTGAACAAGTCCGGG
GGATGGCCGTCAGGGAATTAAATATGAATCTTTTGTCCTGTCCCACCACGACCCCCGGCCTG
CCTCTCTTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT
TCCAAACCCTAGCAGAAGCTACACGCCTCCAACTCCTACCACATCGAAACTTCCCACGATTC
CTGACTGGGATGGCAGAGAAAGAGTGACCCCACCTATTTCTGAACGGATCCAGCTCTCTATC
CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTCACCGTGATGGCATA
CAAACTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTCAGGAGCGCATAG
TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT
TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGC
CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA
CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT
GTGCTGGTGGTCTTGCTCAGCGTCTTTTGCTGGCATATGCACAAAAGGGGCGCTACACCTC
CCAGAAGTGGAAATACAACCGGGGCCGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA
AGGACAACTCCATCCTGGAGATGACAGAAACCAGTTTTCAGATCGTCTCCTTAAATAACGAT
CAACTCCTTAAAGGAGATTTCAGACTGCAGCCCATTTACACCCCAAATGGGGGCATTAATTA
CACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC
ACTGCCATACG**TGA**CAGCCAGAGGCCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA
CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT
TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAAGTG
CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTAAATCTT

FIGURE 13

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP
EGVTVLYLHNNQINNAGFPAELRNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI
SRAALAQLLKLEELHLDDNSISTVGVEDGAFREAISLKLLFLSKNHLSSVPVGLPVDLQELR
VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD
LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN
PWFCDCSIKWVTEWLKYIPSSLNVRGFMCQGPEQVRGMAVRELNMNLLSCPTTTPGLPLFTP
APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVND
TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL
DAFNYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVL
LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG
DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300, 522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

FIGURE 14

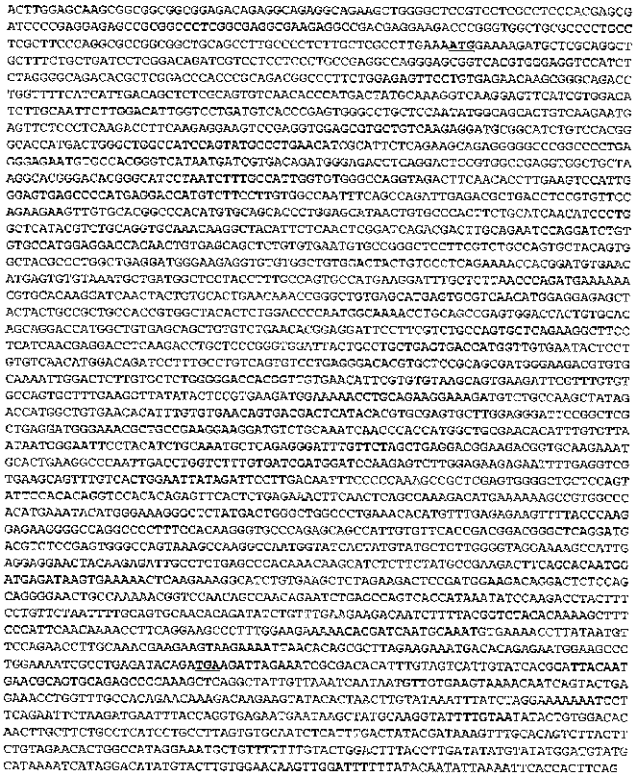


FIGURE 15

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPQTALLESSCENKRADLVFII
DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV
KRMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD
TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHN
CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA
LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKPGC
EHECVNMEESYYCRCHRGYTLDPNGKTCSRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLI
NEDLKTCSRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTCAKLDSCALGDHGCE
HSCVSSEDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLA
EDGKRCRRKDVCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVID
GSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVA
HMKYMGKGSMTGLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN
GITMYAVGVGKAIEEELQEIASEPTNKHLFYAEDFSTMDEISEKLKKGICEALEDSDGRQDS
PAGELPKTVQQPTESEFVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPL
EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMEALENRLRYR

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 115-119, 606-610, 892-896

Casein kinase II phosphorylation site.

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247, 401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784, 781-785, 819-823, 866-870

N-myristoylation site.

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500, 639-645, 690-696, 752-758, 792-798

Amidation site.

amino acids 314-318, 560-564, 601-605

Aspartic acid and asparagine hydroxylation site.

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464, 540-546, 581-587

FIGURE 16

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCG
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTTCCCTCCCGGGGCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG
GCGGTGGAGGCAGGGGAAGTGGTCCTTCCAGCGTGGTACACCTTGCACGGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT
TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG
GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCCAC
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT
TTCCTCTGTCACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT
TGACCCCCACGCCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCCACGACAGAT
GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCG
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCCAGAGTCAAGCTGGCTCTCTGGTATGATGAC
CCCACCACTCATTGGCTAAAGGATTTGGGGTCTCTCCTTCCTATAAGGGTCACCTCTAGCAC
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCCACCTCTC
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTCACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCCAGGC
CCCCTTGATCTGTACCCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC
TTATTAAAACTAACATGAAATATGTGTTGTTTTCATTTGCAAATTTAAATAAAGATACATAA
TGTTTGTATGAAAAA

FIGURE 17

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS
SQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY
SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSCQSPRSKPAVQYQ
WDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA
VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLPWPKSSDTISKNGTL
SSVTSARALRPPHGPPRPGALTPTPSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSR
MGAVPVMVPAQSQAGSLV

Signal peptide:
amino acids 1-29

Transmembrane domain:
amino acids 245-267

N-glycosylation site.
amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.
amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262, 262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 164-175

FIGURE 18

CGCCACCACTGCGGCCACCGCCAATGAAACGCCTCCCGCTCCTAGTGGTTTTTTCCACTTTG
TTGAATTGTTCCTATACTCAAAATTGCACCAAGACACCTTGTCTCCCAAATGCAAAATGTGA
AATACGCAATGGAATTGAAGCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTGTCACAA
TTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGTCCTGTGGCGAAAATGCTAATTGC
ACTAACACAGAAGGAAGTTATTATTGTATGTGTGTACCTGGCTTCAGATCCAGCAGTAACCA
AGACAGGTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCATT
TAGATAATGTCTGTATAGCTGCAAATATTAATAAAACTTTAACAAAAATCAGATCCATAAAA
GAACCTGTGGCTTTGCTACAAGAAGTCTATAGAAATTCTGTGACAGATCTTTCACCAACAGA
TATAATTACATATATAGAAATATTAGCTGAATCATCTTCATTACTAGGTTACAAGAACAACA
CTATCTCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTGAATTTGTAAAAACCGTGAAT
AATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAGAAC
ACATCTTACAAAACTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCCAGAGCTTCC
AAAAGACCACAGAGTTTGATACAAATTCAACGGATATAGCTCTCAAAGTTTTCTTTTTTGAT
TCATATAACATGAAACATATTCATCCTCATATGAATATGGATGGAGACTACATAAATATATT
TCCAAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATA
AGAGTATTGGTCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAAAATTATGAT
AATTCTGAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATGAGCTCAAACCC
ACCCACATTATATGAACTTGAAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATA
GGTATAGGAGTCTATGTGCATTTTGGAATTACTCACCTGATACCATGAATGGCAGCTGGTCT
TCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCT
GACACATTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTC
TTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTT
ACCTTCTGGTTCTTCAGTGAAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTG
TAGCCTATTTCTTGCTGAACTTGTTTTTCTTGTTGGGATCAATACAAATACTAATAAGCTCT
TCTGTTCAATCATTGCCGGACTGCTACACTACTTCTTTTTAGCTGCTTTTGCATGGATGTGC
ATTGAAGGCATACATCTCTATCTCATTGTTGTGGGTGTCATCTACAACAAGGGATTTTTGCA
CAAGAATTTTTATATCTTTGGCTATCTAAGCCCAGCCGTGGTAGTTGGATTTTCGGCAGCAC
TAGGATACAGATATTATGGCACAACCAAAGTATGTTGGCTTAGCACCGAAAACAACTTTATT
TGGAGTTTTATAGGACCAGCATGCCTAATCATTCTTGTTAATCTCTTGGCTTTTGGAGTCAT
CATATACAAAGTTTTTCGTCACACTGCAGGGTTGAAACCAGAAGTTAGTTGCTTTGAGAACA
TAAGGTCTTGTGCAAGAGGAGCCCTCGCTCTTCTGTTCCTTCTCGGCACCACCTGGATCTTT
GGGGTTCTCCATGTTGTGCACGCATCAGTGGTTACAGCTTACCTCTTCACAGTCAGCAATGC
TTTCCAGGGGATGTTCATTTTTTTATTCCTGTGTGTTTATCTAGAAAGATTCAAGAAGAAT
ATTACAGATTGTTCAAAAATGTCCCCTGTTGTTTTGGATGTTTAACGTAAACATAGAGAATG
GTGGATAATTACAACTGCACAAAAATAAAAATTCCAAGCTGTGGATGACCAATGTATAAAAA
TGACTCATCAAATTATCCAATTATTAACTACTAGACAAAAAGTATTTTAAATCAGTTTTTCT
GTTTATGCTATAGGAACTGTAGATAATAAGGTAAAATTATGTATCATATAGATATACTATGT
TTTTCTATGTGAAATAGTTCTGTCAAAATAGTATTGCAGATATTTGGAAAGTAATTGGTTT
CTCAGGAGTGATATCACTGCACCCAAGGAAAGATTTTCTTTCTAACACGAGAAGTATATGAA
TGTCCTGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGTGTTGCCTTTCAAACTAGTCC
CCTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGCAGA
ATATCAAACAGTGAAAAGGGAATGATAAGATGTATTTTGAATGAACTGTTTTTTCTGTAGAC
TAGCTGAGAAATTGTTGACATAAAATAAAGAATTGAAGAAACACATTTTACCATTTTGTGAA
TTGTTCTGAACTTAAATGTCCACTAAAACAACTTAGACTTCTGTTTGCTAAATCTGTTTCTT
TTTCTAATATTCTAAAAAAAAAAAAAAAGGTTTACCTCCACAAATTGAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 19

MKRLPLLVVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC
GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNVCIAA
NINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSSLLGYKNNTISAKDTL
SNSTLTEFVKTVNNFVQRDTFVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSFQKTTEFDT
NSTDIALKVFFFDSYNMKHIHPHMNMDGDYINIPPKRKAAYDSNGNVAVAFLYYKSIGPLLS
SSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTDRYRSLCAF
WNYSPDTMNGSWSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG
IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCSIIAGL
LHYFFLAAFAWMCIEGIHLYLIVVGVIYNKGFLHKNFYIFGYLSPAVVVGFSAALGYRYYGT
TKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGA
LALLFLLGTTWIFGVLHVVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEYYRLFKNV
PCCFGCLR

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636, 648-664

N-glycosylation site.

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181, 188-192, 249-253, 381-385, 395-399

Glycosaminoglycan attachment site.

amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 360-364

Casein kinase II phosphorylation site.

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154, 155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329, 346-350, 608-612

Tyrosine kinase phosphorylation site.

amino acids 36-44, 669-677, 670-678

N-myristoylation site.

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394, 434-440, 480-486, 521-527

Aspartic acid and asparagine hydroxylation site.

amino acids 75-87

FIGURE 20

TGGAAACATATCCTCCCTCATATGAATATGGATGGAGACTACATAAATATATTTCCAAAGNG
AAAAGCCGGCATATGGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATAAGAGTAT
TGGTCCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAAAATTATGATAATTCT
GAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATGAGCTCAAACCCACCCAC
ATTATATGAACTTGAAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATAGGTATA
GGAGTCTATGTGGCATTTTGGAATACTCACCTCATACCATGAATGGCAGCTGGTCTTCAGAG
GGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACA
TTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTCTTACAA
GGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTTACCTTC
TGGTTCTTCAGTGAAATTCAAAGCACCAGGA

FIGURE 21

GCTCCCAGCCAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG
CTAAGCGAGGCCTCCTCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTCACCCCGGCTGGGA
CAAGAAGCCGCCGCCTGCCTGCCCGGGCCCGGGGAGGGGGCTGGGGCTGGGCCCGGAGGCGG
GGTGTGAGTGGGTGTGTGCGGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
TGTCTTGGGCACCTACCCGTGGGGCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGGAG
CCGCCGCGCCGTCAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC
GGCACTCACAGCCCCGCAGCGCATCCCGGTCGCCGCCCAGCCTCCCGCACCCCCATCGCCGG
AGCTGCGCCGAGAGCCCCAGGGAGGTGCC**ATG**CGGAGCGGGTGTGTGGTGGTCCACGTATGC
ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCC
CCACGTGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCCC
ACGGGCTCTCCAGCTGCTTCCTGCGCATCCGTGCCGACGGCGTCGTGGACTGCGCGCGGGGC
CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG
CGTGCACAGCGTGCGGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT
ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGA
TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA
CAGAGGCTTTCTTCCACTCTCTCATTTCCTGCCCATGCTGCCCATGGTCCCAGAGGAGCCTG
AGGACCTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCCTGGAGACCGACAGCATG
GACCCATTTGGGCTTGTCACCGGACTGGAGGCCGTGAGGAGTCCCAGCTTTGAGAAG**TAA**CT
GAGACCATGCCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG
TGCTTCTACAAGAACAGTCCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAA
GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT
CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCTGCCTGGGCCCCCATTCTGCTCCCTCGA
GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC
TCACTTCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC
CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTCAGGAACAGGTGATCCACTCTGTA
AAACAGCAGGTAAACTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG
ACCATTTGCCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCACCAG
GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTCCCC
CTGAGGCCAGTTCTGTCATGGATGCTGTCCTGAGAATAACTTGCTGTCCCGGTGTCACCTGC
TTCCATCTCCCAGCCCACCAGCCCTCTGCCCACCTCACATGCCTCCCCATGGATTGGGGCCT
CCCAGGCCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT
TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC
CCTTTCCCCAGCACTTGGTTTTCCAACATGATATTTATGAGTAATTTATTTTGATATGTACA
TCTCTTATTTTCTTACATTATTTATGCCCCCAAATTATATTTATGTATGTAAGTGAGGTTTG
TTTTGTACATTAAAATGGACTTTGTTTGT

FIGURE 22

MRSGCVVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDPIRLRHLYTSGPHGLSSCFLR1
RADGVVDCARGQSAHSLLEIKAVALRTVAIKGVHSVRYLCMGADGKMQGLLQYSEEDCAFEE
EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD
MFSSPLRTDSMDPFGLVTGLEAVRSPSFEK

Signal peptide:
amino acids 1-22

Casein kinase II phosphorylation site.
amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.
amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 48-59

FIGURE 23

CCCAGAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCA
GAGCAGCCGGCTGCCGCCCCGGGAAGATGGCGAGGAGGAGCCGCCACCGCCTCCTCCTGCTG
CTGCTGCGCTACCTGGTGGTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA
AGAAGACTGTTTCCTCCAGATTAGAGTGGAAGAAACTGGGTCGGAGTGTCTCCTTTGTCTAC
TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCG
GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG
AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT
CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAACTGTGGTAGAGCTACGATGTCAAGA
CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA
ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATACACAATGAATACAAAAACTGGAACT
CTGCAATTTAATACTGTTTCCAAACTGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC
TGTTGGATATCGCAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA
TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTAT
GCTCAGAGGAAAGGCTACTTTTCAAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA
AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGG
CCGCGGCGGGCGGATCACGAGGTCAGGAGTTCTAGACCAGTCTGGCCAATATGGTGAAACCC
CATCTCTACTAAAATACAAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGC
TTGGGAGACAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC
CACTGCAGTCCAGCCTGGGTAACAGAGCAAGATTCCATCTCAAAAAATAAAATAAATAAATA
AATAAATACTGGTTTTTACCTGTAGAATTCTTACAATAAATATAGCTTGATATTC

FIGURE 24

MARRSRHRLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLE
WKKLGRSVSFVYYQQTLQGDFKNRASMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQNLEED
TVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQST
NSSYTMNTKTGTLQFNTVSKLDTGEYSCEARNSVGYRRCPGKRMQVDDLNISGIIAAVVVVA
LVISVCGLGVCYAQRKGYFSKETSFQKSNSSSKATTMSENVQWLTPVIPALWKAAAGGSRGQEF

Signal peptide:
amino acids 1-20

Transmembrane domain:
amino acids 130-144, 238-258

N-glycosylation site.
amino acids 98-102, 187-191, 236-240, 277-281

Casein kinase II phosphorylation site.
amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

N-myristoylation site.
amino acids 182-188, 239-245, 255-261, 257-263, 305-311

Amidation site.
amino acids 226-230

FIGURE 25

GACATCGGAGGTGGGCTAGCACTGAAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAG
AAAGAAGAGGAAGATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCAT
CATGCTGCTATTCCTGCAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA
ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA
TTATATCATTAAGGAAATAGTAACCTTCTCTTCTCCAATATGCATGACATTTTTGGACAATG
CAATTGTGGCACTGGCACTTATTTCAGTGAAGAAAAACTTTGTGGTTCTATGGCATTCATCA
TTTGACAAATGCAAGCATCTTCCTTATCAATCAGCTCCTATTGAACTTACTAGCACTGACTG
TGGAATCCTTAAGGGCCCATTACATTTCTGAAGAAGAAAGCTAAGATGAAGGACATGCCACT
CCGAATTCATGTGCTACTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG
TGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTTGGTTTACACCCAGATCCATTTAT
ATGGAAGCATCTACAGTGGATTGTAATGATTTAGGTCTTTTAACTTTCCCAGCCAGATTGCC
AGCTAACACACAGATTCTTCTCCTACAGACTAACAATATTGCAAAAATTGAATACTCCACAG
ACTTTCCAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAAT
ATTAATGTAAAAAAGATGCCTCAGCTCCTTTCTGTGTACCTAGAGGAAAACAAACTTACTGA
ACTGCCTGAAAAATGTCTGTCCGAACTGAGCAACTTACAAGAACTCTATATTAATCACAACT
TGCTTTCTACAATTTCACCTGGAGCCTTTATTGGCCTACATAATCTTCTTCGACTTCATCTC
AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGAT
TCTGATGATTGGGGAAAATCCAATTATCAGAATCAAAGACATGAACTTTAAGCCTCTTATCA
ATCTTCGCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTT
GGACTGGAAAACTTAGAAAGCATCTCTTTTACGATAACAGGCTTATTAAAGTACCCCATGT
TGCTCTTCAAAAAGTTGTAAATCTCAAATTTTTGGATCTAAATAAAAATCCTATTAATAGAA
TACGAAGGGGTGATTTTAGCAATATGCTACACTTAAAAGAGTTGGGGATAAATAATATGCCT
GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC
TACTAACAACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTCAGACTCCCCAAGCTGG
AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTC
CCAAACCTCAAGGAAATCAGCATACACAGTAACCCCATCAGGTGTGACTGTGTCATCCGTTG
GATGAACATGAACAAAACCAACATTCGATTCATGGAGCCAGATTCACTGTTTTGCGTGGACC
CACCTGAATTCCAAGGTCAGAATGTTCGGCAAGTGCATTTCAGGGACATGATGGAAATTTGT
CTCCCTCTTATAGCTCCTGAGAGCTTTCCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT
TTCCTTTCACTGTAGAGCTACTGCAGAACCACAGCCTGAAATCTACTGGATAACACCTTCTG
GTCAAAAACTCTTGCCTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTA
GATATAAATGGCGTAACTCCCAAAGAAGGGGGTTTATATACTTGTATAGCAACTAACCTAGT
TGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGATCTTTTCCACAAGATAACAATG
GCTCTTTGAATATTAAAATAAGAGATATTCAGGCCAATTCAGTTTTGGTGTCCTGGAAAGCA
AGTTCTAAAATTCTCAAATCTAGTGTTAAATGGACAGCCTTTGTCAAGACTGAAAATTCTCA
TGCTGCGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC
CATCAACTGAGTATAAAATTTGTATTGATATTCCCACCATCTATCAGAAAACAGAAAAAAA
TGTGTAAATGTCACCACCAAAGGTTTGCACCCTGATCAAAAAGAGTATGAAAAGAATAATAC
CACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGATTATTGGTGTGATATGTCTTATCA
GCTGCCTCTCTCCAGAAATGAACTGTGATGGTGGACACAGCTATGTGAGGAATTACTTACAG
AAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAA
AGAAAAAAGTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCCT
AAAAACCACCAAGGAAACCTACTCCAAAAATGAAC

FIGURE 26

MKDMPLRIHVLLGLAITTLVQAVDKKVDCPRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT
FPARLPANTQILLLQTNNIAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE
ENKLTSLPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLELNSNRLQMINSKWFDA
LPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRL
IKVPHVALQKVVNLKFLDLNKNPINRIRRGDFSNMLHLKELGINNMPELISIDSLAVDNLPD
LRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSNPIRC
DCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGQNVRQVHFRDMMEICLPLIAPESPPSNLNV
EAGSYVSFHCRATAEPQPEIYWITPSGQKLLPNTLTDKFYVHSEGTLDINGVTPKEGGLYTC
IATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFV
KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE
YEKNNTTLMACLGGLLGIIGVICLISCLSPEMNCDGGHSYVRNYLQKPTFALGELYPPLIN
LWEAGKEKSTSLKVKATVIGLPTNMS

Signal sequence:

amino acids 1-22

Transmembrane domain:

amino acids 633-650

N-glycosylation site.

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583, 608-612, 624-628, 625-629

Casein kinase II phosphorylation site.

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

Tyrosine kinase phosphorylation site.

amino acids 570-579

N-myristoylation site.

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443, 491-497, 492-498, 634-640, 702-708

Cell attachment sequence.

amino acids 277-280

FIGURE 27

GCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGACACATGTGTTAG
CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC
CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT
GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTA
TGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG
GCTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGG
ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT
GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG
TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC
ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC
AACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGC
CAACGACGCTGACCTTTGTAACCTCCCTAAAAAAACTACCGATTATGCCATGCTGGTCACCA
TGTTTGGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG
GATGCCCGGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA
ACCTGATGATATTAGCACTGTGGTATAGTGTCCAAACTGACTGTCATTGAGAAAGAAAGAAA
GTAGTTTGCGATTGCAGTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAA
ACTTTGTATTTCAGTTTTTTTTGAATTATGCCACTGCTGAACTTTTAACAAACACTACAACA
TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCCGATGGTATATTTCTGAGT
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTT
AATTTAAAAGCAAATAAAAGCTTAACTTTGAACCATGGAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 28

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDL
PPSTVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDLSDNR
IQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFL
NAANDADLCNLPKKTTDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKSLPSRQKK
ADEPDDISTVV

Signal sequence:
amino acids 1-33

Transmembrane domain:
amino acids 205-220

N-glycosylation site.
amino acids 47-51, 94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 199-203

Casein kinase II phosphorylation site.
amino acids 162-166, 175-179

N-myristoylation site.
amino acids 37-43, 45-51, 110-116

FIGURE 29

ACCGAGCCGAGCGGACCGAAGGCGCGCCCGAGATGCAGGTGACCAAGAGGATGCTGGCGGGG
GGCGTGAGGAGCATGCCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCT
GGGCTCAGTGCTGTCAGGCTCGGCCACGGGCTGCCCGCCCCGCTGCGAGTGCTCCGCCCAGG
ACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCATCCCCACCGAG
ACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACGAGTTCGCCAG
CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGCG
CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCCTGAAGCTCATC
CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGAT
CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG
ACAATGACCTCGTCTACATCTCTCACCGCGCCTTCAGCGGCCTCAACAGCCTGGAGCAGCTG
ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGCACGGCCT
CATCGTCCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC
TGTACCGACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATCACACCCAAC
TGCCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCCCTGTGCC
CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGCTTCCTCAACCTCTCCTACAACCCCATCA
GCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGC
GGGCAGCTGGCCGTGGTGGAGCCCTATGCCTTCCGCGGCCTCAACTACCTGCGCGTGCTCAA
TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG
AGACACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTCGGCTCCTGTGGGTGTTCCGG
CGCCGCTGGCGGCTCAACTTCAACCGGCAGCAGCCCACGTGCGCCACGCCCGAGTTTGTCCA
GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACCTGCCGCCGCG
CCCGCATCCGGGACCGCAAGGCCCAGCAGGTGTTTGTGGACGAGGGCCACACGGTGCAGTTT
GTGTGCCGGGCCGATGGCGACCCGCCGCCCGCCATCCTCTGGCTCTCACCCCGAAAGCACCT
GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCT
ACGCCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCAACGCGGGCGGCAACGAC
TCCATGCCCGCCCACCTGCATGTGCGCAGCTACTCGCCCGACTGGCCCCATCAGCCCAACAA
GACCTTCGCTTTCATCTCCAACCAGCCGGGCGAGGGAGAGGCCAACAGCACCCGCGCCACTG
TGCCTTTCCCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTCTTTC
CTGGGCGTCGTCCTCTTCTGCCTGGTGCTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC
AAAGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCG
ACGCGCCCCGCAAGTTCAACATGAAGATGATATGAGGCCGGGGCGGGGGGCAGGGACCCCCG
GGCGGCCGGGCAGGGGAAGGGGCCTGGTCGCCACCTGCTCACTCTCCAGTCCTTCCCACCTC
CTCCCTACCCTTCTACACACGTTCTCTTTCTCCCTCCCGCCTCCGTCCCCTGCTGCCCCCCG
CCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGGGGACCCCA
CCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCA
ATAACTCAATAAAAAAGTTACGAACTTTCTCTGTAACTTGGGTTTCAATAATTATGGATTTT
TATGAAAACTTGAAATAATAAAAAGAGAAAAAAACTAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MQVSKRMLAGGVRSMPSPLLACWQPILLLVLGSVLSGSATGCPPRCECSAQDRAVLCHRKCF
VAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELELNENIVSAVEPGAFNNLFNLRTL
GLRSNRLKLIPLGVFTGLSNLTKQDISENKIVILLDYMFQDLYNLKSLEVGDNDLVYISHRA
FSGLNSLEQLTLEKCNLTSIPTEALSHLHGLIVLRLRHLNINAIRDYSFKRLYRLKVLEISH
WPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLSYNPISTIEGSMLHEL
LRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLEESVFHSVGNLETLILDSNPLA
CDCRLLWVFRRRWRLNFNRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV
FVDEGHTVQFVCRADGDPPPAILWLSPRKHLVSAKSNGRLTVFPDGTLEVRYAQVQDNGTYL
CIAANAGGNDSMPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI
IATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPRKSDAGISSADAPRKFNMKMI

Signal sequence:
amino acids 1-41

Transmembrane domain:
amino acids 556-578

N-glycosylation site.
amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345, 492-496, 505-509, 526-530, 542-546

Casein kinase II phosphorylation site.
amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353, 607-611

Tyrosine kinase phosphorylation site.
amino acids 590-598

N-myristoylation site.
amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143, 262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 32-43

FIGURE 31

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCAAC
ATCGTAGTCCACCCCCTCCCCATCCCCAGCCCCCGGGGATTCAGGCTCGCCAGCGCCCAGCC
AGGGAGCCGGCCGGGAAGCGCGATGGGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCTCCTGC
TGTTCGCCTGCTGCTGGGCGCCCGGCGGGGCCAACCTCTCCCAGGACGACAGCCAGCCCTGG
ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGCTCAAGTGCCAAGTGAAAGATCA
CGAGGACTCATCCCTGCAATGGTCTAACCCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA
GAGCCCTTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGC
ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT
GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCCACAGAAGCCCATCATCACTGGTT
ATAAATCTTCATTACGGGAAAAAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAG
CCTGCAGCCCGGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCAACCCGCAT
ACAGGAAGATCCCAATGGTAAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCC
GGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGAC
AGATCCACCTCTCAACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCC
TCCCCATCCTCGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCC
CCCAGCAGTACCTATGGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGT
GCCCTGATCTTCCCTTTCCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAG
CAACATGGGCAGCTACAAGGCCTACTACACCCTCAATGTTAATGACCCCAGTCCGGTGCCCT
CCTCCTCCAGCACCTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCCTGCTG
CTCATCATGCTCATCTTCCTTGGCCACTACTTGATCCGGCACAAAGGAACCTACCTGACACA
TGAGGCAAAAGGCTCCGACGATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCG
GGCAGTCAGGAGGGGACGACAAGAAGGAATATTTCATCTAGAGGCGCCTGCCCACTTCCTGC
GCCCCCCAGGGGCCCTGTGGGGACTGCTGGGGCCGTCACCAACCCGGACTTGTACAGAGCAA
CCGCAGGGCCGCCCCTCCCGCTTGCTCCCCAGCCCACCCACCCCCCTGTACAGAATGTCTGC
TTTGGGTGCGGTTTTGTACTCGGTTTGGAATGGGGAGGGAGGAGGGCGGGGGGAGGGGAGGG
TTGCCCTCAGCCCTTTCCGTGGCTTCTCTGCATTTGGGTTATTATTATTTTTGTAACAATCC
CAAATCAAATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAAACA
AACAAAAAACA

FIGURE 32

MGAPAASLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLKCQVKDHEDSSLQW
SNPAQQTLYFGEKRALRDNRIQLVTSTPHELSISISNVALADEGEYTCSIFTMPVRTAKSLV
TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQEDPNGK
TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQ
KLLLHCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKA
YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIFLGHYLIRHKGTYLTHEAKGSDD
APDADTAIINAEGGQSGGDDKKEYFI

Signal sequence:
amino acids 1-20

Transmembrane domain:
amino acids 331-352

N-glycosylation site.
amino acids 25-29, 290-294

Casein kinase II phosphorylation site.
amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.
amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304, 306-310, 334-340, 360-364, 385-389, 386-390

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 7-18

FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCCAAACCCTTTTCTTCTCCTTTCCTGG
CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG
TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAA**ATG**CTGCTTTGGATTCTGTT
GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA
TCTGTTCCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAAGGGCTTCACA
AGTCTGCAGCGTTTCACTGCCCCGACTTCCCAGTTTTACCATTTATTTCTGCATGGCAATTC
CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG
AAAACAATGGCTTGCATGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGG
CTGCACATCAACAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA
TCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATATAGACCCGGGGGCCTTCCAGG
ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCCTACCTGCCAAC
GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC
CTATGAGGAGGTCTTGGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCCTT
GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCC
CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC
CACCGAACAGGACTTGTGTCCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCCTG
CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG
GATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT
CAAAATCAGACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCCTTAGCTAACA
GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTCGGGTTTAAAGATGAAC
TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA
GCTTTTCCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA
ACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG
AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA
GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA
TCCTCCCGGGCACTTTCAATGCCATGCCCAAACTGAGGATCCTCATTCTCAACAACAACCTG
CTGAGGTCCCTGCCTGTGGACGTGTTCGCTGGGGTCTCGCTCTCTAAACTCAGCCTGCACAA
CAATTACTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA
TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCA
GAACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACTTCTT
TAGAAAGGATTTCATGCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT
CGCCCACGTTAACTTCGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCGGGACGCACTCC
AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT
GTTTGTCACCTCCGCCTTCACCGTGGTGGGCATGCTCGTGTTTATCCTGAGGAACCGAAAGC
GGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATTCCCTACAGACAGTCTGT
GACTCTTCCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCCACAGAGTGTATGACTG
TGGCTCTCACTCGCTCTCAGAC**TAA**GACCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG
ATACATCCTTCCCCACCGCAGGCACCCCGGGGGCTGGAGGGGCGTGTACCCAAATCCCCGCG
CCATCAGCCTGGATGGGCATAAGTAGATAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT
GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA
GCCAGCTCGCTCTTTGCTGAGAGCCCCTTTTGACAGAAAGCCCAGCACGACCCTCCTGGAAG
AACTGACAGTGCCCTCGCCCTCGGCCCCGGGGCCTGTGGGGTTGGATGCCGCGGTTCTATAC
ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTTCCCCTGTGGATTAG
CCCCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGGCAGTTGCACGAAGGCATGAATGTAT
TGTAAATAAGTAACTTTGACTTCTGAC

FIGURE 34

MLLWILLLETSLCFAAGNVTGDVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYH
LFLHGNSLTRLFPNEFANFYNAVSLHMENNGLHEIVPGAFLGLQLVKRLHINNNKIKSFRKQ
TFLGLDDLEYLQADFNLLRDIDPGAFQDLNKLEVLILNDNLISTLPANVFQYVPITHLDLRG
NRLKTLPYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQ
GKDLNETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFXTNGQEDHATPGSAPNGGTK
IPGNWQIKIRPTAAIATGSSRNKPLANSLPCPGGCSCDHIPGSGLKMNCNNRNVSSLADLKP
KLSNVQELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLLDLRWLYMDSNY
LDTLSRSKFAGLQNLEYLNVEYNAIQLILPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSL
SKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGNPWECSCTIVPFKQWAERLGSEVLMSDLKC
ETPVNFFRKDFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL
VPGLLLVFVTSAFTVVGMLVFILRNRKRSKRRDANSSASEINSLQTVCDSSYWHNGPYNADG
AHRVYDCGSHSLSD

Signal sequence:
amino acids 1-15

Transmembrane domain:
amino acids 618-638

N-glycosylation site.
amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 122-126, 646-650

Casein kinase II phosphorylation site.
amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577, 608-612, 657-661, 666-670, 693-697

N-myristoylation site.
amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349, 354-360, 465-471, 493-499, 598-604, 603-609

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 337-348

FIGURE 35

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCCTGACCCCAGAATAACTCAGGGC
TGCACCGGGCCTGGCAGCGCTCCGCACACATTTCCTGTCGCGGCCTAAGGGAAACTGTTGGC
CGCTGGGCCCGCGGGGGGATTCTTGGCAGTTGGGGGGTCCGTCGGGAGCGAGGGCGGAGGGG
AAGGGAGGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC
AGCTCTGCGTCCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGCGGGGCCTCAG
AGAATGAGGCCGGCGTTCGCCCTGTGCCTCCTCTGGCAGGCGCTCTGGCCCGGGCCGGGCGG
CGGCGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCTCGGGGGCCTGCTACAGCCTGC
ACCACGCTACCATGAAGCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCGCTC
AGCACCGTGCGTGCGGGCGCCGAGCTGCGCGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGG
GCCCGGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTCGCACTGGAGCGCAGGCGTTCCCACT
GCACCCTGGAGAACGAGCCTTTGCGGGGTTTCTCCTGGCTGTCCTCCGACCCCGGCGGTCTC
GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCGCGGAGATGCGC
GGTACTCCAGGCCACCGGTGGGGTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC
GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTGTCCTGCGCCGCGCCCCGGG
GCCGCCTCTAACTTGAGCTATCGCGCGCCCTTCCAGCTGCACAGCGCCGCTCTGGACTTCAG
TCCACCTGGGACCGAGGTGAGTGCGCTCTGCCGGGGACAGCTCCCGATCTCAGTTACTTGCA
TCGCGGACGAAATCGGCGCTCGCTGGGACAAACTCTCGGGCGATGTGTTGTGTCCCTGCCCC
GGGAGGTACCTCCGTGCTGCCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG
CTTTGCCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCA
GTGGGGAAGGACAGCCGACCCTTGGGGGGACCGGGGTGCCCACCAGGCGCCCGCCGGCCACT
GCAACCAGCCCCGTGCCGCAGAGAACATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC
ACCACTTGTCCCTGAACAAGACAATTCAGTAACATCTATTCCTGAGATTCCTCGATGGGGAT
CACAGAGCACGATGTCTACCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC
CCATCAGGGAGCGTGATTTCCAAGTTTAATTCTACGACTTCCTCTGCCACTCCTCAGGCTTT
CGACTCCTCCTCTGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTTGGTGATCT
TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTCACGAAAGCCCCTCTTCCCAGCCA
AGGAAGGAGTCTATGGGCCCGCCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTC
CAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG
CAGAGGGTGCCTTGCTGGCGGAGTCCCCTCTTGGCTCTAGTGATGCATAGGGAAACAGGGGA
CATGGGCACTCCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAACTTAC
TTGTGTAACTGACAATTTCTGCAGAAATCCCCCTTCCTCTAAATTCCCTTTACTCCACTGAG
GAGCTAAATCAGAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGA
TGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTCGGAGAA
TTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATATAATTTACA
TTAAAAAATAATTTCTACCAAAATGGAAAGGAAATGTTCTATGTTGTTCAGGCTAGGAGTAT
ATTGGTTCGAAATCCCAGGGAAAAAAATAAAAATAAAAAATTAAAGGATTGTTGAT

FIGURE 36

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALS
TVRAGAELRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE
SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLCPAPRPGA
ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG
RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSGEGQPTLGGTGVPTRRPPATA
TSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP
SGSVISKFNSTTSSATPQAFDSSSAVVFIFVSTAVVVLVILTMTVLGLVKLCPHESPSSQPR
KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

Signal sequence:
amino acids 1-16

Transmembrane domain:
amino acids 399-418

N-glycosylation site.
amino acids 189-193, 381-385

Glycosaminoglycan attachment site.
amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 98-102, 434-438

Casein kinase II phosphorylation site.
amino acids 275-279, 288-292, 342-346, 445-449

N-myristoylation site.
amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157, 185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469, 477-483

Aspartic acid and asparagine hydroxylation site.
amino acids 262-274

FIGURE 37

CGGACGCGTGGGATTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG
CGTCGAGTCAGACGGCACCATAATCGCCTTTAAAAGTGCCTCCGCCCTGCCGGCCGCGTATC
CCCCGGCTACCTGGGCCGCCCCGCGGCGGTGCGCCCGTGAGAGGGAGCGCGCGGGCAGCCGA
GCGCCGGTGTGAGCCAGCGCTGCTGCCAGTGTGAGCGGCGGTGTGAGCGCGGTGGGTGCGGA
GGGGCGTGTGTGCCGGCGCGCGCGCCGTGGGGTGCAAACCCGAGCGTCTACGCTGCC**ATG**A
GGGGCGCGAACGCCTGGGCGCCACTCTGCCTGCTGCTGGCTGCCGCCACCCAGCTCTCGCGG
CAGCAGTCCCCAGAGAGACCTGTTTTCACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT
TATTGGCAGTGAAGGTTTTCCTGGAGTGTACCCTCCAAATAGCAAATGTACTTGGAAAATCA
CAGTTCCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCATAGACCTCGAGAGTGACAAC
CTGTGCCGCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTGGCCG
CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCAGA
TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA
AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAAC
CCCCAACTGGCCAGACCGGGATTACCCTGCAGGAGTCACTTGTGTGTGGCACATTGTAGCCC
CAAAGAATCAGCTTATAGAATTAAAGTTTGAGAAGTTTGATGTGGAGCGAGATAACTACTGC
CGATATGATTATGTGGCTGTGTTTAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGGAAA
GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAAATGAACTTCTTATTCAGT
TTTTATCAGACTTAAGTTTAACTGCAGATGGGTTTATTGGTCACTACATATTCAGGCCAAAA
AAACTGCCTACAACTACAGAACAGCCTGTCACCACCACATTCCCTGTAACCACGGGTTTAAA
ACCCACCGTGGCCTTGTGTCAACAAAAGTGTAGACGGACGGGGACTCTGGAGGGCAATTATT
GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCACAACCATCACTCGCGATGGGAGTTTG
CACGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTCAGCAGGCGGG
CAAGAACATGAGTGCCAGGCTGACTGTCGTCTGCAAGCAGTGCCCTCTCCTCAGAAGAGGTC
TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCCAGGCAAAATCATGCCAAACAGC
TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG
T**TAA**CAGTGAACTGTGTCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTC
TCTCAGTAGAAAAAAAAATACTTATAAAATTACATACTCTGAAAGAGGATTCCGAAAGATGG
GACTGGTTGACTCTTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAGTTCTT
TGCCTGCTGTCAGAGGAGCAGCTATCTGATTGGAAACCTGCCGACTTAGTGCGGTGATAGGA
AGCTAAAAGTGTCAAGCGTTGACAGCTTGGAAGCGTTTATTTATACATCTCTGTAAAAGGAT
ATTTTAGAATTGAGTTGTGTGAAGATGTCAAAAAAAGATTTTAGAAGTGCAATATTTATAGT
GTTATTTGTTTCACCTTCAAGCCTTTGCCCTGAGGTGTTACAATCTTGTCTTGCGTTTTCTA
AATCAATGCTTAATAAAATATTTTTAAAGGAAAAAAAAAAAA

FIGURE 38

MRGANAWAPLCLLLAAATQLSRQQSPERPVPTCGGILTGESGPIGSEGFPGVYPPNSKCTWK
ITVPEGKVVVLNFRFIDLESDNLCRYDFVDVYNGHANGQRIGRFCGTFRPGALVSSGNKMMV
QMISDANTAGNGFMAMFSAAEPNERGDQYCGGLLDRPSGSFKTPNWPDRDYPAGVTCVWHIV
APKNQLIELKFEKFDVERDNYCRYDYVAVFNGGEVNDARRIGKYCGDSPPAPIVSERNELLI
QFLSDLSLTADGFIGHYIFRPKKLPTTTEQPVTTTPPVTTGLKPTVALCQQKCRTGTLEGN
YCSSDFVLAGTVITTITRDGSLHATVSIINIYKEGNLAIQQAGKNMSARLTVVCKQCPLLRR
GLNYIIMGQVGEDGRGKIMPNSFIMMFKTNQKLLDALKNKQC

Signal sequence:
amino acids 1-23

N-glycosylation site.
amino acids 355-359

Casein kinase II phosphorylation site.
amino acids 64-68, 142-146, 274-278

Tyrosine kinase phosphorylation site.
amino acids 199-208

N-Myristoylation site.
amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295, 305-311, 309-315, 320-326, 330-336

Cell attachment sequence.
amino acids 149-152

FIGURE 39

CGGACGCGTGGGCGGACGCGTGGGCGGCCCACGGCGCCCGCGGGCTGGGGCGGTCGCTTCTT
CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCCATGGCCCCCGAAGG
GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCCTCAACCTCCCAGGACCTATCTGG
CTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTG
CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACTTTGGAG
GTGGAAACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTG
GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT
GACTGAGGAGCCTGGTGGAGAGCTGGTGGTTTCACAAGCAGCAGGAGGCCCCGGACCTCTTCC
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGC
CTTCCCTGTCCTCGGGGAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGG
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCC
AGTGTGGCCTTGGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTCGGCTTGT
TTTGGCCCCTGTGCCCGATGCTCAGGACCTGAGGAATCAAACTGTTTGCAATGCAAGAAGGG
CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACT
GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG
GCCTGCCTAGGCTGCATGGGGGCAGGGCCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCA
GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA
ACAAGCAGTGTGAAAACACCGAGGCCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC
AGAAGACGAGTTGGTGGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA
CGCTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG
ACTGGCTACTGGTTGTCAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATA
ATCGCGGCCACCACCTGTAGGACCTCCTCCCACCCACGCTGCCCCCAGAGCTTGGGCTGCCC
TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTG
CCTTACAGAGCAGCCCAGGTACCCAGGCCCGGCCAGACAAGGCCCCTGGGGTAAAAAGTAGC
CCTGAAGGTGGATACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGA
ATTTCAAAAGTTTTTCCTTAATGGTGGCTGCTAGAGCTTTGGCCCCTGCTTAGGATTAGGTG
GTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGT
TCTGTGTTCACCACATCCCCACACCCCATTGCCACTTATTTATTCATCTCAGGAAATAAAGA
AAGGTCTTGGAAAGTTAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 40

MAPWPPKGLVPAVLWGLSLFLNLPGPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLER
TIRDNPGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQ
QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAG
YGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPEESNCLQCKKGWALHHLKCVDIDE
CGTEGANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECE
TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDELVVLQQMFFG
IIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

Signal sequence:
amino acids 1-29

Transmembrane domain:
amino acids 372-395

N-glycosylation site.
amino acids 79-83, 205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 290-294

Casein kinase II phosphorylation site.
amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

N-myristoylation site.
amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179, 177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289, 326-332, 372-378, 395-401

Aspartic acid and asparagine hydroxylation site.
amino acids 321-333

EGF-like domain cysteine pattern signature.
amino acids 181-193

FIGURE 41

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA
GCACCATGCAGCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCCTGGCCAGCCCCGGG
GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGT
GCCCACCCTGGACAGGGCCGACATGGAGGAGCTGGTCATCCCCACCCACGTGAGGGCCCAGT
ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGGAAAGAGGTTCAGCCAGAGC
TTCCGAGAGGTGGCCGGCAGGTTCCTGGCGTTGGAGGCCAGCACACACCTGCTGGTGTTCGG
CATGGAGCAGCGGCTGCCGCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG
AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGCTGTCCCCGCGCAGCGCCCGGGCC
CGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACCTCCCTCATCGA
CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA
ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTGCTACAGGTGTCGGTGCAGAGG
GAGCATCTGGGCCCGCTGGCGTCCGGCGCCCACAAGCTGGTCCGCTTTGCCTCGCAGGGGCC
GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCCTGGACCTTGGGGACTATGGAG
CTCAGGGCGACTGTGACCCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG
ATGTACATTGACCTGCAGGGGATGAAGTGGGCCGAGAACTGGGTGCTGGAGCCCCCGGGCTT
CCTGGCTTATGAGTGTGTGGGCACCTGCCGGCAGCCCCCGGAGGCCCTGGCCTTCAAGTGGC
CGTTTCTGGGGCCTCGACAGTGCATCGCCTCGGAGACTGACTCGCTGCCCATGATCGTCAGC
ATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGTGCAGAA
GTGCAGCTGTGCCTCGGATGGTGCGCTCGTGCCAAGGAGGCTCCAGCCATAGGCGCCTAGTG
TAGCCATCGAGGGACTTGACTTGTGTGTGTTTCTGAAGTGTTCGAGGGTACCAGGAGAGCTG
GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTC
CCTCTGACAAGTTACCTCACCTAATTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA
GAGCCCTTGCTCAGTTTTCTCTATTCTTATTATTCACTGCACTATATTCTAAGCACTTACAT
GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCANTGTGTCATTGTTTACTTGTCCTGTCAC
TGGATCTGGGCTAAAGTCCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT
TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCT
AAAA

FIGURE 42

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV
ALLQRSHGDRSRGKRFSQSFREVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP
VPKAALHRHGRLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF
WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ
GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQPPEALAFKWPF
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSCASDGALVPRRLQP

Signal sequence:
amino acids 1-18

N-glycosylation site.
amino acids 158-162

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 76-80

Casein kinase II phosphorylation site.
amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-myristoylation site.
amino acids 19-25, 156-162, 225-231, 260-266, 274-280

Amidation site.
amino acids 74-78

TGF-beta family signature.
amino acids 282-298

FIGURE 43

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCGATGGGGACAAA
GGCGCAAGTCGAGAGGAAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
CATTGGGCAGTGTTACAGTGCACTCTTCTGAACCTCAAGTCAGAATTCCTGAGAATAATCCT
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
TGACCTTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCCGG
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT
GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCT
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGGTCATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT
CTTGGTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTCCTGGTGTGAGCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT
CAGGTGCTACCGGACTCTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC
TACACCCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC
ATCCTCCTTCATGCCCTCCCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACTTGTTTAAA
GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGCAAAAATGGCGGGGGTCGCAGGAATCTGCACTCAACTGCCCACCTGGC
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTCTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG
TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGTCCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAACTGGAGGCTGGGCGCAGTGGCTCACGCCTG
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA
ACATGGAGAAACCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAAACTCCAGCTCAAAAAAAAAAAAAAAA

FIGURE 44

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW
KFDQGDTTRLVCYNNKITASYEDRVTFLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWFKDGIVMPTNPKSTRAFSNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVAAVLVTLI
LLGILVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:
amino acids 1-27

Transmembrane domain:
amino acids 238-255

N-glycosylation site.
amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 270-274

Casein kinase II phosphorylation site.
amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158, 193-197, 203-207, 287-291

N-myristoylation site.
amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

FIGURE 45

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGCATGAGCGGCGGTTGGATGGCGCAGGTTGGA
GCGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTGCCGGCCTCGGACTAGGCCT
GGAGGCCGCCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCAGGCCCCAGCTCAG
GCTCGTGCCCACCCACCAAGTTCCAGTGCCGCACCAGTGGCTTATGCGTGCCCCTCACCTGG
CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCC
ATGTACCCAGAAAGGGCAATGCCCACCGCCCCCTGGCCTCCCCTGCCCCTGCACCGGCGTCA
GTGACTGCTCTGGGGGAACTGACAAGAAACTGCGCAACTGCAGCCGCCTGGCCTGCCTAGCA
GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCA
CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG
GGGATGCCACAACCATGGGGCCCCCTGTGACCCTCGAGAGTGTCACCTCTCTCAGGAATGCC
ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCCTCTGTCGGGAATGCCACATCCTC
CTCTGCCGGAGACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC
TCAGTGCAAGCCTGGTCACCGCCACCCTCCTCCTTTTGTCCTGGCTCCGAGCCCAGGAGCGC
CTCCGCCCACTGGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC
CTCGCTGCCCTGAGGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA
GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCC
TGGACACTCCCTATGGAGATCCGGGGAGCTAGCATGGGGAACCTGCCACAGCCAGAACTGAG
GGGCTGGCCCCAGGCAGCTCCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG
CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

FIGURE 46

MSGGWMAQVGAWRTGALGLALLLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR
TSGLCVPLTWRCDRDLDCSDGSDEEECRIEPCTQKGQCPPPPGLPCPCTGVSDCSGGTDKKL
RNCSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVT
LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAAVLSASLVTATLL
LLSWLRAQRLRPLGLLVAMKESLLLSEQKTSLP

Signal sequence:
amino acids 1-30

Transmembrane domain:
amino acids 230-246

N-glycosylation site.
amino acids 126-130, 195-199, 213-217

Casein kinase II phosphorylation site.
amino acids 84-88, 140-144, 161-165, 218-222

N-myristoylation site.
amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218, 224-230, 230-236, 263-269

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 44-55

Leucine zipper pattern.
amino acids 17-39

FIGURE 47

CCCACGCGTCCGGTCTCGCTCGCTCGCCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG
GTTAGACTGGCGGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGA
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG
GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCCAGAGACTCTTG
GCCGTGATCCTGTGGTTTCAGCTGGCCCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTT
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG
GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG
GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA
TAATTCCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA
ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG
TACCCCGACCTACACAATATGGTTTCATTATGTCGCGATGATGGAACGTGGAATAATCTGCC
CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC
TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCCGGTG
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCGTCT
GCCACCCGCGGCCTTGTGAGCGCTACAACCACGGAACTGTGGTGGAGTTTTACTGCGATCCT
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCCTTC
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA
CCACGTGGAAGATTGTGGCGTTCACGGCAACCAGTGTGCTGCTGGTGCTGCTGCTCGTCATC
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCCAGGGGGCCTCCCCGGAG
TTCCAGCAGTGACCCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG
ACGAAGCTGTGAGTGGCGGCTTGAGTGCCTTAGGCCCCGGGTACATGGCCTCTGTGGGCCAG
GGCTGCCCCTTACCCGTGGACGACCAGAGCCCCCAGCATACCCCGGCTCAGGGGACACGGA
CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA
GTCTGTATTCACCTCCCAGGTGCCAAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATA
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCCAGGCATCCATCATGCCCACTGGGT
GTTGTTCCTAAGAAACTGATTGATTAAAAAATTTCCCAAAGTGTCCTGAAGTGTCTCTTCAA
ATACATGTTGATCTGTGGAGTTGATTCCTTTCCTTCTCTTGGTTTTAGACAAATGTAAACAA
AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC
CTGTTTCTTCTTGACACAGACTGATTAAAAATTAAAAGNAAAAAA

FIGURE 48

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGPDDLQVCADPGI
PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHPNGTLGWIPSDNSICVQEDCRI
PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLAS
SNGYVNISELQTSFPVGTVISYRCTPGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLP
PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT
WPSTHETLLTTWKIVAFTATSVLLVLLLVILARMFQTKPKAEPPPRGPPRSSSSDPDFVVVD
GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGDTDTGPGESETCDS
VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEEVASTSPGIHHAHWVLFLRN

Signal sequence:
amino acids 1-41

Transmembrane domain:
amino acids 325-344

N-glycosylation site.
amino acids 104-108, 134-138, 192-196

Casein kinase II phosphorylation site.
amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366, 364-368, 380-384, 467-471, 468-472

N-myristoylation site.
amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424, 478-484

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 394-405

FIGURE 49

CCCACGCGTCCGCTCCGCGCCCTCCCCCCCGCCTCCCGTGCGGTCCGTCGGTGGCCTAGAGA
TGCTGCTGCCGCGGTTGCAGTTGTCGCGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGT
AGCGCCCGAGTGTCGGGGGGCGCACCCGAGTCGGGCC**ATG**AGGCCGGGAACCGCGCTACAGG
CCGTGCTGCTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTCGCCTGCTGAGTGCC
TCGGATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCTTGTTA
TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT
GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAAACTGATAGAA
AAGTTCATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGA
GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATAT
CACAATTTAGGAACTCCTATGTGGATGAGCCGTCCTGCGGCAGCGAGGTCTGCGTGGTCATG
TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCCTACATGTTCCAGTGGAATGATGA
CCGGTGCAACATGAAGAACAATTTCATTTGCAAATATTCTGATGAGAAACCAGCAGTTCCTT
CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG
GAAGAAGATGCCAAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT
CCTAATCCCCAGCATTCCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTT
GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCCTAGCACAAAGAAGCAACACACCATC
TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA
AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCATTCCGAGTGTGTT
CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA
GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA
GTTCTCCCCAGACCAAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG
GTTAT**TAG**GACATATAAAAAACTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAAATC
CTCTTATTTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT
GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTAT
CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTA
GAGTCTCAATAAATGTCACTTGGTTGGTTGTATCTAACTTCTAAGGGACAGAGCTTTACCTG
GCAGTGATAAAGATGGGCTGTGGAGCTTGGAAAACCACCTCTGTTTTCCTTGCTCTATACAG
CAGCACATATTATCATACAGACAGAAAATCCAGAATCTTTTCAAAGCCCACATATGGTAGCACAG
GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTTCAAAGAATAAAATCAAATAAAGA
GCAGGAAAAAAAAA

FIGURE 50

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRL
NFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDL
YAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY
SDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV
VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRPDL
KNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKES
GWVENEIYCY

Signal sequence:
amino acids 1-21

Transmembrane domain:
amino acids 235-254

N-glycosylation site.
amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 296-300

Casein kinase II phosphorylation site.
amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226, 299-303, 306-310, 323-327

N-myristoylation site.
amino acids 18-24, 37-43, 76-82, 146-152

FIGURE 51

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGC
ATCCGCAGGTTCCCGCGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT
GTTTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCACTGGTGT
CTTCAGCATGCGCTTGTGGACCCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACC
AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC
CTGCTGAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCC
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAACTC
AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT
CAATACCATGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT
GCAGCAAATGTTTGCCTTGGGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC
TTTCACCAACCTTCAACCCACAGGAGGTCTTTATTCGTTCCACTAACATTTTTCGGAATCTG
GAGTCCACCCGTTGTTTGCTGGCTGGGCTTTTCCAGTGTCAGAAAGAAGGACCCATCATCAT
CCACACTGATGAAGCAGATTCAGAAGTCTTGTATCCCAACTACCAAAGCTGCTGGAGCCTGA
GGCAGAGAACCAGAGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTG
AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTTG
CACCGATGATCGAACAGAGAGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGG
GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC
CATGGACTCTGCCACTGCCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG
TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTGACCACAAATGGCCACCGTTTGCT
GTTGACCTGACCATGGAACTTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA
TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGT
TCTTGAATGCCATGTCAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA
ACTCAGGTGATGGAAGTTGGAAATGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATT
TTAAAATAAAGTGCCTTTATACAATG

FIGURE 52

MITGVPSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVPRHGAR
SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMPAGQL
TKVGMQQMPALGERLRKNYVEDIPPLSPTFNPQEVFIRSTNIPRNLESTRCLLAGLFQCQKE
GPIIIHTDEADSEVLYPNYQSCWSLRQRTRGRRQTASLQPGISEDLKKVKDRMGIDSSDKVD
PFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFLHILES
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEW
FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMSVYTLSPEKYHALCSQTQVMEVGNEE

Signal sequence:
amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 218-222

Casein kinase II phosphorylation site.
amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site.
amino acids 280-288

N-myristoylation site.
amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.
amino acids 216-220

Leucine zipper pattern.
amino acids 10-32

Histidine acid phosphatases phosphohistidine signature.
amino acids 50-65

FIGURE 53

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT
TAAATTTCAGCTCATCACCTTCACCTGCCTTGGTCATGGCTCTGCTATTCTCCTTGATCCTT
GCCATTTGCACCAGACCTGGATTCCTAGCGTCTCCATCTGGAGTGCGCCTGGTGGGGGGCCT
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG
ACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCA
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG
ATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCTTTCTCCCCA
GTCCCAGAGGGTGTCAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA
GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCCGCAAAGGTGG
TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCA
GGATTGCCCTTCTGGGCCTTGGGGGAACAACACCTGCAACCATGATGAAGACACGTGGGTCG
AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGA
CCAGGTGGTATGCAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTCGTTGCTCAGGGGAGGAG
CAGTCCCTGGAGCAGTGCCAGCACAGATTTTGGGGGTTTCACGACTGCACCCACCAGGAAGA
TGTGGCTGTCATCTGCTCAGTGTAGGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA
GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT
TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCCTGATTCTCAGGCGTTCAGAGTTGG
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACACAGTTTGGAACTACATCA
CCACCTTTCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACACACATA
CACCATTTGTCCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA
TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACTTTATTTACAATAATAAAGATAGCAC
TATGTGTTCAAA

FIGURE 54

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC
RELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEEVYDCSHDEDAGASC
ENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCQTGWSLRAAKVVCRQLGCGRAVL
TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDLRLVG
GDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL
DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

Signal sequence:
amino acids 1-15

Casein kinase II phosphorylation site.
amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238, 267-271, 294-298, 316-320, 336-340

N-myristoylation site.
amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143, 180-186, 263-269, 286-292

Amidation site.
amino acids 196-200

Speract receptor repeated domain signature.
amino acids 29-67, 249-287

FIGURE 55

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG
GTGAAGGCCATGGACTTCATCACCTCCACAGCCATCCTGCCCCTGCTGTTCGGCTGCCTGGG
CGTCTTCGGCCTCTTCCGGCTGCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCGGAATGCTG
TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT
GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA
ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC
TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC
GACATACTTGTCAACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA
TGTGGACAAGAGGGTCATGGAGACAAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC
TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG
ATGAGCATTCCTTTTCGATCAGCATATGCAGCCTCCAAGCACGCAACCCAGGCTTTCTTTGA
CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA
TCCACACCAACCTCTCTGTAAATGCCATCACCGCGGATGGATCTAGGTATGGAGTTATGGAC
ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTTCCTTGGCTGTTTATCTTCGAA
CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC
AAGAACTCCTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC
TTACTCTACAAGGGACAGTTGCATTTGCTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG
AAAGACTGAAGAAACACATCTCGTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA
AGCTTCTTCCCAGGGTGAGGGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACACT
AAAAACTAGAAATAAACATCTCAAACAGTAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAG
AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTAC

FIGURE 56

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA
KLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL
VNNAGISYRGTIMDTTVDVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAISSIQGKMSI
PFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDTTT
AQGRSPVEVAQDVLAAVGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSRNS

Signal sequence:
amino acids 1-21

Transmembrane domain:
amino acids 104-120, 278-292

N-glycosylation site.
amino acids 228-232

Glycosaminoglycan attachment site.
amino acids 47-51

Casein kinase II phosphorylation site.
amino acids 135-139, 139-143, 253-257

Tyrosine kinase phosphorylation site.
amino acids 145-153, 146-153

N-myristoylation site.
amino acids 44-50, 105-111, 238-244, 242-248, 291-297

Amidation site.
amino acids 265-269

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 6-17

FIGURE 57

CCCACGCGTCCGCTGGTGTTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA
AAAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGATGAAATTTCTTCTGGACATCCTC
CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCCTAA
GAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTACAGGAGCTGGGCATGGAATTGGGA
GACTGACTGCCTATGAATTTGCTAAACTTAAAAGCAAGCTGGTTCTCTGGGATATAAATAAG
CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTCATACCTTTGT
GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG
GAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTCTATACATCAGATTTGTTTGCTACA
CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA
GGCATTTCTTCCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG
CTGGACATGTCTCGGTCCCCTTCTTACTGGCTTACTGTTCAAGCAAGTTTGCTGCTGTTGGA
TTTCATAAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAAACAACATG
TCTGTGTCCTAATTTCGTAAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA
CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGGATTCTGACTGAGCAGAAGATG
ATTTTTATTCCATCTTCTATAGCTTTTTTAACAACATTGGAAAGGATCCTTCCTGAGCGTTT
CCTGGCAGTTTTAAAACGAAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA
AAGCGCAATAAGCACCTAGTTTTCTGAAAACTGATTTACCAGGTTTAGGTTGATGTCATCTA
ATAGTGCCAGAATTTTAATGTTTGAACTTCTGTTTTTTTCTAATTATCCCCATTTCTTCAATA
TCATTTTTGAGGCTTTGGCAGTCTTCATTTACTACCACTTGTTCTTTAGCCAAAAGCTGATT
ACATATGATATAAACAGAGAAATACCTTTAGACGTGACTTTAAGGAAAATGAAGAAAAAGAA
CCAAAATGACTTTATTAAAATAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA
AATTTGTACCATAACCGTTTATTTAACATATATTTTTATTTTTGATTGCACTTAAATTTTGT
ATAATTTGTGTTTCTTTTTCTGTTCTACATAAAATCAGAAACTTCAAGCTCTCTAAATAAAA
TGAAGGACTATATCTAGTGGTATTTCACAATGAATATCATGAACTCTCAATGGGTAGGTTTC
ATCCTACCCATTGCCACTCTGTTTCCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT
GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG
GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 58

MKFLLDILLLLPLLIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK
LVLWDINKHGLEETAAKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAEIGDVSILVNNAGVV
YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC
SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVVNRLMH
GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMKAQ

Signal sequence:
amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 30-34, 283-287

Casein kinase II phosphorylation site.
amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.
amino acids 43-49, 72-78, 122-128, 210-216

FIGURE 59

CCCACGCGTCCGCGGACGCGTGGGTCGACTAGTTCTAGATCGCGAGCGGCCGCCCGCGGCTC
AGGGAGGAGCACCGACTGCGCCGCACCCTGAGAGATGGTTGGTGCCATGTGGAAGGTGATTG
TTTCGCTGGTCCTGTTGATGCCTGGCCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT
GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC
TGGGAAGATCCAAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCCAGGACTGAACATGA
AGAGTTATGCCGGCTTCCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGGTTC
TTCCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG
AGGTTCATCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCACAAGTAACATGA
CCTTGCGTGACAGAGACTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA
GTGGGCACAGGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT
AGCACGGGATTTATACAGTGCACTAATTCAGTTTTTCCAGATATTTCCTGAATATAAAAATA
ATGACTTTTATGTCACTGGGGAGTCTTATGCAGGGAAATATGTGCCAGCCATTGCACACCTC
ATCCATTCCCTCAACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA
TGGATATTCTGATCCCGAATCAATTATAGGGGGCTATGCAGAATTCCTGTACCAAATTGGCT
TGTTGGATGAGAAGCAAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC
AGGAAGCAGAACTGGTTTGAGGCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC
AAGTGATCCTTCTTACTTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTTTGCGGT
GCACGGAACCTGAGGATCAGCTTTACTATGTGAAATTTTTGTCACTCCCAGAGGTGAGACAA
GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACTATAGTTGAAAAGTACTTGCGAGA
AGATACAGTACAGTCAGTTAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA
TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGC
ATGGACTGGAAAGGATCCCAGGAATACAAGAAGGCAGAAAAAAAAGTTTGGAAGATCTTTAA
ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGACTTCCATCAGGTAATTATTC
GAGGTGGAGGACATATTTTACCCTATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGA
TTCATTTATGGAAAAGGATGGGATCCTTATGTTGGATAAACTACCTTCCCAAAAGAGAACAT
CAGAGGTTTTCATTGCTGAAAAGAAAATCGTAAAAACAGAAAATGTCATAGGAATAAAAAAA
TTATCTTTTCATATCTGCAAGATTTTTTTCATCAATAAAAATTATCCTTGAAACAAGTGAGC
TTTTGTTTTTGGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA
TTATTTAACTTAAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA
AATTTTAGGGTCTTGAATAGGAAGTTTTAATTTCTTCTAAGAGTAAGTGAAAAGTGCAGTTG
TAACAAACAAAGCTGTAACATCTTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT
GTTTGGAAATATTATTGGATAAGAATAGCTCAATTATCCCAAATAAATGGATGAAGCTATAA
TAGTTTTGGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA
AAAATATTATATATAAAAGTAAAAAAAAAA

FIGURE 60

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDSGQPLFLTPYIEAGKIQKGRELSL
VGPFPGLNMKSYAGFLTVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH
GPYVVTSNMTLRDRDFPWTTTLSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQF
FQIFPEYKNNDFYVTGESYAGKYVPAIAHLIHSLNPVREVKINLNGIAIGDGYSDPESIIGG
YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNVTG
CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT
EIMNNYKVLIYNGQLDIIVAAALTERSLMGMDWKGSQEYKKAEKKVWKIFKSDSEVAGYIRQ
AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

Signal sequence:
amino acids 1-22

N-glycosylation site.
amino acids 81-85, 132-136, 307-311, 346-350

Casein kinase II phosphorylation site.
amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352, 353-357, 424-428

Tyrosine kinase phosphorylation site.
amino acids 423-432

N-myristoylation site.
amino acids 22-28, 110-116, 156-162, 232-238

Serine carboxypeptidases, serine active site.
amino acids 200-208

Crystallins beta and gamma 'Greek key' motif signature.
amino acids 375-391

FIGURE 61

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTTGGCTACAACAT
TTTTCCCTTTCCTAACAAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT
GCTGGAGAAGAAAGGGCTGAGGGCAGAGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTG
CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG
TCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC
CTGGGGGAGGGCCTGCCTAACAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT
AAGACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCTAATATCAAATTGACTGGCTGGG
TGAACTTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAA
TAGAGATGCTTTGTAAAATAAAATTTTAAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA
GACCAAAATAGATAACAGGATTCCCTGAACATTCCTAAGAGGGAGAAAGTATGTTAAAAATA
GAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACCCTGGGTC
AGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTGGTCTGACCACTCTGCCTTGTGTTTT
GCAGAATCATGTGAGGGCCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCT
CCTCACCCCCGCCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG
ACAGCCGCTCTGTGGTCCTGCTCTCAGTGGTCTGGGTGCTGCTGGCCCCCCCAGCAGCCGGC
ATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT
CCACCAAGGGACGGGGGCCGTCTATGTGGGGGCCATCAACCGGGTCTATAAGCTGACAGGCA
ACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG
CCCCTCATCGTGCAGCCCTGCAGCGAAGTGCTCACCCTCACCAACAATGTCAACAAGCTGCT
CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCA
AGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTGGAGCCATCCCACAAGAAGGAGCACTAC
CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA
TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCCA
GCCGGAAGCTGCCCCGAGACCCTGAGTCCTCAGCCATGCTCGACTATGAGCTACACAGCGAT
TTTGTCTCCTCTCTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACAT
CTTCTACATCTACGGCTTTGCTAGTGGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCGAGA
CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGCGG
CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGCTGCACCCGGGC
CGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC
AGGCCTTCAATATCACCAGCCAGGACGATGTACTCTTGCCATCTTCTCCAAAGGGCAGAAG
CAGTATCACCACCCGCCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACTT
GCAGATCAAGGAGCGCCTGCAGTCCTGCTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC
TGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCATCGATGATAACTTCTGTGGA
CTGGACATCAACCAGCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTGTACACCAC
CAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTG
TGGGGACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCC
ATTCACCTCCTCAGCAAAGAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTTAACTATAG
GCAACTTTATTTTCTTGGGGAACAAACGTGAAATGGGGAGCTAAGAAGGGGTTAATTTTGTG
ACTTAGCTTCTAGCTACTTCCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA
TTTCAATATTTCCCAAACTTTAAGAAAAAACTTTAAGAAGGTACATCTGCAAAAGCAAA

FIGURE 62

MGTLGQASLFAPPGNYFWSDHSALCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP
RALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTFHSENRDWTFNHLTVHQGTGAVYVGAINRV
YKLTGNLTIQVAHKTGPEEDNKSRYPPLIVQPCSEVLTLTNNVNKLLIIDYSENRLLACGSL
YQGVCKLLRLDDLFILVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDY
FPTLSSRKLPRDPESSAMLDYELHSDFVSSLIKIPSDTIALVSHFDIFYIYGFASGGFVYFL
TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP
GDSLAQAFNITSQDDVLFAIFSKGQKQYHHPPDDSALCAFPIRAINLQIKERLQSCYQGEGN
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG
YSVVFVGTKSGKLKKVRVYEFRCSNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

Signal sequence:
amino acids 1-32

Transmembrane domain:
amino acids 71-87

N-glycosylation site.
amino acids 130-134, 145-149, 217-221, 381-385

Casein kinase II phosphorylation site.
amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387, 384-388, 471-475, 481-485, 530-534

N-myristoylation site.
amino acids 220-226, 319-325, 353-359, 460-466, 503-509

FIGURE 63

AGGCTCCCGCGCGCGGCTGAGTGCGGACTGGAGTGGGAACCCGGGTCCCCGCGCTTAGAGAACACGCGATGACCA
CGTGGAGCCTCCGGCGGAGCCGGCCCGCACGCTGGGACTCCTGCTGCTGGTCGTCTTGGGCTTCCTGGTGCTCC
GCAGCCTGGACTGGAGCACCCTGGTCCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGCCCAAGGGCTGGA
ACTTCATGCTGGAGGATTCCACCTTCTGGATCTTCGGGGGCTCCATCCACTATTTCCGTGTGCCCAGGGAGTACT
GGAGGGACCGCCTGCTGAAGATGAAGGCCTGTGGCTTGAACACCCTCACCACCTATGTTCCGTGGAACCTGCATG
AGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCTTCGTCCTGATGGCCGCAGAGATCG
GGCTGTGGGTGATTCTGCGTCCAGGCCCCTACACCTGCAGTGAGATGGACCTCGGGGGCTTGCCCAGCTGGCTAC
TCCAAGACCCTGGCATGAGGCTGAGGACAACTTACAAGGGCTTCACCGAAGCAGTGGACCTTTATTTTGACCACC
TGATGTCCAGGGTGGTGCCACTCCAGTACAAGCGTGGGGGACCTATCATTGCCGTGCAGGTGGAGAATGAATATG
GTTCCTATAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGGAACTGC
TCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGATTGTCCAGGGAGTCTTGGCCACCATCAACTTGCAGT
CAACACACGAGCTGCAGCTACTGACCACCTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGATGGTGATGGAGT
ACTGGACGGGGTGGTTTGACTCGTGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGT
CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCACCAACTTTGGCTTCATGAATG
GAGCCATGCACTTCCATGACTACAAGTCAGATGTCACCAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG
ATTACACGGCCAAGTACATGAAGCTTCGAGACTTCTTCGGCTCCATCTCAGGCATCCCTCTCCCTCCCCCACCTG
ACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGTCTTGTACCTGTCTCTGTGGGACGCCCTCAAGTACC
TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCCT
TCGGGTACATTCTCTATGAGACCAGCATCACCTCGTCTGGCATCCTCAGTGGCCACGTGCATGATCGGGGGCAGG
TGTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT
ACACCGTGCTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGAGAATATTGATGACCAGCGCAAAG
GCTTAATTGGAAATCTCTATCTGAATGATTCACCCCTGAAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA
GCTTCTTTCAGAGGTTCGGCCTGGACAAATGGNGTTCCCTCCCAGAAACACCCACATTACCTGCTTTCTTCTTGG
GTAGCTTGTCCATCAGCTCCAGGCCTTGTGACACCTTTCTGAAGCTGGAGGGCTGGGAGAAGGGGGTTGTATTCA
TCAATGGCCAGAACCTTGGACGTTACTGGAACATTGGACCCCAGAAGACGCTTTACCTCCCAGGTCCCTGGTTGA
GCAGCGGAATCAACCAGGTCATCGTTTTTGAGGAGACGATGGCGGGCCCTGCATTACAGTTCACGGAAACCCCCC
ACCTGGGCAGGAACCAGTACATTAAGTGAGCGGTGGCACCCCCTCCTGCTGGTGCCAGTGGGAGACTGCCGCCTC
CTCTTGACCTGAAGCCTGGTGGCTGCTGCCCCACCCCTCACTGCAAAAGCATCTCCTTAAGTAGCAACCTCAGGG
ACCGGGGGCTACAGTCTGCCCCTGTCTCAGCTCAAAACCCTAAGCCTGCAGGGAAAGGTGGGATGGCTCTGGGCC
TGGCTTTGTTGATGATGGCTTTCCTACAGCCCTGCTCTTGTGCCGAGGCTGTCGGGCTGTCTCTAGGGTGGGAGC
AGCTAATCAGATCGCCCAGCCTTTGGCCCTCAGAAAAAGTGCTGAAACGTGCCCTTGCACCGGACGTCACAGCCC
TGCGAGCATCTGCTGGACTCAGGCGTGCTCTTTGCTGGTTCCTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT
TTTATCCCCGAAATCCTGGGTGTGTCACCAGTGTAGAGGGTGGGGAAGGGGTGTCTCACCTGAGCTGACTTTGTT
CTTCCTTCACAACCTTCTGAGCCTTCTTTGGGATTCTGGAAGGAACTCGGCGTGAGAAACATGTGACTTCCCCTT
TCCCTTCCCACTCGCTGCTTCCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCCTGCGTCTTCC
CAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATGTGAGTCCTGGCAGAAGCCATGGCCCACGTCTGCA
CATCCAGGGAGGAGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCC
AGGGAGGAGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGA
GGAGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGG
ACAGAAGGCCCAGCTCAGTGGCCCCCGCTCCCCACCCCCCACGCCCGAACAGCAGGGGCAGAGCAGCCCTCCTTC
GAAGTGTGTCCAAGTCCGCATTTGAGCCTTGTTCTGGGGCCCAGCCCAACACCTGGCTTGGGCTCACTGTCCTGA
GTTGCAGTAAAGCTATAACCTTGAATCACAA

FIGURE 64

MTTWSLRRRPARTLGLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW
IFGGSIHYFRVPREYWRDRLLKMKACGLNTLTTYVPWNLHEPERGKPDFSGNLDLEAFVLMA
AEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMSRVVPLQ
YKRGGPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLLTSDNKDGLSKGIVQGVLAT
INLQSTHELQLLTTFLFNVQGTQPKMVMEYWTGWFDSWGGPHNILDSSEVLKTVSAIVDAGS
SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTAKYMKLRDFFGSISGIP
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSFGYILYE
TSITSSGILSGHVHDRGQVFVNTVSIGFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN
IDDQRKGLIGNLYLNDSPLKNFRIYSLDMKKSFFQRFGLDKWXSLPETPTLPAFFLGSLSIS
STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLSSGINQVIVFEETMAGPA
LQFTETPHLGRNQYIK

Signal sequence:
amino acids 1-27

Casein kinase II phosphorylation site.
amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

N-myristoylation site.
amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315, 320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

FIGURE 65

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC
CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG
GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC
CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACCCC**ATG**GCTCCC
AAGAAGCTGTCCTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCA
GGCAGACACTCGGTCGTTCGTAGTGGATAGGGGTCATGACCGGTTTCTCCTAGACGGGGCCC
CGTTCCGCTATGTGTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC
CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACTA
CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA
ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG
TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTC
AGATCCAGACTTCCTTGCCGCAGTGGACTCCTGGTTCAAGCTCTTGCTGCCCAAGATATATC
CATGCCTTTATCACAATGGGGGCAACATCATTAGCATTCAGGTGGAGAATGAATATGGTAGC
TACAGAGCCTGTGACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCCGTGCACTGCTAGG
AGAAAAGATCTTGCTCTTCACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGG
GACTCTATACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT
CGGAAGTATGAACCCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA
CTGGGGCCAGAATCACTCCACACGGTCTGTGTCAGCTGTAACCAAAGGACTAGAGAACATGC
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG
AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC
TATATCTGAAGCAGGGGACCCCACACCTAAGCTTTTTGCTCTTCGAGATGTCATCAGCAAGT
TCCAGGAAGTTCCTTTGGGACCTTTACCTCCCCCGAGCCCCAAGATGATGCTTGGACCTGTG
ACTCTGCACCTGGTTGGGCATTTACTGGCTTTCCTAGACTTGCTTTGCCCCCGTGGGCCCAT
TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC
GAACCTATATGACCCATACCATTTTTGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC
CATGACCGTGCCTATGTGATGGTGGATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAG
AGACAAACTATTTTTGACGGGGAAACTGGGGTCCAAACTGGATATCTTGGTGGAGAACATGG
GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG
GGGCAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG
GTGGTTTCCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT
ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG
ACCAAGGGCCAAGTCTGGATCAATGGGTTTAACTTGGGCCGGTACTGGACAAAGCAGGGGCC
ACAACAGACCCTCTACGTGCCAAGATTCCTGCTGTTTCCTAGGGGAGCCCTCAACAAAATTA
CATTGCTGGAACTAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTTGGATAAGCCTATC
CTCAATAGCACTAGTACTTTGCACAGGACACATATCAATTCCCTTTCAGCTGATACACTGAG
TGCCTCTGAACCAATGGAGTTAAGTGGGCAC**TGA**AAGGTAGGCCGGCATGGTGGCTCATGC
CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTCAGGACTTCAAGA
CCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGCCGGGCGTG
ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC
AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA
GACACTCCATCTCAAAAAAAAAAAA

FIGURE 66

MAPKKLSCLRSLLLPLSLTLLLPQADTRSPVVDRGHDRFLLDGAPFRYVSGSLHYFRVPRVL
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYI
CAEWEMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE
YGSYRACDFSYMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF
TLLRKYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNMYMFHGGTNF
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPPSPKMMI
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVPN
NGVHDRAYVMVDGVPQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKP
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFYSKTFPILGSVGDTFLYL
PGWTKGQVWINGFNLGRYWTKQGPQQTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLD
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

Signal sequence:
amino acids 1-27

N-glycosylation site.
amino acids 97-101, 243-247, 276-280, 486-490, 625-629

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 4-8

Casein kinase II phosphorylation site.
amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554, 603-607, 644-648

Tyrosine kinase phosphorylation site.
amino acids 191-198

N-myristoylation site.
amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233, 231-237, 274-280, 296-300, 307-313, 447-453, 484-490

FIGURE 67

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC
ACCCACAATATGGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT
TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT
CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT
GCGTTCCTTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT
CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAACTAGTTTGAACCATGAGTGGACATTTG
AAAAACTCAGGCAGCACATTTCACGCAACGCCCAGGACAAGCAGGAGTTGCATCTGTTCATG
CTGTCGGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAACT
AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC
TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA
TGCCTTCACGTGAAGTTCACTGATGTGGCTGAAATTCCTGCCTGGGTGTATTTGCTCAAAAA
CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG
AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT
CCCTCCAACATTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTCATAATGACGGCAC
TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTCGCTGAGCTGGAACTCCAGA
ACTGTGAGCTAGAGAGAATCCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT
TTAAAGTCCAATAACATTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT
GACTTGTTTAAAATTATGGCATAACAAAATTGTTACTATTCCTCCCTCTATTACCCATGTCA
AAAACTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT
AGTTTACAGAAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA
AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTCTGC
CAAAACAATTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC
TCACTCCCAGAGAAAGTTGGTCAGCTCTCCCAGCTCACTCAGCTGGAGCTGAAGGGGAACTG
CTTGGACCGCCTGCCAGCCCAGCTGGGCCAGTGTCGGATGCTCAAGAAAAGCGGGCTTGTTG
TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA
AATATTCCCTTTGCAAATGGGATTTAAACTAAGATAATATATGCACAGTGATGTGCAGGAAC
AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG
ATACATCTTTTAAAATAAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT
GTTCAATGTTTGTAGGGTTTTAAGTCATTCATTTCCAAATCATTTTTTTTTTTCTTTGGGG
AAAGGGAAGGAAAAATTATAATCACTAATCTTGGTTCTTTTTAAATTGTTTGTAACTTGGAT
GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT
TTCTTACTAAAAAAAAAAAAAAAAA

FIGURE 68

MAYMLKKLLISYISIICVYGFICLYTLFWLFRIPLKEYSFEKVREESSFSDIPDVKNDFAFL
LHMVDQYDQLYSKRFGVFLSEVSENKLREISLNHEWTFEKLRQHISRNAQDKQELHLFMLSG
VPDAVFDLTDLDVLKLELIPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRDHLRCLH
VKFTDVAEIPAWVYLLKNLRELYLIGNLNSENNKMIGLESLRELRHLKILHVKSNLTKVPSN
ITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAELELQNCELERIPHAIFSLSNLQELDLKS
NNIRTIEEIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLESLYFSNNKLESLPVAVFSLQ
KLRCLDVSYNNISMIPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTLNLGQNCITSLP
EKVGQLSQLTQLELKGNCLDRLPAQLGQCRMLKKSGLVVEDHLFDTLPLEVKEALNQDINIP
FANGI

Signal sequence:
amino acids 1-20

N-glycosylation site.
amino acids 241-245, 248-252, 383-387

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 326-330

Casein kinase II phosphorylation site.
amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

Tyrosine kinase phosphorylation site.
amino acids 349-355, 375-381

N-myristoylation site.
amino acids 78-84, 124-130, 212-218, 392-398

FIGURE 69

CCCACGCGTCCGGCCTTCTCTCTGGACTTTGCATTTCCATTCCTTTTCATTGACAAACTGACTTTTTTTATTTCT
TTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGACATTTGTGTTTTACACACATAAGGAT
CTGTGTTTGGGGTTTCTTCTTCCTCCCCTGACATTGGCATTGCTTAGTGGTTGTGTGGGGAGGGAGACCACGTGG
GCTCAGTGCTTGCTTGCACTTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC
ATCGCTGGTGGTATCCTGGCGGCCTTGCTCCTGCTGATAGTTGTCGTGCTCTGTCTTTACTTCAAAATACACAAC
GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAAGGTGTGGTGGGCCAAG
AACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGTCCTGCCCTGCAGTGCTGTGAAGGATATAGAATGTGTGCC
AGTTTTGATTCCCTGCCACCTTGCTGTTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAAA
GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG
CTCCCCATCAGTTTCATGGAAAATAACTCAGTGCCTGCTGGGAACCAGCTGCTGGAGATCCCTACAGAGAGCTTC
CACTGGGGGCAACCCTTCCAGGAAGGAGTTGGGGAGAGAGAACCCTCACTGTGGGGAATGCTGATAAACCAGTCA
CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCGTGGCTGGAACTGACGTTTCCCTGGAGGTGTCCAGAAA
GCTGATGTAACACAGAGCCTATAAAAGCTGTCGGTCCTTAAGGCTGCCCAGCGCCTTGCCAAAATGGAGCTTGTA
AGAAGGCTCATGCCATTGACCCTCTTAATTCTCTCCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAAT
GCAAGCTGCACAGTCAGTCTAGGGGGTGCCAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC
AGTGAGAACTGCACCTGGACAATAGAAAGACCAGAAAACAAAAGCATCAGAATTATCTTTTCCTATGTCCAGCTT
GATCCAGATGGAAGCTGTGAAAGTGAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCTCTGCTAGGG
CAAGTCTGCAGTAAAAACGACTATGTTCCTGTATTTGAATCATCATCCAGTACATTGACGTTTCAAATAGTTACT
GACTCAGCAAGAATTCAAAGAACTGTCTTTGTCTTCTACTACTTCTTCTCTCCTAACATCTCTATTCCAAACTGT
GGCGGTTACCTGGATACCTTGGAAGGATCCTTCACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT
TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAAACTAAACTTCAAAGAGATTTTCCTAGAAATAGAC
AAACAGTGCAAATTTGATTTTCTTGCCATCTATGATGGCCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT
GGCCGTGTGACTCCCACCTTCGAATCGTCATCAAACTCTCTGACTGTCGTGTTGTCTACAGATTATGCCAATTCT
TACCGGGGATTTTCTGCTTCCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT
GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA
GACCCAACTTGCAGACCAAAATTATCAAATGTTGTGGAATTTTCTGTCCTCTTAATGGATGTGGTACAATCAGA
AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC
CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACA
GAAGATGATGTAATACAAAGTCAAAATGCACTGGGCAAATATAACACCAGCATGGCTCTTTTTGAATCCAATTCA
TTTGAAAAGACTATACTTGAATCACCATATTATGTGGATTTGAACCAAACTCTTTTTCTTCAAGTTAGTCTGCAC
ACCTCAGATCCAAATTTGGTGGTGTTTCTTGATACCTGTAGAGCCTCTCCCACCTCTGACTTTGCATCTCCAACC
TACGACCTAATCAAGAGTGGATGTAGTCGAGATGAAACTTGTAAGGTGTATCCCTTATTTGGACACTATGGGAGA
TTCCAGTTTAATGCCTTTAAATTCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAAGTTTTGATATGTGAT
AGCAGTGACCACCAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCATATAAATGG
AAAACAGATTCCATCATAGGACCCATTCGTCTGAAAAGGGATCGAAGTGCAAGTGGCAATTCAGGATTTCAGCAT
GAAACACATGCGGAAGAAACTCCAAACCAGCCTTTCAACAGTGTGCATCTGTTTTCCTTCATGGTTCTAGCTCTG
AATGTGGTGACTGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG
CAGAACTATTAACTAACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGT
GGCTACACATATTATGAATAAATGAGGAAGGGCCTGAAAGTGACACACAGGCCTGCATGTAAAAAAA

FIGURE 70

MELVRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI
ERPENKSIRIIPSYVQLDPDGSCESENIKVFDGTSSNGPLLGQVCSKNDYVPVFESSSSTLT
FQIVTDSARIQRTVFVFYYFPSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYCVWHIQV
EKDYKIKLNFKEIFLEIDKQCKFDFLAIYDGPSTNSGLIGQVCGRVTPTFESSSNSLTVVLS
TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIISKSYLEAFNSNGNNLQLKDPTCRP
KLSNVVEPSVPLNGCGTIRKVEDQSITYTNIITPSASSTSEVITRQKLQIIVKCEMGHNST
VEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN
LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRPQFNAFKFLRSMSSVYL
QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDSIIGPIRLKRDRSASGNSGFQHETHA
EETPNQPFNSVHLFSFMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

Signal sequence:

amino acids 1-24

Transmembrane domain:

amino acids 571-586

N-glycosylation site.

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374, 394-398, 419-423

Casein kinase II phosphorylation site.

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383, 408-412, 463-467, 520-524, 556-560

Tyrosine kinase phosphorylation site.

amino acids 172-180, 407-415, 407-416, 519-528

N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 71

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG
GGACATGCGGCCCCAGGAGCTCCCCAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGTTGCTGC
TGCTGCCGCCGCCGCCGTGCCCTGCCCACAGCGCCACGCGCTTCGACCCCACCTGGGAGTCC
CTGGACGCCCGCCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG
GGGAGTGTTTTCCGTGCCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCAAAAGGAAA
AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCCTCCTAGTTTCAAATATGAAGAT
TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAGTGGGCAGATATTTTTCAGGC
CTCTGGTGCCAAATACATTGTCTTAACTTCCAAACATCATGAAGGCTTTACCTTGTGGGGGT
CAGAATATTCGTGGAACTGGAATGCCATAGATGAGGGGCCCAAGAGGGACATTGTCAAGGAA
CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTGGACTGTACTATTCCCTTTTTGA
ATGGTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG
TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGG
TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT
ATATAATGAAAGCCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA
TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA
CATAAATGGGAAAACTGCATGACAATAGACAAACTGTCCTGGGGCTATAGGAGGGAAGCTGG
AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG
GAGGAAATCTTTTGATGAATATTGGGCCCACACTAGATGGCACCATTTCTGTAGTTTTTGAG
GAGCGACTGAGGCAAGTGGGGTCCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA
TACCTGGCGATCCCAGAATGACACTGTCACCCCAGATGTGTGGTACACATCCAAGCCTAAAG
AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCCACATCAGGACAGCTGTTCCTTGGCCAT
CCCAAAGCTATTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAACTG
GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTCATCAGATGC
CGTGTAAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG
ATGCTGCAAGTTATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGA
AAGCAATGTAAACTGGATAAGAAAATTATTTGGCAGTTCAGCCCTTTCCCTTTTTCCCACTA
AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCCATTAAAGTC
TCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTCACATTATAGTAG
CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATG
ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAAT
TTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACA
CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG
AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG
GTACAGTAAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG
CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCTAGGACATTA
TTGAACACTGCCAGACGTTATAAATACTGTATGCTTACGCTACACTACATTTATAAAAAAAA
GTTTTTCTTTCTTCAATTATAAATTAACATAAGTGTACTGTAACTTTACAAACGTTTTAATT
TTTAAAACCTTTTTGGCTCTTTTGTAATAACACTTAGCTTAAAACATAAACTCATTGTGCAA
ATGTAA

FIGURE 72

MRPQELPRLAFPLLLLLLLLLPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFIHWG
VFSVPSFGSEWFWWYWQKEKIPKYVEFMKDNYPPSFKYEDFGPLFTAKFFNANQWADIFQAS
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLFEW
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKWGWALALTNVI

Signal sequence:
amino acids 1-28

N-glycosylation site.
amino acids 171-175, 239-243, 377-381

Casein kinase II phosphorylation site.
amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319, 375-375

Tyrosine kinase phosphorylation site.
amino acids 361-369, 389-397

N-myristoylation site.
amino acids 143-149, 178-184, 255-261, 272-278, 428-434

Leucine zipper pattern.
amino acids 410-432

Alpha-L-fucosidase putative active site.
amino acids 283-295

FIGURE 73

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT
TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC
TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACG**ATG**GCCAGGTGCTTCAGCCTGGTGTTG
CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT
TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGC
AGCTGAATTTCACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG
GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTTGGAGA
TGGATTCGTGGTCATCTCTAGGATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTG
TCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACTCATCTGAT
ACTTGGACTAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA
AACTGCAACACAAACAACAGAATTTATTGTCAGTGACAGTACCTACTCGGTGGCATCCCCTT
ACTCTACAATACCTGCCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG
AGAAAAAAATTGATTTGTGTCACAGAAGTTTTTATGGAAACTAGCACCATGTCTACAGAAAC
TGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCC
CCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGCTGCAGCTGGTCTTGGATTTTGC
TATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAATCAGCAGAAGGAAATGAT
CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA
AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCGATGCCTGGAA
GCTGAAGTT**TAG**ATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTTCATGCTCC
TTACCCTGCCCCAGCTGGGGAAATCAAAAGGGCCAAAGAACCAAAGAAGAAAGTCCACCCTT
GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC
CCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCC
TTTCTAGCCTGGCTATGTCCTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAA
GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG
TGGGTTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGAC
CCTTTCTTCAGCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTA
AGAGCAAAAGAATGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG
ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT
GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA
ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT
AGGAAATATACTTTTACAAGTAACAAAAATAAAAACTCTTATAAATTTCTATTTTTATCTGA
GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA
ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT
TGAATATTATTCCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTTTCA
GTTTTGATATTTCTAGCTTATCTACTTCCAAACTAATTTTTATTTTTGCTGAGACTAATCTT
ATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG
TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAACAAATGTATCACTA
GCCCTCCTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA
AGCATTTAGAAAACTT

FIGURE 74

MARCPSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACR
LLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQF
AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP
APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFF
GAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSP
SKTTVRCLEAEV

Signal sequence:
amino acids 1-16

Transmembrane domain:
amino acids 235-254

N-glycosylation site.
amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.
amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.
amino acids 79-88

N-myristoylation site.
amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 75

AGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTCGGTGCCGCGACTTTCACGATGG
CTCGCCCAACCTTACTACCTTCTGTCGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAA
ACTGCCGCCGCTCTGCCACGGTCTGCCCACCCAACGCGAAGACGGTAACCCGTGTGACTTTG
ACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA
TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT
TCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC
TGATGACGTGCAAACCCCCCCTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAA
ACCATTGATGAGGAACTAGAACGGCACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA
TTGGTCTAATGACTGCCAATCATTTCCCCCTATCTATGCTGACCTCTCCCTTAAATACAACT
GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC
AAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAA
GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG
AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAAACTATCA
AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTCAGA
TGGGGAAAACAAGAAGGATAAATAACATCCTCACTTTGGCAGTGCTTCCTCTCCTGTCAATT
CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGG
CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAG
GCACCCTACAGGAAGGCCTGCCATGCTGTGGCCAACTGTTTCACTGGAGCAAGAAAGAGATC
TCATAGGACGGAGGGGGAAATGGTTTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATC
AGCTATTCAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCCTTCTTAG
TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAA
CGCTAAGAATTTTCCCCCAAGGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTC
TTCATTAAAAGTATAAGCCTAACTTTGTCGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG
TTTTTATCATTGAAGACAATATTGAACAACCCCCTATTTTGTGGGGATTGAGAAGGGGTGAA
TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCAC
TAACCCTCTGACATACTCCCCACACCCAGTTGATGGCTTTCCGTAATAAAAAGATTGGGATT
TCCTTTTG

FIGURE 76

MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRKLPPLCHGLPTQREDGNPCDFD
WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL
MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNC
TGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLFTLILFQGGKEAMRRFQIDKKGRAVSWTFSE
ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

Signal sequence:
amino acids 1-48

Transmembrane domain:
amino acids 111-125

N-glycosylation site.
amino acids 165-169, 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 154-158, 265-269

Casein kinase II phosphorylation site.
amino acids 51-55, 145-149, 245-249, 286-290, 288-292

N-myristoylation site.
amino acids 188-194, 225-231

Myb DNA-binding domain repeat signature 1.
amino acids 244-253

FIGURE 77

GGACAGCTCGCGGCCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGACGTTTGCCCTG
GGGCCCCAGCCTGGCCCGGGTCACCCTGGCATGAGGAGATGGGCCTGTTGCTCCTGGTCCCA
TTGCTCCTGCTGCCCGGCTCCTACGGACTGCCCTTCTACAACGGCTTCTACTACTCCAACAG
CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG
TGGTGGAGACACCCGAGGAGACCCTGTTCACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC
CCCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAGCT
GTCGGAGAACGGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT
TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG
ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCATTGACGGGCTGGA
GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCGGGGTGTGGTCTTTCCTTACCAGTCCCCCA
ACGGGCGCTACCAGTTCAACTTCCACGAGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG
GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCTGGGAGGAGGGCCTGGACTGGTGCAACGCGGG
CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCCGGCAGCCCTGCGGTGGCC
CAGGCCTGGCACCTGGCGTGCGAAGCTACGGCCCCCGCCACCGCCGCCTGCACCGCTATGAT
GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCCTGAGAAGCT
GACGCTGACAGAGGCAAGGGAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC
AGCTCTTTGCCGCCTGGAAGTTCCATGGCCTGGACCGCTGCGACGCTGGCTGGCTGGCAGAT
GGCAGCGTCCGCTACCCTGTGGTTCACCCGCATCCTAACTGTGGGCCCCCAGAGCCTGGGGT
CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTGTACGGTGTTTACTGCTACCGCCAGC
ACTAGGACCTGGGGCCCTCCCCTGCCGCATTCCCTCACTGGCTGTGTATTTATTGAGTGGTT
CGTTTTCCCTTGTGGGTTGGAGCCATTTTAACTGTTTTTATACTTCTCAATTTAAATTTTCT
TTAAACATTTTTTTACTATTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG
GATGCCCCACTCCAGGAATCATGCTTGCTCCCCTGGGCCATTTGCGGTTTTGTGGGCTTCTG
GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTCCCTTAAGGAGGTTGGTGCCCAGAGTGGGC
GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCCT
CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGGAGCTGGGCTTTGGGCCTCTCCTGCC
CACCTCTACTTCTCTGTGAAGCCGCTGACCCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA
GCCAGTTCTAGGCTTCCAGGCGAAATCTGAGGGAAGGAAGAAACTCCCCTCCCCGTTCCCCT
TCCCCTCTCGGTTCCAAAGAATCTGTTTTGTTGTCATTTGTTTCTCCTGTTTCCCTGTGTGG
GGAGGGGCCCTCAGGTGTGTGTACTTTGGACAATAAATGGTGCTATGACTGCCTTCCGCCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 78

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNGHGKDLLNGVKLVVETPEETLFTYQ
GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQD
KEHDVSLEIQDLRLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ
VCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR
HRRLHRYDVFCFATALKGRVYYLEHPEKIQTLTEAREACQEDDATIAKVGQLFAAWKFHGLDR
CDAGWLADGSVRYPVVHPHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

Signal sequence:
amino acids 1-17

Casein kinase II phosphorylation site.
amino acids 29-33, 53-57, 111-115, 278-282

Tyrosine kinase phosphorylation site.
amino acids 137-145

N-myristoylation site.
amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

FIGURE 79

GGAGAGCGGAGCGAAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC
TTCTGTTGCTACTGAGGCACGGGGCCCAGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC
CAGGGGAGGGTGCACCAGGCGGCCCCCCTGAGCGACGCTCCCCATGATGACGCCCACGGGAA
CTTCCAGTACGACCATGAGGCTTTCCTGGGACGGGAAGTGGCCAAGGAATTCGACCAACTCA
CCCCAGAGGAAAGCCAGGCCCGTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGAC
GGCGACGGCTGGGTGTCGCTGGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCA
CATACGGGACTCGGTGAGCGCGGCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGG
GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCGGTGAAGAATTTCATGAC
GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGGCGTTTCCGGGTGGC
CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGAGCTGACAGCCTTCCTGCACCCCGAGG
AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAAA
GATGGCTATGTCCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGA
GGAGCCGGCGTGGGTGCAGACGGAGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGG
ATGGGCACCTGGATGGGAGTGAGGTGGGCCACTGGGTGCTGCCCCCTGCCCAGGACCAGCCC
CTGGTGGAAGCCAACCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA
AGCGGAAATCCTGGGTAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGG
ACCTGACCCGGCACCACGATGAGCTGTGAGCACCGCGCACCTGCCACAGCCTCAGAGGCCCG
CACAATGACCGGAGGAGGGCCGCTGTGGTCTGGCCCCCTCCCTGTCCAGGCCCCGCAGGAG
GCAGATGCAGTCCCAGGCATCCTCCTGCCCCTGGGCTCTCAGGGACCCCCTGGGTCGGCTTC
TGTCCCTGTCACACCCCCAACCCCAGGGAGGGGCTGTCATAGTCCCAGAGGATAAGCAATAC
CTATTTCTGACTGAGTCTCCCAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTCAGCTCTAA
GAACCGCCCCAACCCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCCT
GGCCCCAGCCCTCTCCTGCCTGGCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCAATAA
AAGCCAGCGCCGGGACCTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 80

MMWRPSVLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDAPHDDAHGNFQYDHEAFLGR
EVAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT
YDTDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGDSMATRE
ELTAFLHPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPAWVQTERQQ
FRDFRDLNKDGHLDGSEVGHWVLPPAQDQPLVEANHLLHESDTDKDGRLSKAEILGNWNMFV
GSQATNYGEDLTRHHDEL

Signal sequence:
amino acids 1-20

N-glycosylation site.
amino acids 140-144

Casein kinase II phosphorylation site.
amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293, 291-295, 298-302

N-myristoylation site.
amino acids 263-269, 311-317

Endoplasmic reticulum targeting sequence.
amino acids 325-330

FIGURE 81

GGGGCCTTGCCTTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG
GCGGCGGGCGCGGGTGCGAGGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTCGCTCCCAG
CCTGTCTGTCGTCGTTTGGCGCCCCCGCCTCCCCGCGGTGCGGGGTTGCACACCGATCCTG
GGCTTCGCTCGATTTGCCGCCGAGGCGCCTCCCAGACCTAGAGGGGCGCTGGCCTGGAGCAG
CGGGTCGTCTGTGTCCTCTCTCCTCTGCGCCGCGCCCGGGGATCCGAAGGGTGCGGGGCTCT
GAGGAGGTGACGCGCGGGGCCTCCCGCACCCTGGCCTTGCCCGCATTCTCCCTCTCTCCCAG
GTGTGAGCAGCCTATCAGTCACCATGTCCGCAGCCTGGATCCCGGCTCTCGGCCTCGGTGTG
TGTCTGCTGCTGCTGCCGGGGCCCGCGGGCAGCGAGGGAGCCGCTCCCATTGCTATCACATG
TTTTACCAGAGGCTTGGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCCAGGGGGCTGCC
CTCTTGAGGAATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGG
GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC
TGGTCGAGAAAACTATTCCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGAT
GGTCTGCTTCTTTCACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA
GTGTCCACAGCACATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACTGG
CAATAAAGATTGTAAAGCAGACATTGCATTTCTGATTGATGGAAGCTTTAATATTGGGCAGC
GCCGATTTAATTTACAGAAGAATTTTGTTGGAAAAGTGGCTCTAATGTTGGGAATTGGAACA
GAAGGACCACATGTGGGCCTTGTTCAAGCCAGTGAACATCCCAAAATAGAATTTTACTTGAA
AAACTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGTA
ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCTGGA
GTAAGAAAAGGGATCCCCAAAGTGGTGGTGGTATTTATTGATGGTTGGCCTTCTGATGACAT
CGAGGAAGCAGGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCA
AGCCTATCCCTGAAGAACTGGGGATGGTTCAGGATGTCACATTTGTTGACAAGGCTGTCTGT
CGGAATAATGGCTTCTTCTCTTACCACATGCCCAACTGGTTTGGCACCACAAAATACGTAAA
GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT
CAGTGAACATTGCCTTTCTAATTGATGGCTCCAGCAGTGTTGGAGATAGCAATTTCCGCCTC
ATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGAT
AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCAGTTTCACTGACTATAGCACCA
AAGAGAATGTCCTAGCTGTCATCAGAAACATCCGCTATATGAGTGGTGGAACAGCTACTGGT
GATGCCATTTCCTTCACTGTTAGAAATGTGTTTGGCCCTATAAGGGAGAGCCCCAACAAGAA
CTTCCTAGTAATTGTCACAGATGGGCAGTCCTATGATGATGTCCAAGGCCCTGCAGCTGCTG
CACATGATGCAGGAATCACTATCTTCTCTGTTGGTGTGGCTTGGGCACCTCTGGATGACCTG
AAAGATATGGCTTCTAAACCGAAGGAGTCTCACGCTTTCTTCACAAGAGAGTTCACAGGATT
AGAACCAATTGTTTCTGATGTCATCAGAGGCATTTGTAGAGATTTCTTAGAATCCCAGCAAT
AATGGTAACATTTTGACAACTGAAAGAAAAAGTACAAGGGGATCCAGTGTGTAAATTGTATT
CTCATAATACTGAAATGCTTTAGCATACTAGAATCAGATACAAAACTATTAAGTATGTCAAC
AGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGT
ACTTTGTTAAAAACACTGCTGAGGCTTCATAATCATGGCTCTTAGAAACTCAGGAAAGAGGA
GATAATGTGGATTAAAACCTTAAGAGTTCTAACCATGCCTACTAAATGTACAGATATGCAAA
TTCCATAGCTCAATAAAAGAATCTGATACTTAGACCAAAAAAAAAAA

FIGURE 82

MSAAWIPALGLGVCLLLLPGPACSEGAAPIAITCFTRGLDIRKEKADVLCPGGCPLEEFSVY
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASFTVT
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRRFNLQKN
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLKNFTSAKDVLFAIKEVGFRGGNSNTGKAL
KHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG
MVQDVTFVDKAVCRNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI
DGSSSVGDSNFRLMLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAHDAGITI
PSVGVAWAPLDDLKDMASKPKESHAFFTREFTGLEPIVSDVIRGICRDFLESQQ

Signal sequence:
amino acids 1-24

N-glycosylation site.
amino acids 100-104, 221-225

Casein kinase II phosphorylation site.
amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424, 425-429, 478-482, 528-532

N-myristoylation site.
amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211, 239-245, 240-246, 261-267, 403-409, 442-448, 443-449

Amidation site.
amino acids 145-149

FIGURE 83

CGCCGCGCTCCCGCACCCGCGGCCCGCCCACCGCGCCGCCTCCCGCATCTGCACCCGCAGCCC
GGCGGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCG
GGGCGGCGGCTGCGGGCGCAGAGCGGAG**ATG**CAGCGGCTTGGGGCCACCCTGCTGTGCCTGC
TGCTGGCGGCGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTC
AAGCCCGGCCCGGCTCTCAGCTACCCGCAGGAGGAGCCCACCCTCAATGAGATGTTCCGCGA
GGTTGAGGAACTGATGGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGG
CAGAAGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCCAGCTAT
CACAATGAGACCAACACAGACACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAAT
TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTG
TGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGC
ATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCT
CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAA
TGGCCACCAGGGGCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTG
TGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA
GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATG
GAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG
GTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC
CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG
AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGCT
GCACTGCTGGGAGGGGAAGAGATT**TAG**ATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA
ATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTTCCTACA
TCTTCTTCCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTCAGC
TCCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAAAC
TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAG
ACAGCCGTTTGTTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGG
AGTCTCCCTCTGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAA
CCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTG
TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTCACCCTGCATTACATGTGTTTATTCATCC
AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC
CCTCTCTCAGCACAGCCTGGGGAGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG
GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT
CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCCACACCAGCCTTGGTGCCACCAA
AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTTCTTGAGGCATGCACATCTGGAATTAAG
GTCAAACTAATTCTCACATCCCTCTAAAAGTAAACTACTGTTAGGAACAGCAGTGTTCTCAC
AGTGCGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGT
TGCATTAGTAACTTTGAAAGGTATATGACTGAGCGTAGCATACAGGTTAACCTGCAGAAACA
GTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAAC
TGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTT
GTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATGTTTTCAGGTGTCA
TGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTA
TAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGC
ATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 84

MQRLGATLLCLLLAAAVPTAPAPAPTATSAFVKPGPALSYPQEEATLNEMFREVEELMEDTQ
HKLRSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTG
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMLCTRDSECCG
DQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRL
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEEI

Signal sequence:
amino acids 1-19

N-glycosylation site.
amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.
amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316, 327-331

N-myristoylation site.
amino acids 202-208, 217-223

Amidation site.
amino acids 140-144

FIGURE 85

AAGGAGGCTGGGAGGAAAGAGGTAAGAAAGGTTAGAGAACCTACCTCACATCTCTCTGGGCTCAGAAGGACTCTG
AAGATAACAATAATTTCAGCCCATCCACTCTCCTTCCCTCCCAAACACACATGTGCATGTACACACACACATACA
CACACATACACCTTCCTCTCCTTCACTGAAGACTCACAGTCACTCACTCTGTGAGCAGGTCATAGAAAAGGACAC
TAAAGCCTTAAGGACAGGCCTGGCCATTACCTCTGCAGCTCCTTTGGCTTGTTGAGTCAAAAAACATGGGAGGGG
CCAGGCACGGTGACTCACACCTGTAATCCCAGCATTTTGGGAGACCGAGGTGAGCAGATCACTTGAGGTCAGGAG
TTCGAGACCAGCCTGGCCAACATGGRGAAACCCCCATCTCTACTAAAAATACAAAAATTAGCCAGGAGTGGTGGC
AGGTGCCTGTAATCCCAGCTACTCAGGTGGCTGAGCCAGGAGAATCGCTTGAATCCAGGAGGCGGAGGATGCAGT
CAGCTGAGTGCACOGCTGCACTCCAGCCTGGGTGACAGAATGAGACTCTGTCTCAAACAAACAAACACGGGAGGA
GGGGTAGATACTGCTTCTCTGCAACCTCCTTAACTCTGCATCCTCTTCTTCCAGGGCTGCCCCTGATGGGGCCTG
GCAATGACTGAGCAGGCCCAGCCCCAGAGGACAAGGAAGAGAAGGCATATTGAGGAGGGCAAGAAGTGACGCCCG
GTGTAGAATGACTGCCCTGGGAGGGTGGTTCCTTGGGCCCTGGCAGGGTTGCTGACCCTTACCCTGCAAAACACA
AAGAGCAGGAGTCCAGACTCTCCTTGTGAATGGTCCCCTGCCCTGCAGCTCCACCATGAGGCTTCTCGTGGCCCC
ACTCTTGCTAGCTTGGGTGGCTGGTGCCACTGCCACTGTGCCCGTGGTACCCTGGCATGTTCCCTGCCCCCCTCA
GTGTGCCTGCCAGATCCGCCCTGGTATACGCCCCGCTCGTCCTACCGCGAGGCTACCACTGTGGACTGCAATGA
CCTATTCCTGACGGCAGTCCCCCCGGCACTCCCCGCAGGCACACAGACCCTGCTCCTGCAGAGCAACAGCATTGT
CCGTGTGGACCAGAGTGAGCTGGGCTACCTGGCCAATCTCACAGAGCTGGACCTGTCCCAGAACAGCTTTTCGGA
TGCCCGAGACTGTGATTTCCATGCCCTGCCCCAGCTGCTGAGCCTGCACCTAGAGGAGAACCAGCTGACCCGGCT
GGAGGACCACAGCTTTGCAGGGCTGGCCAGCCTACAGGAACTCTATCTCAACCACAACCAGCTCTACCGCATCGC
CCCCAGGGCCTTTTCTGGCCTCAGCAACTTGCTGCGGCTGCACCTCAACTCCAACCTCCTGAGGGCCATTGACAG
CCGCTGGTTTGAAATGCTGCCCAACTTGGAGATACTCATGATTGGCGGCAACAAGGTAGATGCCATCCTGGACAT
GAACTTCCGGCCCCTGGCCAACCTGCGTAGCCTGGTGCTAGCAGGCATGAACCTGCGGGAGATCTCCGACTATGC
CCTGGAGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTCTATGACAACCAGCTGGCCCGGGTGCCCAGGCGGGCACT
GGAACAGGTGCCCGGGCTCAAGTTCCTAGACCTCAACAAGAACCCGCTCCAGCGGGTAGGGCCGGGGGACTTTGC
CAACATGCTGCACCTTAAGGAGCTGGGACTGAACAACATGGAGGAGCTGGTCTCCATCGACAAGTTTGCCCTGGT
GAACCTCCCCGAGCTGACCAAGCTGGACATCACCARTAACCCACGGCTGTCCTTCATCCACCCCCGCCTTCCA
CCACCTGCCCCAGATGGAGACCCTCATGCTCAACAACAACGCTCTCAGTGCCTTGCACCAGCAGACGGTGGAGTC
CCTGCCCAACCTGCAGGAGGTAGGTCTCCACGGCAACCCCATCCGCTGTGACTGTGTCATCCGCTGGGCCAATGC
CACGGGCACCCGTGTCCGCTTCATCGAGCCGCAATCCACCCTGTGTGCGGAGCCTCCGCACCTCCAGCGCCTCCC
GGTCCGTGAGGTGCCCTTCCGGGAGATGACGGACCACTGTTTGCCCCTCATCTCCCCACGAAGCTTCCCCCCAAG
CCTCCAGGTAGCCAGTGGAGAGAGCATGGTGCTGCATTGCCGGGCACTGGCCGAACCCGAACCCGAGATCTACTG
GGTCACTCCAGCTGGGCTTCGACTGACACCTGCCCATGCAGGCAGGAGGTACCGGGTGTACCCCGAGGGGACCCT
GGAGCTGCGGAGGGTGACAGCAGAAGAGGCAGGGCTATACACCTGTGTGGCCCAGAACCTGGTGGGGGCTGACAC
TAAGACGGTTAGTGTGGTTGTGGGCCGTGCTCTCCTCCAGCCAGGCAGGGACGAAGGACAGGGCTGGAGCTCCG
GGTGCAGGAGACCCACCCCTATCACATCCTGCTATCTTGGGTCACCCCACCCAACACAGTGTCCACCAACCTCAC
CTGGTCCAGTGCCTCCTCCCTCCGGGGCCAGGGGGCCACAGCTCTGGCCCCGCCTGCCTCGGGGAACCCACAGCTA
CAACATTACCCGCCTCCTTCAGGCCACGGAGTACTGGGCCTGCCTGCAAGTGGCCTTTGCTGATGCCCACACCCA
GTTGGCTTGTGTATGGGCCAGGACCAAAGAGGCCACTTCTTGCCACAGAGCCTTAGGGGATCGTCCTGGGCTCAT
TGCCATCCTGGCTCTCGCTGTCCTTCTCCTGGCAGCTGGGCTAGCGGCCCACCTTGGCACAGGCCAACCCAGGAA
GGGTGTGGGTGGGAGGCGGCCTCTCCCTCCAGCCTGGGCTTTCTGGGGCTGGAGTGCCCCTTCTGTCCGGGTTGT
GTCTGCTCCCCTCGTCCTGCCCTGGAATCCAGGGAGGAAGCTGCCCAGATCCTCAGAAGGGGAGACACTGTTGCC
ACCATTGTCTCAAAATTCTTGAACCTCAGCCTGTTCTCAGCAGTAGAGAAATCACTAGGACTACTTTTTACCAAA
AGAGAAGCAGTCTGGGCCAGATGCCCTGCCAGGAAAGGGACATGGACCCACGTGCTTGAGGCCTGGCAGCTGGGC
CAAGACAGATGGGGCTTTGTGGCCCTGGGGGTGCTTCTGCAGCCTTCAAAAAGTTGCCCTTACCTCCTAGGGTCA
CCTCTGCTGCCATTCTGAGGAACATCTCCAAGGAACAGGAGGGACTTTGGCTAGAGCCTCCTGCCTCCCCATCTT
CTCTCTGCCCAGAGGCTCCTGGGCCTGGCTTGGCTGTCCCCTACCTGTGTCCCCGGGCTGCACCCCTTCCTCTTC
TCTTTCTCTGTACAGTCTCAGTTGCTTGCTCTTGTGCCTCCTGGGCAAGGGCTGAAGGAGGCCACTCCATCTCAC
CTCGGGGGGCTGCCCTCAATGTGGGAGTGACCCCAGCCAGATCTGAAGGACATTTGGGAGAGGGATGCCCAGGAA
CGCCTCATCTCAGCAGCCTGGGCTCGGCATTCCGAAGCTGACTTTCTATAGGCAATTTTGTACCTTTGTGGAGAA
ATGTGTCACCTCCCCCAACCCGATTCACTCTTTTCTCCTGTTTTGTAAAAAATAAAAATAAATAATAACAATAAA
AAAA

FIGURE 86

MRLLVAPLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLFLTA
VPPALPAGTQTLLLQSNSIVRVDQSELGYLANLTELDLSQNSFSDARDCDFHALPQLLSLHL
EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE
MLPNLEILMIGGNKVDAILDMNFRPLANLRSLVLAGMNLREISDYALEGLQSLESLSFYDNQ
LARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELVSIDKFALVNLP
ELTKLDITNNPRLSFIHPRAFHHLPQMETLMLNNNALSALHQQTVESLPNLQEVGLHGNPIR
CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVDFREMTDHCLPLISPRSFPPSLQ
VASGRSMVLHCRALAEPEPEIYWVTPAGLRLTPAHAGRYRVYPEGTLELRRVTAEEAGLYT
CVAQNLVGADTKTVSVVVGRALLQPGRDEGQGLELRVQETHPYHILLSWVTPPNTVSTNLTW
SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS
CHRALGDRPGLIAILALAVLLLAAGLAAHLGTGQPRKGVGGRRDLPPAWAFWGWSAPSVRVV
SAPLVLPWNPGRKLPRSSEGETLLPPLSQNS

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 629-648

N-glycosylation site.

amino acids 94-98, 381-385, 555-559, 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 485-489

Casein kinase II phosphorylation site.

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146, 243-247, 313-317, 488-492, 700-704

Tyrosine kinase phosphorylation site.

amino acids 532-540

N-myristoylation site.

amino acids 15-21, 493-499, 566-572

Amidation site.

amino acids 470-474, 660-664, 692-696

FIGURE 87

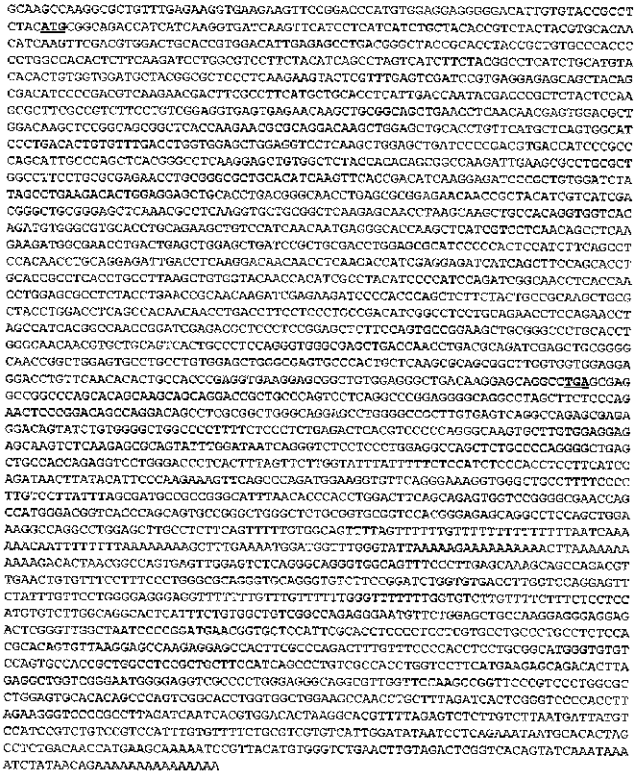


FIGURE 88

MRQTIIKVIKFILIICYTVYYVHNIKPDVDCTVDIESLTGYRTYRCAHPLATLPKILASFYI
SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK
RFAVFLSEVSENKLRQLNLNNEWTLDKLRQRLTKNAQDKLELHLFMLSGIPDTVFDLVELEV
LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLRENLRALHIKFTDIKEIPLWI
YSLKTLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI
NNEGTKLIVLNSLKKMANLTELELIRCDLERIPESIFSLHNLQEIDLKDNNLKTIEEIISFQ
HLHRLTCLKLWYNHIAYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKLRYLDLSENNLT
FLPADIGLLQNLQNLAITANRIETLPPELFQCRKLRALHLGNNVLQSLPSRVGELTNLTQIE
LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

Transmembrane domain:
amino acids 51-75 (type II)

N-glycosylation site.
amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 85-89

Casein kinase II phosphorylation site.
amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368, 398-402, 493-497

N-myristoylation site.
amino acids 173-179, 261-267, 395-401, 441-447

FIGURE 89

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCATGGACCTGGCACTGCGGCGCTCTCCCGT
CCCGCGGTGGTTGCTGCTGCTGCCGCTGCTGCTGGGCCTGAACGCAGGAGCTGTCATTGACT
GGCCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG
TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCTGGTCAT
GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTGGAAACTTTGAGGAAATTGGGC
CCCTTGACAGTGATCTCAAACCACGGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT
GTGGATAATCCCGTGGGCACTGGGTTCAGTTATGTGAATGGTAGTGGTGCCTATGCCAAGGA
CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG
AATTCCAGACAGTTCCATTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGC
ATTGGTCTAGAGCTTTATAAGGCCATTCAGCGAGGCACCATCAAGTGCAACTTTGCGGGGGT
TGCCTTGGGTGATTCCTGGATCTCCCCTGTTGATTCGGTGCTCTCCTGGGGACCTTACCTGT
ACAGCATGTCTCTTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA
CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT
GATCATTGAACAGAACACAGATGGGGTGAACTTCTATAACATCTTAACCAAAAGCACTCCCA
CGTCTACAATGGAGTCGAGTCTAGAATTCACACAGAGCCACCTAGTTTGTCTTTGTCAGCGC
CACGTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAA
GCTCAAAATTATTCCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA
TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC
AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTCAGGAGGCCTG
GGTGCGGAAACTGAAGTGGCCAGAACTGCCTAAATTCAGTCAGCTGAAGTGGAAGGCCCTGT
ACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTC
TACTGGATTCTGAAAGCTGGTCATATGGTTCCTTCTGACCAAGGGGACATGGCTCTGAAGAT
GATGAGACTGGTGACTCAGCAAGAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT
TGGGGCACAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT
AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT
GGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAAACCTAAGATTTTTTAAAAAATTGAT
TTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

FIGURE 90

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEDGKEVWDYVTVRKDAYMFWWLYYATNSC
KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSDLKPRKTTWLQAASLLFVDNPVGTGFSY
VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLELYKAIQR
GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVNKGLYRE
ATELWGKAEMIIEQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQRDALS
QLMNGPIRKFLKIIPEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL
IVDTMGQEAWVRKLKWPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP
SDQGDMALKMMRLVTQQR

Signal sequence:
amino acids 1-25

N-glycosylation site.
amino acids 64-68, 126-130, 362-366

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 101-105

Casein kinase II phosphorylation site.
amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

N-myristoylation site.
amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175, 187-193, 195-201, 331-337, 332-338, 360-366

FIGURE 91

GGCCGCGGGAGAGGAGGCCATGGGCGCGCGCGGGGCGCTGCTGCTGGCGCTGCTGCTGGCTC

GGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGGCGGCGCCGTTATCAGGACCATGCGGCCGA

CGGGTCATCACGTCGCGCATCGTGGGTGGAGAGGACGCCGAACTCGGGCGTTGGCCGTGGCA

GGGGAGCCTGCGCCTGTGGGATTCCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGG

CACTCACGGCGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG

GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCG

TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCTATGACATTG

CCTTGGTGAAGCTGTCTGCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAG

GCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAA

AGAGGATGAGGCACTGCCATCTCCCCACACCCTCCAGGAAGTTCAGGTCGCCATCATAAACA

ACTCTATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCCCAAGGACATCTTTGGAGACATG

GTTTGTGCTGGCAACGCCCAAGCCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCTT

GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTGGGCTGTG

GTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAAG

CTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTCTTTTTCCCTCT

TCTCTGGGCTCTCCCACTCCTGGGGCCGGTCTGAGCCTACCTGAGCCCATGCAGCCTGGGGC

CACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTTTGGTAATAAACACATTCCAGTTGA

TGCCTTGCAGGGCATTCTTCAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 92

MGARGALLLALLLARAGLRKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSLRLW
DSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYPVSNI
YLSPRYLGNSPYDIALVKLSAPVTYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP
SPHTLQEVQVAIINNSMCNHLFLKYSFRKDIPGDMVCAGNAQGGKDACFGDSGGPLACNKNG
LWYQIGVVSWGVGCGRPNRPGVYTNISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL
LGPV

Signal sequence:
amino acids 1-18

N-glycosylation site.
amino acids 167-171, 200-204, 273-277

Casein kinase II phosphorylation site.
amino acids 86-90, 134-138, 161-165, 190-194, 291-295

N-myristoylation site.
amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245, 259-265, 269-275

Amidation site.
amino acids 33-37

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 252-263,

Serine proteases, trypsin family, histidine active site.
amino acids 78-84

FIGURE 93

CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGAATGGGACTCCAAGCCTGCCTCCTAGGGCT
CTTTGCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGACCAGCGGAGGACGC
TGCCCCCAGGCTGGGTGTCCCTGGGCCGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTT
GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGCAGGCTGTGTCGGATCCCAG
CTCTCCTCAATACGGAAAATACCTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC
CACTGACCCTCCACACGGTGCAAAAATGGCTCTTGGCAGCCGGAGCCCAGAAGTGCCATTCT
GTGATCACACAGGACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCC
TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCCAC
ATCCCTACCAGCTTCCACAGGCCTTGCCCCCCCATGTGGACTTTGTGGGGGGACTGCACCGT
TTTCCCCCAACATCATCCCTGAGGCAACGTCCTGAGCCGCAGGTGACAGGGACTGTAGGCCT
GCACCTGGGGGTAACCCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAAGACGTGG
GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCCAGTTCCTGGAGCAGTATTTCCATGAC
TCAGACCTGGCTCAGTTCATGCGCCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGC
CCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCCGGGATTGAGGCCAGTCTAGATGTGCAGT
ACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG
GGACAGGAGCCCTTCCTGCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGT
GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA
ACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCCTGCTCTTCGCCTCAGGTGACAGT
GGGGCCGGGTGTTGGTCTGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAG
CCCCTATGTCACCACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCACAAATGAAA
TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCCACGGCCTTCATACCAGGAG
GAAGCTGTAACGAAGTTCCTGAGCTCTAGCCCCCACCTGCCACCATCCAGTTACTTCAATGC
CAGTGGCCGTGCCTACCCAGATGTGGCTGCACTTTCTGATGGCTACTGGGTGGTCAGCAACA
GAGTGCCCATTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTTTGGGGGGATCCTA
TCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCCTCTTGGCTTTCTCAACCCAAG
GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTC
TGGATGAAGAGGTAGAGGGCCAGGGTTTCTGCTCTGGTCCTGGCTGGGATCCTGTAACAGGC
TGGGGAACACCAACTTCCCAGCTTTGCTGAAGACTCTACTCAACCCCTGACCCTTTCCTATC
AGGAGAGATGGCTTGTCCCCTGCCCTGAAGCTGGCAGTTCAGTCCCTTATTCTGCCCTGTTG
GAAGCCCTGCTGAACCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCTGAAA
TGCTGTGAGCTTGACTTGACTCCCAACCCTACCATGCTCCATCATACTCAGGTCTCCCTACT
CCTGCCTTAGATTCCTCAATAAGATGCTGTAACTAGCATTTTTTGAATGCCTCTCCCTCCGC
ATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAGACTCTGTGCACTA
TTTCACTTGATATTCATTCCCCAATTCACTGCAAGGAGACCTCTACTGTCACCGTTTACTCT
TTCCTACCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATG
GCCTTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACTTCTCTG
ACTACTCTTGTCTTCCTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGC
TCCATTTGTAGATTTTTGCTCTTCTCAGTTTACTCATTGTCCCCTGGAACAAATCACTGACA
TCTACAACCATTACCATCTCACTAAATAAGACTTTCTATCCAATAATGATTGATACCTCAAA
TGTAAAAAA

FIGURE 94

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQYGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWL
SIRQAELLLPGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGGLHRFPPTSSLRQRP
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG
RPPLGFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPTSQLC

Signal sequence:
amino acids 1-16

N-glycosylation site.
amino acids 210-214, 222-226, 286-290, 313-317, 443-447

Glycosaminoglycan attachment site.
amino acids 361-365, 408-412, 538-542

Casein kinase II phosphorylation site.
amino acids 212-216, 324-328, 392-396, 420-424, 525-529

N-myristoylation site.
amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225, 248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488, 521-527, 533-539, 549-555

FIGURE 95

GCCGCGCGCTCTCTCCCGGCGCCCACACCTGTCTGAGCGGCGCAGCGAGCCGCGGCCCGGGC

GGGCTGCTCGGCGCGGAACAGTGCTCGGCATGGCAGGGATTCCAGGGCTCCTCTTCCTTCTC

TTCTTTCTGCTCTGTGCTGTTGGGCAAGTGAGCCCTTACAGTGCCCCCTGGAAACCCACTTG

GCCTGCATACCGCCTCCCTGTCGTCTTGCCCCAGTCTACCCTCAATTTAGCCAAGCCAGACT

TTGGAGCCGAAGCCAAATTAGAAGTATCTTCTTCATGTGGACCCCAGTGTCATAAGGGAACT

CCACTGCCCACTTACGAAGAGGCCAAGCAATATCTGTCTTATGAAACGCTCTATGCCAATGG

CAGCCGCACAGAGACGCAGGTGGGCATCTACATCCTCAGCAGTAGTGGAGATGGGCCCCAAC

ACCGAGACTCAGGGTCTTCAGGAAAGTCTCGAAGGAAGCGGCAGATTTATGGCTATGACAGC

AGGTTCAGCATTTTTGGGAAGGACTTCCTGCTCAACTACCCTTTCTCAACATCAGTGAAGTT

ATCCACGGGCTGCACCGGCACCCTGGTGGCAGAGAAGCATGTCCTCACAGCTGCCCACTGCA

TACACGATGGAAAAACCTATGTGAAAGGAACCCAGAAGCTTCGAGTGGGCTTCCTAAAGCCC

AAGTTTAAAGATGGTGGTCGAGGGGCCAACGACTCCACTTCAGCCATGCCCGAGCAGATGAA

ATTTCAGTGGATCCGGGTGAAACGCACCCATGTGCCCAAGGGTTGGATCAAGGGCAATGCCA

ATGACATCGGCATGGATTATGATTATGCCCTCCTGGAACTCAAAAAGCCCCACAAGAGAAAA

TTTATGAAGATTGGGGTGAGCCCTCCTGCTAAGCAGCTGCCAGGGGGCAGAATTCACTTCTC

TGGTTATGACAATGACCGACCAGGCAATTTGGTGTATCGCTTCTGTGACGTCAAAGACGAGA

CCTATGACTTGCTCTACCAGCAATGCGATGCCCAGCCAGGGGCCAGCGGGTCTGGGGTCTAT

GTGAGGATGTGGAAGAGACAGCAGCAGAAGTGGGAGCGAAAAATTATTGGCATTTTTTCAGG

GCACCAGTGGGTGGACATGAATGGTTCCCCACAGGATTTCAACGTGGCTGTCAGAATCACTC

CTCTCAAATATGCCCAGATTTGCTATTGGATTAAAGGAAACTACCTGGATTGTAGGGAGGGG

TGACACAGTGTTCCCTCCTGGCAGCAATTAAGGGTCTTCATGTTCTTATTTTAGGAGAGGCC

AAATTGTTTTTTGTCATTGGCGTGCACACGTGTGTGTGTGTGTGTGTGTGTGTAAGGTGT

CTTATAATCTTTTACCTATTTCTTACAATTGCAAGATGACTGGCTTTACTATTTGAAAACTG

GTTTGTGTATCATATCATATATCATTTAAGCAGTTTGAAGGCATACTTTTGCATAGAAATAA

AAAAAATACTGATTTGGGGCAATGAGGAATATTTGACAATTAAGTTAATCTTCACGTTTTTG

CAAACTTTGATTTTTATTTCATCTGAACTTGTTTCAAAGATTTATATTAAATATTTGGCATA

CAAGAGATATGAAAAAAAAAAAAAAA

FIGURE 96

MAGIPGLLFLLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS
SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIYILSSSGDGAQHRDSGSSGKS
RRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG
TQKLRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYA
LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD
AQPGASGSGVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW
IKQNYLDCREG

Signal sequence:
amino acids 1-19

N-glycosylation site.
amino acids 93-97, 207-211

Glycosaminoglycan attachment site.
amino acids 109-113, 316-320

Casein kinase II phosphorylation site.
amino acids 77-81, 95-99, 108-112, 280-284, 351-355

N-myristoylation site.
amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.
amino acids 171-177

FIGURE 97

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTGCTCCCCCGCCCCACCAGCCATGGTGGTTT
CTGGAGCGCCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG
GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA
GCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA
GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC
ACTGCTGCCCACTGTTTCAAGGACAACCTGAACAAACCATACCTGTTCTCTGTGCTGCTGGG
GGCCTGGCAGCTGGGGAACCCTGGCTCTCGGTCCCAGAAGGTGGGTGTTGCCTGGGTGGAGC
CCCACCCTGTGTATTCCTGGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCGTCTCGAG
CGCTCCATACAGTTCTCAGAGCGGGTCCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT
CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC
CCCACCCTCAGACCCTGCAGAAGCTGAAGGTTCCTATCATCGACTCGGAAGTCTGCAGCCAT
CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT
GGAGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG
GCGCCTGGCTGCTGGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCC
GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA
GCTCCGCGGGCGCGCTCAGGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG
CCGCGCGCTCCTAGGGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGGCGGCCAGATCCACA
TCTGGATCTGGATCTGCGGCGGCCTCGGGCGGTTTCCCCCGCCGTAAATAGGCTCATCTACC
TCTACCTCTGGGGGCCCGGACGGCTGCTGCGGAAAGGAAACCCCCTCCCCGACCCGCCCGAC
GGCCTCAGGCCCCCCTCCAAGGCATCAGGCCCCGCCCAACGGCCTCATGTCCCCGCCCCCAC
GACTTCCGGCCCCGCCCCCGGGCCCCAGCGCTTTGTGTATATAAATGTTAATGATTTTTAT
AGGTATTTGTAACCCTGCCCACATATCTTATTTATTCCTCCAATTTCAATAAATTATTTATT
CTCCAAAAAAAAAA

FIGURE 98

>></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318

><subunit 1 of 1, 317 aa, 1 stop

><MW: 33732, pI: 7.90, NX(S/T): 1

MVVSGAPPALGGGCLGTFTSLLLLASTAILNAARIPVPPACGKPQQLNRVVGGEDSTDSEWP
WIVSIQKNGTHHCAGSLLTSRWVITAAHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVA
WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPPNTHCWISGWGSIQDG
VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAGQGPITEDMLCAGYLEGERDACLGDSGGPLMC
QVDGAWLLACIISWGEGCAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG
SGAAARS

Signal sequence:
amino acids 1-32

N-glycosylation site.
amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459, 628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.
amino acids 826-830

Casein kinase II phosphorylation site.
amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268, 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653, 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962, 1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.
amino acids 607-615

N-myristoylation site.
amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

FIGURE 99

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTA
CTGCTGGTGGCCACCACAGGCCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGT
GGAGCTGCACAACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA
GATGGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCGTGTGGGGCCAC
AACAAGGAGCGCGGGCGCCGCGGCGAGAATCTGTTCGCCATCACAGACGAGGGCATGGACGT
GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT
GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC
GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACT
GGTGTGCAACTATGAGCCTCCGGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC
CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC
CCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGGGCGACTGA
AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT
TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAAACCCAGGCC
CCAACTTCCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC
AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA
CCTTCCCCAAATCGACCCATGTTCCTATCCCAAAATCAGCAGACAAAGTGACAGACAAAACA
AAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG
GGAACTCCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG
AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC
CACACGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAA
TGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCGTCCTTGCCAGGTGCAGAGGGCCCTGACA
AGCCTAGCGTTGTGTCAGGGCTGAACTCGGGCCCTGGTCATGTGTGGGGCCCTCTCCTGGGA
CTACTGCTCCTGCCTCCTCTGGTGTTGGCTGGAATCTTCTGAATGGGATACCACTCAAAGGG
TGAAGAGGTCAGCTGTCCTCCTGTCATCTTCCCCACCCTGTCCCCAGCCCCTAAACAAGATA
CTTCTTGGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC
ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGCTCAGGAGGCCGCCTGAGGACTGCACACC
GGGCCCACACCTCTCCTGCCCCTCCCTCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCT
CACTGCCTACCTGGCCTGGGGCTGTCTGCCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG
TGTAGCTGGGGATGGGGATTCCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC
TTTGAGTGGGGGAGGCAGGGACGAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAAACCT
GTCCAACCTGTGAAA

FIGURE 100

MHGSCSFLMLLLPLLLLLVATTGPVGALTDEEKRLMVELHNLYRAQVSPTASDMLHMRWDEE
LAAFAKAYARQCVWGHNKERGRRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ
MCGHYTQVVWAKTERIGCGSHFCEKLQGVEETNIELLVCNYEPPGNVKGKRPYQEGTPCSQC
PSGYHCKNSLCEPIGSPEDAQDLPYLVTEAPSFRATEASDSRKMGTPSSLATGIPAFLVTEV
SGSLATKALPAVETQAPTSLATKDPPSMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKS
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEAELPPSSEVLAS
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV
SGLNSGPGHVWGPLLGLLLLPPLVLAGIF

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.

amino acids 439-443

Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237, 250-256

Amidation site.

amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.

amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

FIGURE 101

GTAACTGAAGTCAGGCTTTTCATTTGGGAAGCCCCCTCAACAGAATTCGGTCATTCTCCAAGTTATGGTGGACGT
ACTTCTGTTGTTCTCCCTCTGCTTGCTTTTTCACATTAGCAGACCGGACTTAAGTCACAACAGATTATCTTTCAT
CAAGGCAAGTTCCATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCC
AAATCTGGGACCAGTCTCGGCAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGA
ACATCTGAAAGAGTTTCAGTCCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAAACTGCATT
TCCAGCCCTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAA
TTTGGCCAACACACTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAAACT
GCCCCAACTGCAACATCTGGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGG
TGCTCTGAAGTCTCTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTTGGGGCTGAGCAA
CATGGAAATTTTGCAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCT
GCAGGAACTTCATCTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAG
TGAGCTGGACCTAACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATAC
ACTGCACATTGGGAACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTT
GGATCTGAAGAACAATGAAATTTCCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAAACTGAG
GCGACTGATACTCCAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCA
TCTAGACCTGAGTGACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAAACTGCAACAATT
GCATTTAAATACATCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCCCACAGTGGGTGGCGGAAAACAACTT
TCAGAGCTTTGTAAATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTTGCTGTTAGCCCAGA
TGGCTTTGTGTGTGATGATTTTCCCAAACCCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTC
CAATTTGAGTTTCATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTTGCTTGGAAAAAAGACAATGA
ACTACTGCATGATGCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCAC
CATCCTTCGGCTGCGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCATCTCCAATCACTTTGGTTC
ATCCTACTCTGTCAAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCAT
CCGAGCTGGGGCCATGGCACGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCCAGATAGCCTGGCAGAAGGA
TGGGGGCACAGACTTCCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCGAGGATGACGTGTTCTTTATCGT
GGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTCAGCAAATGC
AACTCTGACTGTCCTAGAAACACCATCATTTTTGCGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGC
CGTCCTACAGTGCATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGT
AACCGAGAGGCACTTTTTTGCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAA
ATACACATGTGAGATGTCTAACACCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAAC
CTGCGACTCCCCTCAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGC
CGTGGTTTGCTGTGTGGTGGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAAGGCGGAGGAATGAAGA
TTGCAGCATTACCAACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCATCTCAGGGAACGTT
AGCTGACAGGCAGGATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACCAGTTTGTCACATCTTCAGGTGCTGG
ATTTTTCTTACCACAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGC
CACAGATCTGTTCCTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGA
TCCTTTTGAAACATATCATACAGGTTGCAGTCCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTA
CATAAAGAAAAAGGAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATC
GTGGCCTTCACATGTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTG
TCTAAACAAGTCCTCTTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGG
TACCTTTGGAAAAGCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCAGCC
AAGAGCCTTTTATTTGAAAGCTCATTCTTCCCCAGACTTGGACTCTGGGTCAGAGGAAGATGGGAAAGAAAGGAC
AGATTTTCAGGAAGAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACTACAGGACTCCAAATTTTCAGTC
TTATGACTTGGACACATAGACTGAATGAGACCAAAGGAAAAGCTTAACATACTAGCTCAAGTGAACTTTTATTTA
AAAGAGAGAGAATCTTATGTTTTTTAAATGGAGTTATGAATTTTAAAAGSATAAAAATGCTTTATTTATACAGAT
GAACCAAAATTACAAAAAGTTATGAAAATTTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAAACTA
TTTTTTAACTTTGTTTTATGCAAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTT
TTATAATGCCAGATTTCTTTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTT
TTAAATAGAAGTTACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAA

FIGURE 102

MVDVLLLFSLCLLFHISRPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSAN
ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEP
GYFDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKM
QRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWE
FCQKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNE
ISWTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ
MKKLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCD
DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQG
GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGA
MARLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN
SAGSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTER
HFFAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDG
WATVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTLAD
RQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGP
MYLKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISW
PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTPGKALRRPHLDA
YSSFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQS
YDLDT

Signal sequence:

amino acids 1-19

Transmembrane domain:

amino acids 746-765

N-glycosylation site.

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459, 628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268, 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653, 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962, 1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

FIGURE 103

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAATG
AAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGAAC
GAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATGAAA
TAAACCAGAGTTAGACCCGCGGGGGTTGGTGTGTTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCCCTCC
CCACCCCCAAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCATTTTCTC
TATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT
GGTGTGGTGGTGTTTTCCTTCTTTTTGAATTTCCCACAAGAGGAGAGGAAATTAATAATACATCTGCAAAGAAA
TTTCAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCACAGTTGGA
TTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTTTTTTTAAAT
TTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTTAACCACCTGGATTTCCATCTGGATGTTGCT
GTGATCAGTCTGAAATACAACTGTTTGAATTCCAGAAGGACCAACACCAGATAAATTATGAATGTTGAACAAGAT
GACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGGGCCCTATTTGACCCCCTGCTTGTGGTGCT
GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA
CCAGTTCAGCAAGGTGATTTGTGTTCGGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT
GCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCACTTGGAAATCCT
ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGAACCTCAACACTCTGGA
ACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAACTGAAGGAGCTCTGGTT
GCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTTGCCCCGACTAGACTTAGG
GGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACTTGAGGTATTTGAACCTTGC
CATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACTAGATGAGCTGGATCTTTCTGGGAATCA
TTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAACTGTGGATGATACAGTCCCAGAT
TCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTGGCACACAATAATCTAAC
ATTACTGCCCCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACATTTACATCACAACCCTTGGAACTG
TAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTTGTGCCCGGTG
TAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGT
GATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAGCTGAAATGTCGGGCCTCCACATC
CCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAAGTGCGGATAGCTGT
GCTCAGTGATGGTACGTTAAATTTCACAAATGTAACTGTGCAAGATACAGGCATGTACACATGTATGGTGAGTAA
TTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCAGCAACCACTACTCCTTTCTCTTACTTTTC
AACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCACAGATAACAATGTGGGTCCCACTCC
AGTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAGGTCGACAGAGAAAACCTT
CACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGATGAGGTCATGAAGACTACCAAAATCATCAT
TGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGCACCA
TCGGCAAAACCATCACGCCCAACAAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACC
CATGGAAAGCCACCTGCCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT
CAACCACACAACAACAGTTAACACAATAAATTCAATACACAGTTCAGTGCATGAACCGTTATTGATCCGAATGAA
CTCTAAAGACAATGTACAAGAGACTCAAATCTAAAACATTTACAGAGTTACAAAAAACAAACAATCAAAAAAAAA
GACAGTTTATTAAAAATGACACAAATGACTGGGCTAAATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAAACAA
AAAAGAAAAGAAATTTATTTATTAAAAATTCTATTGTGATCTAAAGCAGACAAAAA

FIGURE 104

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGISTNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSNSVGN
TTASATLNVTAATTTPFSYPSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

Signal sequence:
amino acids 1-44

Transmembrane domain:
amino acids 523-543

N-glycosylation site.
amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 183-187

Casein kinase II phosphorylation site.
amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.
amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397, 422-428, 433-439, 531-537

FIGURE 105

AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCCTGCTGTTTGGGGGCA
TGAAAGGGCTTCGCCGCCGGGAGTAAAAGAAGGAATTGACCGGGCAGCGCGAGGGAGGAGCGCGCACGCGACCGC
GAGGGCGGGCGTGCACCCTCGGCTGGAAGTTGTGCCGGGCCCCGAGCGCGCGCCGGCTGGGAGCTTCGGGTAGA
GACCTAGGCCGCTGGACCGCGATGAGCGCGCCGAGCCTCCGTGCGCGCGCCGCGGGGTTGGGGCTGCTGCTGTGC
GCGGTGCTGGGGCGCGCTGGCCGGTCCGACAGCGGCGGTCGCGGGGAACTCGGGCAGCCCTCTGGGGTAGCCGCC
GAGCGCCCATGCCCCACTACCTGCCGCTGCCTGGGGGACCTGCTGGACTGCAGTCGTAAGCGGCTAGCGCGTCTT
CCCGAGCCACTCCCGTCCTGGGTCGCTCGGCTGGACTTAAGTCACAACAGATTATCTTTCATCAAGGCAAGTTCC
ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCA
GTCTCGGCAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG
TTTCAGTCCCTTGAAACTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAAACTGCATTTCCAGCCCTACAG
CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCAACACA
CTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAAACTGCCCCAACTGCAA
CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGGTGCTCTGAAGTCT
CTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTTGGGGGCTGAGCAACATGGAAATTTTG
CAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCTGCAGGAACTTCAT
CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA
ACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATACACTGCACATTGGG
AACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAAC
AATGAAATTTCCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGGCTTGACAAACTGAGGCGACTGATACTC
CAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGT
GACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAAACTGCAACAATTGCATTTAAATACA
TCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCCCACAGTGGGTGGCGGAAAACAACTTTCAGAGCTTTGTA
AATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTTGCTGTTAGCCCAGATGGCTTTGTGTGT
GATGATTTTCCCAAACCCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTTC
ATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTTGCTTGGAAAAAAGACAATGAACTACTGCATGAT
GCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCACCATCCTTCGGCTG
CGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCATCTCCAATCACTTTGGTTCATCCTACTCTGTC
AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCATCCGAGCTGGGGCC
ATGGCACGCTTGGAGTGTGCTGCTGTGGGCACCCAGCCCCCCAGATAGCCTGGCAGAAGGATGGGGGCACAGAC
TTCCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCGAGGATGACGTGTTCTTTATCGTGGATGTGAAGATA
GAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTCAGCAAATGCAACTCTGACTGTC
CTAGAAACACCATCATTTTTGCGGCCACTGTTGGACCGAACTGTACCAAGGGAGAAACAGCCGTCCTACAGTGC
ATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGTAACCGAGAGGCAC
TTTTTTGCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAAATACACATGTGAG
ATGTCTAACACCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAACCTGCGACTCCCCT
CAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGCTGT
GTGGTGGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAAGGCGGAGGAATGAAGATTGCAGCATTACC
AACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCATCTCAGGGAACGTTAGCTGACAGGCAG
GATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACCAGTTTGTCACATCTTCAGGTGCTGGATTTTCCTTACCA
CAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTC
CTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGATCCTTTTGAAACA
TATCATACAGGTTGCAGTCCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTACATAAAGAAAAAG
GAGTGCTACCCATGTTCTCATCCTTCAGAAGAACCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACAT
GTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCC
TCTTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGGTACCTTTGGAAAA
GCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCAGCCAAGAGCCTTTTAT
TTGAAAGCTCATTCTTCCCCAGACTTGGACTCTGGGTCAGAGGAAGATGGGAAAGAAAGGACAGATTTTCAGGAA
GAAAATCACATTTGTACCTTAAACAGACTTAGAAAACTACAGGACTCCAAATTTTCACTCTTATGACTTGGAC
ACATAGACTGAATGAGACCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACTTCTATTTAAAAGAGAGAGAAT
CTTATGTTTTTTAAATGGAGTTATGAATTTTAAAAGGATAAAAATGCTTTATTTATACAGATGAACCAAAATTAC
AAAAAGTTATGAAAATTTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAAACTATTTTTAACTTTG
TTTTATGCAAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTTTATAATGCCAGA
TTTCTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTTTAAATAGAAGTT
ACTTCATTATATTTGCACATTATATTTAATAAAATGTGTCAATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 106

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAAERPCPTTCRCLGDLLDCSR
KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT
LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY
FDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKMQR
NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWEFC
QKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEIS
WTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK
KLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCDDF
PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQGGE
VMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGAMA
RLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQNSA
GSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHF
FAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDGWA
TVGVVIIAVVCCVVGTSLVWVVIIYHTRRNEDCSITNTDETNLPADIPSYLSSQGTLADRQ
DGYVSSESGSHHQFVTSSGAGFPLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGPMY
LKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISWPS
HVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDAYS
SFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQSYDLDT

Signal sequence:
amino acids 1-27

Transmembrane domain:
amino acids 808-828

N-glycosylation site.
amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519, 688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.
amino acids 886-890

Casein kinase II phosphorylation site.
amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378, 383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735, 799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022, 1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.
amino acids 667-675

N-myristoylation site.
amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433, 513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.
amino acids 58-80, 65-87

FIGURE 107

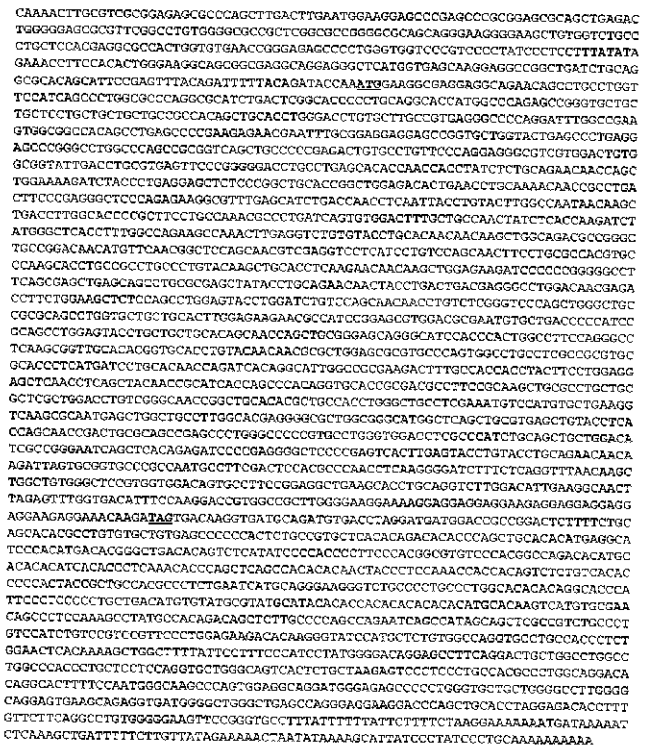


FIGURE 108

MEGEEAEQPAWFHQPWRPGASDSAPPAGTMAQSRVLLLLLLLPPQLHLGPVLAVRAPGFGRS
GGHSLSPEENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCGGIDLREFPGDLP
EHTNHLSLQNNQLEKIYPEELSRLHRLETLNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNK
LTLAPRFLPNALISVDFAANYLTKIYGLTPGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNV
EVLILSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN
ETPWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLLHSN
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMILHNQITGIGREDFATTYF
LEELNLSYNRITSPQVHRDAFRKLRLLRSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALA
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKD
RGRLGKEKEEEEEEEEEEEETR

Signal sequence:
amino acids 1-48

N-glycosylation site.
amino acids 243-247, 310-314, 328-332, 439-443

Casein kinase II phosphorylation site.
amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-myristoylation site.
amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341, 477-483, 498-502, 539-545, 548-554

Leucine zipper pattern.
amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493, 535-557

FIGURE 109

GGGAGGGGGCTCCGGGCGCCGCGCAGCAGACCTGCTCCGGCCGCGCGCCTCGCCGCTGTCCTCCGGGAGCGGCAG
CAGTAGCCCGGGCGGCGAGGGCTGGGGGTTCCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCCACCACCC
CAACCTGTTCCTCGCGCGCCACTGCGCTGCGCCCCAGGACCCGCTGCCCAACATGGATTTTCTCCTGGCGCTGGT
GCTGGTATCCTCGCTCTACCTGCAGGCGGCCGCCGAGTTCGACGGGAGGTGGCCCAGGCAAATAGTGTCATCGAT
TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTCAGCC
TGTGTGCCAACCACGATGCAAACATGGTGAATGTATCGGCCAAACAATGTCAAGTGTCATCCTGGTTATGCTGG
AAAAACCTGTAATCAAGATCTAAATGAGTGTGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA
CGGCAGCTACAAGTGCTACTGTCTCAACGGATATATGCTCATGCCGGATGGTTCCTGCTCAAGTGCCCTGACCTG
CTCCATGGCAAACTGTCAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCTGGCCT
GCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT
TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTCATAAAGGCTTCGATCTCATGTATATTGGAGG
CAAATATCAATGTCATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA
CGTACGTGGGTCCTACAAGTGCAAATGTAAAGAAGGATACCAGGGTGATGGACTGACTTGTGTGTATATCCCAAA
AGTTATGATTGAACCTTCAGGTCCAATTCATCTACCAAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAA
TAATAATTGGATTCCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAA
CAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAAGCCAACACCAATTCCTACTCCACCACCACCACCACC
CCTGCCAACAGACCTCAGAACACCTCTACCACCTACAACCCCAGAAAGGCCAACCACCGGACTGACAACTATAGC
ACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCCTCAGAAACCCAGAGG
AGATGTGTTCAGTGTTCTGGTACACAGTTGTAATTTTGACCATGGACTTTGTGATGGATCAGGGAGAAAGACAA
TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTCGGCAGCCAAAGCCCCAGG
GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCCTCATGCATTCAGGGGACCTGTGCCTGTCATTCAGGCA
CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAACACGGTGCCCACGGAGCAGCCCTGTG
GGGAAGAAATGGTGGCCATGGCTGGAGGCAAACACAGATCACCTTGCGAGGGCTGACATCAAGAGCGAATCACA
AAGATGATTAAAGGGTTGGAAAAAAAGATCTATGATGGAAAATTAAAGGAACTGGGATTATTGAGCCTGGAGAAG
AGAAGACTGAGGGGCAAACCATTGATGGTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTG
TTCTCCATATGCACTAAGAATAGAACAAGAGGAAACTGGCTTAGACTAGAGTATAAGGGAGCATTTCTTGGCAGG
GGCCATTGTTAGAATACTTCATAAAAAAAGAAGTGTGAAAATCTCAGTATCTCTCTCTTTCTAAAAAATTAGA
TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACAAATTATAGAATTTCCCAAA
AGATGTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG
CATTTCCCTCTTGACCTCCTAATGGAGAGGGATTGAAAGGGGAAGAGCCCACCAAATGCTGAGCTCACTGAAATA
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAAGAAAAAAGGAAACTATTTATTCCAAATGAGAGTATGATGGAC
AGATATTTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTCTTCATGGTAAAGGTATAAGCC
TTTCATTTGTTCAATGGATGATGTTTCAGATTTTTTTTTTTTAAGAGATCCTTCAAGGAACACAGTTCAGAGAG
ATTTTCATCGGGTGCATTCTCTCTGCTTCGTGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC
ACACCGGCAGACCTTTCCTTCACCTCATCAGTATGATTCAGTTTCTCTGTTATCAATTGGACTCTCCCAGGTTCCAC
AGAACAGTAATATTTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCATTTAACCTCGTAAAGGCAGGGCTGG
AGGGGGAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTTAATGGTTCATT
TCCTTTATGGTCATATAACTGCACAGCTGAAGATGAAAGGGGAAAATAAATGAAAACTTTACTTTTCGATGCCAA
TGATACATTGCACTAAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTTTATTATTTAATGTTTT
CTAAAATAAAAAATGTTAGTGGTTTTCCAAATGGCCTAATAAAACAATTATTTGTAAATAAAAACACTGTTAGTAAT

FIGURE 110

MDFLLALVLVSSLYLQAAAEPDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQP
RCKBGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD
GSCSSALTCSMANCQYGCDVVKGQIRCQCPSPGLHLAPDGRTCVDVDECATGRASCPRFRQC
VNTPGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQG
DGLTCVYIPKVMIEPSGPIHVPKGNGTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRP
TSKPTTRPTPKPTPIPTPPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDN
RVQTDPQKPRGDVFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGG
KAARLVLPLGRLMHSGDLCLSFRHKVTGLHSGTLQVFVRKHGAIGAALWGRNGGHGWRQTQI
TLRGADIKSESQR

Signal sequence:
amino acids 1-17

N-glycosylation site.
amino acids 273-277

Casein kinase II phosphorylation site.
amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.
amino acids 199-206

N-myristoylation site.
amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242, 421-427, 433-439, 462-468, 476-482

Aspartic acid and asparagine hydroxylation site.
amino acids 104-116, 186-198, 231-243

Cell attachment sequence.
amino acids 382-385

EGF-like domain cysteine pattern signature.
amino acids 75-87

FIGURE 111

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCCTTTAGATTGTGA
AATGTGGCTCAAGGTCTTCACAACTTTCCTTTCCTTTGCAACAGGTGCTTGCTCGGGGCTGA
AGGTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTCAGGCCCTCTACCTACCCGTC
CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA
CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAATACC
AACACAAGTTCACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCT
GATGAAGGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACTCTATCTGCCAGTCA
GAAGATACAAGTCACGGTTGATGATCCTGTCACAAAGCCAGTGGTGCAGATTCATCCTCCCT
CTGGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAGGGGGCACTCGG
CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC
TCCCCAAAACAATACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT
GCCTGGTGAGGAACCCTGTCAGTGAAATGGAAAGTGATATCATTATGCCCATCATATTATAT
GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT
TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCCAACACCTACT
CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTT
GCATCTGAGAAAGTAGCCCAGAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC
CGGCAGGCAAGATGAAACTCATTTCACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG
CACAGAAAGGAAAATCATTGTCACCTTTAGCAAGTATAACTGGAATATCACTATTTTTGATT
ATATCCATGTGTCTTCTCTTCCTATGGAAAAAATATCAACCCTACAAAGTTATAAAACAGAA
ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATG
CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG
ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT
GTATGAAGTTATTCAGCACATCCCTGCCCAGCAGCAAGACCATCCAGAGTGAACTTTCATGG
GCTAAACAGTACATTCGAGTGAAATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGT
ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA
TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT
GTGCAACAGAAAAACATGTTGGGGAAATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGG
GGAGAACGAAAGTGACAGGGGTTTCCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA
ATTAGTTCTACTCTACACCTTCACTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA
TGTGGAAACTTTACATTGTTCGATTTTCAGCAGACTTTGTTTTATTAAATTTTTATTAGTG
TTAAGAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA
CAAAGTAATAAGGATGGTTGTCACAAAAACAAAACTATGCCTTCTCTTTTTTTTCAATCACC
AGTAGTATTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTA
TTTTTTTCAAGGAAAGATGGATTCAAATAAATTATTCTGTTTTTGCTTTTAAAAAAAAAAAAA

FIGURE 112

MWLKVFTTFLSFATGACSGLKVTVPSHTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH
TMPKYLLGSVNKSVVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQGNGTLSASQ
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMCLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFS
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTV
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMDYVCCAYNNIT
GRQDETHFTVIITSVGLEKLAQKGKSLSPLASITGISLFLIISMCLLFLWKKYQPYKVIKQK
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRSVPASDCVSGQDLHSTV
YEVIQHIPAQQQDHPE

Signal sequence:
amino acids 1-18

Transmembrane domain:
amino acids 341-359

N-glycosylation site.
amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208, 276-280, 308-312

Casein kinase II phosphorylation site.
amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.
amino acids 272-280

N-myristoylation site.
amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237, 239-245

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 7-18

FIGURE 113

GCAAGCGGCGAAATGGCGCCCTCCGGGAGTCTTGCAGTTCCCCTGGCAGTCCTGGTGCTGTT
GCTTTGGGGTGCTCCCTGGACGCACGGCCGCCGGAGCAACGTTCGCGTCATCACGGACGAGA
ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCCTGCT
TGTCAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAA
TATTGCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTC
TTCCTACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG
AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTCATCATG
GTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATGTGGA
TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT
GTTTTTGCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC
AGATTGCCTTTGTCCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCCTTCAAAAAAAT
TATTATCAGAATCTGCACAACCTTTGAAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAA
GATGTTTCAGAAGAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGACTTTCCACAGAATGC
CATAAGACAACGCCTCTGGGTCCATCATTGGCCACAGATAAATCCTAGTTAAATTTTATAG
TTATCTTAATATTATGATTTTGATAAAAACAGAAGATTGATCATTTTGTTTGGTTTGAAGTG
AACTGTGACTTTTTTGAATATTGCAGGGTTCAGTCTAGATTGTCATTAAATTGAAGAGTCTA
CATTCAGAACATAAAAGCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATG
GTTTAAATAGTTCTCTAATTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGT
TTAATAATAACCTATTTCAAGTCTGAGTTTTGAAAATTTACATTTCCCAAGTATTGCATTAT
TGAGGTATTTAAGAAGATTATTTTAGAGAAAAATATTTCTCATTTGATATAATTTTTCTCTG
TTTCACTGTGTGAAAAAAAGAAGATATTTCCCATAAATGGGAAGTTTGCCCATTGTCTCAAG
AAATGTGTATTTCAGTGACAATTTCGTGGTCTTTTTAGAGGTATATTCCAAAATTTCCTTGT
ATTTTTAGGTTATGCAACTAATAAAAACTACCTTACATTAATTAATTACAGTTTTCTACACA
TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTTTAAGTTCATGGTATTCTCTTGATTC
CAACAAAGTTTGATTTTCTCTTGTATTTTTCTTACTTACTATGGGTTACATTTTTTATTTTT
CAAATTGGATGATAATTTCTTGGAAACATTTTTTATGTTTTAGTAAACAGTATTTTTTTGTT
GTTTCAAACTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAAATT
TTGGCCACTTTTTTCAGATTTACATCATTCTTGCTGAACTTCAACTTGAAATTGTTTTTTT
TTTCTTTTTGGATGTGAAGGTGAACATTCCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTA
TTTTACATTTTGAAAATTCAAAGAAGCTTAATATAAAAGTTTGCATTCTACTCAGGAAAAAG
CATCTTCTTGTATATGTCTTAAATGTATTTTTGTCCTCATATACAGAAAGTTCTTAATTGAT
TTTACAGTCTGTAATGCTTGATGTTTTAAAATAATAACATTTTTATATTTTTTAAAAGACAA
ACTTCATATTATCCTGTGTTCTTTCCTGACTGGTAATATTGTGTGGGATTTCACAGGTAAAA
GTCAGTAGGATGGAACATTTTAGTGTATTTTTACTCCTTAAAGAGCTAGAATACATAGTTTT
CACCTTAAAAGAAGGGGGAAAATCATAAATACAATGAATCAACTGACCATTACGTAGTAGAC
AATTTCTGTAATGTCCCCTTCTTTCTAGGCTCTGTTGCTGTGTGAATCCATTAGATTTACAG
TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAAATATTGTACCATT
AAAGAGTTTGGATGTGTAACTTGTGATGCCTTAGAAAAATATCCTAAGCACAAAATAAACCT
TTCTAACCACTTCATTAAAGCTGAAAAAAAAAAAAAAAAA

FIGURE 114

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL
QPEWESFAEWGEDLEVNIAKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI
NFISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVPAL
ATLFSGLLLGLCMIFVADCLCPSKRRRPQPYPYPSKKLLSESAQPLKKVEEEQEADEEDVSE
EEAESKEGTNKDFPQNAIRQRSLGPSLATDKS

Signal sequence:
amino acids 1-26

Transmembrane domain:
amino acids 182-201

Casein kinase II phosphorylation site.
amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.
amino acids 107-115

N-myristoylation site.
amino acids 20-26, 192-198

Amidation site.
amino acids 25-29

FIGURE 115

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTTAACTAATTCAACAAACGGGACCCTT
CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAGAGCGGG
AAGGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAAACTCTGGGTGTCTGC
ATTGCTGATGGCCTGGTTTGGTGTCCTGAGCTGTGTGCAGGCCGAATTCTTCACCTCTATTG
GGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC
CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC
TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCCTGTGAATGCCTACAAACTGG
TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCCTGCAGGACTCAGCTGCA
GGTTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGG
AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCA
GAGGGGAACTTCCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG
GGCCGCTCGGCCTACAATGAAGGGGACTATTATCATACGGTGTTGTGGATGGAGCAGGTGCT
AAAGCAGCTTGATGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCA
GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCCTGCTC
TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGA
GGAAGAGAGAGAAAAAACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCCAGAAGGCA
TCTATGAGAGGCCTGTGGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCGTGGG
GAGGGTGTCAAACTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAA
CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACA
TCGTCAGGTACTACGATGTCATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA
CCTAAACTTGCACGAGCCACCGTTCGTGATCCCAAGACAGGAGTCCTCACTGTCGCCAGCTA
CCGGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATC
GTCGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT
TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG
CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG
GTGGTGCCACCGTCTTCCCTGATCTGGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTG
TTCTGGTACAACCTCTTGCGGAGCGGGGAAGGTGACTACCGAACAAGACATGCTGCCTGCCC
TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT
TGAGACCTTGTGGATCAACAGAAGTTGACTGACATCCTTTTCTGTCCTTCCCCTTCCTGGTC
CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCCTTTGTATGTTCCTATCAGGCT
GATTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT
GTGACTGAAGTCCCAGCCCTTCCATTCAGCCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA
AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTTGTACCTCAGGTGTT
TTAGGTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACCTTGTTTACATGTTTGTTTTTAT
GGCATTTCTATCTATTGTGGCTTTACCAAAAAATAAAATGTCCCTACCAGAAAAAAAAAA

FIGURE 116

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA
NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALEDLVLQDSAAGFIANLSVQRQFFP
TDEDEIGAAKALMRLQDTYRLDPGTISRGELPGTKYQAMLSVDDCFGMGRSAYNEGDYYHTV
LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLLSLDPSHERAGGNLR
YFEQLLEEEREKTLTNQTEAELATPEGIYERPVDYLPERDVYESLCRGEGVKLTPRRQKRLF
CRYHHGNRAPQLLIAPFKEEDEWDSPHIVRYYDVMSDEEIERIKEIAKPKLARATVRDPKTG
VLTVASYRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS
RRPFDSGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGEGDYR
TRHAACPVLVGCKWVSNKWFHERGQEFLRPCGSTEVD

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270, 346-350, 365-369, 385-389, 457-461, 530-534

Tyrosine kinase phosphorylation site.

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

FIGURE 117

GCAGTATTGAGTTTTACTTCCTCCTCTTTTTAGTGGAAGACAGACCATAATCCCAGTGTGAGTGAAATTGATTGT
TTCATTTATTACCGTTTTGGCTGGGGGTTAGTTCCGACACCTTCACAGTTGAAGAGCAGGCAGAAGGAGTTGTGA
AGACAGGACAATCTTCTTGGGGATGCTGGTCCTGGAAGCCAGCGGGCCTTGCTCTGTCTTTGGCCTCATTGACCC
CAGGTTCTCTGGTTAAAACTGAAAGCCTACTACTGGCCTGGTGCCCATCAATCCATTGATCCTTGAGGCTGTGCC
CCTGGGGCACCCACCTGGCAGGGCCTACCACCATGCGACTGAGCTCCCTGTTGGCTCTGCTGCGGCCAGCGCTTC
CCCTCATCTTAGGGCTGTCTCTGGGGTGCAGCCTGAGCCTCCTGCGGGTTTCCTGGATCCAGGGGGAGGGAGAAG
ATCCCTGTGTCGAGGCTGTAGGGGAGCGAGGAGGGCCACAGAATCCAGATTCGAGAGCTCGGCTAGACCAAAGTG
ATGAAGACTTCAAACCCCGGATTGTCCCCTACTACAGGGACCCCAACAAGCCCTACAAGAAGGTGCTCAGGACTC
GGTACATCCAGACAGAGCTGGGCTCCCGTGAGCGGTTGCTGGTGGCTGTCCTGACCTCCCGAGCTACACTGTCCA
CTTTGGCCGTGGCTGTGAACCGTACGGTGGCCCATCACTTCCCTCGGTTACTCTACTTCACTGGGCAGCGGGGGG
CCCGGGCTCCAGCAGGGATGCAGGTGGTGTCTCATGGGGATGAGCGGCCCGCCTGGCTCATGTCAGAGACCCTGC
GCCACCTTCACACACACTTTGGGGCCGACTACGACTGGTTCTTCATCATGCAGGATGACACATATGTGCAGGCCC
CCCGCCTGGCAGCCCTTGCTGGCCACCTCAGCATCAACCAAGACCTGTACTTAGGCCGGGCAGAGGAGTTCATTG
GCGCAGGCGAGCAGGCCCGGTACTGTCATGGGGGCTTTGGCTACCTGTTGTCACGGAGTCTCCTGCTTCGTCTGC
GGCCACATCTGGATGGCTGCCGAGGAGACATTCTCAGTGCCCGTCCTGACGAGTGGCTTGGACGCTGCCTCATTG
ACTCTCTGGGCGTCGCCTGTGTCTCACAGCACCAGGGGCAGCAGTATCGCTCATTTGAACTGGCCAAAAATAGGG
ACCCTGAGAAGGAAGGGAGCTCGGCTTTCCTGAGTGCCTTCGCCGTGCACCCTGTCTCCGAAGGTACCCTCATGT
ACCGGCTCCACAAACGCTTCAGCGCTCTGGAGTTGGAGCGGGCTTACAGTGAAATAGAACAACTGCAGGCTCAGA
TCCGGAACCTGACCGTGCTGACCCCCGAAGGGGAGGCAGGGCTGAGCTGGCCCGTTGGGCTCCCTGCTCCTTTCA
CACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTCACAGAGCAGCACACCTTCTCCTGTGCAGATGGGG
CTCCCAAGTGCCCACTACAGGGGGCTAGCAGGGCGGACGTGGGTGATGCGTTGGAGACTGCCCTGGAGCAGCTCA
ATCGGCGCTATCAGCCCCGCCTGCGCTTCCAGAAGCAGCGACTGCTCAACGGCTATCGGCGCTTCGACCCAGCAC
GGGGCATGGAGTACACCCTGGACCTGCTGTTGGAATGTGTGACACAGCGTGGGCACCGGCGGGCCCTGGCTCGCA
GGGTCAGCCTGCTGCGGCCACTGAGCCGGGTGGAAATCCTACCTATGCCCTATGTCACTGAGGCCACCCGAGTGC
AGCTGGTGCTGCCACTCCTGGTGGCTGAAGCTGCTGCAGCCCCGGCTTTCCTCGAGGCGTTTGCAGCCAATGTCC
TGGAGCCACGAGAACATGCATTGCTCACCCTGTTGCTGGTCTACGGGCCACGAGAAGGTGGCCGTGGAGCTCCAG
ACCCATTTCTTGGGGTGAAGGCTGCAGCAGCGGAGTTAGAGCGACGGTACCCTGGGACGAGGCTGGCCTGGCTCG
CTGTGCGAGCAGAGGCCCCTTCCCAGGTGCGACTCATGGACGTGGTCTCGAAGAAGCACCCTGTGGACACTCTCT
TCTTCCTTACCACCGTGTGGACAAGGCCTGGGCCCGAAGTCCTCAACCGCTGTCGCATGAATGCCATCTCTGGCT
GGCAGGCCTTCTTTCCAGTCCATTTCCAGGAGTTCAATCCTGCCCTGTCACCACAGAGATCACCCCCAGGGCCCC
CGGGGGCTGGCCCTGACCCCCCTCCCCTCCTGGTGCTGACCCCTCCCGGGGGGCTCCTATAGGGGGAGATTTG
ACCGGCAGGCTTCTGCGGAGGGCTGCTTCTACAACGCTGACTACCTGGCGGCCCGAGCCCGGCTGGCAGGTGAAC
TGGCAGGCCAGGAAGAGGAGGAAGCCCTGGAGGGGCTGGAGGTCATGGATGTTTTCCTCCGGTTCTCAGGGCTCC
ACCTCTTTCGGGCCGTAGAGCCAGGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGCAGCCCACGGCTCAGTGAAG
AACTCTACCACCGCTGCCGCCTCAGCAACCTGGAGGGGCTAGGGGGCCGTGCCCAGCTGGCTATGGCTCTCTTTG
AGCAGGAGCAGGCCAATAGCACTTAGCCCGCCTGGGGGCCCTAACCTCATTACCTTTCCTTTGTCTGCCTCAGCC
CCAGGAAGGCAAGGCAAGATGGTGGACAGATAGAGAATTGTTGCTGTATTTTTTAAATATGAAAATGTTATTAA
ACATGTCTTCTGCC

FIGURE 118

MRLSSLLALLRPALPLILGLSLGCSLSLLRVSWIQGEGEDPCVEAVGERGGPQNPDSRARLD
QSDEDFKPRIVPYYRDPNKPYKKVLRTRYIQTELGSRERLLVAVLTSRATLSTLAVAVNRTV
AHHFPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETLRHLHTHFGADYDWFPIMQDDTY
VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRSLLLRLRPHLDGCRG
DILSARPDEWLGRCLIDSLGVGCVSQHGQQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE
GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPEGEAGLSWPVGLPAPFTPHSRFEV
LGWDYFTEQHTFSCADGAPKCPLQGASRADVGDALETALEQLNRRYQPRLRFQKQRLLNGYR
RFDPARGMSYTLDLLLECVTQRGHRRALARRVSLLRPLSRVEILPMPYVTEATRVQLVLPLL
VAEAAAAPAFLEAFAANVLEPREHALLTLLLVYGPREGGRGAPDPFLGVKAAAAELERRYPG
TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLTTVWTRPGPEVLNRCRMNAISGWQAFFP
VHFQEFNPALSPQRSPPGPPGAGPDPPSPPGADPSRGAPIGGRFDRQASAEGCFYNADYLAA
RARLAGELAGQEEEEEALEGLEVMDVFLRPSGLHLFRAVEPGLVQKPSLRDCSPRLSEELYHR
CRLSNLEGLGGRAQLAMALFEQEQANST

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 489-507

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-678

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

FIGURE 119

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCCTGTCCCCAAGCC
GTTCTAGACGCGGGAAAAATGCTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG
AAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA
TGCACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTCA
GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAAAACC
CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG
AGTTCTTCAGTTCTGAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATG
TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACTG
GTTCTTCCTTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTGTTAA
AAAAGGATCCATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT
GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT
CAATATCCCAGAAAAGTGTCCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC
AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA
AAAGATGTATTTAATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCC
CAACCAGGTAGTAGAAGGCTGTTGTTCAGATATGGCTGTTACTTTTAATGGACTGACTCCAA
ATCAGATGCATGTGATGATGTATGGGGTATACCGCCTTAGGGCATTTGGGCATATTTTCAAT
GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAGTGGTAGAAAAGCG
TGAATATGATCTTTGTATAGGACGTGTGTTGTCATTATTTGTAGTAGTAACTACATATCCAA
TACAGCTGTATGTTTCTTTTTCTTTTCTAATTTGGTGGCACTGGTATAACCACACATTAAAG
TCAGTAGTACATTTTTAAATGAGGGTCGTTTTTTTCTTTAAAACACATGAACATTGTAAATG
TGTTGGAAAGAAGTGTTTTAAGAATAATAATTTTGCAAATAAACTATTAATAAATATTATAT
GTGATAAATTCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTT
AAAAAATTTTAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT
TGTGATTAAAGTAAAACTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT
AAGCCTCCCCAAGTTCCAATGGATTTGCCTTCTCAAAATGTACAACTAAGCAACTAAAGAAA
ATTAAAGTGAAAGTTGAAAAAT

FIGURE 120

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME
LSKSFRVYCIILVKPKDVSLWAAVKETWTKHCDKAEPFSSENVKVFESINMDTNDMWLMMRK
AYKYAPDKYRDQYNWFFLARPTTFAIIENLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMEGG
IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNT
KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFL
PPNGSDND

Signal sequence:
amino acids 1-33

N-glycosylation site.
amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 319-323, 464-468

Casein kinase II phosphorylation site.
amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.
amino acids 736-743

N-myristoylation site.
amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 14-25

Cell attachment sequence.
amino acids 247-250

FIGURE 121

CCCACGCGTCCGATCTTACCAACAAAACACTCCTGAGGAGAAAGAAAGAGAGGGAGGGAGAG
AAAAAGAGAGAGAGAAACAAAAAACCAAAGAGAGAGAAAAAATGAATTCATCTAAATCAT
CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTTCCCAAATGTTCTTATGGACTGTTGCT
GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTCGCAT
CTTTCAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTCACAGAGCTCTCCTGCT
ACAATTATGGATCAGGTTCAGTCAAGAATTGTTGTCCATTGAACTGGGAATATTTTCAATCC
AGCTGCTACTTCTTTTCTACTGACACCATTTCCTGGGCGTTAAGTTTAAAGAACTGCTCAGC
CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCCTACAAGA
AACCTAAAATGAGAGAGTTTTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAGTGGCAA
TGGGTGGACGGCACACCTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA
CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA
ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG
AACAAAGGAAAATCTCTTTAAGAACAGAAGGCACAACTCAAATGTGTAAAGAAGGAAGAGCA
AGAACATGGCCACACCCACCGCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTC
ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAA

FIGURE 122

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIPQTCDEKFQLPEN
FTELSCYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSAMGAHLVVINSQEEQ
EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCATMRDSS
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

Signal sequence:
amino acids 1-42

N-glycosylation site.
amino acids 2-6, 62-66, 107-111

Casein kinase II phosphorylation site.
amino acids 51-55, 120-124, 163-167, 175-179, 181-185

N-myristoylation site.
amino acids 15-21, 74-80, 155-161

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 27-38

【国際公開パンフレット（コレクトバージョン）】

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/47, C12N 15/62, C07K 16/18 — A3

(11) International Publication Number: WO 00/15796

(43) International Publication Date: 23 March 2000 (23.03.00)

(21) International Application Number: PCT/US99/21090

(22) International Filing Date: 15 September 1999 (15.09.99)

(30) Priority Data:
PCT/US98/19330 16 September 1998 (16.09.98) US

(71) Applicant *(for all designated States except US)*: GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: CHEN, Jian [CN/US]; 121 York Drive, Princeton, NJ 08540 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). GURNEY, Austin, L. [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). HILLAN, Kenneth [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). PENNICA, Diane [US/US]; 2417 Hale Drive, Burlingame, CA 94010 (US). WOOD, William, I. [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US). YUAN, Jean [CN/US]; 176 West 37th Avenue, San Mateo, CA 94403 (US).

(74) Agents: KRESNAK, Mark, T. et al.; Genentech Inc., 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report:
24 August 2000 (24.08.00)

(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

(57) Abstract

The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

FOR THE PURPOSES OF INFORMATION ONLY

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

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

【手続補正書】
【提出日】平成12年6月14日(2000.6.14)
【手続補正１】
【補正対象書類名】明細書
【補正対象項目名】請求項４
【補正方法】変更
【補正の内容】
【請求項４】
登録番号ＡＴＣＣ２０９２５８、ＡＴＣＣ２０９２５６、ＡＴＣＣ２０９２６４、ＡＴＣＣ２０９２５０、ＡＴＣＣ２０９３７５、ＡＴＣＣ２０９３７８、ＡＴＣＣ２０９３８４、ＡＴＣＣ２０９３９６、ＡＴＣＣ２０９４２０、ＡＴＣＣ２０９４８０、ＡＴＣＣ２０９２６５、ＡＴＣＣ２０９２５７、ＡＴＣＣ２０９２６２、ＡＴＣＣ２０９２５３、ＡＴＣＣ２０９４０２、ＡＴＣＣ２０９４０１、ＡＴＣＣ２０９３９７、ＡＴＣＣ２０９４００、ＡＴＣＣ２０９３８５、ＡＴＣＣ２０９３６７、ＡＴＣＣ２０９４３２、ＡＴＣＣ２０９２６３、ＡＴＣＣ２０９２５１、ＡＴＣＣ２０９２５５、ＡＴＣＣ２０９２５２、ＡＴＣＣ２０９３７７、ＡＴＣＣ２０９３７０、ＡＴＣＣ２０９５２３、ＡＴＣＣ２０９３７２、ＡＴＣＣ２０９３７４、ＡＴＣＣ２０９３７３、ＡＴＣＣ２０９３８２、ＡＴＣＣ２０９３８３、ＡＴＣＣ２０９４０３、ＡＴＣＣ２０９３９８、ＡＴＣＣ２０９３９９、ＡＴＣＣ２０９３９２、ＡＴＣＣ２０９３８７、ＡＴＣＣ２０９３８８、ＡＴＣＣ２０９３９４、ＡＴＣＣ２０９４２１、ＡＴＣＣ２０９３９３、ＡＴＣＣ２０９４１８、ＡＴＣＣ２０９４８５、ＡＴＣＣ２０９４８３、ＡＴＣＣ２０９４８２、ＡＴＣＣ２０９４９１、ＡＴＣＣ２０９４８１、ＡＴＣＣ２０９４３８、ＡＴＣＣ２０９９２７、ＡＴＣＣ２０９４３９、ＡＴＣＣ２０９４８９、ＡＴＣＣ２０９４３３、ＡＴＣＣ２０９４８８、ＡＴＣＣ２０９４３４、ＡＴＣＣ２０９３９５、ＡＴＣＣ２０９４８６、ＡＴＣＣ２０９４９０、ＡＴＣＣ２０９４８４又はＡＴＣＣ２０９３７１で寄託されたＤＮＡの全長コード化配列を含む単離された核酸。
【手続補正２】
【補正対象書類名】明細書
【補正対象項目名】請求項１３
【補正方法】変更
【補正の内容】
【請求項１３】
登録番号ＡＴＣＣ２０９２５８、ＡＴＣＣ２０９２５６、ＡＴＣＣ２０９２６４、ＡＴＣＣ２０９２５０、ＡＴＣＣ２０９３７５、ＡＴＣＣ２０９３７８、ＡＴＣＣ２０９３８４、ＡＴＣＣ２０９３９６、ＡＴＣＣ２０９４２０、ＡＴＣＣ２０９４８０、ＡＴＣＣ２０９２６５、ＡＴＣＣ２０９２５７、ＡＴＣＣ２０９２６２、ＡＴＣＣ２０９２５３、ＡＴＣＣ２０９４０２、ＡＴＣＣ２０９４０１、ＡＴＣＣ２０９３９７、ＡＴＣＣ２０９４００、ＡＴＣＣ２０９３８５、ＡＴＣＣ２０９３６７、ＡＴＣＣ２０９４３２、ＡＴＣＣ２０９２６３、ＡＴＣＣ２０９２５１、ＡＴＣＣ２０９２５５、ＡＴＣＣ２０９２５２、ＡＴＣＣ２０９３７７、ＡＴＣＣ２０９３７０、ＡＴＣＣ２０９５２３、ＡＴＣＣ２０９３７２、ＡＴＣＣ２０９３７４、ＡＴＣＣ２０９３７３、ＡＴＣＣ２０９３８２、ＡＴＣＣ２０９３８３、ＡＴＣＣ２０９４０３、ＡＴＣＣ２０９３９８、ＡＴＣＣ２０９３９９、ＡＴＣＣ２０９３９２、ＡＴＣＣ２０９３８７、ＡＴＣＣ２０９３８８、ＡＴＣＣ２０９３９４、ＡＴＣＣ２０９４２１、ＡＴＣＣ２０９３９３、ＡＴＣＣ２０９４１８、ＡＴＣＣ２０９４８５、ＡＴＣＣ２０９４８３、ＡＴＣＣ２０９４８２、ＡＴＣＣ２０９４９１、ＡＴＣＣ２０９４８１、ＡＴＣＣ２０９４３８、ＡＴＣＣ２０９９２７、ＡＴＣＣ２０９４３９、ＡＴＣＣ２０９４８９、ＡＴＣＣ２０９４３３、ＡＴＣＣ２０９４８８、ＡＴＣＣ２０９４３４、ＡＴＣＣ２０９３９５、ＡＴＣＣ２０９４８６、ＡＴＣＣ２０９４９０、ＡＴＣＣ

２０９４８４又はＡＴＣＣ ２０９３７１で寄託されたヌクレオチドによってコードされるアミノ酸配列に対して少なくとも８０％の配列同一性を有する単離されたＰＲＯポリペプチド。
【手続補正３】
【補正対象書類名】明細書
【補正対象項目名】０２８１
【補正方法】変更
【補正の内容】
【０２８１】
実施例５１：ヒトＰＲＯ２４４をコードするｃＤＮＡクローンの単離
上記実施例１に記載したように、ｐｈｒａｐを用いて他のＥＳＴ配列に対してコンセンサスＤＮＡ配列を組み立てた。このコンセンサス配列に基づいて、対象とする配列を含むｃＤＮＡライブラリをＰＣＲにより同定するため、及びＰＲＯ２４４の全長コード化配列のクローンを単離するプローブとして使用するために、オリゴヌクレオチドを合成した。
一対のＰＣＲプライマー（正及び逆）を合成した：
５’－ＴＴＣＡＧＣＴＴＣＴＧＧＧＡＴＧＴＡＧＧＧ－３’（３０９２３．ｆ１）（配列番号：３７８）
５’－ＴＡＴＴＣＣＴＡＣＣＡＴＴＴＣＡＣＡＡＡＴＣＣＧ－３’（３０９２３．ｒ１）（配列番号：３７９）
また、プローブも合成した。
５’－ＧＧＡＧＧＡＣＴＧＴＧＣＣＡＣＣＡＴＧＡＧＡＧＡＣＴＣＴＴＣＡＡＡＣＣＣＡＡＧＧＣＡＡＡＡＴＴＧＧ－３’（３０９２３．ｐ１）（配列番号：２３７）
全長クローンの供給源について幾つかのライブラリをスクリーニングするために、ライブラリからのＤＮＡを上で同定したＰＣＲプライマー対でのＰＣＲ増幅によりスクリーニングした。次いで、ポジティブライブラリを、プローブオリゴヌクレオチド及びＰＣＲプライマーの一方を用いてＰＲＯ２４４遺伝子をコードするクローンを単離するのに使用した。
ｃＤＮＡライブラリの作成のためのＲＮＡはヒト胎児腎臓組織から単離した。
上記のように単離したクローンのＤＮＡ配列決定により、全長ＤＮＡ配列及びＰＲＯ２４４の誘導タンパク質配列が得られた。
ＰＲＯ２４４の全ヌクレオチド配列をＦｉｇ１２１（配列番号：３７６）に示す。クローンＤＮＡ３５６６８－１１７１は、単一のオープンリーディングフレームを含み、ヌクレオチド位置１０６－１０８に見かけの翻訳開始部位を持つ（Ｆｉｇ１２１）。予測されるポリペプチド前駆体は２１９アミノ酸長である。クローンＤＮＡ３５６６８－１１７１はＡＴＣＣに寄託され（ＤＮＡ３５６６３－１１７１と命名）、ＡＴＣＣ寄託番号２０９３７１が付与された。タンパク質は細胞質ドメイン（ａａ１－２０）、膜貫通ドメイン（ａａ２１－４６）、及び細胞外ドメイン（ａａ４７－２１９）を、ａａ５５－２０６のＣレクチンドメインと共に有している。
全長配列のＢＬＡＳＴ及びＦａｓｔＡ配列アラインメント分析法に基づいて、ＰＲＯ２４４は、肝臓レクチンニワトリ（４３％）、ＨＩＣｈｐ１２０結合Ｃ型レクチン（４２％）、マクロファージレクチン２（ＨＵＭＨＭＬ２－１，４１％）、及び配列ＰＲ３２１８８（４４％）と顕著なアミノ酸配列同一性を示す。
【手続補正４】
【補正対象書類名】明細書
【補正対象項目名】０３２６
【補正方法】変更
【補正の内容】
【０３２６】
材料の寄託
次の材料をアメリカン・タイプ・カルチャー・コレクション，１２３０１ パークローン

ドライブ，ロックヴィル，ＭＤ，ＵＳＡ（ＡＴＣＣ）に寄託した：

材料	ATCC寄託番号	寄託日
DNA32292-1131	ATCC 209258	1997年9月16日
DNA33094-1131	ATCC 209256	1997年9月16日
DNA33223-1136	ATCC 209264	1997年9月16日
DNA34435-1140	ATCC 209250	1997年9月16日
DNA27864-1155	ATCC 209375	1997年10月16日
DNA36350-1158	ATCC 209378	1997年10月16日
DNA32290-1164	ATCC 209384	1997年10月16日
DNA35639-1172	ATCC 209396	1997年10月17日
DNA33092-1202	ATCC 209420	1997年10月28日
DNA49435-1219	ATCC 209480	1997年11月21日
DNA35638-1141	ATCC 209265	1997年9月16日
DNA32298-1132	ATCC 209257	1997年9月16日
DNA33089-1132	ATCC 209262	1997年9月16日
DNA33786-1132	ATCC 209253	1997年9月16日
DNA35918-1174	ATCC 209402	1997年10月17日
DNA37150-1178	ATCC 209401	1997年10月17日
DNA38260-1180	ATCC 209397	1997年10月17日
DNA39969-1185	ATCC 209400	1997年10月17日
DNA32286-1191	ATCC 209385	1997年10月16日
DNA33461-1199	ATCC 209367	1997年10月15日
DNA40628-1216	ATCC 209432	1997年11月7日
DNA33221-1133	ATCC 209263	1997年9月16日
DNA33107-1135	ATCC 209251	1997年9月16日
DNA35557-1137	ATCC 209255	1997年9月16日
DNA34434-1139	ATCC 209252	1997年9月16日
DNA33100-1159	ATCC 209377	1997年10月16日
DNA35600-1162	ATCC 209370	1997年10月16日
DNA34436-1238	ATCC 209523	1997年12月10日

DNA33206-1165	ATCC 209372	1997年10月16日
DNA35558-1167	ATCC 209374	1997年10月16日
DNA35599-1168	ATCC 209373	1997年10月16日
DNA36992-1168	ATCC 209382	1997年10月16日
DNA34407-1169	ATCC 209383	1997年10月16日
DNA35841-1173	ATCC 209403	1997年10月17日
DNA33470-1175	ATCC 209398	1997年10月17日
DNA34431-1177	ATCC 209399	1997年10月17日
DNA39510-1181	ATCC 209392	1997年10月17日
DNA39423-1182	ATCC 209387	1997年10月17日
DNA40620-1183	ATCC 209388	1997年10月17日
DNA40604-1187	ATCC209394	1997年10月17日
DNA38268-1188	ATCC 209421	1997年10月28日
DNA37151-1193	ATCC 209393	1997年10月17日
DNA35673-1201	ATCC 209418	1997年10月28日
DNA40370-1217	ATCC 209485	1997年11月21日
DNA42551-1217	ATCC 209483	1997年11月21日
DNA39520-1217	ATCC 209482	1997年11月21日
DNA41225-1217	ATCC 209491	1997年11月21日
DNA43318-1217	ATCC 209481	1997年11月21日
DNA40587-1231	ATCC 209438	1997年11月7日
DNA41338-1234	ATCC 209927	1998年6月2日
DNA40981-1234	ATCC 209439	1997年11月7日
DNA37140-1234	ATCC 209489	1997年11月21日
DNA40982-1235	ATCC 209433	1997年11月7日
DNA41379-1236	ATCC 209488	1997年11月21日
DNA44167-1243	ATCC 209434	1997年11月7日
DNA39427-1179	ATCC 209395	1997年10月17日
DNA40603-1232	ATCC 209486	1997年11月21日
DNA43466-1225	ATCC 209490	1997年11月21日
DNA43046-1225	ATCC 209484	1997年11月21日
DNA35668-1171	ATCC 209371	1997年10月16日

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/21090

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C12N15/62 C07K16/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL - EMROD [Online] Entry CG48852, Acc.no. U48852, 18 March 1996 (1996-03-18) CHEN, H. ET AL.: "Cricetulus griseus HT protein mRNA, complete cds." XP002131426	1
Y	the whole document	2-4,6,8, 10-13, 17,18
	--- -/--	

[X] Further documents are listed in the continuation of box C. [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 February 2000	14 06 2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Smalt, R

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/21090

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL - EMEST6 [Online] Entry/Acc.no. AI066551, 3 August 1998 (1998-08-03) STRAUSBERG, R.: "col4f12.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:1566191 3' similar to TR:Q60438 Q60438 HT PROTEIN. ;, mRNA sequence." XP002131427	1,5,7,9
Y	the whole document	2-4,6,8, 10-13, 17,18
A	KLEIN R D ET AL: "Selection for genes encoding secreted proteins and receptors" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, no. 93, 1 July 1996 (1996-07-01), pages 7108-7113, XP002077277 ISSN: 0027-8424 the whole document	
A	YOKOYAMA-KOBAYASHI M ET AL: "A signal sequence detection system using secreted protease activity as an indicator" GENE,NL,ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, vol. 163, no. 2, 3 October 1995 (1995-10-03), pages 193-196, XP004041983 ISSN: 0378-1119 the whole document	
A	WO 97 07198 A (GENETICS INST) 27 February 1997 (1997-02-27) the whole document	
A	EP 0 834 563 A (SMITHKLINE BEECHAM CORP) 8 April 1998 (1998-04-08) the whole document	
P,X	WO 99 14328 A (CHEN JIAN ;GENENTECH INC (US); PENNICA DIANE (US); YUAN JEAN (US);) 25 March 1999 (1999-03-25) the whole document	1,3-18

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/21090

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

see additional sheet, subject1.

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet (1)) (July 1998)

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 99/21090

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: Invention 1 - claims 1-18, all partially

Nucleic acid with seq.ID.1, encoding a polypeptide comprising the amino acid sequence as respresented in seq.ID.2 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.2 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, and an antibody against said polypeptide.

2. Claims: Inventions 2-60: claims 1-18, all partially

Inventions 2-60: Idem as subject 1 but limited to each of the protein sequences as in Seq.ID's 2,4,12,18,23,28,34,39,49,59,64,69,71,73,84,91,96,104,109,114,119,127,132,137,142,148,153,159,164,170,175,177,185,190,195,201,207,213,221,227,236,245,250,255,257,259,261,263,285,290,292,294,310,315,320,325,332,339,341, and 377, and the corresponding nucleic acid sequences 3,11,17,22,27,33,38,48,58,63,68,70,72,83,90,95,103,108,113,118,126,131,136,141,147,152,158,163,169,174,176,184,189,194,200,206,212,220,226,235,244,249,254,256,258,260,262,284,289,291,293,309,314,319,324,331,338,340, and 376, whereby invention 2 is limited to Seq.ID's 3 and 4, invention 3 is limited to Seq.ID:11 and 12,....., and invention 60 is limited to Seq.ID:376 and 377.

For the sake of conciseness, the first subject matter is explicitly defined, the other subject matters are defined by analogy thereto.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/21090

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9707198	A	27-02-1997	US	5707829 A	13-01-1998
			AU	6712396 A	18-02-1997
			AU	6768596 A	12-03-1997
			CA	2227220 A	06-02-1997
			CA	2229208 A	27-02-1997
			EP	0839196 A	06-05-1998
			EP	0851875 A	08-07-1998
			JP	11510045 T	07-09-1999
			US	6043344 A	28-03-2000
			WO	9704097 A	06-02-1997
			US	5969093 A	19-10-1999
EP 0834563	A	08-04-1998	JP	10179178 A	07-07-1998
			US	5824504 A	20-10-1998
WO 9914328	A	25-03-1999	AU	9317898 A	05-04-1999
			ZA	9808460 A	19-03-1999
			AU	9312198 A	05-04-1999
			AU	9484398 A	05-04-1999
			WO	9914327 A	25-03-1999
			WO	9914234 A	25-03-1999
			AU	9395998 A	05-04-1999
			WO	9914241 A	25-03-1999
			AU	9317498 A	05-04-1999
			AU	1126099 A	17-05-1999
			AU	1288399 A	17-05-1999
			WO	9921998 A	06-05-1999
			WO	9921999 A	06-05-1999
			AU	1532499 A	15-06-1999
			WO	9927098 A	03-06-1999
			AU	1703399 A	15-06-1999
			WO	9927100 A	03-06-1999

Form PCT/ISA/210 (patent family annex) (July 1992)

フロントページの続き

(51)Int.Cl.[7]	FI		テーマコード（参考）
Ｃ１２Ｎ 1/21	Ｃ１２Ｎ 1/21		
Ｃ１２Ｎ 5/10	Ｃ１２Ｐ 21/02	C	
Ｃ１２Ｐ 21/02	Ｃ１２Ｐ 21/08		
Ｃ１２Ｐ 21/08	Ｃ１２Ｎ 5/00	B	

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

(72)発明者 ゴッダード，オードリー
アメリカ合衆国 カリフォルニア ９４１３１，サンフランシスコ，コンゴ ストリート １１０
(72)発明者 ガーニー，オースティン エル.
アメリカ合衆国 カリフォルニア ９４００２，ベルモント，デビー レーン １
(72)発明者 ヒラン，ケネス
アメリカ合衆国 カリフォルニア ９４１１４，サンフランシスコ，スワード ストリート ６４
(72)発明者 ペニカ，ダイアン
アメリカ合衆国 カリフォルニア ９４０１０，バーリンガム，ホール ドライブ ２４１７
(72)発明者 ウッド，ウィリアム アイ.
アメリカ合衆国 カルフォルニア ９４０１０，ヒルスボロー，サウスダウン コート ３５
(72)発明者 ユアン，ジーン
アメリカ合衆国 カルフォルニア ９４０００３，サンマテオ，ウェスト ３７ティーエイチ アヴェニュー １７６

Ｆターム(参考) 4B024 AA01 AA11 BA02 BA03 BA21 BA31 BA41 BA43 BA63 CA04 DA02 DA06 DA12 EA02 EA04 GA11 HA01 HA15
4B064 AG02 AG13 AG20 AG27 CA02 CA06 CA10 CA19 CC24 DA01 DA13
4B065 AA26X AA72X AA90X AA93X AA93Y AB01 AC14 BA02 CA24 CA25 CA44 CA46
4H045 AA10 AA11 AA30 BA10 CA40 DA01 DA20 DA50 DA76 EA22 EA28 EA51 EA54 FA74