

(19) 日本国特許庁(JP)

再公表特許(A1)

(11) 国際公開番号

WO2003/057887

発行日 平成17年5月19日(2005.5.19)

(43) 国際公開日 平成15年7月17日(2003.7.17)

(51) Int. Cl. ⁷	F I		
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A	
A 6 1 K 38/00	A 6 1 P 35/00		
A 6 1 P 35/00	C 1 2 M 1/00	A	
C 1 2 M 1/00	C 1 2 N 1/15		
C 1 2 N 1/15	C 1 2 N 1/19		
	審査請求 未請求 予備審査請求 有 (全 115 頁) 最終頁に続く		

出願番号	特願2003-558181 (P2003-558181)	(71) 出願人	301021533 独立行政法人産業技術総合研究所 東京都千代田区霞が関1-3-1
(21) 国際出願番号	PCT/JP2003/000008	(71) 出願人	597145779 アマシャム バイオサイエンス株式会社 東京都新宿区百人町3丁目25番1号 サンケンビルディング
(22) 国際出願日	平成15年1月6日(2003.1.6)	(71) 出願人	000237204 富士レビオ株式会社 東京都中央区日本橋浜町2丁目62番5号
(31) 優先権主張番号	特願2001-401507 (P2001-401507)	(74) 代理人	100089705 弁理士 社本 一夫
(32) 優先日	平成13年12月28日(2001.12.28)	(74) 代理人	100076691 弁理士 増井 忠武
(33) 優先権主張国	日本国(JP)		
(31) 優先権主張番号	特願2002-163832 (P2002-163832)		
(32) 優先日	平成14年6月5日(2002.6.5)		
(33) 優先権主張国	日本国(JP)		
(31) 優先権主張番号	特願2002-203696 (P2002-203696)		
(32) 優先日	平成14年7月12日(2002.7.12)		
(33) 優先権主張国	日本国(JP)		

最終頁に続く

(54) 【発明の名称】 新規UDP-N-アセチル-D-ガラクトサミン：ポリペプチドN-アセチルガラクトサミン転移酵素及びこれをコードする核酸

(57) 【要約】

コアタンパク質やペプチド配列のセリン、スレオニン残基の水酸基にN-アセチルガラクトサミンを 1 結合で転移する活性を有する酵素を単離し、その遺伝子の構造を明らかにすることにより、該酵素等の遺伝子工学的な生産、該酵素に基づくGalNAc修飾ペプチドあるいはGalNAc修飾タンパク質などの基質としての生産、該酵素に基づく未知O-結合型糖鎖の付加したタンパク質あるいは未知ペプチドのO-結合型糖鎖修飾位置の同定や、該遺伝子等に基づく疾患の診断を可能にすること。

配列表の配列番号1-7に示されるアミノ酸配列又は該アミノ配列において1若しくは複数個のアミノ酸が置換し若しくは欠失し、若しくは該アミノ配列に1若しくは複数個のアミノ酸が挿入され若しくは付加されたアミノ酸配列を有し、基質となるコアタンパク質あるいはペプチドのセリン、スレオニン残基の水酸基にN-アセチルガラクトサミン(GalNAc)を 1 結合で転移する活性を有するタンパク質および該タンパク質をコードする核酸を用いる。

【特許請求の範囲】**【請求項 1】**

配列表の配列番号 1 - 7 に示されるアミノ酸配列又は該アミノ配列において 1 若しくは複数個のアミノ酸が置換し若しくは欠失し、若しくは該アミノ配列に 1 若しくは複数個のアミノ酸が挿入され若しくは付加されたアミノ酸配列を有し、タンパク質あるいはペプチドのセリン、スレオニン残基の水酸基に N - アセチルガラクトサミン (GalNAc) を 1 結合で転移する活性を有するタンパク質。

【請求項 2】

前記タンパク質は、配列番号 1 - 7 に示されるアミノ酸配列と 70% 以上の同一性を有する請求項 1 記載のタンパク質。

10

【請求項 3】

前記タンパク質は、配列番号 1 - 7 に示されるアミノ酸配列と 90% 以上の同一性を有する請求項 1 記載のタンパク質。

【請求項 4】

前記タンパク質は、配列番号 1 に示されるアミノ酸配列又は該配列において 1 若しくは数個のアミノ酸が置換し若しくは欠失し、若しくは該アミノ配列に 1 若しくは数個のアミノ酸が挿入され若しくは付加されたアミノ酸配列を有する請求項 1 記載のタンパク質。

【請求項 5】

前記タンパク質は、配列番号 1 - 7 に示されるアミノ酸配列を有する請求項 4 記載のタンパク質。

20

【請求項 6】

請求項 1 ないし 5 のいずれか 1 項に記載のタンパク質をコードする核酸。

【請求項 7】

配列番号 8 - 14 に示される塩基配列を有する核酸とストリンジェントな条件下でハイブリダイズする、請求項 1 記載のタンパク質をコードする核酸。

【請求項 8】

配列表の配列番号 8 に示される塩基配列の 1 nt ~ 1629 nt、配列番号 9 に示される塩基配列の 1 nt ~ 1659 nt、配列番号 10 に示される塩基配列の 1 nt ~ 1812 nt、配列番号 11 に示される塩基配列の 1 nt ~ 1746 nt、配列番号 12 に示される塩基配列の 1 nt ~ 1824 nt、配列番号 13 に示される塩基配列の 1 nt ~ 1797 nt および配列番号 14 に示される塩基配列の 1 nt ~ 1920 nt までの塩基配列を有する請求項 6 記載の核酸。

30

【請求項 9】

請求項 6 ないし 8 のいずれか 1 項に記載の核酸を含み、宿主細胞中で該核酸を発現することができる組換えベクター。

【請求項 10】

請求項 6 ないし 8 のいずれか 1 項に記載の核酸が導入され、該核酸を発現する細胞。

【請求項 11】

請求項 6 ないし 8 のいずれか 1 項に記載の核酸、又は当該核酸の塩基配列と相補的な塩基配列からなる核酸にストリンジェントな条件下でハイブリダイズすることを特徴とする測定用核酸。

40

【請求項 12】

請求項 8 記載の核酸中の部分領域と相補的な配列を有する請求項 11 記載の測定用核酸。

【請求項 13】

プローブ又はプライマーである請求項 11 又は 12 記載の測定用核酸。

【請求項 14】

塩基数が 15 塩基以上である請求項 13 記載の測定用核酸。

【請求項 15】

請求項 11 ないし 14 のいずれか 1 項に記載の測定用核酸のうち少なくとも一つ含むマイクロアレイ。

50

【請求項 16】

請求項 11 ないし 14 のいずれか 1 項に記載の測定用核酸及び測定説明書を含む測定用キット。

【請求項 17】

請求項 1 ないし 5 のいずれか 1 項に記載のタンパク質を含む糖鎖合成キット。

【請求項 18】

請求項 1 ないし 5 のいずれか 1 項に記載のタンパク質を用いて合成した Tn 抗原 (GalNAc-Ser/Thr) を持つ多抗原性の糖ペプチドを用いた免疫感作療法。

【請求項 19】

請求項 1 ないし 5 のいずれか 1 項に記載のタンパク質を用いた、未知 O-結合型糖鎖の付加したタンパク質あるいは未知ペプチドの O-結合型糖鎖修飾位置の同定法。

10

【発明の詳細な説明】

技術分野

本発明は、コアタンパク質、ペプチド配列のセリン、スレオニン残基の水酸基に N-アセチルガラクトサミン (GalNAc) を 1 結合で転移する活性を有する新規な酵素及びそれをコードする核酸、並びに該核酸を測定するための核酸および測定用キットに関する。

背景技術

これまでに、コアタンパク質やペプチド配列のセリン、スレオニン残基の水酸基に N-アセチルガラクトサミンを 1 結合で転移する活性を有するヒトの酵素としては 8 種類が知られている [末尾の参考文献 1 - 8 を参照]。しかし、それぞれの遺伝子で発現組織や転移可能なアミノ酸配列や周囲のアミノ酸への糖修飾のパターンにより、使用される酵素は異なっている。即ち、全てのヒトのタンパク質に N-アセチルガラクトサミン修飾を施すには、あらゆる種類の UDP-N-アセチル-D-ガラクトサミン：ポリペプチド N-アセチルガラクトサミン転移酵素が必要になると考えられる。さらに、コアタンパク質やペプチド配列のセリン、スレオニン残基の水酸基に N-アセチルガラクトサミンを 1 結合で転移する活性を有する酵素は O-結合型糖鎖の根元酵素であり、この酵素の変化による細胞表面やタンパク質上の糖鎖密度の変化や糖鎖構造の欠如により、多数の疾患が引き起こされることが知られている。具体的な例としては、癌細胞の糖鎖修飾は正常細胞と比較して異なることが多い。最も特徴的な癌細胞の糖鎖変化は、ムチン型 (O-結合型) 糖鎖において起こり、これらの変化は癌細胞の浸潤や転移に関連し、癌細胞の脱分化や成長速度の上昇を反映する。その O-結合型糖鎖の根元には GalNAc が存在する。細胞表面の GalNAc の密度あるいは露出度の変化は Gal/GalNAc 特異的な Ca²⁺ 依存性レクチンをもつマクロファージの癌細胞の認識や接着度合と関連している [末尾の参考文献 9 を参照]。さらに、合成した Tn 抗原を持つ多抗原性の O-結合型糖ペプチドは、免疫感作療法に使用され、腫瘍耐性マウスの生存率が增加するという報告もなされている [末尾の参考文献 10 を参照]。このように、糖転移酵素の解析は、疾患の診断や治療においても重要であると考えられる。

20

30

発明の開示

コアタンパク質やペプチド配列のセリン、スレオニン残基の水酸基に N-アセチルガラクトサミンを 1 結合で転移する活性を有する酵素を単離し、その遺伝子の構造を明らかにすることにより、該酵素等の遺伝子工学的な生産、該酵素に基づく GalNAc 修飾ペプチドあるいは GalNAc 修飾タンパク質などの基質としての生産、該酵素に基づく未知 O-結合型糖鎖の付加したタンパク質あるいは未知ペプチドの O-結合型糖鎖修飾位置の同定や、該遺伝子等に基づく疾患の診断が可能になる。しかしながら、従来は UDP-N-アセチル-D-ガラクトサミン：ポリペプチド N-アセチルガラクトサミン転移酵素では受容体基質としないにもかかわらず、生体内に糖鎖構造が存在することから、未だ分離精製もされていない該酵素が存在し、かつ、該酵素の単離及び遺伝子の同定の手がかりはない。そのために、該酵素に対する抗体も作製されていない。

40

発明の詳細な説明

50

本発明は、配列表の配列番号 1 - 7 に示されるアミノ酸配列又は該アミノ配列において 1 若しくは複数個のアミノ酸が置換し若しくは欠失し、若しくは該アミノ配列に 1 若しくは複数個のアミノ酸が挿入され若しくは付加されたアミノ配列を有しコアタンパク質、ペプチド配列のセリン、スレオニン残基の水酸基に N - アセチルガラクトサミンを 1 結合で転移する活性を有するタンパク質を提供する。また、本発明は、該タンパク質をコードする核酸を提供する。さらに、本発明は、該核酸を含み、宿主細胞中で該核酸を発現することができる組換えベクターを提供する。さらに、本発明は、該組換えベクターにより形質転換され、前記核酸を発現する細胞を提供する。さらに、本発明は、核酸と特異的にハイブリダイズする、該核酸の測定用核酸及び測定用キットを提供する。さらに、本発明は、ペプチドコア糖鎖構造を合成するため、他の糖転移酵素の基質及びその調製方法を提供する。本発明はまた、本発明のタンパク質を含む糖鎖合成キットを提供し、本発明キットは糖鎖合成装置及び糖鎖合成試薬とともに使用することができる。さらに、該酵素を用いて合成した Tn 抗原 (GalNAc - Ser / Thr) を持つ多抗原性の糖ペプチドの免疫感作療法及び試薬、キット類を提供する。さらに、該酵素を用いた、未知 O - 結合型糖鎖の付加したタンパク質あるいは未知ペプチドの O - 結合型糖鎖修飾位置の同定法及び試薬、キット類を提供する。

10

発明を実施するための最良の形態

(1) タンパク質

下記実施例において詳述する方法によりクローニングされた、本発明のタンパク質をコードする核酸は、配列表の配列番号 8 - 14 に示される塩基配列を有し、それがコードする推定アミノ酸配列が、該塩基配列の下に記載されている。配列番号 1 - 7 には、該アミノ酸配列のみを取り出して示す。

20

下記実施例で得られた本発明のタンパク質は、次の性質を有する酵素である。なお、各性質及びその測定方法は下記実施例において詳述されている。

作用： コアタンパク質とポリペプチドのセリン、スレオニン残基の水酸基に N - アセチルガラクトサミンを 1 結合で転移する (EC 2 . 4 . 1 . 4 1、UDP - N - acetyl - D - galactosamine : polypeptide GalNAc - transferase)。触媒する反応を反応式で記載すると、

$$\text{UDP - N - acetyl - D - galactosamine} + \text{peptide (S/T)} \rightleftharpoons$$

30

$$\text{UDP} + \text{N - acetyl - beta - D - galactosamine - 1 - S/T - peptide}$$

基質特異性： コアタンパク質のポリペプチドのセリン、スレオニン残基の水酸基。

なお、一般に、酵素のような生理活性を有するタンパク質において、そのアミノ酸配列のうち、1 若しくは複数個のアミノ酸が置換し若しくは欠失し、若しくは該アミノ酸配列に 1 若しくは複数個のアミノ酸が挿入され若しくは付加された場合であっても、該生理活性が維持されることがあることは周知である。従って、配列番号 1 - 7 に示されるアミノ酸配列において 1 若しくは複数個のアミノ酸が置換し若しくは欠失し、若しくは該アミノ酸配列に 1 若しくは複数個のアミノ酸が挿入され若しくは付加されたアミノ酸配列を有し、コアタンパク質、ペプチド中のセリン、スレオニン残基の水酸基に N - アセチルガラクトサミンを 1 結合で活性を有するタンパク質 (以下、便宜的に「修飾タンパク質」) も本発明の範囲に含まれる。

40

ここで、「複数個」とは、好ましくは 1 - 200 個、より好ましくは 1 - 100 個、さらにより好ましくは 1 - 50 個、最も好ましくは 1 - 20 個である。一般的には、部位特異的な変異によってアミノ酸が置換された場合に、元々のタンパク質が有する活性は保持される程度に置換が可能なアミノ酸の個数は、好ましくは 1 - 10 個である。

このような修飾タンパク質のアミノ酸配列は、配列番号 1 - 7 に示されるアミノ酸配列と 70 % 以上、好ましくは 90 % 以上、さらに好ましくは 95 % 以上の相同性を有することが好ましい。なお、アミノ酸配列の相同性は、FASTA のような周知のコンピューターソフトを用いて容易に算出することができ、このようなソフトはインターネットによって

50

も利用に供されている。さらに、該修飾タンパク質としては、配列番号1 - 7に示されるアミノ酸配列又は該配列において1若しくは数個のアミノ酸が置換し若しくは欠失し、若しくは該アミノ酸配列に1若しくは数個のアミノ酸が挿入され若しくは付加されたアミノ配列を有するものが特に好ましい。

本発明のタンパク質は、クローニングされた核酸の塩基配列からの推定に基づいて、配列番号1 - 7のアミノ酸配列を有するが、その配列を有するタンパク質のみに限定されるわけではなく、本明細書中に記載した特性を有する限り全ての相同タンパク質を含むことが意図される。

本願明細書において、同一性のパーセントは、例えば、Altschulら(Nucl. Acids. Res. 25, p. 3389 - 3402, 1997)に記載されているBLASTプログラム、あるいはPearsonら(Proc. Natl. Acad. Sci. USA, p. 2444 - 2448, 1988)に記載されているFASTAを用いて配列情報と比較し決定することが可能である。当該プログラムは、インターネット上でNational Center for Biotechnology Information (NCBI)、あるいはDNA Data Bank of Japan (DDBJ)のウェブサイトから利用することが可能である。各プログラムによる同一性検索の各種条件(パラメーター)は同サイトに詳しく記載されており、一部の設定を適宜変更することが可能であるが、検索は通常デフォルト値を用いて行う。なお、当業者に用いられる、配列比較の他のプログラムもまた、使用可能である。

一般的に、同様の性質を有するアミノ酸同士の置換(例えば、ある疎水性アミノ酸から別の疎水性アミノ酸への置換、ある親水性アミノ酸から別の親水性アミノ酸への置換、ある酸性アミノ酸から別の酸性アミノ酸への置換、あるいはある塩基性アミノ酸から別の塩基性アミノ酸への置換)を導入した場合、得られる修飾タンパク質はもとのタンパク質と同様の性質を有することが多い。遺伝子組換え技術を使用して、このような所望の変異を有する組換えタンパク質を作製する手法は当業者に周知であり、このような修飾タンパク質も本発明の範囲に含まれる。

本発明のタンパク質は、例えば、後述の実施例に従って、本発明の核酸による配列番号2に記載のDNA配列を大腸菌、酵母、昆虫、または動物細胞に、それぞれの宿主で増幅可能な発現ベクターを用いて導入および発現させることにより、当該タンパク質を大量に得ることができる。

本発明によって、このタンパク質のアミノ酸配列およびそれをコードするDNA配列が開示されれば、当該配列またはその一部を利用して、ハイブリダイゼーション、PCR等の核酸増幅反応等の遺伝子工学的手法を用いて、他の生物種から同様の生理活性を有するタンパク質をコードする遺伝子を容易に単離することができる。このような場合、それらの遺伝子がコードする新規タンパク質も本発明の範囲に含まれる。

なお、本発明のタンパク質は、そのアミノ酸配列が上述した通りのものであり、前記酵素活性を有するものであれば、タンパク質に糖鎖が結合していてもよい。

(2) 核酸

本発明は、配列番号1 - 7で示されるアミノ酸配列をコードする核酸及び上記修飾タンパク質のアミノ酸配列をコードする核酸も提供する。本発明の核酸は、一本鎖および二本鎖型両方のDNA、およびそのRNA相補体も含む。DNAには、例えば、天然由来のDNA、組換えDNA、化学合成したDNA、PCRによって増幅されたDNA、およびそれらの組み合わせが含まれる。核酸としてはDNAが好ましい。なお、周知の通り、コドンには縮重があり、1つのアミノ酸をコードする塩基配列が複数存在するアミノ酸もあるが、上記アミノ酸配列をコードする塩基配列であれば、いずれの塩基配列を有するものも本願発明の範囲に含まれる。なお、下記実施例において実際にクローニングされたcDNAの塩基配列が配列番号2に示されている。配列2に示す塩基配列を有する核酸とストリンジェントな条件下(例えば、5 x Denhardt's reagent, 6 x SSC, 0.5% SDS又は0.1% SDSといった一般的なハイブリダイゼーション溶液を用いて50 ~ 65 で反応を行なう)において、ハイブリダイズし、かつ、上記

修飾タンパク質をコードする核酸も本発明の範囲内に入る。

本明細書において、「ストリンジェントな条件下」とは、中程度又は高程度なストリンジェントな条件においてハイブリダイズすることを意味する。具体的には、中程度のストリンジェントな条件は、例えば、DNAの長さに基づき、一般の技術を有する当業者によって、容易に決定することが可能である。基本的な条件は、Sambrookら、Molecular Cloning: A Laboratory Manual、第3版、Vol. 1、7.42-7.45 Cold Spring Harbor Laboratory Press, 2001に示され、そしてニトロセルロースフィルターに関し、5×SSC、0.5% SDS、1.0mM EDTA (pH 8.0)の前洗浄溶液、約40-50での、約50%ホルムアミド、2×SSC-6×SSC(又は約42での約50%ホルムアミド中の、スターク溶液(Stark's solution)などの他の同様のハイブリダイゼーション溶液)のハイブリダイゼーション条件、および約60、0.5×SSC、0.1% SDSの洗浄条件の使用が含まれる。高ストリンジェントな条件もまた、例えばDNAの長さに基づき、当業者によって、容易に決定することが可能である。一般的に、こうした条件は、中程度にストリンジェントな条件よりも高い温度及び/又は低い塩濃度でのハイブリダイゼーション及び/又は洗浄を含み、例えば上記のようなハイブリダイゼーション条件、及びおよそ68、0.2×SSC、0.1% SDSの洗浄を伴うと定義される。当業者は、温度および洗浄溶液塩濃度は、プローブの長さ等の要因に従って、必要に応じて調整可能であることを認識するであろう。

10

本発明の核酸は、例えば以下の方法により調製することが可能である。

20

公知のUDP-N-アセチルガラクトサミン転移酵素遺伝子の塩基配列をクエリーとして塩基配列の検索を行い、EST配列を得ることができた。さらに、この配列から見出したゲノム配列又はその一部を利用して、ハイブリダイゼーションや核酸増幅反応等の遺伝子工学の基本的手法を用いてcDNAライブラリーなどから本発明核酸(例えば本発明DNA)を調製することができる。

ここで、核酸増幅反応は、例えば、ポリメラーゼ連鎖反応(PCR)[Saiki R. K. et al., Science, 230, 1350-1354(1985)], ライゲース連鎖反応(LCR)[Wu D. Y. et al., Genomics, 4, 560-569(1989); Barringer K. J. et al., Gene, 89, 117-122(1990); Barany F., Proc. Natl. Acad. Sci. USA, 88, 189-193(1991)], 及び転写に基づく増幅[Kwoh D. Y. et al., Proc. Natl. Acad. Sci. USA, 86, 1173-1177(1989)]等の温度循環を必要とする反応、並びに鎖置換反応(SDA)[Walker G. T. et al., Proc. Natl. Acad. Sci. USA, 89, 392-396(1992); Walker G. T. et al., Nuc. Acids Res., 20, 1691-1696(1992)], 自己保持配列複製(3SR)[Guatelli J. C., Proc. Natl. Acad. Sci. USA, 87, 1874-1878(1990)], 及びQレプリカーゼシステム[リザイルディら、BioTechnology, 6, 1197-1202(1988)]等の恒温反応を含む。また、欧州特許第0525882号に記載されている標的核酸と変異配列の競合増幅による核酸配列に基づく増幅(Nucleic Acid Sequence Based Amplification: NASABA)反応等も利用可能である。好ましくはPCR法である。

30

後述する実施例に示すように、各種プライマーを使用した場合、PCR産物としてDNA断片が得られるので、これを例えばアガロースゲル電気泳動等の分子量によりDNA断片を篩い分ける方法で分離し、特定のバンドを切り出す方法等の常法に従って単離して本発明の核酸を得ることができる。

40

上記のようなハイブリダイゼーション、核酸増幅反応等を使用してクローニングされる相同な核酸は、配列表の配列番号1に記載の塩基配列に対して少なくとも50%以上、好ましくは60%以上、より好ましくは70%以上、さらに好ましくは80%以上、さらにな

50

お好ましくは90%以上、最も好ましくは95%以上の同一性を有する。

同一性パーセントは、視覚的検査および数学的計算によって決定することが可能である。あるいは、2つの核酸配列の同一性パーセントは、Devereuxら, *Nucl. Acids Res.*, 12, 387 (1984)に記載され、そしてウイスコンシン大学遺伝学コンピューターグループ(UWCG)より入手可能なGAPコンピュータープログラム、バージョン6.0を用いて、配列情報を比較することによって、決定可能である。GAPプログラムの好ましいデフォルトパラメーターには:(1)ヌクレオチドに関する単一(unary)比較マトリックス(同一に対し1および非同に対して0の値を含む)、並びにSchwartz及びDayhoff監修, *Atlas of Protein Sequence and Structure*, pp. 353-358, National Biomedical Research Foundation (1979)に記載されるような、Gribskov及びBurgess, *Nucl. Acids Res.*, 14, 6745 (1986)の加重比較マトリックス;(2)各ギャップに対する3.0のペナルティおよび各ギャップ中の各記号に対しさらに0.10のペナルティ;及び(3)末端ギャップに対するペナルティなし、が含まれる。当業者に用いられる、配列比較の他のプログラムもまた、使用可能である。

なお、本発明のタンパク質は、そのアミノ酸配列が上記した通りのものであり、上記した酵素活性を有するものであれば、タンパク質に糖鎖が結合していてもよい。すなわち、本発明の「タンパク質」は「糖タンパク質」をも包含する。

(3) 組換えベクターと形質転換体

本発明によれば、単離した本発明の核酸を含む組換えベクターが提供される。プラスミド等のベクターに本発明核酸のDNA断片を組込む方法としては、例えば、Sambrook, J.ら, *Molecular Cloning, A Laboratory Manual* (3rd edition), Cold Spring Harbor Laboratory, 1.1 (2001)に記載の方法などが挙げられる。簡便には、市販のライゲーションキット(例えば、宝酒造製等)を用いることもできる。このようにして得られる組換えベクター(例えば、組換えプラスミド)は、宿主細胞(例えば、大腸菌DH5、TB1、LE392、XL-392、又はXL-1Blue等)に導入される。

プラスミドを宿主細胞に導入する方法としては、Sambrook, J.ら, *Molecular Cloning, A Laboratory Manual* (3rd edition), Cold Spring Harbor Laboratory, 16.1 (2001)に記載の塩化カルシウム法または塩化カルシウム/塩化ルビジウム法、リン酸カルシウム法、リポフェクション法、エレクトロポレーション法、エレクトロインジェクション法、PEGなどの化学的な処理による方法、遺伝子銃などを用いる方法などが挙げられる。

ベクターは、簡単には当業界において入手可能な組換え用ベクター(例えば、プラスミドDNA等)に所望の遺伝子を常法により連結することによって調製することができる。用いられるベクターの具体例としては、大腸菌由来のプラスミドとして、例えば、pFLAG-CMV1、pDONR201、pBluescript、pUC18、pUC19、pBR322等が例示されるが、これらに限定されない。

当業者であれば制限末端は発現ベクターに適合するように適宜選択することが可能である。発現ベクターは、「本発明酵素を発現させたい宿主細胞」に適したものを当業者であれば適宜選択することができる。このように本発明発現ベクターは上記の本発明核酸が目的の宿主細胞中で発現しうるように遺伝子発現に關与する領域(プロモーター領域、エンハンサー領域、オペレーター領域等)が適切に配列されており、さらに本発明核酸が適切に発現するように構築されていることが好ましい。また、発現ベクターの構築は、制限処理及び連結作業を必要としない、Gatewayシステム(インビトロジェン社)を用いることもできる。Gatewayシステムとは、PCR産物の方向性を維持したままクローニングができ、また、DNA断片を適切に改変した発現ベクターにサブクローニングを可能にした部位特異的な組換えを利用したシステムである。具体的には、PCR産物とドナ

10

20

30

40

50

ーベクターとから部位特異的な組換え酵素であるBPクロナーゼによってエントリークローンを作成し、その後、このクローンと別の組換え酵素であるLBクロナーゼによって組換え可能なデスティネーションベクターにPCR産物を移入することにより、発現系に対応した発現クローンを調製するものである。最初にエントリークローンを作成すれば、制限酵素やりガーゼで作業する手間の係るサブクロニングステップが不要である点を特徴の一つとする。

発現ベクターの種類は、原核細胞および/または真核細胞の各種の宿主細胞中で所望の遺伝子を発現し、所望のタンパク質を生産する機能を有するものであれば特に限定されないが、例えば、哺乳類用発現ベクターとしてpFLAG-CMV1、pcDNA3.1、pGreenLanternなどが好ましく、大腸菌用発現ベクターとしてpQE-30、pQE-60、pMAL-C2、pMAL-p2、pSE420などが好ましく、酵母用発現ベクターとしてpYES2(サッカロマイセス属)、pPIC3.5K、pPIC9K、pA0815(以上ピキア属)、昆虫用発現ベクターとしてpFastBacp、pBacPAK8/9、pBK283、pVL1392、pBlueBac4.5などが好ましい。

上記「本発明発現ベクター」を宿主細胞に組み込み、形質転換体を得ることができる。上記「宿主細胞」として真核細胞(哺乳類細胞、酵母、昆虫細胞等)であっても原核細胞(大腸菌、枯草菌等)であっても使用することができる。本発明の形質転換体をえるための宿主細胞は、特に限定されず、さらに、または、ヒト(例えば、HeLa、293T、SH-SY5Y)、マウス(例えば、Neuro2a、NIH3T3)、サル(例えば、COS-1)由来の培養細胞でもよい。これらはいずれも公知であり、市販されているか(例えば、大日本製薬社)、あるいは公共の研究機関(例えば、理研セルバンク)より入手可能である。あるいは、胚、器官、組織若しくは非ヒト個体も使用可能である。

ところで、「本発明の核酸」は、ヒトゲノムライブラリーから発見された核酸であるため、本発明においては真核細胞を本発明の形質転換体の宿主細胞として用いるとより天然物に近い性質を有した「本発明酵素」が得られる(例えば糖鎖が付加された態様など)と考えられる。従って、「宿主細胞」としては真核細胞、特に哺乳類細胞を選択することが好ましい。動物細胞としてはマウス由来、アフリカツメガエル由来、ラット由来、ハムスター由来、サル由来またはヒト由来の細胞若しくはそれらの細胞から樹立した培養細胞株などが例示される。

宿主細胞として細菌、特に大腸菌を用いる場合、一般に発現ベクターは少なくとも、プロモーター/オペレーター領域、開始コドン、所望のタンパク質をコードする遺伝子、終止コドン、ターミネーターおよび複製可能単位から構成される。

宿主細胞として酵母、植物細胞、動物細胞または昆虫細胞を用いる場合には、一般に発現ベクターは少なくとも、プロモーター、開始コドン、所望のタンパク質をコードする遺伝子、終止コドン、ターミネーターを含んでいることが好ましい。またシグナルペプチドをコードするDNA、エンハンサー配列、所望の遺伝子の5'側および3'側の非翻訳領域、選択マーカー領域または複製可能単位などを適宜含んでもよい。

本発明のベクターにおいて、好適な開始コドンとしては、メチオニンコドン(ATG)が例示される。また、終止コドンとしては、常用の終止コドン(例えば、TAG、TGA、TAAなど)が例示される。

複製可能単位とは、宿主細胞中でその全DNA配列を複製することができる能力をもつDNAを意味し、天然のプラスミド、人工的に修飾されたプラスミド(天然のプラスミドから調製されたプラスミド)および合成プラスミド等が含まれる。好適なプラスミドとしては、E.coliではプラスミドpQE30、pET又はpCAL若しくはそれらの人工的修飾物(pQE30、pET又はpCALを適当な制限酵素で処理して得られるDNAフラグメント)が、酵母ではプラスミドpYES2若しくはpPIC9Kが、また昆虫細胞ではプラスミドpBacPAK8/9等があげられる。

エンハンサー配列、ターミネーター配列については、例えば、それぞれSV40に由来するもの等、当業者において通常使用されるものを用いることができる。

選択マーカーとしては、通常使用されるものを常法により用いることができる。例えばテトラサイクリン、アンピシリン、またはカナマイシンもしくはネオマイシン、ハイグロマイシンまたはスペクチノマイシン等の抗生物質耐性遺伝子などが例示される。

発現ベクターは、少なくとも、上述のプロモーター、開始コドン、所望のタンパク質をコードする遺伝子、終止コドン、およびターミネーター領域を連続的かつ環状に適当な複製可能単位に連結することによって調製することができる。またこの際、所望により制限酵素での消化やT4 DNAリガーゼを用いるライゲーション等の常法により適当なDNAフラグメント（例えば、リンカー、他の制限酵素部位など）を用いることができる。

本発明の発現ベクターの宿主細胞への導入〔形質転換（形質移入）〕は従来公知の方法を用いて行うことができる。

例えば、細菌（*E. coli*, *Bacillus subtilis*等）の場合は、例えばCohenらの方法〔*Proc. Natl. Acad. Sci. USA*, 69, 2110 (1972)〕、プロトプラスト法〔*Mol. Gen. Genet.*, 168, 111 (1979)〕やコンピテント法〔*J. Mol. Biol.*, 56, 209 (1971)〕によって、*Saccharomyces cerevisiae*の場合は、例えばHinnenらの方法〔*Proc. Natl. Acad. Sci. USA*, 75, 1927 (1978)〕やリチウム法〔*J. B. Bacteriol.*, 153, 163 (1983)〕によって、植物細胞の場合は、例えばリーフディスク法〔*Science*, 227, 129 (1985)〕、エレクトロポレーション法〔*Nature*, 319, 791 (1986)〕によって、動物細胞の場合は、例えばGrahamの方法〔*Virology*, 52, 456 (1973)〕、昆虫細胞の場合は、例えばSummerらの方法〔*Mol. Cell Biol.*, 3, 2156-2165 (1983)〕によってそれぞれ形質転換することができる。

なお、組換えベクターの構築及びそれを用いて本発明の核酸を宿主細胞に導入する方法の具体例が下記実施例2に詳述されている。

(4) タンパク質の単離・精製

近年、遺伝子工学的手法として、形質転換体を培養、生育させて、その培養物、生育物から目的物質を単離・精製する手法が確立されている。

本発明のタンパク質は、上記の如く調整された発現ベクターを含む形質転換細胞を栄養培地で培養することによって発現（生産）することができる。栄養培地は、宿主細胞（形質転換体）の生育に必要な炭素源、無機窒素源もしくは有機窒素源を含んでいることが好ましい。炭素源としては、たとえばグルコース、デキストラン、可溶性デンプン、ショ糖、メタノールなどが、例示される。無機窒素源もしくは有機窒素源としては、例えばアンモニウム塩類、硝酸塩類、アミノ酸、コーンステープ・リカー、ペプトン、カゼイン、肉エキス、大豆粕、バレイショ抽出液などが例示される。また、所望により他の栄養素（例えば無機塩（例えば、塩化ナトリウム、塩化カルシウム、リン酸二水素ナトリウム、塩化マグネシウム）、ビタミン類、抗生物質（例えばテトラサイクリン、ネオマイシン、アンピシリン、カナマイシン等）など）を含んでいてもよい。培養は、当業界において知られている方法により行われる。培養条件、例えば温度、培地のpH及び培養時間は、本発明のタンパク質が大量に生産されるように適宜選択される。

本発明のタンパク質は、上記培養により得られる培養物より以下のようにして取得することができる。すなわち、本発明のタンパク質が宿主細胞内に蓄積する場合には、遠心分離やろ過などの操作により宿主細胞を集め、これを適当な緩衝液（例えば濃度が10～100 mM程度のトリス緩衝液、リン酸緩衝液、HEPES緩衝液、MES緩衝液などの緩衝液。pHは用いる緩衝液によって異なるが、pH5.0～9.0の範囲が望ましい）に懸濁した後、用いる宿主細胞に適した方法で細胞を破壊し、遠心分離により宿主細胞の内容物を得る。一方、本発明のタンパク質が宿主細胞外に分泌される場合には、遠心分離やろ過などの操作により宿主細胞と培地を分離し、培養液を得る。宿主細胞破壊液、あるいは培養液はそのまま、または硫酸沈殿と透析を行なった後に、本発明のタンパク質の単離・精製に供することができる。単離・精製の方法としては、以下の方法が挙げることが

10

20

30

40

50

できる。即ち、当該タンパクに6×ヒスチジンやGST、マルトース結合タンパクといったタグを付けている場合には、一般に用いられるそれぞれのタグに適したアフィニティークロマトグラフィーによる方法を挙げることができる。一方、そのようなタグを付けずに本発明のタンパク質を生産した場合には、例えば後述する実施例に詳しく述べられている方法、即ちイオン交換クロマトグラフィーによる方法を挙げることもできる。また、これに加えてゲルろ過や疎水性クロマトグラフィー、等電点クロマトグラフィーなどを組み合わせる方法も挙げることもできる。

本発明のタンパク質を、セリン、スレオニン残基を有するコアタンパク質、ペプチド配列、糖タンパク質、オリゴ糖または多糖等に作用させることにより、N-アセチルガラクトサミンが1結合で転移される。従って、本発明のタンパク質は、糖タンパク質の糖鎖の修飾や、糖類の合成に用いることができる。さらに、このタンパク質を免疫原として動物に投与することにより、該タンパク質に対する抗体を作製することができ、該抗体を用いて免疫測定法により該タンパク質を測定することが可能になる。従って、本発明のタンパク質およびこれをコードする核酸は、このような免疫原の作製に有用である。

10

本発明のタンパク質はまた、精製および同定を容易にするために添加されるペプチドを含んでもよい。こうしたペプチドには、例えば、ポリ-Hisまたは米国特許第5,011,912号およびHoppら, Bio/Technology, 6:1204, 1988に記載される抗原性同定ペプチドが含まれる。こうしたペプチドの1つはFLAG(登録商標)ペプチド、Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys(配列番号41)であり、該ペプチドは非常に抗原性であり、そして特異的なモノクローナル抗体が可逆的に結合するエピトープを提供し、発現された組換えタンパク質の迅速なアッセイおよび容易な精製を可能にする。4E11と称されるネズミハイブリドーマは、本明細書に援用される米国特許第5,011,912号に記載されるように、特定の二価金属陽イオンの存在下で、FLAG(登録商標)ペプチドに結合するモノクローナル抗体を産生する。4E11ハイブリドーマ細胞株は、寄託番号HB 9259下に、アメリカン・タイプ・カルチャー・コレクション(American Type Culture Collection)に寄託されている。FLAG(登録商標)ペプチドに結合するモノクローナル抗体は、Eastman Kodak Co., Scientific Imaging Systems Division、コネチカット州ニューヘブンより入手可能である。

20

30

具体的には、後述される実施例2に記載されるように、本発明のタンパク質を発現する発現ベクターにFLAGのcDNAを挿入し、FLAG標識したタンパク質を発現させ、抗FLAG抗体によって、該タンパク質の発現を確認することができる。

(5) 測定用核酸

本発明により、本発明の新規酵素のcDNAの塩基配列が明らかになったので、該酵素のmRNA又はcDNAと特異的にハイブリダイズする、前記本発明の測定用核酸(以下、単に「測定用核酸」)が本発明により提供された。ここで、「特異的」とは、検査対象となる細胞中に存在する他の核酸とハイブリダイズせず、上記本発明の核酸とのみハイブリダイズするという意味である。測定用核酸は、上記本発明の核酸に示される塩基配列を有する核酸中の部分領域と相同的な配列を有することが一般的に好ましいが、1~2塩基程度の不一致があっても差し支えないことが多い。尚、「測定」には、検出、増幅、定量、および半定量のいずれもが包含される。

40

(a) プライマー

本発明測定用核酸を核酸増幅反応用のプライマーとして使用する場合、本発明測定用核酸は、オリゴヌクレオチドであって、

配列番号1-7に示すタンパク質をコードする遺伝子の塩基配列から以下の条件を満たすように2つの領域を選択し：

- 1) 各領域の長さが15-50塩基であること；
- 2) 各領域中のG+Cの割合が40-70%であること；

上記領域と同じ塩基配列若しくは上記領域に相補的な塩基配列を有する一本鎖DNAを製

50

造し、または、上記一本鎖DNAによってコードされるアミノ酸残基を変化させないように遺伝子暗号の縮重を考慮した一本鎖DNAの混合物を製造し、さらに必要であれば上記タンパク質をコードする遺伝子の塩基配列に対する結合特異性を失わないように修飾した上記一本鎖DNAを製造する

ことを含む方法により製造された当該オリゴヌクレオチドが提供される。

本発明のプライマーは、本発明核酸の部分領域と相同的な配列を有することが好ましいが、1または2塩基の不一致があっても差し支えない。

なお、本発明のプライマーの塩基数は15塩基以上、好ましくは18塩基以上、50塩基以下である。

(b) プローブ

本発明測定用核酸をプローブとして使用する場合、本発明測定用核酸は、配列番号8-14に記載の塩基配列の全体又は部分領域と相同的な配列を有することが好ましい。本発明の測定用核酸が、cDNAプローブの場合、塩基数は、15塩基以上、好ましくは20塩基以上で、それぞれ遺伝子のコード領域の全長以下が好ましい。配列番号8-14に記載した塩基配列又はその相補的な塩基配列と20%以下、好ましくは10%以下の不一致があっても、プローブとしての機能を果たし得る。また、本発明の測定用核酸が、合成オリゴヌクレオチドの場合、塩基数は15塩基以上、好ましくは20塩基以上である。合成オリゴヌクレオチドの場合、長さによるが、配列番号8-14に記載した塩基配列又はその相補的な塩基配列と1または2塩基程度の不一致があってもプローブとしての機能を果たし得る。また、前述のプライマーも本発明のプローブとして使用し得る。

本発明のプローブには、該プローブが標的配列とハイブリダイズしたことを検出または確認するために、蛍光標識、放射標識、ビオチン標識等の標識を付した標識プローブが含まれる。被検核酸またはその増幅物を固相化し、標識プローブとハイブリダイズさせ、洗浄後、固相に結合された標識を測定することにより、検体中に被検核酸が存在するかを決定することができる。

(c) マイクロアレイ

本発明測定用核酸は、マイクロアレイとして使用することができる。マイクロアレイは、ゲノム機能の大量で迅速な解析を可能にする手段である。具体的には、固相基盤、例えばガラス上に高密度に固定した多数の異なる核酸プローブに、標識した核酸をハイブリダイズさせ、各々のプローブからのシグナルを検出し、得られたデータを解析するものである。本明細書において「マイクロアレイ」というときは、固相基盤、例えば、メンブラン、フィルター、チップ、ガラス上に、本発明測定用核酸を配列したものをいう。

(d) 被検核酸の測定

被検核酸の部分領域と相補的な配列を有する核酸をPCRのような遺伝子増幅法のプライマー、又はプローブとして用いて被検核酸を測定する方法自体は周知であり、下記実施例3には、ヒト細胞中の本発明のタンパク質のmRNAをノーザンプロット及びインサイチュハイブリダイゼーションにより測定した方法が具体的に詳述されている。また、被検核酸の測定には、前記マイクロアレイを使用することもできる。

PCRのような核酸増幅法自体は、この分野において周知であり、そのための試薬キット及び装置も市販されているので容易に行うことができる。上記した本発明の測定用核酸の一对をプライマーとして用い、被検核酸を鋳型として用いて核酸増幅法を行なうと、被検核酸が増幅されるのに対し、検体中に被検核酸が含まれない場合には増幅が起きないので、増幅産物を検出することにより検体中に被検核酸が存在するか否かを知ることができる。増幅産物の検出は、増幅後の反応溶液を電気泳動し、バンドをエチジウムブロミド等で染色する方法や、電気泳動後の増幅産物をナイロン膜等の固相に不動化し、被検核酸と特異的にハイブリダイズする標識プローブとハイブリダイズさせ、洗浄後、該標識を検出することにより行うことができる。また、クエンチャー蛍光色素とレポーター蛍光色素を用いたいわゆるリアルタイム検出PCRを行うことにより、検体中の被検核酸の量を定量することも可能である。なお、リアルタイム検出PCR用のキットも市販されているので、容易に行うことができる。さらに、電気泳動バンドの強度に基づいて被検核酸を半定量

10

20

30

40

50

することも可能である。なお、被検核酸は、mRNAでも、mRNAから逆転写したcDNAであってもよい。被検核酸としてmRNAを増幅する場合には、上記一对のプライマーを用いたNASBA法(3SR法、TMA法)を採用することもできる。NASBA法自体は周知であり、そのためのキットも市販されているので、上記一对のプライマーを用いて容易に実施することができる。

プローブとしては、上記測定用核酸に蛍光標識、放射標識、ビオチン標識等の標識を付した標識プローブを用いることができる。被検核酸又はその増幅物を固相化し、標識プローブとハイブリダイズさせ、洗浄後、固相に結合された標識を測定することにより、検体中に被検核酸が存在するか否かを調べることができる。あるいは、測定用核酸を固相化し、被検核酸をハイブリダイズさせ、固相に結合した被検核酸を標識プローブ等で検出することも可能である。このような場合、固相に結合した測定用核酸もプローブと呼ばれる。

本発明の酵素を、セリン、スレオニン残基を有するペプチド、コアタンパク質、糖タンパク質等に作用させることにより、N-アセチルガラクトサミンを1結合させる。従って、本発明の酵素は、糖タンパク質の糖鎖の修飾や、糖類の合成に用いることができる。さらに、この酵素を免疫原として動物に投与することにより、該酵素に対する抗体を作製することができ、該抗体を用いて免疫測定法により該酵素を測定することが可能になる。従って、本発明の酵素及びこれをコードする核酸は、このような免疫原の作製に有用である。

実施例

以下、本発明を実施例に基づきより具体的に説明する。もっとも、本発明は下記実施例に限定されるものではない。

実施例 1

遺伝子データベースの検索とGalNAc-T11の塩基配列決定

既存のUDP-N-アセチルガラクトサミン転移酵素遺伝子(X85018、X85019、X92689、Y08564、Y08565、AJ002744、AB032959、AB040672)と類似遺伝子の高い相同性を有するアミノ酸配列領域を用いて、遺伝子データベースから類似遺伝子の検索を行った。用いた配列はUDP-N-アセチルガラクトサミン転移酵素遺伝子における配列番号：X85018、X85019、X92689、Y08564、Y08565、AJ002744、AB032959、AB040672である。また検索は、Blast[Altschul et al., J. Mol. Biol. 215, 403-410(1990)]等のプログラムを利用した。

その結果、EST配列Gene Bank Accetion No. AB032956が見出された。AB032956の配列において、そのうちのヌクレオチド118-1745に新規遺伝子のORFが見出された。次に、表1で示すプライマー(配列番号15-17)を用いて、後述の実施例8で述べる方法で長さ1629塩基の遺伝子をクローニングし、塩基配列を決定した。後述する実施例11及び12により、この新規遺伝子の翻訳産物がUDP-N-アセチルガラクトサミン転移酵素活性を示したことから、この新規UDP-N-アセチルガラクトサミン転移酵素遺伝子をGalNAcT-11と命名した(配列番号8)。

実施例 2

遺伝子データベースの検索とGalNAc-T12の塩基配列決定

実施例1と同様の検索法を用いて、EST配列Gene Bank Accetion No. AK022753が見出された。AK022753の配列において、ヌクレオチド62-1735に新規遺伝子のORF領域が見出された。表1で示すプライマー(配列番号18-20)を用いて、後述の実施例8で述べる方法で遺伝子をクローニングし、塩基配列を決定したところ、長さ1659塩基の遺伝子が得られた。この遺伝子はGenBank No. AK022753と異なるAlternative splicing formであることが判明した。後述する実施例11及び12により、この新規遺伝子の翻訳産物がUDP-N-アセチルガラクトサミン転移酵素活性を示したことから、この新規UDP-N-アセチルガラクトサミン転移酵素遺伝子をGalNAcT-12と命名した

(配列番号9)。

表1. 実施例8で各遺伝子クローニングに用いたプライマー配列とクローニングに用いたcDNAの由来(組織または細胞)

遺伝子	プライマー-F1 配列	プライマー-F2 配列	プライマー-R 配列	クローニングに用いたcDNAの由来(組織または細胞)
GalNAcT-11	5'- GGGGACAAGTTTGTACA AAAAAGCAGGCTTCGAA GGAGATAGAACCATGAG GAAGATCCGCGCCAATG CCA-3' (配列番号15)	5'- GGGGACAAGTTTGTACAA AAAAAGCAGGCTTCGGGG CGCGCAGAGGGCAGGCAG GA-3' (配列番号16)	5'- GGGGACCACTTTGTACA AGAAAGCTGGGTCCCAC AGATCCTCGCTCTGAAG GGC-3' (配列番号17)	ヒト胎児脳 Marathon-Ready cDNA (CLONTECH Cat. 7402-1)
GalNAcT-12	5'- GGGGACAAGTTTGTACA AAAAAGCAGGCTTCGAA GGAGATAGAACCATGCG GCGCCTGACTCGTCGGC TG-3' (配列番号18)	5'- GGGGACAAGTTTGTACAA AAAAAGCAGGCTTCGTGCA GACCCCTAAGCCTTCGGA CGCTG-3' (配列番号19)	5'- GGGGACCACTTTGTACA AGAAAGCTGGGTCTTAA GAGCTCACCATGTCCCA GTGCTG-3' (配列番号20)	MKN45細胞 胃
GalNAcT-13	5'- GGGGACAAGTTTGTACA AAAAAGCAGGCTTCGAA GGAGATAGAACCATGAG GCGGAAGGAGAAGCGGC TC-3' (配列番号21)	5'- GGGGACAAGTTTGTACAA AAAAAGCAGGCTTCGGCA GCCCCGACGGCACCCCTGG GGGA-3' (配列番号22)	5'- GGGGACCACTTTGTACA AGAAAGCTGGGTCTCAG TTCTATGAATTTTTC CAA-3' (配列番号23)	KHM-1B IgA ミエローマ (Health Science Research Resources Bank, JCRB0133)
GalNAcT-14	5'- GGGGACAAGTTTGTACA AAAAAGCAGGCTTCGAA GGAGATAGAACCATGTG GGGGCGCACGGCGCGGC GG-3' (配列番号24)	5'- GGGGACAAGTTTGTACAA AAAAAGCAGGCTTCGGCG CGAGCCGGTCATGCCGCG GC-3' (配列番号25)	5'- GGGGACCACTTTGTACA AGAAAGCTGGGTCTCAT AACATGCGCTCTTTGAA GAACCA-3' (配列番号26)	ヒト肺 Marathon-Ready cDNA (CLONTECH Cat. 7408-1)
GalNAcT-15	5'- GGGGACAAGTTTGTACA AAAAAGCAGGCTTCGAA GGAGATAGAACCATGGT GTGCACCAGGAAGACCA AAACT-3' (配列番号27)	5'- GGGGACAAGTTTGTACAA AAAAAGCAGGCTTCGGGG GCAGGAGCCGGCCCGCA CAAG-3' (配列番号28)	5'- GGGGACCACTTTGTACA AGAAAGCTGGGTCTCAG GACGCGAGGCTCCTCAG GACG-3' (配列番号29)	NCI-H929 IgA ミエローマ (ATCC, CRL-9068)
GalNAcT-16	5'- GGGGACAAGTTTGTACA AAAAAGCAGGCTTCATG GCTTCACTGAGAAGAGT CAAAG-3' (配列番号30)	5'- GGGGACAAGTTTGTACAA AAAAAGCAGGCTTCGCAG CGGAGACGCCTTCCACGA G-3' (配列番号31)	5'- GGGGACCACTTTGTACA AGAAAGCTGGGTCTTAC TTGATGGAGTCTTAAT GGTCCAC-3' (配列番号32)	ヒト脳 (whole), Marathon-Ready cDNA (CLONTECH, Cat. 7400-1)
GalNAcT-17	5'- GGGGACAAGTTTGTACA A AAAAAGCAGGCTTCATGC TCCTAAGGAAGCGATAC AGGCAC-3' (配列番号33)	5'- GGGGACAAGTTTGTACAA AAAAAGCAGGCTTCGCCAG GTACCCCTGGACTTTGG GG-3' (配列番号34)	5'- GGGGACCACTTTGTACA AGAAAGCTGGGTCTTCA TCGTTTCATCCACAGCAT TGATCT-3' (配列番号35)	ヒト脳 (whole), Marathon-Ready cDNA (CLONTECH, Cat. 7400-1)

実施例3

遺伝子データベースの検索とGalNAc-T13の塩基配列決定

実施例 1 で述べた同様な検索法を用いて、EST 配列 Gene Bank Accetion No. BG772195、BF359671、BF359677、BF359680、AK023782 が見出された。これらの EST は GENETYX (SOFTWARE DEVELOPMENT CO., LTD) を用いた塩基配列のアラインメント分析から同一遺伝子由来の EST であることが判明した。そのうち BG772195 のヌクレオチド 139 - 141 に開始コドン、AK023782 のヌクレオチド 1276 - 1278 に終止コドンを発見した。そして、表 1 で示すプライマー (配列番号 21 - 23) を用いて、後述の実施例 8 で述べる方法で遺伝子をクローニングし、塩基配列を決定したところ、長さ 1812 塩基の遺伝子が得られた。後述する実施例 11 及び 12 により、この新規遺伝子の翻訳産物を UDP - N - アセチルガラクトサミン転移酵素活性が示したことから、この新規 UDP - N - アセチルガラクトサミン転移酵素遺伝子を GalNAcT - 13 と命名した (配列番号 10)。

10

実施例 4

遺伝子データベースの検索と GalNAc - T14 の塩基配列決定

実施例 1 と同様の検索法を用いて、Human EST 配列 GenBank Accetion No. AK024865、BE677813、AI800923、または Human Genome 配列 GenBank Accetion No. AL136084 が見出された。GENETYX を用いた塩基配列のアラインメント解析から、AL136084 のヌクレオチド 5286 - 5658 の塩基配列、EST の AK024865、BE677813、AI800923 の塩基配列は同一遺伝子由来の配列であることが判明した。そのうち、AL136084 のヌクレオチド 5286 - 5288 に開始コドンが存在し、AK024865 のヌクレオチド 1317 - 1319 に終止コドンが存在していることを発見した。そして、表 1 で示すプライマー (配列番号 24 - 26) を用いて、後述の実施例 8 で述べる方法で遺伝子をクローニングし、塩基配列を決定したところ、長さ 1746 bps の遺伝子が得られた。実施例 11 及び 12 により、この新規遺伝子の翻訳産物が UDP - N - アセチルガラクトサミン転移酵素活性を示したことから、この新規 UDP - N - アセチルガラクトサミン転移酵素遺伝子を GalNAcT - 14 と命名した (配列番号 11)。

20

実施例 5

遺伝子データベースの検索と GalNAc - T15 の塩基配列決定

実施例 1 と同様の検索法を用いて、Human EST 配列 GenBank Accetion No. BG420528、BG675527 が見出された。以上の EST 配列は Macaca fascicularis の脳由来の GenBank Accetion No. AB050513 とそれぞれの部分に相同性を示したので、同一の Human 遺伝子由来の塩基配列であると考えられた。そのうち BG420528 においては、ヌクレオチド 19 - 21 に翻訳開始コドンが存在した。しかし、この遺伝子の翻訳終止コドンが不明のため、CLONTECH 社の肺の Marathon cDNA を用いて 3' RACE (rapid amplification of cDNA ends) により情報を取得した。

30

具体的には、Marathon Lung cDNA (Clontech, Cat. 7408 - 1) 付属の AP1 プライマーと (DNA 断片の両側に AP1、AP2 のアダプターがついている)、見出した BG675527 配列のヌクレオチド 731 - 755 までの部分に設定したプライマー GP151 (CCGGCTCATCGAATGCCGCTACGCC) (配列番号 36) で PCR (94 5 秒、68 4 分、を 40 サイクル) を行った。常法により精製し、塩基配列を決定したところ終止コドンを発見した (配列番号 12)。

40

そして、表 1 で示すプライマー (配列番号 27 - 29) を用いて、後述の実施例 8 で述べる方法で遺伝子をクローニングし、塩基配列を決定したところ、長さ 1824 塩基の遺伝子が得られた。後述する実施例 11 及び 12 により、この新規遺伝子の翻訳産物が UDP - N - アセチルガラクトサミン転移酵素活性を示したことから、この新規 UDP - N - ア

50

セチルガラクトサミン転移酵素を GalNAcT-15 と命名した (配列番号 12)。

実施例 6

遺伝子データベースの検索と GalNAc-T16 の塩基配列決定

実施例 1 と同様の検索法を用いて、Human EST 配列 GenBank Accetion No. BF527569、AU121279 が見出された。さらに、ゲノム配列 GenBank accetion No. NT-007799.4 が見出され、上記 2 つの EST が同一遺伝子であることを見出した。しかし、翻訳開始点は不明であったので、CLONTECH 社の胎児脳の Marathon cDNA を用いて 5' RACE ならびにクローニングを行った。

具体的には、Marathon Fetal brain cDNA 付属の AP1 プライマーと (DNA 断片の両側に AP1、AP2 のアダプターがついている)、見出した配列部分に設定したプライマー GP157 (5' - AGAGCATTCCTCTTGGTGTAGAAGCC - 3') (配列番号 37) で PCR (94 5 秒、68 4 分、を 40 サイクル) を行った。さらに、Marathon cDNA 付属の AP2 プライマーと、配列部分に設定したプライマー GP-158 プライマー (5' - TTCCGCTCAATGTGGGCCACCCGTG - 3') (配列番号 38) で nested PCR (94 5 秒、68 4 分、を 40 サイクル) を行った。目的とするフラグメントをゲルから切りだし、常法により精製し、塩基配列を決定した。

さらに、決定した配列の情報をもとに、再度 5' RACE を行った。具体的には AP1 プライマーと GP-158 プライマーで PCR (94 5 秒、68 4 分、を 40 サイクル) を行った。その後、AP2 プライマーと O6-R3 プライマー (5' - GGTCCTCCGGCGTCCACCAAG - 3') (配列番号 39) を用いて nested PCR (94 5 秒、68 4 分、を 40 サイクル) を行った。目的とするフラグメントをゲルから切りだし、常法により精製し、塩基配列を決定した。

最終的に、表 1 で示すプライマー (配列番号 30 - 32) を用いて、後述の実施例 8 で述べる方法で遺伝子をクローニングし、塩基配列を決定したところ、長さ 1797 塩基の遺伝子が得られた。後述する実施例 11 及び 12 により、この新規遺伝子の翻訳産物が UDP-N-アセチルガラクトサミン転移酵素活性を示したことから、この新規 UDP-N-アセチルガラクトサミン転移酵素遺伝子を GalNAcT-16 と命名した (配列番号 13)。

実施例 7

遺伝子データベースの検索と GalNAc-T17 の塩基配列決定

実施例 1 と同様の検索法を用いて、Human EST 配列 GenBank Accetion No. BG699346, BG714178, BG715977 が見出された。これらの EST 配列はヒトゲノム配列 GenBank accetion No. AC090935 と相同性が見出され、Genscan DNA 解析予測ソフトにより、ヌクレオチド 51185 - 51304 の領域に、翻訳開始コドンを含むエキソンとヌクレオチド 103445 - 103533 の領域に翻訳終止コドンを含むエキソンを見出した。この情報に従い、表 1 で示すプライマー (配列番号 33 - 35) を用いて、後述の実施例 8 で述べる方法で遺伝子をクローニングし、塩基配列を決定したところ、長さ 1920 塩基の遺伝子が得られた。後述する実施例 11 及び 12 により、この新規遺伝子の翻訳産物が UDP-N-アセチルガラクトサミン転移酵素活性を示したことから、この新規 UDP-N-アセチルガラクトサミン転移酵素遺伝子を GalNAcT-17 と命名した (配列番号 14)。

実施例 8

7 種類の GalNAc-T 遺伝子の発現ベクターへの組み込み

7 種類の GalNAc-T 遺伝子の発現系を作成するため、まず 7 種類の GalNAc-T 遺伝子全長あるいは遺伝子の一部をインビトロジェン社の Gateway システムの pDONR (商標) 201 および pFastBac 由来のデスティネーションベクターに組み込み、さらにインビトロジェン社の Bac-to-Bac システムによる Bacmid を

作成した。以下に詳細を述べる。

(1) GatewayシステムによるpFastBacへの組み込み

1 エントリークローンの作成

上記の実施例1-7で得られた各遺伝子配列により、表1で示したプライマーセット及びDNAポリメラーゼとしてExpand High Fidelity PCR system (Roche Cat. No. 1146 173)を用いて、94 15秒、68 3分を30サイクルの反応条件で各DNA断片を増幅した。目的の断片をアガロースゲル電気泳動後のゲルから切りだし、精製後BP Clonase反応によってpDONR201へ組込んだ。各遺伝子に対するプライマーF1とRで増幅されたDNA断片は各遺伝子の全長ORF、各遺伝子に対するプライマーF2とRで増幅されたDNA断片は活性測定用トランケート型遺伝子としてpDONR201へ組み込み、「エントリークローン」を作成した。反応は目的とするDNA断片5 μ l、pDONR201 1 μ l (150ng)、反応緩衝液2 μ l、BP Clonase mix 2 μ lを25 $^{\circ}$ Cで1時間インキュベートして行った。Proteinase Kを1 μ l加えて37 $^{\circ}$ Cで10分おき反応を停止した。

10

その後上記mix全量(11 μ l)をコンピテントセル(大腸菌DH5 α)100 μ lと混合し、ヒートショック法の後、カナマイシンを含むLBプレートにまいた。翌日コロニーをとり、直接PCRで目的DNAを確認し、ベクターを抽出・精製した。

2 発現クローンの作成

上記エントリークローンは挿入部位の両側にラムダファージが大腸菌から切り出される際の組換え部位であるattLを持つもので、LR Clonase(ラムダファージの組換え酵素Int、IHF、Xisを混合したもの)とデスティネーションベクターと混合することで、挿入部位がデスティネーションベクターに移り、発現クローンが作成される。具体的工程は以下のとおりである。

20

まずエントリークローン1 μ l、pFBIFを0.5 μ l(75ng)、LR反応緩衝液2 μ l、TE4.5 μ l、LR Clonase mix 2 μ lを25 $^{\circ}$ Cで1時間反応させ、Proteinase Kを1 μ l加えて37 $^{\circ}$ C 10分インキュベートして反応を終了させた(この組換え反応でpFBIF-GalNAc-Tが生成される)。pFBIFは、pFastBac1にIg(MHFQVQIFSFLLISASVIMSRG)(配列番号40)とFLAGタグ(DYKDDDDK)(配列番号41)のシーケンスを加えたもので、OT3(5'-GATCATGCATTTTTC AAGTGCAGATTTTCAGCTTTCCTGCTAATCAGTGCCTCAGTCATAATGTCACTGGAGATTACAAGGACGACGATGACAAG-3')(配列番号42)を鋳型に、プライマーOT20(5'-CGGGA TCCAT GCATT TTCAA GTGCA G-3')(配列番号43)と、OT21(5'-GGAAT TCTTGT CATCG TCGTC CTTG-3')(配列番号44)を用いてPCRを行い、得られたDNA断片をBam H1とEco R1で挿入して得られたものである。Igは発現タンパク質を分泌型にするため、FLAGタグは精製のため挿入した。

30

(2) Bac-to-BacシステムによるBacmidの作成

続いてBac-to-Bacシステム(インビトロジェン社)を用いて上記pFBIF-GalNAc-TとpFastBacとの間で組換えをさせ、昆虫細胞中で増殖可能なBacmidにGalNAc-Tその他の配列を挿入した。このシステムはTn7の組換え部位を利用して、Bacmidを含む大腸菌(DH10BAC)に目的遺伝子を挿入させたpFastBacを導入するだけで、ヘルパープラスミドから産生される組換えタンパク質によって目的とする遺伝子がBacmidへとりこまれるというものである。またBacmidにはlacZ遺伝子が含まれており、古典的な青(挿入なし)-白コロニー(挿入あり)による選択が可能である。

40

即ち、上記精製ベクター(pFBIF-GalNAc-T)をコンピテントセル(大腸菌DH10BAC)50 μ lと混合し、ヒートショック法の後、カナマイシン、ゲンタマイシン、テトラサイクリン、Bluo-gal、及びIPTGを含むLBプレートにまき、

50

翌日白い単独コロニーをさらに培養し、Bacmidを回収した。

実施例 9

Bacmidの昆虫細胞への導入

上記白コロニーから得られたBacmidに目的配列が挿入していることを確認した後、このBacmidを昆虫細胞Sf21に導入した。即ち35mmのシャーレにSf21 cell 9×10^5 cell / 2ml (抗生物質を含むSf-900II)を加え、27で1時間培養して細胞を接着した。(Solution A)精製したbacmid DNA 5 μ lに抗生物質を含まないSf-900II 100 μ lを加えた。(Solution B) Cell FECTIN Reagent 6 μ lに抗生物質を含まないSf-900II 100 μ lを加えた。その後、Solution AおよびSolution Bを丁寧に混合して15~45分間、室温でインキュベートした。細胞が接着したことを確認して、培養液を吸引して抗生物質を含まないSf-900II 2mlを加えた。Solution AとSolution Bを混合して作製した溶液(lipid-DNA complexes)に抗生物質を含まないSf900II 800 μ lを加えて丁寧に混和した。細胞から培養液を吸引し、希釈したlipid-DNA complexes溶液を細胞に加え、27で5時間インキュベートした。その後、Transfection mixturesを除き、抗生物質を含むSf-900II培養液2mlを加えて27で72時間インキュベーションした。Transfectionから72時間後にピペティングにより細胞を剥がし、細胞と培養液を回収した。これを3000rpm, 10分間遠心し、上清を別のチューブに保存した(この上清が一次ウイルス液となる)。

T75培養フラスコにSf21細胞 1×10^7 cell / 20ml Sf-900II (抗生物質入り)を入れて、27で1時間incubationした。細胞が接着したら一次ウイルスを800 μ l添加して、27で48時間培養した。48時間後にピペティングにより細胞を剥がし、細胞と培養液を回収した。これを3000rpm, 10分間遠心し、上清を別のチューブに保存する(この上清を二次ウイルス液とした)。

さらに、T75培養フラスコにSf21細胞 1×10^7 cell / 20ml Sf-900II (抗生物質入り)を入れて、27で1時間インキュベートした。細胞が接着したら二次ウイルス液1000 μ lを添加して、27で72~96時間培養した。培養後にピペティングにより細胞を剥がし、細胞と培養液を回収した。これを3000rpm, 10分間遠心し、上清を別のチューブに保存した(この上清を三次ウイルス液とした)。

加えて、100ml用スピナーフラスコにSf21細胞 6×10^5 cells / ml濃度で100mlを入れ、三次ウイルス液を1ml添加して27で約96時間培養した。培養後に、細胞及び培養液を回収した。これを3000rpm, 10分間遠心し、上清を別のチューブに保存した(この上清を四次ウイルス液とした)。

一次から三次までのセルペレットをソニケーションし(ソニケーション緩衝液: 20mM HEPES pH7.5, 2% Triton X-100, 1x Protease inhibitor cocktail)細胞粗抽出液をH₂Oで20倍にし、常法によりSDS-PAGEによる電気泳動についてウエスタンブロッティングを行ない、それぞれ目的とする各GalNAc-Tタンパク質の発現を確認した。FLAG配列のついたGalNAc-T (FLAG-GalNAc-T)に対する検出用抗体としては、抗FLAG M2-ペルオキシダーゼ(A-8592, SIGMA社)を用いた。各FLAG-GalNAc-Tはアミノ酸配列から計算した分子量とほぼ同じ位置にバンドが検出された。

実施例 10

各GalNAc-Tのレジン精製

上記三次感染のボトルからさらに四次感染をさせ、両ボトルからペレットと上清を回収し、遠心分離後(5000rpm 10分を2回)ペレットをソニケーションした(ソニケーション緩衝液: 20mM HEPES pH7.5, 2% Triton X-100、

1x Protease inhibitor cocktail)、ペレット粗抽出液と上清についてタンパク質定量(BIO-RAD社DC Protein Assay Kit)し、タンパク量を調整の上SDS-PAGE、ウエスタンブロッティングでGalNAc-T10の発現を確認した。この結果から精製には、もっとも相対的発現量の多い、FLAG-GalNAc-Tの上清を使用することとした。

四次感染の各FLAG-GalNAc-T上清20mlにNaN₃(0.05%)、NaCl(150mM)、CaCl₂(2mM)、抗M1レジン(Sigma社)(100μl)を混合し、4℃で一夜攪拌した。翌日遠心して(3000rpm 5分4秒)ペレットを回収し、2mMのCaCl₂・TBSを900μl加えて再度遠心分離(2000rpm 5分4秒)し、ペレットを200μlの1mM CaCl₂・TBSに浮遊させ活性測定サンプル(各GalNAc-T酵素液)とした。

10

実施例11

GalNAc-Tの酵素活性の検出方法と受容体基質の調製

GalNAc-TによるUDP-N-アセチル-D-ガラクトサミン:ポリペプチドN-アセチルガラクトサミン転移酵素としての活性測定には、受容体基質として下記表2に記載したFITC(またはFAM)標識した合成ペプチド、および供与体基質(donor substrate)としてUDP-GalNAcを使用した。

反応液は、受容体基質(50pmol)、Tris-HCl緩衝液(pH7.4)(25mM)、MnCl₂(5mM)、UDP-GalNAc(200nM)から成り、これにGalNAc-T酵素液を2.5μl加えて、さらにH₂Oを加えて全量を20μlとした(カッコ内は最終濃度)。

20

上記反応混合液を37℃で6時間反応させ、反応終了後、H₂Oを40μl加え、軽く遠心後、上清を取得した。得られた上清をMillipore社の簡易フィルター(Ultrafree-MC)を通して、20μl-40μlをHPLC分析に供した。HPLCカラムはCOSMOSIL 5C18-AR(ナカライテスク社Code No.378-66)を用いた。展開バッファーAは0.05%TFAを含むイオン交換水、展開バッファーBとしては受容体基質によって0.05%TFAの2-プロパノール:アセトニトリル=7:3と0.05%TFAのアセトニトリルを用いた。分離にはバッファーBの濃度勾配を用い、30分で50%まで上がるように勾配をかけた。検出条件は励起波長:492nm、蛍光波長:520nmとした。

30

表2. 各 GalNAc-T 酵素活性測定用受容体基質のアミノ酸配列と由来
(FITC または FAM は N 末端或いは C 末端に標識した)

合成ペプチド	アミノ酸配列	由来
Muc1a'	FITC-AHGVTSAPDTR (配列番号 45)	ヒトムチン型糖タンパク質 Muc1
MUC2-F	PTTTPITTTTTVTPTPTPTGTQTK-FITC (配列番号 46)	ヒト小腸由来ムチン MUC2
MUC5AC	FITC-GTTPSPVPTTSTTSA (配列番号 47)	ヒト respiratory gland ムチン MUC5AC
MUC7	PTPSATTPAPPSSAPPETTAAK-FITC (配列番号 48)	ヒト唾液腺ムチン MUC7
EA2	PTTDSTTPAPTTK-FITC (配列番号 49)	ラット 顎下腺ムチン
HPK-F	VPSTPPTPSPSTPPTPSPSK-FITC (配列番号 50)	ヒト IgA1 ヒンジ領域
SDC284	GPTEVAQTPTPETFLTTIK-FITC (配列番号 51)	ヒト シンデカン-3
Muc13-58	ANTPSFPTATSPAPPIK-FAM (配列番号 74)	ヒトムチン型糖タンパク質 Muc13

ここで、例えば、FITC (または FAM) 標識した合成ペプチドである表2中のEA2を受容体基質として用いた場合の既知のUDP-N-アセチル-D-ガラクトサミン：ポリペプチドN-アセチルガラクトサミン転移酵素 (GalNAc-T10) の活性について検討した。上記の実験手法に従って得られた反応液をHPLCで分析し、転移酵素と反応させていない受容体基質EA2ピーク (保持時間20.27分) と、これとは0.72分の差を示した転移酵素反応後の単一の産物ピーク (保持時間19.55分) を分取した。次いで、凍結乾燥後、MALDI-TOF質量分析装置 (BRUKER社、機種：REFLEXIII) で分析した。質量分析の結果、基質EA2 (理論質量数：m/z 1675.37；実測質量数：m/z 1676.98) に対して (図1参照)、HPLC分析で保持時間19.55分ピークの産物の実測質量数は1879.80 m/z であり、基質との差は202.82 m/z であった。この質量数の差は、一つのN-アセチル-D-ガラクトサミン (GalNAc、理論質量数：203) に相当し、従って、基質EA2に1ヶ所GalNAcが付加されたことを示している (図2参照)。このようにGalNAc-T10によって受容体基質であるEA2にGalNAcが付加した産物をEA2-T10と命名した。

30

上記の具体的な実験手法に従えば、各種酵素と各種受容体基質との反応は、反応前後のHPLC分取、および主要なピークのMALDI-TOFへの供することによって確認することができる。さらに、反応後のGalNAc-Tが付加した受容体基質は、後述するような二次的GalNAc-T活性を測定するために使用することができる。

40

実施例12

各GalNAc-TのUDP-N-アセチル-D-ガラクトサミン：ポリペプチドN-アセチルガラクトサミン転移酵素活性の確認

本明細書ではGalNAc-T活性を以下の定義に従い2つに分類した。

GalNAcが結合していないペプチドを受容体基質として反応し、いくつかのGalNAcがペプチドと結合させる触媒活性を「一次GalNAc-T活性」と定義する。一方、すでに一ヶ所あるいは複数箇所にGalNAcが結合しているペプチドを受容体基質として、さらに新しいGalNAcを結合させる触媒活性を「二次GalNAc-T活性」と

50

定義する。

上記の実施例のHPLC分析により、受容体基質ピークとは異なり、各GalNAc-Tとの反応後は新たな反応産物由来のピークが検出された。その具体的なピークの保持時間、受容体基質ピークの保持時間との差、さらに反応産物の生産量を表すピーク面積の比(パーセンテージ)を表3にまとめた。この結果により、受容体基質MUC1a'を用いた実験において、GalNAc-T11、GalNAc-T12、GalNAc-T14、GalNAc-T16、GalNAc-T17が一次GalNAcT活性を有することが明らかになった。受容体基質MUC2-Fを用いることにより、GalNAc-T11、GalNAc-T12、GalNAc-T14が一次GalNAcT活性を有することが明らかになった。受容体基質MUC7、MUC5AC、EA2を用いることにより、すべてのGalNAc-Tが一次GalNAcT活性を有することが明らかになった。受容体基質HPK-Fにおいて、GalNAc-T11、GalNAc-T12が一次GalNAcT活性を有することが明らかになった。受容体基質SDC284において、GalNAc-T11、GalNAc-T12、GalNAc-T13、GalNAc-T14、GalNAc-T17が一次GalNAcT活性を有することが明らかになった。さらに、例えば、GalNAc-T6と反応させて得られた、GalNAcが一カ所付加されたEA2-T6受容体基質を用いると、GalNAc-T11、GalNAc-T12、GalNAc-T13、GalNAc-T16、GalNAc-T17との反応後は受容体基質ピークとは異なる保持時間に新たな反応産物由来のピークが検出された。その具体的なピークの保持時間、受容体基質ピークの保持時間との差、さらに反応産物の生産量を表すピーク面積の比(パーセンテージ)を表3にまとめた。この結果により、GalNAc-T11、GalNAc-T12、GalNAc-T13、GalNAc-T16、GalNAc-T17には二次GalNAcT活性があることが認められた。

表3. HPLCによる各GalNAc-T酵素の活性測定結果分析

	複数産物	受容体基質 (min)	産物(min)	保持時間の差 (min)	産物面積(%)	活性
Muc1 a'						
T11		19.48	18.78	0.70	20.1	+
T12		19.46	18.77	0.69	10.8	+
T13		19.47	18.77	0.70	9.5	+
T14		20.21	19.47	0.74	35.2	+
T15		19.74	ND			-
T16		19.45	18.76	0.70	13.2	+
T17		19.48	18.78	0.70	82.1	+
MUC2-F						
T11		21.85	21.34	0.52	12.8	+
T12	P1	21.84	21.32	0.53	47.0	+
	P2	21.84	20.93	0.92	4.9	
T13		21.80	ND			-
T14		21.73	21.23	0.50	5.7	+
T15		21.84	ND			-
T16		21.85	ND			-
T17		21.83	20.80			-

Muc7						
T11		20.27	19.88	0.39	8.9	+
T12	P1	20.26	19.89	0.37	48.1	+
	P2	20.26	19.61	0.65	44.3	
T13		20.27	19.67	0.60	21.9	+
T14		20.33	19.95	0.38	8.6	+
T15		20.27	19.88	0.40	5.7	+
T16		20.27	19.89	0.39	10.6	+
T17		20.45	20.04	0.40	56.0	+
Muc5AC						
T11	P1	22.40	21.48	0.92	25.7	+
	P2	22.40	20.81	1.59	25.0	
	P3	22.40	20.45	1.95	7.4	
T12	P1	22.40	22.01	0.38	39.9	+
	P2	22.40	21.48	0.92	13.4	
	P3	22.40	21.16	1.24	8.2	
T13	P1	22.41	20.82	1.58	39.9	+
	P2	22.41	20.55	1.86	6.7	
T14		22.40	21.47	0.93	38.2	+
T15		22.40	21.60	0.79	3.4	+
T16	P1	22.41	21.49	0.92	34.2	+
	P2	22.41	20.83	1.58	5.1	
T17	P1	22.23	21.31	0.92	16.1	+
	P2	22.23	20.65	1.58	40.3	
	P3	22.23	20.28	1.95	8.2	
EA2						
T11		20.28	19.56	0.71	13.5	+
T12		20.27	19.57	0.70	22.9	+
T13		20.28	19.14	1.14	6.8	+
T14		20.28	19.55	0.73	25.1	+
T15		20.27	19.57	0.70	23.2	+
T16		20.31	19.60	0.71	12.6	+
T17		21.20	20.53	0.67	16.0	+
HPK-F						
T11	P1	18.15	17.80	0.35	16.3	+
	P2	18.15	17.46	0.69	11.8	
T12	P1	18.15	17.84	0.31	8.3	+
	P2	18.15	17.46	0.68	8.2	
T13		18.15	ND			-
T14		18.21	ND			-
T15		18.16	ND			-
T16		18.16	ND			-
T17		20.88	ND			-
SDC284						
T11		22.25	20.52	1.73	23.0	+
T12		22.24	20.49	1.76	16.9	+

T13		22.26	20.48	1.78	20.8	+
T14		22.09	20.31	1.78	8.7	+
T15		22.23	ND			-
T16		22.23	ND			-
T17	P1	22.15	20.41	1.74	15.1	+
	P2	22.15	15.53	6.62	23.9	
	P3	22.15	14.97	7.18	13.6	
EA2-T6						
T11	P1	19.60	19.19	0.41	12.2	+
	P2	19.60	18.70	0.90	3.5	
	P3	19.60	17.29	2.31	9.1	
T12		19.58	19.13	0.45	4.2	+
T13		19.57	19.18	0.38	100.0	+
T14		19.53	ND	-		-
T15		19.58	ND	-		-
T16		19.57	19.17	0.41	5.5	+
T17	P1	19.62	19.22	0.40	4.9	+
	P2	19.62	18.69	0.92	60.3	

実施例 13

20

GalNAc-TのFAM(5-carboxyfluorescein succinimidyl ester: 5-カルボキシフルオレセインスクシンイミジルエステル)ラベルした受容体基質の調製

Muc1、Muc2、Muc5AC、Muc7、およびEA2由来のオリゴペプチドにFAMラベルしたものを株式会社サワディ(東京、日本)から入手し、受容体基質として用いた。一つのGalNAc基が付加したMuc1a、Muc2、Muc5AC、Muc7、およびEA2は、GalNAc-T2[末尾の参考文献11を参照]、-T6[末尾の参考文献12を参照]、または-T10によるGalNAcの付加後、HPLCにより精製して、GalNAc-TアッセイのGalNAc-糖ペプチド受容体基質とした。調製した基質を表4に示す。

30

表4. GalNAc-T酵素活性測定用受容体基質

グリコシル化されていないペプチド	一つのGalNAcでグリコシル化されたペプチド	ペプチド配列
Muc1a	GalNAc-Muc1a	FAM-AHGVT*SAPDTR (配列番号52)
Muc2	GalNAc-Muc2	PTTTPITTTTIVTPTPTGTQTK-FAM (配列番号53)
Muc5AC	GalNAc-Muc5AC	FAM-GTTPSPVPTTSTSA (配列番号54)
Muc7	GalNAc-Muc7	PTPSATT*PAPPSSSAPPETTAAK-FAM (配列番号55)
EA2	GalNAc-EA2	PTTDSTT*PAPTTK-FAM (配列番号56)

*: 一つのGalNAcが結合するサイトを示す。印を付けていないのは同定されていないことを意味する。すべての糖ペプチドはマスマスペクトロメトリーにより同定した。

実施例 1 4G a l N A c - T の F A M ラベルした受容体基質への活性の検出法

G a l N A c - T アッセイは、2 5 m M T r i s - H C l (p H 7 . 4)、5 m M M n C l 2、0 . 2 % T r i t o n X - 1 0 0、0 . 2 5 m M U D P - G a l N A c (S i g m a)、適当な 5 - カルボキシフルオレセインスクシンイミジルエステル (F A M) - ラベルした受容体基質、および精製した酵素を含む 2 0 μ l の混合物中で 3 7 ° で行った。図 3 および図 4 では 5 0 p m o l の受容体基質および 1 2 5 n g の酵素を経時反応に用い、図 5 および図 6 ではその 1 / 1 0 の基質および酵素を一次 G a l N A c - T 活性および二次 G a l N A c - T 活性を検出する反応に用いた。反応は 3 分間煮沸することで停止させ、そして以前に示したように [末尾の参考文献 1 3 を参照] 逆相 C 1 8 カラム (C o s m o s i l 5 C 1 8 - A R、4 . 6 x 2 5 0 m m ; ナカライテスク社) を用いて高速液体クロマトグラフィー (H P L C) で受容体基質および反応生成物のピークを検出・分析した。

10

実施例 1 5マトリックス支援レーザー脱離イオン化 (M A L D I - T O F) マススペクトロメトリーによる糖ペプチド生成物のアミノ酸配列解析法

G a l - T 1 4 の反応の生成物のグリコシル化部位を決定するために、H P L C により単離した糖ペプチドは、R E F L E X (登録商標) I V (ブルカーダルトニクス) を用いて M A L D I - T O F マススペクトロメトリーで分子量解析すると共に、プロテインシーケンサー P P S Q - 2 3 A (島津製作所) 上でエドマン分解を行うことによりアミノ酸配列を解析した。マススペクトロメトリー (M A L D I - T O F) およびアミノ酸配列決定の方法は以前に示した [末尾の参考文献 1 3 を参照] 。

20

実施例 1 6G a l N A c - T 1 4 の U D P - N - アセチル - D - ガラクトサミン : ポリペプチド N - アセチルガラクトサミン転移酵素活性の確認

実施例 1 2 と同様に G a l N A c - T を二つに分類して定義した。G a l N A c - T 1 4 の酵素活性の測定は実施例 1 4 に示したように行い、基質には実施例 1 3 に示したものをを用いた。

M u c 1 a ペプチド誘導体 (F A M - A H G V T S A P D T R) (配列番号 5 2) を G a l N A c - T 1 4 と共にインキュベートし、H P L C により解析した。図 3 に示したように、G a l N A c - T の反応のインキュベーション時間と共に基質のピーク (S で示した) が徐々に消費されて他のピーク (P で示した) が現れ、G a l N A c - T 1 4 が G a l N A c を M u c 1 a ペプチドに転移したことを示した。G a l N A c - T 1 4 とのインキュベーションにより M u c 1 a ペプチドにおいてはシフトしたピークが一つだけ現れた。これは一つの S e r / T h r 部位のみが G a l N A c 修飾されたことを示している。生成物を H P L C により単離し、実施例 1 5 に示したように M A L D I - T O F マススペクトロメトリーおよびプロテインシーケンシングにかけ、生成物は 5 番目のトレオニンに G a l N A c を一つ有することを同定した。

30

一方、図 4 に示したように、M u c 5 A C ペプチド誘導体 (F A M - G T T P S P V P T T S T T S A) (配列番号 5 4) に対する反応においては、生成物は複数ピークとして現れた。G a l N A c - T 1 4 との短いインキュベーションの後には、H P L C 上で M u c 5 A C ペプチドのシフトした高いピークが一つ現れたが、このピークは徐々に消費され、インキュベーション時間が増えるにつれて複数のピークが現れた。この結果は、G a l N A c - T 1 4 は M u c 5 A C ペプチドの複数の部位に G a l N A c を転移したことを示すものである。

40

さらに、p p - G a l N A c - T 1 4 は E A 2 ペプチドには酵素活性を示したが、M u c 2 および M u c 7 についてはごくわずかな活性しか示さなかった (図 5) 。

G a l N A c - T 1 4 が M u c 5 A C ペプチドの複数の部位に G a l N A c を転移したことは、G a l N A c - T 1 4 は二次 G a l N A c - T 活性を有することを示している。このことを確認するために、実施例 1 3 に示した G a l N A c が一つ転移したペプチドを基

50

質として用いてGalNAc-Tアッセイを行った。図6に示したように、GalNAc-T14とのインキュベーションにより、GalNAc-Muc5ACペプチドを基質としたときのみシフトしたピークが生じ、その他のGalNAcが転移したEA2、Muc1a、Muc2、およびMuc7を基質としたときにはシフトしたピークが現れないか、または現れてもごくわずかであった。

実施例17

リアルタイムPCRによるヒト組織および細胞系列におけるGalNAc-T14転写物の定量解析による分布の解析

GalNAc-T14転写物の定量解析には、TaqMan(登録商標)Universal PCR Master MixおよびABI PRISM 7700 Sequence Detection System(アプライドバイオシステム社)を用い、定量リアルタイムPCRを行った。種々のヒト組織のMarathon Ready(登録商標)cDNAはクロンテックより購入した。ヒト組織cDNAとしては、脳全体、大脳皮質、小脳、胎児脳、骨髄、甲状腺、胸腺、脾臓、白血球、心臓、骨格筋、肺、肝臓、食道、胃、小腸、大腸、膵臓、腎臓、副腎、乳腺、子宮、胎盤、卵巣、精巣、および前立腺に由来するものを用いた。種々のガン細胞系列からのcDNAは以前に示した方法で調製した[末尾の参考文献14を参照]。ガン細胞系列のcDNAとしては、GOTO(神経芽腫)、SCCH-26(神経芽腫)、T98G(膠芽腫)、U25I(膠芽腫)、白血病プロミエロサイト、黒色腫、HL-60(前黒色腫)、K562(白血病)、Daudi(B細胞(Burkitt's))、Namalwa(B細胞白血病)、PC-1(肺扁平上皮細胞)、EBC-1(肺扁平上皮細胞)、PC7(肺腺癌)、HepG2(肝臓)、MKN45(胃)、KATO III(胃)、HSC43(胃)、Colo205(大腸)、HCT15(大腸)、LSC(大腸)、LSB(大腸)、SW480(大腸)、およびSW1116(大腸)に由来するものを用いた。

GalNAc-T14のプライマーセットおよびプローブは以下のものである；フォワードプライマー、5'-CGTGCCCGCTTGGAAC-3'(配列番号57)、リバースプライマー、5'-TGTTCTGGAGCATCCCGAAG-3'(配列番号58)、およびプローブ、マイナーグループバインダーを持つ5'-TGGGGATGTTGACAGAGAGGAAGCAGCT-3'(配列番号59)[末尾の参考文献15を参照]。PCR条件は、50℃、2分間を1サイクル、95℃、10分間を1サイクル、95℃、15秒間、および60℃、1分間を50サイクルで行った。GalNAc-T14の転写物の相対量は、同じcDNA中のGADPH(glyceraldehyde-3-phosphate dehydrogenase:グリセルアルデヒド-3-ホスフェートデヒドロゲナーゼ)転写物の量により標準化した。

リアルタイムPCR法により、種々のヒト組織および細胞系列におけるGalNAc-T14転写物の発現レベルを決定した。GalNAc-T14は、そのレベルは異なるものの、殆どの組織中(図7)、および種々の細胞系列中(図8)で発現した。かなり高いレベルの発現がみられた組織は、小腸、胃、膵臓、および大腸であった。中程度の発現が、精巣、甲状腺、および脾臓で観察された。試験したその他の組織の発現レベルは、非常に低い、または検出不可能であった。GalNAc-T14は試験した細胞系列においては比較的低いレベル、または検出不可能である程度に発現した。大腸癌および胃癌細胞系列におけるGalNAc-T14の発現レベル(図7)は、それぞれ大腸および胃のそれら(図8)と比較して低かった。

実施例18

GalNAc-T13のUDP-N-アセチル-D-ガラクトサミン:ポリペプチドN-アセチルガラクトサミン転移酵素活性の確認

実施例12と同様にGalNAc-Tを二つに分類して定義した。GalNAc-T13の酵素活性の測定は実施例14に示したように行い、基質には実施例13に示したものを用いた。

Muc1a、Muc2、Muc5AC、Muc7、およびEA2由来のオリゴペプチドに

10

20

30

40

50

FAMラベルしたペプチド誘導体（実施例13参照）をGalNAc-T13と共にインキュベートし、HPLCにより解析した。図9に示したように、GalNAc-T13の反応前（上段）と反応後（下段）で、Muc1、Muc5AC、Muc7、およびEA2の基質ピーク（Sで示した）が消費されて他のピーク（Pで示した）が現れ、GalNAc-T13がGalNAcを各ペプチドに有意に転移したことを示した。しかしながら、Muc2を基質とした場合には、GalNAc-T13の転移活性はごくわずかであった。

さらに、一次GalNAc-T13活性を評価するために、Muc1をGalNAc-T13と共に反応時間を変えてインキュベートし、反応混合物をHPLCにより解析した。図10に示すように、GalNAc-T13はMuc1aペプチドを効率的に利用し、インキュベーション時間が延長されるにつれてシフトした一本の生成物ピーク" P "が現れた。EA2ペプチドは、7個の可能な糖鎖修飾位置を含み、ラットの顎下腺ムチンリピート由来である。GalNAc-T13とEA2ペプチドとの反応生成物は独特のパターンの複数ピークを示した（図11）。生成物ピークP1及びP2は、基質ピークに近接し、反応時間を通して非常に低かった。これに対しピークP3、P4及びP5は逐次的に現れ、より多くの糖鎖修飾ペプチドであると推定でき、また、P1及びP2より高かった。

以上の結果は、GalNAc-T13がEA2ペプチドに複数個のGalNAcを転移したことを示唆している。そこで、GalNAc-T13によるEA2ペプチドに転移したGalNAcの数を決定する試みを行った。まず、生成物ピークをHPLCで精製し、画分を質量分析アッセイに供した。図12パネルBに示したように、ピークP3、P4およびP5（T13で生じた）、並びにP1'およびP2'（T6で生じた）に対応する生成物について測定した分子量は、それぞれFAMラベルしたEA2ジ-、トリ-、テトラ-およびモノ-糖ペプチドと予想される分子量に合致していた。GalNAc-T13によって生じたP1及びP2の生成量は非常に小さく、質量分析用として回収することはできなかった。P2及びP1'の保持時間が同じであったので、以上の結果は、GalNAc-T13によって生じたP1及びP2の両方とも一つのGalNAc残基を異なった部位に結合したことを強く示唆している。同時に、この結果は、GalNAc-T13がEA2ペプチドのThr又はSer部位に最初に一つのGalNAcを効率の悪い方法で転移し、その後、すぐに他の部位に別のGalNAcを転移することを示した。これは、GalNAc-T13がEA2に対して強い二次的なGalNAc糖ペプチド転移活性、及び相対的に弱い一次的なポリペプチドGalNAc-転移活性を有することを示唆した。

GalNAc-T13の二次活性を確かめるために、GalNAc-T13アッセイ用に受容体基質としてGalNAcを転移したEA2モノ-糖ペプチドを調製した。図13パネルBに示すように、基質ピーク（S"）は急速に消費され、さらにシフトしたピークP1"がGalNAc-T13との短いインキュベーション後に現れた。その後、P1"は徐々に消費され、さらにシフトしたピークがインキュベーションの時間経過とともに現れた。

また、GalNAc-T13の酵素反応8時間後に残存したEA2基質の量は、たった10分間の酵素反応におけるGalNAc-EA2基質の残存量と殆ど同等であった（図13）。この結果は、GalNAc-T13がEA2よりモノ-GalNAc-EA2に対してより選択的な活性を有することを示唆している。即ち、GalNAc-T13は、弱い一次活性と強い二次活性を併せ持つ糖転移酵素であることを示している。

これまでの研究では、*in vitro*のO-結合型糖鎖修飾の速度論において、近接および遠隔のO-グリカンの、ネガティブ及びポジティブな相対的效果を示している[末尾の参考文献16-18を参照]。EA2へのGalNAcの付加は、ペプチド構造を変化させる可能性があり、GalNAc-T13に対する特定の受容体部位に接近することに影響している。従来刊行物では、遠隔部位と同様に糖鎖修飾部位に近いペプチド骨格構造に対するO-型糖鎖修飾の効果を同定している[末尾の参考文献19-21を参照]。他に可能な作用機序としては、GalNAc-T13のレクチンドメインがGalNAcを付加したペプチドに結合し、EA2糖ペプチドの糖鎖修飾を触媒するようGalNAc

10

20

30

40

50

- T13の能力に貢献するということが挙げられる。これは、ヒトGalNAc-T4に見出される作用機序と同様である[末尾の参考文献22を参照]。

GalNAc-T13はまたGalNAc-Muc5ACに対しても強力にGalNAc-糖ペプチド転移酵素活性を示した。

GalNAc-T13は、ラットGalNAc-T9[末尾の参考文献23を参照]のオルソログ異性体である。しかしながら、ヒトGalNAc-T13はラットGalNAc-T9とは異なった基質特異性を示した。ラットGalNAc-T9は、Muc1a、EA2、Muc2及びMuc5ACを含む未修飾ペプチド基質に対しては、GalNAcを転移しないことが報告されている[末尾の参考文献16を参照]。ところが、ヒトGalNAc-T13は未修飾ポリペプチド基質に対して明らかな触媒活性を示した。GalNAc-T13及びラット-T9の両方とも、EA2及びMuc5AC糖ペプチドに対して強力な二次活性を示した。以上の結果は、GalNAc-T13およびラット-T9が*in vitro*において異なった受容体基質特異性を有することを示している。

実施例19

リアルタイムPCRによるヒト組織および細胞系列におけるGalNAc-T13転写物の定量解析による分布の解析

GalNAc-T13転写物の定量解析は、実施例17に記載の方法に準じて行った。ヒト組織cDNAは、脳全体、小脳、大脳皮質、胎児脳、副腎、甲状腺、胸腺、肺、心臓、乳腺、肝臓、食道、胃、小腸、大腸、膵臓、脾臓、腎臓、子宮、胎盤、精巣、卵巣、骨格筋、および骨髄に由来するものを用いた。ガン細胞系列のcDNAとしては、SCCH-26(神経芽腫)、GOTO(神経芽腫)、T98G(膠芽腫)、U25I(膠芽腫)、YKG-1(膠芽腫)、M118-MG(膠芽腫)、A172(膠芽腫)、PC-1(肺扁平上皮細胞)、EBC-1(肺扁平上皮細胞)、PC7(肺腺癌)、HSC43(胃)、MKN45(胃)、KATO III(胃)、A431(扁平上皮)、Colo205(大腸)、LSC(大腸)、LSB(大腸)、SW480(大腸)、SW116(大腸)、HepG2(肝臓)、Capan-2(膵臓)、KHM-1B(IgA骨髄腫)、Namalwa(B細胞白血病)、Daudi(B細胞Burkitt's)、K562(白血病)、U266(骨髄腫)、HL-60(前骨髄腫白血病)、および黒色腫に由来するものを用いた。また、GalNAc-T13のプライマーセットおよびプローブは以下のものである；フォワードプライマー、5'-CGCGTAGGAAATGGAGAAACA-3'(配列番号60)、リバースプライマー、5'-TCGGTATGCCTGATCCACTCT-3'(配列番号61)、およびプローブ、5'-AGACCTTACCCCATGACCGATGCTGA-3'(配列番号62)。

GalNAc-T13は試験した全ての組織中に発現していたが、発現レベルは組織に応じて異なっていた(図14)。転写物の発現レベルが最も高い組織は小腸であり、中程度レベルのものは、胃、膵臓、卵巣、胸腺および脾臓であった。上記以外の組織では、微量ないしより低いレベルであった。ヒトGalNAc-T異性体、即ち、GalNAc-T1、-T2、-T4、-T7および-T8はいたるところに発現している[末尾の参考文献5-8を参照]。複数のGalNAc-Tが一つの組織に発現している理由はまだ明らかにされていない。酵素は基質特異性を有するので[末尾の参考文献24-26を参照]、細胞内では受容体基質全体の糖鎖修飾に対して、GalNAc-T間で競争および協調があるのかもしれない。即ち、組織内で複数のGalNAc-Tの発現は、タンパク質の一連のプロセッシングに必要であり、または適切なタンパク質の糖鎖修飾を確実にするために過剰に提供しているのかもしれない。

また、GalNAc-T13の発現レベルは、試験したほとんどの腫瘍細胞系列において低く(図15)、細胞系列間の特異的な発現パターンは見出すことができなかった。

実施例20

GalNAc-T12転写物のAlternatively Spliced Isoformの同定

各発現組織または細胞系列におけるGalNAc-T12転写物の多数のAlterna

tively Spliced Isoformを決定するために、2つのプライマー (si - 3 : 5' - GTCGGCTGGTTCTGCCAGTCTTCG - 3' (配列番号 63) および si - 4 : 5' - TGGAAAGGTCCGTGCAATACACCAAG - 3' (配列番号 64)) を用いて RT - PCR を行った。PCR プログラムは、95 5分での最初の变性、続いて 95 30秒、58 30秒、72 30秒の 50 サイクル、および 72 5分での最終のインキュベーションからなる。PCR 産物を 4つの制限エンドヌクレアーゼ : Hae II、Hinf I、Hae III および Nci I によって消化し、次いで、10% PAGE で電気泳動によって解析し、エチジウムブロマイド (0.5 μg/ml) によって染色した。

GalNAc - T12 の cDNA 配列と Genome Project Database に寄託されたゲノム配列との比較によって、2つの酵素、GalNAc - T12 および - T2 の染色体位置、並びにゲノム構造を決定した (図 16 パネル A)。GalNAc - T12 (受入れ番号 NT005194) のゲノム配列は染色体 2 に位置し、228 kb の長さであり、少なくとも 17 個のエキソンを含み、各エキシンの長さは GalNAc - T2 (受入れ番号 NT021973) とほぼ同等であった。本発明者らが取得した新規配列 (番号 AB078144) は、登録された配列 (No. AK022753) および 図 16 パネル A - B とは異なっていたが、これらは GalNAc - T12 の 2つの Alternately spliced isoform であることは明らかである。第一の isoform (No. AB078144) は 15 個のエキソンを含み、第二の isoform (No. AK022753) は 16 個のエキソンを含む (図 16 パネル B)。

実施例 2 1

GalNAc - T12 の UDP - N - アセチル - D - ガラクトサミン : ポリペプチド N - アセチルガラクトサミン転移酵素活性の確認

GalNAc - T12 の酵素活性の測定は実施例 1 4 に示したように行い、基質には Muc5AC、Muc2、および Muc13 - 58 由来のペプチド (表 2 を参照) を用いた。上記ペプチドを GalNAc - T12 及び GalNAc - T2 とともに 48 時間インキュベートし、反応生成物をそれぞれ異なる時間で HPLC によって解析した (図 17)。図 17 パネル A、カラム I に示したように、Muc5AC の 50% が反応によって消化された場合、GalNAc - T12 および - T2 の両方とも 3本のピーク (GalNAc - T12 に対しては P1、P2 および P3 (上段)、GalNAc - T2 に対しては P1'、P2' および P3' (下段)) を同じ保持時間 (20.33分、20.92分および 20.60分) が現れた。GalNAc - T12 における主要な生成物ピーク (P2)、および GalNAc - T2 における主要な生成物ピーク (P1') は異なる保持時間で現れた。その後、はじめの Muc5AC の 80% が反応によって消費された場合、保持時間が延長されるにつれて (図 17 パネル A、カラム II)、それぞれの主要な生成物ピークが保持された。ここで、新規な糖ペプチド種を HPLC によって単離し、MALDI - TOF 質量分析によって確認した。付加された GalNAc 残基の数は、P1 および P1' は一つ、P2 および P2' もまた一つであったが、P3 および P3' は二つの GalNAc が付加していることが分かる。以上の結果より、P2 および P1' は異なる保持時間で現れたが、一つの GalNAc が付加されているので、GalNAc - T12 および - T2 は Muc5AC の異なる位置に最初の GalNAc を主に付加したことを示している。また、図 17 パネル A、カラム III では、全ての元々の基質がちょうどすっかり消失した場合、2つの酵素の反応生成物パターンは非常に異なり、GalNAc - T12 は多種の GalNAc - T 付加産物を生成していることが分かる。

Muc2 に対する GalNAc - T12 および - T2 の活性を HPLC で比較すると、生成物ピーク P1 (T12 について) と P1' (T2 について) ; P2 (T12) と P2' (T2) は、それぞれ同じ保持時間 : 20.99分と 20.58分に現れていることが分かる (図 17 パネル B)。MALDI - TOF 質量分析によって、P1 と P1' はモノ - GalNAc 糖ペプチドであり、P2 と P2' はジ - GalNAc 糖ペプチドであった。GalNAc - T2 は、GalNAc - T12 より速く Muc2 に第二の GalNAc を

付加する傾向にあることが図17パネルBから見出すことができる。従って、GalNAc-T12は、GalNAc-T2とは異なる保持時間を示し、GalNAc-T2とは異ったGalNAc付加産物を生成するという特徴を有している。

これは、50%および20%の基質が消化された場合に、P2はまだ検出できないが、一方、P2'は現われ、そのピークが増大したことによる。全ての基質がちょうど消化された場合、P2は非常に小さいが、P2'は高レベルであり、そして保持時間がP2'より小さい、多くのピーク(三つのGalNAcが含まれることがMALDI-TOFによって確かめられているが、データは示さず)がGalNAc-T2について現れた。GalNAc-T12については、最初の基質が完全に消化された後に、反応生成物を試験した(データ示さず)。保持時間が延長されるにつれて、P1は減少し、その代りにP2が増大し、そして、意図時間がP2よりも短い、より多くのピーク(3個のGalNAc残基を含むことがMALDI-TOFによって確かめられているが、データは示さず)が現れた。また、GalNAc-T12は、Muc2により多くのGalNAcを付加させることができる。

10

Muc13は、上皮細胞及び血液細胞に発現するヒト細胞表面ムチンである。本発明者が設計したペプチド基質であるMuc13-58はTSPAPボックスを有し、これはムチンボックス(XTPXP)に類似しているが少し異なる。Muc13-58に対するGalNAc-T12および-T2の活性は、ほぼ同等であった(図17パネルC)。P1(T12について)およびP1'(T2について); P2(T12)およびP2'(T2)は、それぞれ同じ保持時間: 22.96分および23.29分で現れた。各生成物ピークの大きさは、消化される最初の基質量に応じて互いに均等であった。

20

実施例22

リアルタイムPCRによるヒト組織および細胞系列におけるGalNAc-T12転写物の定量解析による分布の解析

多種のヒト正常組織および細胞系列におけるGalNAc-T12の発現レベルを定量的リアルタイムPCRを用いて調べた。

GalNAc-T13転写物の定量解析は、実施例17に記載の方法に準じて行った。ヒト組織cDNAは、脳全体、大脳皮質、胸腺、肺、心臓、乳腺、肝臓、胃、小腸、大腸、膵臓、脾臓、腎臓、膀胱、子宮、胎盤、精巣、卵巣、骨格筋、白血球、B細胞、骨髄、胎児脳、胎児胸腺、胎児肺、胎児肝臓、胎児小腸、胎児脾臓、胎児腎臓、胎児骨格筋、および胎児に由来するものを用いた。ガン細胞系列のcDNAとしては、NAGAI、NB-9、SCCH-26、IMR32、SK-N-SH、GOTO、NT2-R10、SK-N-MC、A172、KG-1-C、YKG-1、T98G、U25I、PC-1、EBC-1、PC-7、SBC-1、MKN-45、KATO III、HSC43、Colo205、HCT15、LSC、LSB、SW480、SW1116、HepG2、Capan-2、HL60、K562、Namalwa、Jurkat、U937、Raji、およびRamosに由来するものを用いた。

30

また、GalNAc-T12のプライマーセットおよびプローブは以下のものである; フォワードプライマー(エキソン5にコードされる): 5'-TGAGATGCACACTGCTGGTGTATTG-3'(配列番号65)、リバープライマー(エキソン5-6にコードされる): 5'-GGGTGCGGTTTAATACACTGCG-3'(配列番号66)、およびプローブ(エキソン5における): 5'-AACGAGGCCCGCTCCACGCTGCT-3'(配列番号67)。PCRは、95 30秒、60 15秒、72 30秒を50サイクルで行った。

40

GalNAc-T12転写物は、試験した全てのヒト成人および胎児組織によって普遍的に発現している。特に、腎臓と胎児腎臓において発現が高かった(図18パネルA)。しかしながら、試験した細胞系列における発現パターンは、組織よりも制限され、神経芽腫、ある種の腺癌(PC-7)、胃癌(MKN45)、白血病(HL60、Namalwa、Jurkat)、およびバーキット病(Ramos、Raji、Dudi)に限定される(図18パネルB)。続いて、6種の組織(腎臓、胎児腎臓、脳、胎児脳、骨髄および

50

胎児肝臓)、および6種の細胞系列(HL-60、MKN45、GOTO、Namalwa、Daudi、SCCH-26)においてGalNAc-T12転写物の2つのAlternatively spliced isoformの存在を確認した。そこで、エキソン2、3または4を含んでいるはずの断片を増幅した。エキソン2、3および4は、それぞれHae III、Hinf I、またはHae IIIとNci I部位を含むため、isoformを確認するために上記4つの制限エンドヌクレアーゼを用いて増幅した断片を消化した。図16パネルCに見られるように、増幅されたバンドは、Hae IIIおよびNci Iによってのみ消化されたが、しかしながら、Hae IIIまたはHinf Iによって消化された断片は検出できなかった。上記の結果は、isoform 1が主な転写物であり、一方、isoform 2が副産物であることを示している。

10

実施例 2 3

GalNAc-T11のUDP-N-アセチル-D-ガラクトサミン：ポリペプチドN-アセチルガラクトサミン転移酵素活性の確認

GalNAc-T11の酵素活性の測定は実施例14に示したように行い、基質にはMuc1a、Muc2、Muc5AC、Muc7、Muc13(58)、SDC284、およびEA2由来のペプチドを用いた。

図19は、各受容体基質に対するGalNAc-T11の反応前(パネル上段)および反応後(パネル下段)のHPLCチャートを示している。酵素反応時間はそれぞれ6時間であるが、Muc5Acについては48時間反応させたものについても示した。

実施例 2 4

リアルタイムPCRによるヒト組織および細胞系列におけるGalNAc-T11転写物の定量解析による分布の解析

多種のヒト正常組織および細胞系列におけるGalNAc-T12の発現レベルを定量的リアルタイムPCRを用いて調べた。

GalNAc-T13転写物の定量解析は、実施例17に記載の方法に準じて行った。ヒト組織cDNAは、脳全体、大脳皮質、小脳、胎児脳、骨髄、甲状腺、胸腺、脾臓、白血球、心臓、細胞骨格、肺、肝臓、食道、胃、小腸、大腸、膵臓、腎臓、副腎、乳腺、子宮、胎盤、卵巣、精巣、および前立腺に由来するものを用いた(図20参照)。

また、GalNAc-T11のプライマーセットおよびプローブは以下のものである; フォワードプライマー: 5'-AGCGTCAATCAATCACCTTCCA-3'(配列番号68)、リバースプライマー: 5'-TTCGGTTCAGGACACTCTTCAAC-3'(配列番号69)、およびプローブ: 5'-CCGTTCCACCCTGCGCGC-3'(配列番号70)。

GalNAc-T11とGAPDHに対する標準曲線を各プラスミドDNAの連続的な希釈によって作成した。GalNAc-T11転写物の発現レベルは、同じcDNAで測定したGAPDH転写物の発現レベルによって標準化した。発現レベルは、全RNAの1μgにおける標的遺伝子のコピー数として表した(図20)。なお、データは3回の実験から得たものであり、平均値±標準偏差として示した。

実施例 2 5

GalNAc-T17のUDP-N-アセチル-D-ガラクトサミン：ポリペプチドN-アセチルガラクトサミン転移酵素活性の確認

GalNAc-T17の酵素活性の測定は実施例14に示したように行い、基質にはEA2、Muc1a、Muc2、Muc13(58)、Muc5AC、SDC284、およびHRP由来のペプチドを用いた。

図21は、各受容体基質に対するGalNAc-T17の反応前(パネル上段)および反応後(パネル下段)のHPLCチャートを示している。酵素反応時間はそれぞれ6時間である。

実施例 2 6

リアルタイムPCRによるヒト組織および細胞系列におけるGalNAc-T17転写物の定量解析による分布の解析

20

30

40

50

多種のヒト正常組織および細胞系列における G a l N A c - T 1 7 の発現レベルを定量的リアルタイムPCRを用いて調べた(図22)。

G a l N A c - T 1 3 転写物の定量解析は、実施例17に記載の方法に準じて行った。ヒト組織cDNAは、実施例17に記載した組織に由来するものを用いた(図22パネルA参照)。ガン細胞系列のcDNAとしては、実施例17に記載したほかに、U266(骨髄腫)、A431(扁平上皮)、SW116(大腸)、およびCapan-2(膵臓)に由来するものを用いた(図22パネルB参照)。

また、G a l N A c - T 1 1 のプライマーセットおよびプローブは以下のものである；フォワードプライマー：5' - G G A T G T G C T C G T C T T C A T G G A - 3' (配列番号71)、リバースプライマー：5' - G A T A C C A C T C G G C T C C T G T C A - 3' (配列番号72)、およびプローブ：5' - C T G G A G C C C C T C C T C A G C A G A A T A G C - 3' (配列番号73)。

参考文献

1. White, T.W., R. Bruzzone, and D.L. Paul, The connexin family of intercellular channel forming proteins. *Kidney Int*, 1995. 48(4): p. 1148-57.
2. Toba, S., et al., Brain-specific expression of a novel human UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase (GalNAc-T9). *Biochim Biophys Acta*, 2000. 1493(1-2): p. 264-8.
3. Bennett, E.P., H. Hassan, and H. Clausen, cDNA cloning and expression of a novel human UDP-N-acetyl-alpha-D-galactosamine. Polypeptide N-acetylgalactosaminyltransferase, GalNAc-t3. *J Biol Chem*, 1996. 271(29): p. 17006-12.
4. Bennett, E.P., et al., Cloning and characterization of a close homologue of human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase-T3, designated GalNAc-T6. Evidence for genetic but not functional redundancy. *J Biol Chem*, 1999. 274(36): p. 25362-70.
5. Bennett, E.P., et al., Cloning of a human UDP-N-acetyl-alpha-D-Galactosamine:polypeptide N-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. *J Biol Chem*, 1998. 273(46): p. 30472-81.
6. White, K.E., et al., Molecular cloning of a novel human UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase, GalNAc-T8, and analysis as a candidate autosomal dominant hypophosphatemic rickets (ADHR) gene. *Gene*, 2000. 246(1-2): p. 347-56.
7. Bennett, E.P., et al., A novel human UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase, GalNAc-T7, with specificity for partial GalNAc-glycosylated acceptor substrates. *FEBS Lett*, 1999. 460(2): p. 226-30.

8. White, T., et al., Purification and cDNA cloning of a human UDP-N-acetyl- α -D-galactosamine:polypeptide N-acetylgalactosaminyltransferase. *J Biol Chem*, 1995. 270(41): p. 24156-65.
9. Sakamaki, T., Y. Imai, and T. Irimura, Enhancement in accessibility to macrophages by modification of mucin-type carbohydrate chains on a tumor cell line: role of a C-type lectin of macrophages. *J Leukoc Biol*, 1995. 57(3): p. 407-14.
10. Moyana, T.N. and J. Xiang, Expression of tumor-associated polymorphic epithelial mucin and carcinoembryonic antigen in gastrointestinal carcinoma tumors. Implications for immunodiagnosis and immunotherapy. *Cancer*, 1995. 75(12): p. 2836-43.
11. White, T., E. P. Bennett, K. Takio, T. Sorensen, N. Bonding, and H. Clausen, Purification and cDNA Cloning of a Human UDP-N-acetyl- α -D-galactosamine: polypeptide N-Acetylgalactosaminyltransferase., *J. Biol. Chem.*, 1995. 270, p.24156-65
12. Bennett, E. P., H. Hassan, U. Mandel, M. A. Hollingsworth, N. Akisawa, Y. Ikematsu, G. Merckx, A. G. van Kessel, S. Olofsson, and H. Clausen, Cloning and Characterization of a Close Homologue of Human UDP-N-acetyl- α -D-galactosamine:Polypeptide N-Acetylgalactosaminyltransferase-T3, Designated GalNAc-T6., *J. Biol. Chem.*, 1999. 274(36), p.25362-70
13. Iwai, T., N. Inaba, A. Naundorf, Y. Zhang, M. Gotoh, H. Iwasaki, T. Kudo, A. Togayachi, Y. Ishizuka, H. Nakanishi, and H. Narimatsu, Molecular Cloning and Characterization of a Novel UDP-GlcNAc:GalNAc-peptide β 1,3-N-Acetylglucosaminyltransferase (β 3Gn-T6), an Enzyme Synthesizing the Core 3 Structures of O-Glycan. *J. Biol. Chem.*, 2002. 277(15): p. 12802-9
14. Togayachi, A., T. Akashima, R. Ookubo, T. Kudo, S. Nishihara, H. Iwasaki, A. Natsume, H. Mio, J. Inokuchi, T. Irimura, K. Sasaki, and H. Narimatsu, Molecular Cloning and Characterization of UDP-GlcNAc:Lactosyl

lceramide β 1,3-N-Acetylglucosaminyltransferase (β 3Gn-T5), an Essential Enzyme for the Expression of HNK-1 and Lewis X Epitopes on Glycolipids, *J. Biol. Chem.*, 2001. 276, p.22032-40

15. Kutuyavin, I.V., I.A. Afonina, A. Mills, V.V. Gorn, E.A. Lukhtanov, E.S. Belousov, M.J. Singer, D.K. Walburger, S.G. Lokhov, A.A. Gall, R. Dempsy, M.W. Reed, R.B. Meyer and J. Hedgpeth, 3'-Minor Groove Binder-DNA Probes Increases Sequence Specificity at PCR Extension Temperatures. *Nucleic Acids Res*, 2000. 28(2), p. 655-61

16. Ten Hagen, K.G., G.S. Bedi, D. Teraert, P.D. Kingsley, F.K. Hagen, M.M. Balys, T.M. Beres, P. Degand and L.A. Tabak, Cloning and Characterization of a Ninth Member of the UDP-GalNAc:polypeptide N-Acetylgalactosaminyltransferase Family, ppGaNTase-T9. *J. Biol. Chem.*, 2001. 276, p.17395-404

17. Hanisch, F.G., S. Muller, H. Hassan, H. Clausen, N. Zachara, A.A. Gooley, H. Paulsen, K. Alving and J. Peter-Katalinic, Dynamic Epigenetic Regulation of Initial O-Glycosylation by UDP-N-Acetylgalactosamine:Peptide N-Acetylgalactosaminyltransferases. Site-Specific Glycosylation of MUC1 Repeat Peptide Influences the Substrate Qualities at Adjacent or Distant Ser/Thr Positions. *J. Biol. Chem.*, 1999. 274, p.9946-54

18. Brockhausen, I., D. Toki, J. Brockhausen, S. Peters, T. Bielfeldt, A. Kleen, H. Paulsen, M. Meldal, F. Hagen and L.A. Tabak, Specificity of O-Glycosylation by Bovine Colostrum UDP-GalNAc:Polypeptide Alpha-N-Acetylgalactosaminyltransferase Using Synthetic Glycopeptide Substrates. *Glycoconj. J.*, 1996. 13, p.849-56

19. Kirnarsky, L., O. Prakash, S.M. Vogen, M. Nomoto, M.A. Hollingsworth and S. Sherman, Structural Effects of O-Glycosylation on a 15-Residue Peptide from the Mucin (MUC1) Core Protein. *Biochemistry*, 2000. 39, p. 12076-82

20. Huang, X., J.J. Barchi Jr., F.D. Lung, P.P. Roller, P.L. Nara, J. Musc

hik and R.R. Garrity, Glycosylation Affects Both the Three-Dimensional Structure and Antibody Binding Properties of the HIV-1III_B GP120 Peptide RP135. *Biochemistry* 1997. 36, p.10846-56

21. Huang, X., M.C. Smith, J.A. Berzofsky and J.J. Barchi, Jr., Structural Comparison of a 15 Residue Peptide from the V3 Loop of HIV-1III_B and an O-Glycosylated Analogue. *FEBS Lett.*, 1996. 393, p.280-6

22. Hassan, H., C.A. Reis, E.P. Bennett, E. Mirgorodskaya, P. Roepstorff, M.A. Hollingsworth, J. Burchell, J. Taylor-Papadimitriou and H. Clausen, The Lectin Domain of UDP-N-Acetyl-D-Galactosamine: Polypeptide N-Acetylgalactosaminyltransferase-T4 Directs its Glycopeptide Specificities. *J. Biol. Chem.*, 2000. 275, p.38197-205

23. Ten Hagen, K.G., G.S. Bedi, D. Tetaert, P.D. Kingsley, F.K. Hagen, M.M. Balys, T.M. Beres, P. Degand, L.A. Tabak, Cloning and characterization of a ninth member of the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family, ppGaNTase-T9., *J. Biol. Chem.*, 2001. 276, p.17395-404

24. Wandall, H.H., H. Hassan, E. Mirgorodskaya, A.K. Kristensen, P. Roepstorff, E.P. Bennett, P.A. Nielsen, M.A. Hollingsworth, J. Burchell, J. Taylor-Papadimitriou and H. Clausen, Substrate Specificities of Three Members of the Human UDP-N-Acetyl-Alpha-D-Galactosamine:Polypeptide N-Acetylgalactosaminyltransferase Family, GalNAc-T1, -T2, and -T3. *J. Biol. Chem.*, 1997. 272, p.23503-14

25. Clausen, H. and E.P. Bennett, A Family of UDP-GalNAc:Polypeptide N-Acetylgalactosaminyl-transferases Control the Initiation of Mucin-Type O-Linked Glycosylation. *Glycobiology*, 1996. 6, p.635-46

26. Sorensen, T., T. White, H.H. Wandall, A.K. Kristensen, P. Roepstorff and H. Clausen, UDP-N-Acetyl-Alpha-D-Galactosamine:Polypeptide N-Acetylgalactosaminyltransferase. Identification and Separation of Two Distinct Transferase Activities. *J. Biol. Chem.*, 1995. 270, p.24166-73

産業上の利用の可能性

本発明により、コアタンパク質やペプチド配列のセリン、スレオニン残基の水酸基にN-アセチルガラクトサミンを 1 結合で転移する活性を有する酵素が単離され、その遺伝子の構造が明らかにされた。したがって、該酵素等の遺伝子工学的な生産、該酵素に基づく GalNAc 修飾ペプチドあるいは GalNAc 修飾タンパク質などの基質としての生産

、該酵素に基づく未知 O - 結合型糖鎖の付加したタンパク質あるいは未知ペプチドの O - 結合型糖鎖修飾位置の同定や、該遺伝子らに基づく疾患の診断が可能になった。

【配列表】

SEQUENCE LISTING

<110> Amersham Biosciences K.K.

National Institute of Advanced Industrial Science and Technology

<120> Novel UDP-N-Acetyl-D-Galactosamin:Polypeptide

N-Acetylgalactosaminyl Transferases and Nucleotic Acids Encoding Them

<130> YCT-791

<150> JP2001-401507

<151> 2001-12-28

<150> JP2002-163832

<151> 2002-06-05

<150> JP2002-203696

<151> 2002-07-12

<160> 74

<210> 1

<211> 542

<212> PRT

<213> Homo sapiens

<400> 1

Met Arg Lys Ile Arg Ala Asn Ala Ile Ala Ile Leu Thr Val Ala Trp

1

5

10

15

Ile Leu Gly Thr Phe Tyr Tyr Leu Trp Gln Asp Asn Arg Ala His Ala

20

25

30

Ala Ser Ser Gly Gly Arg Gly Ala Gln Arg Ala Gly Arg Arg Ser Glu

35

40

45

Gln Leu Arg Glu Asp Arg Thr Ile Pro Leu Ile Val Thr Gly Thr Pro

50

55

60

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

Tyr Asn Phe Pro Glu Gly Asn Ala Leu Thr Tyr Ile Arg Asn Thr Lys
 355 360 365
 Arg Thr Ala Glu Val Trp Met Asp Glu Tyr Lys Gln Tyr Tyr Tyr Glu
 370 375 380
 Ala Arg Pro Ser Ala Ile Gly Lys Ala Phe Gly Ser Val Ala Thr Arg
 385 390 395 400
 Ile Glu Gln Arg Lys Lys Met Asn Cys Lys Ser Phe Arg Trp Tyr Leu
 405 410 415
 Glu Asn Val Tyr Pro Glu Leu Thr Val Pro Val Lys Glu Ala Leu Pro
 420 425 430
 Gly Ile Ile Lys Gln Gly Val Asn Cys Leu Glu Ser Gln Gly Gln Asn
 435 440 445
 Thr Ala Gly Asp Phe Leu Leu Gly Met Gly Ile Cys Arg Gly Ser Ala
 450 455 460
 Lys Asn Pro Gln Pro Ala Gln Ala Trp Leu Phe Ser Asp His Leu Ile
 465 470 475 480
 Gln Gln Gln Gly Lys Cys Leu Ala Ala Thr Ser Thr Leu Met Ser Ser
 485 490 495
 Pro Gly Ser Pro Val Ile Leu Gln Met Cys Asn Pro Arg Glu Gly Lys
 500 505 510
 Gln Val Ser Leu Leu Ala Ser Gly Pro Glu Ala Gln Gln Pro Glu Gly
 515 520 525
 Pro Cys Leu Arg Val Ala Asp Leu Gly Arg Arg Ala Pro Asp
 530 535 540

<210> 2

<211> 552

<212> PRT

<213> Homo sapiens

<400> 2

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

	35		40		45										
Trp	Asp	Asp	Leu	Trp	Asp	Gln	Phe	Asp	Glu	Arg	Arg	Tyr	Leu	Asn	Ala
	50				55							60			
Lys	Lys	Trp	Arg	Val	Gly	Asp	Asp	Pro	Tyr	Lys	Leu	Tyr	Ala	Phe	Asn
	65				70						75				80
Gln	Arg	Glu	Ser	Glu	Arg	Ile	Ser	Ser	Asn	Arg	Ala	Ile	Pro	Asp	Thr
				85						90				95	
Arg	His	Leu	Arg	Cys	Thr	Leu	Leu	Val	Tyr	Cys	Thr	Asp	Leu	Pro	Pro
				100						105				110	
Thr	Ser	Ile	Ile	Ile	Thr	Phe	His	Asn	Glu	Ala	Arg	Ser	Thr	Leu	Leu
		115						120						125	
Arg	Thr	Ile	Arg	Ser	Val	Leu	Asn	Arg	Thr	Pro	Thr	His	Leu	Ile	Arg
	130					135						140			
Glu	Ile	Ile	Leu	Val	Asp	Asp	Phe	Ser	Asn	Asp	Pro	Asp	Asp	Cys	Lys
	145				150						155				160
Gln	Leu	Ile	Lys	Leu	Pro	Lys	Val	Lys	Cys	Leu	Arg	Asn	Asn	Glu	Arg
				165						170				175	
Gln	Gly	Leu	Val	Arg	Ser	Arg	Ile	Arg	Gly	Ala	Asp	Ile	Ala	Gln	Gly
			180							185				190	
Thr	Thr	Leu	Thr	Phe	Leu	Asp	Ser	His	Cys	Glu	Val	Asn	Arg	Asp	Trp
		195						200						205	
Leu	Gln	Pro	Leu	Leu	His	Arg	Val	Lys	Glu	Asp	Tyr	Thr	Arg	Val	Val
	210					215						220			
Cys	Pro	Val	Ile	Asp	Ile	Ile	Asn	Leu	Asp	Thr	Phe	Thr	Tyr	Ile	Glu
	225				230						235				240
Ser	Ala	Ser	Glu	Leu	Arg	Gly	Gly	Phe	Asp	Trp	Ser	Leu	His	Phe	Gln
				245						250				255	
Trp	Glu	Gln	Leu	Ser	Pro	Glu	Gln	Lys	Ala	Arg	Arg	Leu	Asp	Pro	Thr
			260							265				270	
Glu	Pro	Ile	Arg	Thr	Pro	Ile	Ile	Ala	Gly	Gly	Leu	Phe	Val	Ile	Asp
		275						280						285	
Lys	Ala	Trp	Phe	Asp	Tyr	Leu	Gly	Lys	Tyr	Asp	Met	Asp	Met	Asp	Ile
	290					295					300				
Trp	Gly	Gly	Glu	Asn	Phe	Glu	Ile	Ser	Phe	Arg	Val	Trp	Met	Cys	Gly
	305				310						315				320
Gly	Ser	Leu	Glu	Ile	Val	Pro	Cys	Ser	Arg	Val	Gly	His	Val	Phe	Arg

	325		330		335										
Lys	Lys	His	Pro	Tyr	Val	Phe	Pro	Asp	Gly	Asn	Ala	Asn	Thr	Tyr	Ile
	340		345		350										
Lys	Asn	Thr	Lys	Arg	Thr	Ala	Glu	Val	Trp	Met	Asp	Glu	Tyr	Lys	Gln
	355		360		365										
Tyr	Tyr	Tyr	Ala	Ala	Arg	Pro	Phe	Ala	Leu	Glu	Arg	Pro	Phe	Gly	Asn
	370		375		380										
Val	Glu	Ser	Arg	Leu	Asp	Leu	Arg	Lys	Asn	Leu	Arg	Cys	Gln	Ser	Phe
385			390		395										400
Lys	Trp	Tyr	Leu	Glu	Asn	Ile	Tyr	Pro	Glu	Leu	Ser	Ile	Pro	Lys	Glu
	405		410		415										
Ser	Ser	Ile	Gln	Lys	Gly	Asn	Ile	Arg	Gln	Arg	Gln	Lys	Cys	Leu	Glu
	420		425		430										
Ser	Gln	Arg	Gln	Asn	Asn	Gln	Glu	Thr	Pro	Asn	Leu	Lys	Leu	Ser	Pro
	435		440		445										
Cys	Ala	Lys	Val	Lys	Gly	Glu	Asp	Ala	Lys	Ser	Gln	Val	Trp	Ala	Phe
	450		455		460										
Thr	Tyr	Thr	Gln	Gln	Ile	Leu	Gln	Glu	Glu	Leu	Cys	Leu	Ser	Val	Ile
465			470		475										480
Thr	Leu	Phe	Pro	Gly	Ala	Pro	Val	Val	Leu	Val	Leu	Cys	Lys	Asn	Gly
	485		490		495										
Asp	Asp	Arg	Gln	Gln	Trp	Thr	Lys	Thr	Gly	Ser	His	Ile	Glu	His	Ile
	500		505		510										
Ala	Ser	His	Leu	Cys	Leu	Asp	Thr	Asp	Met	Phe	Gly	Asp	Gly	Thr	Glu
	515		520		525										
Asn	Gly	Lys	Glu	Ile	Val	Val	Asn	Pro	Cys	Glu	Ser	Ser	Leu	Met	Ser
	530		535		540										
Gln	His	Trp	Asp	Met	Val	Ser	Ser								
545			550												

<210> 3

<211> 603

<212> PRT

<213> Homo sapiens

<400> 3

Met Arg Arg Lys Glu Lys Arg Leu Leu Gln Ala Val Ala Leu Val Leu
 1 5 10 15
 Ala Ala Leu Val Leu Leu Pro Asn Val Gly Leu Trp Ala Leu Tyr Arg
 20 25 30
 Glu Arg Gln Pro Asp Gly Thr Pro Gly Gly Ser Gly Ala Ala Val Ala
 35 40 45
 Pro Ala Ala Gly Gln Gly Ser His Ser Arg Gln Lys Lys Thr Phe Phe
 50 55 60
 Leu Gly Asp Gly Gln Lys Leu Lys Asp Trp His Asp Lys Glu Ala Ile
 65 70 75 80
 Arg Arg Asp Ala Gln Arg Val Gly Asn Gly Glu Gln Gly Arg Pro Tyr
 85 90 95
 Pro Met Thr Asp Ala Glu Arg Val Asp Gln Ala Tyr Arg Glu Asn Gly
 100 105 110
 Phe Asn Ile Tyr Val Ser Asp Lys Ile Ser Leu Asn Arg Ser Leu Pro
 115 120 125
 Asp Ile Arg His Pro Asn Cys Asn Ser Lys Arg Tyr Leu Glu Thr Leu
 130 135 140
 Pro Asn Thr Ser Ile Ile Ile Pro Phe His Asn Glu Gly Trp Ser Ser
 145 150 155 160
 Leu Leu Arg Thr Val His Ser Val Leu Asn Arg Ser Pro Pro Glu Leu
 165 170 175
 Val Ala Glu Ile Val Leu Val Asp Asp Phe Ser Asp Arg Glu His Leu
 180 185 190
 Lys Lys Pro Leu Glu Asp Tyr Met Ala Leu Phe Pro Ser Val Arg Ile
 195 200 205
 Leu Arg Thr Lys Lys Arg Glu Gly Leu Ile Arg Thr Arg Met Leu Gly
 210 215 220
 Ala Ser Val Ala Thr Gly Asp Val Ile Thr Phe Leu Asp Ser His Cys
 225 230 235 240
 Glu Ala Asn Val Asn Trp Leu Pro Pro Leu Leu Asp Arg Ile Ala Arg
 245 250 255
 Asn Arg Lys Thr Ile Val Cys Pro Met Ile Asp Val Ile Asp His Asp
 260 265 270
 Asp Phe Arg Tyr Glu Thr Gln Ala Gly Asp Ala Met Arg Gly Ala Phe
 275 280 285

Asp Trp Glu Met Tyr Tyr Lys Arg Ile Pro Ile Pro Pro Glu Leu Gln
 290 295 300
 Lys Ala Asp Pro Ser Asp Pro Phe Glu Ser Pro Val Met Ala Gly Gly
 305 310 315 320
 Leu Phe Ala Val Asp Arg Lys Trp Phe Trp Glu Leu Gly Gly Tyr Asp
 325 330 335
 Pro Gly Leu Glu Ile Trp Gly Gly Glu Gln Tyr Glu Ile Ser Phe Lys
 340 345 350
 Val Trp Met Cys Gly Gly Arg Met Glu Asp Ile Pro Cys Ser Arg Val
 355 360 365
 Gly His Ile Tyr Arg Lys Tyr Val Pro Tyr Lys Val Pro Ala Gly Val
 370 375 380
 Ser Leu Ala Arg Asn Leu Lys Arg Val Ala Glu Val Trp Met Asp Glu
 385 390 395 400
 Tyr Ala Glu Tyr Ile Tyr Gln Arg Arg Pro Glu Tyr Arg His Leu Ser
 405 410 415
 Ala Gly Asp Val Ala Val Gln Lys Lys Leu Arg Ser Ser Leu Asn Cys
 420 425 430
 Lys Ser Phe Lys Trp Phe Met Thr Lys Ile Ala Trp Asp Leu Pro Lys
 435 440 445
 Phe Tyr Pro Pro Val Glu Pro Pro Ala Ala Ala Trp Gly Glu Ile Arg
 450 455 460
 Asn Val Gly Thr Gly Leu Cys Ala Asp Thr Lys His Gly Ala Leu Gly
 465 470 475 480
 Ser Pro Leu Arg Leu Glu Gly Cys Val Arg Gly Arg Gly Glu Ala Ala
 485 490 495
 Trp Asn Asn Met Gln Val Phe Thr Phe Thr Trp Arg Glu Asp Ile Arg
 500 505 510
 Pro Gly Asp Pro Gln Tyr Thr Lys Lys Phe Cys Phe Asp Ala Ile Ser
 515 520 525
 His Thr Ser Pro Val Thr Leu Tyr Asp Cys His Ser Met Lys Gly Asn
 530 535 540
 Gln Leu Trp Lys Tyr Arg Lys Asp Lys Thr Leu Tyr His Pro Val Ser
 545 550 555 560
 Gly Ser Cys Met Asp Cys Ser Glu Ser Asp His Arg Ile Phe Met Asn
 565 570 575

Thr Cys Asn Pro Ser Ser Leu Thr Gln Gln Trp Leu Phe Glu His Thr
 580 585 590
 Asn Ser Thr Val Leu Glu Lys Phe Asn Arg Asn
 595 600

<210> 4
 <211> 581
 <212> PRT
 <213> Homo sapiens

<400> 4
 Met Trp Gly Arg Thr Ala Arg Arg Arg Cys Pro Arg Glu Leu Arg Arg
 1 5 10 15
 Gly Arg Glu Ala Leu Leu Val Leu Leu Ala Leu Leu Ala Leu Ala Gly
 20 25 30
 Leu Gly Ser Val Leu Arg Ala Gln Arg Gly Ala Gly Ala Gly Ala Ala
 35 40 45
 Glu Pro Gly Pro Pro Arg Thr Pro Arg Pro Gly Arg Arg Glu Pro Val
 50 55 60
 Met Pro Arg Pro Pro Val Pro Ala Asn Ala Leu Gly Ala Arg Gly Glu
 65 70 75 80
 Ala Val Arg Leu Gln Leu Gln Gly Glu Glu Leu Arg Leu Gln Glu Glu
 85 90 95
 Ser Val Arg Leu His Gln Ile Asn Ile Tyr Leu Ser Asp Arg Ile Ser
 100 105 110
 Leu His Arg Arg Leu Pro Val Arg Trp Asn Pro Leu Cys Lys Glu Lys
 115 120 125
 Lys Tyr Asp Tyr Asp Asn Leu Pro Arg Thr Ser Val Ile Ile Ala Phe
 130 135 140
 Tyr Asn Glu Ala Trp Ser Thr Leu Leu Arg Thr Val Tyr Ser Val Leu
 145 150 155 160
 Glu Thr Ser Pro Asp Ile Leu Leu Glu Glu Val Ile Leu Val Asp Asp
 165 170 175
 Tyr Ser Asp Arg Glu His Leu Lys Glu Arg Leu Ala Asn Glu Leu Ser
 180 185 190
 Gly Leu Pro Lys Val Arg Leu Ile Arg Ala Asn Lys Arg Glu Gly Leu

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

	485		490		495
Arg Tyr Asn Thr His Gln Pro Glu Gly Cys Ile Ala Val Glu Ala Gly					
	500		505		510
Met Asp Thr Leu Ile Met His Leu Cys Glu Glu Thr Ala Pro Glu Asn					
	515		520		525
Gln Lys Phe Ile Leu Gln Glu Asp Gly Ser Leu Phe His Glu Gln Ser					
	530		535		540
Lys Lys Cys Val Gln Ala Ala Arg Lys Glu Ser Ser Asp Ser Phe Val					
545		550		555	560
Pro Leu Leu Arg Asp Cys Thr Asn Ser Asp His Gln Lys Trp Phe Phe					
	565		570		575
Lys Glu Arg Met Leu					
	580				

<210> 5

<211> 607

<212> PRT

<213> Homo sapiens

<400> 5

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

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

Val Tyr Met Ala Trp Asn Ile Pro Gln Glu Asp Ser Gly Ile Asp Ile
 420 425 430
 Gly Asp Ile Thr Ala Arg Lys Ala Leu Arg Lys Gln Leu Gln Cys Lys
 435 440 445
 Thr Phe Arg Trp Tyr Leu Val Ser Val Tyr Pro Glu Met Arg Met Tyr
 450 455 460
 Ser Asp Ile Ile Ala Tyr Gly Val Leu Gln Asn Ser Leu Lys Thr Asp
 465 470 475 480
 Leu Cys Leu Asp Gln Gly Pro Asp Thr Glu Asn Val Pro Ile Met Tyr
 485 490 495
 Ile Cys His Gly Met Thr Pro Gln Asn Val Tyr Tyr Thr Ser Ser Gln
 500 505 510
 Gln Ile His Val Gly Ile Leu Ser Pro Thr Val Asp Asp Asp Asp Asn
 515 520 525
 Arg Cys Leu Val Asp Val Asn Ser Arg Pro Arg Leu Ile Glu Cys Ser
 530 535 540
 Tyr Ala Lys Ala Lys Arg Met Lys Leu His Trp Gln Phe Ser Gln Gly
 545 550 555 560
 Gly Pro Ile Gln Asn Arg Lys Ser Lys Arg Cys Leu Glu Leu Gln Glu
 565 570 575
 Asn Ser Asp Leu Glu Phe Gly Phe Gln Leu Val Leu Gln Lys Cys Ser
 580 585 590
 Gly Gln His Trp Ser Ile Thr Asn Val Leu Arg Ser Leu Ala Ser
 595 600 605

<210> 6

<211> 598

<212> PRT

<213> Homo sapiens

<400> 6

Met Ala Ser Leu Arg Arg Val Lys Val Leu Leu Val Leu Asn Leu Ile
 1 5 10 15
 Ala Val Ala Gly Phe Val Leu Phe Leu Ala Lys Cys Arg Pro Ile Ala
 20 25 30

Arg Thr Pro Ala Met Ile Gly Cys Ser Phe Val Val Asn Arg Lys Phe
 325 330 335
 Phe Gly Glu Ile Gly Leu Leu Asp Pro Gly Met Asp Val Tyr Gly Gly
 340 345 350
 Glu Asn Ile Glu Leu Gly Ile Lys Val Trp Leu Cys Gly Gly Ser Met
 355 360 365
 Glu Val Leu Pro Cys Ser Arg Val Ala His Ile Glu Arg Lys Lys Lys
 370 375 380
 Pro Tyr Asn Ser Asn Ile Gly Phe Tyr Thr Lys Arg Asn Ala Leu Arg
 385 390 395 400
 Val Ala Glu Val Trp Met Asp Asp Tyr Lys Ser His Val Tyr Ile Ala
 405 410 415
 Trp Asn Leu Pro Leu Glu Asn Pro Gly Ile Asp Ile Gly Asp Val Ser
 420 425 430
 Glu Arg Arg Ala Leu Arg Lys Ser Leu Lys Cys Lys Asn Phe Gln Trp
 435 440 445
 Tyr Leu Asp His Val Tyr Pro Glu Met Arg Arg Tyr Asn Asn Thr Val
 450 455 460
 Ala Tyr Gly Glu Leu Arg Asn Asn Lys Ala Lys Asp Val Cys Leu Asp
 465 470 475 480
 Gln Gly Pro Leu Glu Asn His Thr Ala Ile Leu Tyr Pro Cys His Gly
 485 490 495
 Trp Gly Pro Gln Leu Ala Arg Tyr Thr Lys Glu Gly Phe Leu His Leu
 500 505 510
 Gly Ala Leu Gly Thr Thr Thr Leu Leu Pro Asp Thr Arg Cys Leu Val
 515 520 525
 Asp Asn Ser Lys Ser Arg Leu Pro Gln Leu Leu Asp Cys Asp Lys Val
 530 535 540
 Lys Ser Ser Leu Tyr Lys Arg Trp Asn Phe Ile Gln Asn Gly Ala Ile
 545 550 555 560
 Met Asn Lys Gly Thr Gly Arg Cys Leu Glu Val Glu Asn Arg Gly Leu
 565 570 575
 Ala Gly Ile Asp Leu Ile Leu Arg Ser Cys Thr Gly Gln Arg Trp Thr
 580 585 590
 Ile Lys Asn Ser Ile Lys
 595

<210> 7

<211> 639

<212> PRT

<213> Homo sapiens

<400> 7

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

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

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

<210> 8

<211> 1629

<212> DNA

<213> Homo sapiens

<400> 8

atg agg aag atc cgc gcc aat gcc atc gcc atc ctg acc gta gcc tgg	48		
Met Arg Lys Ile Arg Ala Asn Ala Ile Ala Ile Leu Thr Val Ala Trp			
1	5	10	15
atc ctg ggc act ttc tac tac tta tgg cag gac aac cga gcc cac gca	96		
Ile Leu Gly Thr Phe Tyr Tyr Leu Trp Gln Asp Asn Arg Ala His Ala			
	20	25	30
gca tcc tcc ggc ggc cgg ggc gcg cag agg gca ggc agg agg tcg gag	144		
Ala Ser Ser Gly Gly Arg Gly Ala Gln Arg Ala Gly Arg Arg Ser Glu			
	35	40	45
cag ctc cgc gag gac cgc acc atc ccg ctc att gtg aca gga act ccc	192		
Gln Leu Arg Glu Asp Arg Thr Ile Pro Leu Ile Val Thr Gly Thr Pro			

50	55	60	
tcg aaa ggc ttt gat gag aag gcc tac ctg tgc gcc aag cag cig aag			240
Ser Lys Gly Phe Asp Glu Lys Ala Tyr Leu Ser Ala Lys Gln Leu Lys			
65	70	75	80
gct gga gag gac ccc tac aga cag cac gcc ttc aac cag ctg gag agt			288
Ala Gly Glu Asp Pro Tyr Arg Gln His Ala Phe Asn Gln Leu Glu Ser			
85	90	95	
gac aag ctg agc cca gac cgg ccc atc cgg gac acc cgc cat tac agc			336
Asp Lys Leu Ser Pro Asp Arg Pro Ile Arg Asp Thr Arg His Tyr Ser			
100	105	110	
tgc cca tct gtg tcc tac tcc tgc gac ctg cca gcc acc agc gtc atc			384
Cys Pro Ser Val Ser Tyr Ser Ser Asp Leu Pro Ala Thr Ser Val Ile			
115	120	125	
atc acc ttc cac aat gag gcc cgt tcc acc ctg ctg cgc aca gig aag			432
Ile Thr Phe His Asn Glu Ala Arg Ser Thr Leu Leu Arg Thr Val Lys			
130	135	140	
agt gtc ctg aac cga act cct gcc aac ttg atc cag gag atc att tta			480
Ser Val Leu Asn Arg Thr Pro Ala Asn Leu Ile Gln Glu Ile Ile Leu			
145	150	155	160
gtg gat gac ttc agc tca gat ccg gaa gac tgt cta ctc ctg acc agg			528
Val Asp Asp Phe Ser Ser Asp Pro Glu Asp Cys Leu Leu Leu Thr Arg			
165	170	175	
atc ccc aag gtc aag tgc ctg cgc aat gat cgg cgg gaa ggg ctg atc			576
Ile Pro Lys Val Lys Cys Leu Arg Asn Asp Arg Arg Glu Gly Leu Ile			
180	185	190	
cgg tcc cga gtg cgt ggg gcg gac gtg gct gca gct acc gtt ctc acc			624
Arg Ser Arg Val Arg Gly Ala Asp Val Ala Ala Ala Thr Val Leu Thr			

195	200	205	
ttt ctg gat agc cac tgc gaa gtg aac acc gag tgg ctg ccg ccc atg			672
Phe Leu Asp Ser His Cys Glu Val Asn Thr Glu Trp Leu Pro Pro Met			
210	215	220	
ctg cag cgg gtg aag gag gac cac acc cgc gtg gtg agt ccc atc att			720
Leu Gln Arg Val Lys Glu Asp His Thr Arg Val Val Ser Pro Ile Ile			
225	230	235	240
gat gtc atc agt ctg gat aat ttt gcc tac ctt gca gca tct gct gac			768
Asp Val Ile Ser Leu Asp Asn Phe Ala Tyr Leu Ala Ala Ser Ala Asp			
	245	250	255
ctt cgt gga ggg ttc gac tgg agc ctg cat ttc aag tgg gag cag atc			816
Leu Arg Gly Gly Phe Asp Trp Ser Leu His Phe Lys Trp Glu Gln Ile			
	260	265	270
cct ctt gag cag aag atg acc cgg aca gac ccc acc agg ccc ata agg			864
Pro Leu Glu Gln Lys Met Thr Arg Thr Asp Pro Thr Arg Pro Ile Arg			
	275	280	285
acg cct gtc ata gct gga gga atc ttc gtg atc gac aag tcc tgg ttt			912
Thr Pro Val Ile Ala Gly Gly Ile Phe Val Ile Asp Lys Ser Trp Phe			
	290	295	300
aac cac ttg gga aag tat gat gcc cag atg gac atc tgg ggg gga gag			960
Asn His Leu Gly Lys Tyr Asp Ala Gln Met Asp Ile Trp Gly Gly Glu			
305	310	315	320
aat ttt gag ctc tcc ttc agg gtg tgg atg tgt ggt ggc agt ctg gag			1008
Asn Phe Glu Leu Ser Phe Arg Val Trp Met Cys Gly Gly Ser Leu Glu			
	325	330	335
atc gtc ccc tgc agc cgg gtg ggc cat gtc ttc agg aaa cgg cac ccc			1056
Ile Val Pro Cys Ser Arg Val Gly His Val Phe Arg Lys Arg His Pro			

340	345	350	
tac aac ttc cct gag ggt aat gcc ctc acc tac atc agg aat act aag			1104
Tyr Asn Phe Pro Glu Gly Asn Ala Leu Thr Tyr Ile Arg Asn Thr Lys			
355	360	365	
cgc act gca gaa gtg tgg atg gat gaa tac aag caa tac tac tat gag			1152
Arg Thr Ala Glu Val Trp Met Asp Glu Tyr Lys Gln Tyr Tyr Tyr Glu			
370	375	380	
gcc cgg ccc tcg gcc atc ggg aag gcc ttc ggc agt gtg gct acg cgg			1200
Ala Arg Pro Ser Ala Ile Gly Lys Ala Phe Gly Ser Val Ala Thr Arg			
385	390	395	400
ata gag cag agg aag aag atg aac tgc aag tcc ttc cgc tgg tac ctg			1248
Ile Glu Gln Arg Lys Lys Met Asn Cys Lys Ser Phe Arg Trp Tyr Leu			
405	410	415	
gag aac gtc tac cca gag ctc acg gtc ccc gtg aag gaa gca ctc ccc			1296
Glu Asn Val Tyr Pro Glu Leu Thr Val Pro Val Lys Glu Ala Leu Pro			
420	425	430	
ggc atc att aag cag ggg gtg aac tgc tta gaa tct cag ggc cag aac			1344
Gly Ile Ile Lys Gln Gly Val Asn Cys Leu Glu Ser Gln Gly Gln Asn			
435	440	445	
aca gct ggt gac ttc ctg ctt gga atg ggg atc tgc aga ggg tct gcc			1392
Thr Ala Gly Asp Phe Leu Leu Gly Met Gly Ile Cys Arg Gly Ser Ala			
450	455	460	
aag aac ccg cag ccc gcc cag gca tgg ctg ttc agt gac cac ctc atc			1440
Lys Asn Pro Gln Pro Ala Gln Ala Trp Leu Phe Ser Asp His Leu Ile			
465	470	475	480
cag cag cag ggg aag tgc ctg gct gcc acc tcc acc tta atg tcc tcc			1488
Gln Gln Gln Gly Lys Cys Leu Ala Ala Thr Ser Thr Leu Met Ser Ser			

485	490	495	
cct gga tcc cca gtc ata ctg cag atg tgc aac cct aga gaa ggc aag			1536
Pro Gly Ser Pro Val Ile Leu Gln Met Cys Asn Pro Arg Glu Gly Lys			
500	505	510	
cag gtg agt ctc ctt gcc tct ggc cca gag gcc cag cag ccc gag ggg			1584
Gln Val Ser Leu Leu Ala Ser Gly Pro Glu Ala Gln Gln Pro Glu Gly			
515	520	525	
cca tgc ctc agg gtg gca gac ctt ggc agg aga gcc cct gat tga			1629
Pro Cys Leu Arg Val Ala Asp Leu Gly Arg Arg Ala Pro Asp			
530	535	540	
<210> 9			
<211> 1659			
<212> DNA			
<213> Homo sapiens			
<400> 9			
atg cgg cgc ctg act cgt cgg ctg gtt ctg cca gtc ttc ggg gtg ctc			48
Met Arg Arg Leu Thr Arg Arg Leu Val Leu Pro Val Phe Gly Val Leu			
1	5	10	15
tgg atc acg gtg ctg ctg ttc ttc tgg gta acc aag agg aag ttt gag			96
Trp Ile Thr Val Leu Leu Phe Phe Trp Val Thr Lys Arg Lys Leu Glu			
20	25	30	
gtg cgg acg gga cct gaa gtg cag acc cct aag cct tcg gac gct gac			144
Val Pro Thr Gly Pro Glu Val Gln Thr Pro Lys Pro Ser Asp Ala Asp			
35	40	45	
tgg gac gac ctg tgg gac cag ttt gat gag cgg cgg tat ctg aat gcc			192
Trp Asp Asp Leu Trp Asp Gln Phe Asp Glu Arg Arg Tyr Leu Asn Ala			
50	55	60	

aaa aag tgg cgc gtt ggt gac gac ccc tat aag ctg tat gct ttc aac	240
Lys Lys Trp Arg Val Gly Asp Asp Pro Tyr Lys Leu Tyr Ala Phe Asn	
65	70 75 80
cag cgg gag agt gag cgg atc tcc agc aat cgg gcc atc ccg gac act	288
Gln Arg Glu Ser Glu Arg Ile Ser Ser Asn Arg Ala Ile Pro Asp Thr	
85	90 95
cgc cat ctg aga tgc aca ctg ctg gtg tat tgc acg gac ctt cca ccc	336
Arg His Leu Arg Cys Thr Leu Leu Val Tyr Cys Thr Asp Leu Pro Pro	
100	105 110
act agc atc atc atc acc ttc cac aac gag gcc cgc tcc acg ctg ctc	384
Thr Ser Ile Ile Ile Thr Phe His Asn Glu Ala Arg Ser Thr Leu Leu	
115	120 125
agg acc atc cgc agt gta tta aac cgc acc cct acg cat ctg atc cgg	432
Arg Thr Ile Arg Ser Val Leu Asn Arg Thr Pro Thr His Leu Ile Arg	
130	135 140
gaa atc ata tta gtg gat gac ttc agc aat gat cct gat gac tgt aaa	480
Glu Ile Ile Leu Val Asp Asp Phe Ser Asn Asp Pro Asp Asp Cys Lys	
145	150 155 160
cag ctc atc aag ttg ccc aag gtg aaa tgc ttg cgc aat aat gaa cgg	528
Gln Leu Ile Lys Leu Pro Lys Val Lys Cys Leu Arg Asn Asn Glu Arg	
165	170 175
caa ggt ctg gtc cgg tcc cgg att cgg ggc gct gac atc gcc cag ggc	576
Gln Gly Leu Val Arg Ser Arg Ile Arg Gly Ala Asp Ile Ala Gln Gly	
180	185 190
acc act ctg act ttc ctc gac agc cac tgt gag gtg aac agg gac tgg	624
Thr Thr Leu Thr Phe Leu Asp Ser His Cys Glu Val Asn Arg Asp Trp	
195	200 205

ctc cag cct ctg ttg cac agg gtc aaa gag gac tac acg cgg gtg gtg 672
 Leu Gln Pro Leu Leu His Arg Val Lys Glu Asp Tyr Thr Arg Val Val
 210 215 220

tgc cct gtg atc gat atc att aac ctg gac acc ttc acc tac atc gag 720
 Cys Pro Val Ile Asp Ile Ile Asn Leu Asp Thr Phe Thr Tyr Ile Glu
 225 230 235 240

tct gcc tcg gag ctc aga ggg ggg ttt gac tgg agc ctc cac ttc cag 768
 Ser Ala Ser Glu Leu Arg Gly Gly Phe Asp Trp Ser Leu His Phe Gln
 245 250 255

tgg gag cag ctc tcc cca gag cag aag gct cgg cgc ctg gac ccc acg 816
 Trp Glu Gln Leu Ser Pro Glu Gln Lys Ala Arg Arg Leu Asp Pro Thr
 260 265 270

gag ccc atc agg act cct atc ata gct gga ggg ctc ttc gtg atc gac 864
 Glu Pro Ile Arg Thr Pro Ile Ile Ala Gly Gly Leu Phe Val Ile Asp
 275 280 285

aaa gct tgg ttt gat tac ctg ggg aaa tat gat atg gac atg gac atc 912
 Lys Ala Trp Phe Asp Tyr Leu Gly Lys Tyr Asp Met Asp Met Asp Ile
 290 295 300

tgg ggt ggg gag aac ttt gaa atc tcc ttc cga gtg tgg atg tgc ggg 960
 Trp Gly Gly Glu Asn Phe Glu Ile Ser Phe Arg Val Trp Met Cys Gly
 305 310 315 320

ggc agc cta gag atc gtc ccc tgc agc cga gtg ggg cac gtc ttc cgg 1008
 Gly Ser Leu Glu Ile Val Pro Cys Ser Arg Val Gly His Val Phe Arg
 325 330 335

aag aag cac ccc tac gtt ttc cct gat gga aat gcc aac acg tat ata 1056
 Lys Lys His Pro Tyr Val Phe Pro Asp Gly Asn Ala Asn Thr Tyr Ile
 340 345 350

aag aac acc aag cgg aca gct gaa gtg tgg atg gat gaa tac aag caa	1104
Lys Asn Thr Lys Arg Thr Ala Glu Val Trp Met Asp Glu Tyr Lys Gln	
355	360
365	
tac tat tac gct gcc cgg cca ttc gcc ctg gag agg ccc ttc ggg aat	1152
Tyr Tyr Tyr Ala Ala Arg Pro Phe Ala Leu Glu Arg Pro Phe Gly Asn	
370	375
380	
gtt gag agc aga ttg gac ctg agg aag aat ctg cgc tgc cag agc ttc	1200
Val Glu Ser Arg Leu Asp Leu Arg Lys Asn Leu Arg Cys Gln Ser Phe	
385	390
395	400
aag tgg tac ctg gag aat atc tac cct gaa ctc agc atc ccc aag gag	1248
Lys Trp Tyr Leu Glu Asn Ile Tyr Pro Glu Leu Ser Ile Pro Lys Glu	
405	410
415	
tec tcc atc cag aag ggc aat atc cga cag aga cag aag tgc ctg gaa	1296
Ser Ser Ile Gln Lys Gly Asn Ile Arg Gln Arg Gln Lys Cys Leu Glu	
420	425
430	
tct caa agg cag aac aac caa gaa acc cca aac cta aag ttg agc ccc	1344
Ser Gln Arg Gln Asn Asn Gln Glu Thr Pro Asn Leu Lys Leu Ser Pro	
435	440
445	
tgt gcc aag gtc aaa ggc gaa gat gca aag tcc cag gta tgg gcc ttc	1392
Cys Ala Lys Val Lys Gly Glu Asp Ala Lys Ser Gln Val Trp Ala Phe	
450	455
460	
aca tac acc cag cag atc ctc cag gag gag ctg tgc ctg tca gtc atc	1440
Thr Tyr Thr Gln Gln Ile Leu Gln Glu Glu Leu Cys Leu Ser Val Ile	
465	470
475	480
acc ttg ttc cct ggc gcc cca gtg gtt ctt gtc ctt tgc aag aat gga	1488
Thr Leu Phe Pro Gly Ala Pro Val Val Leu Val Leu Cys Lys Asn Gly	
485	490
495	

gat gac cga cag caa tgg acc aaa act ggt tcc cac atc gag cac ata 1536
 Asp Asp Arg Gln Gln Trp Thr Lys Thr Gly Ser His Ile Glu His Ile
 500 505 510

gca tcc cac ctc tgc ctc gat aca gat atg ttc ggt gat ggc acc gag 1584
 Ala Ser His Leu Cys Leu Asp Thr Asp Met Phe Gly Asp Gly Thr Glu
 515 520 525

aac ggc aag gaa atc gtc gtc aac cca tgt gag tcc tca ctc atg agc 1632
 Asn Gly Lys Glu Ile Val Val Asn Pro Cys Glu Ser Ser Leu Met Ser
 530 535 540

cag cac tgg gac atg gtg agc tct tga 1659
 Gln His Trp Asp Met Val Ser Ser
 545 550

<210> 10

<211> 1812

<212> DNA

<213> Homo sapiens

<400> 10

atg agg cgg aag gag aag cgg ctc ctg cag gcg gtg gcg ctg gtg ctg 48
 Met Arg Arg Lys Glu Lys Arg Leu Leu Gln Ala Val Ala Leu Val Leu
 1 5 10 15

gcg gcc ctg gtc ctc ctg ccc aac gtg ggg ctt tgg gcg ctg tac cgc 96
 Ala Ala Leu Val Leu Leu Pro Asn Val Gly Leu Trp Ala Leu Tyr Arg
 20 25 30

gag cgg cag ccc gac ggc acc cct ggg gga tcg ggg gcg gcg gtg gcg 144
 Glu Arg Gln Pro Asp Gly Thr Pro Gly Gly Ser Gly Ala Ala Val Ala
 35 40 45

ccg gcg gcg gga cag ggc tca cac agt cga caa aag aaa acg ttt ttc 192
 Pro Ala Ala Gly Gln Gly Ser His Ser Arg Gln Lys Lys Thr Phe Phe

50	55	60	
tig gga gat ggg cag aag ctg aag gac tgg cat gac aag gag gcc atc			240
Leu Gly Asp Gly Gln Lys Leu Lys Asp Trp His Asp Lys Glu Ala Ile			
65	70	75	80
cgg agg gac gct cag cgc gta gga aat gga gaa caa gga aga cct tac			288
Arg Arg Asp Ala Gln Arg Val Gly Asn Gly Glu Gln Gly Arg Pro Tyr			
	85	90	95
ccc atg acc gat gct gag aga gtg gat cag gca tac cga gaa aat gga			336
Pro Met Thr Asp Ala Glu Arg Val Asp Gln Ala Tyr Arg Glu Asn Gly			
	100	105	110
ttt aac atc tac gtc agt gat aaa atc tcc ttg aat cgc tct ctc cca			384
Phe Asn Ile Tyr Val Ser Asp Lys Ile Ser Leu Asn Arg Ser Leu Pro			
	115	120	125
gat atc cgg cac cca aac tgc aac agc aag cgc tac ctg gag aca ctt			432
Asp Ile Arg His Pro Asn Cys Asn Ser Lys Arg Tyr Leu Glu Thr Leu			
	130	135	140
ccc aac aca agc atc atc atc ccc ttc cac aac gag ggc tgg tcc tcc			480
Pro Asn Thr Ser Ile Ile Ile Pro Phe His Asn Glu Gly Trp Ser Ser			
	145	150	155
ctc ctc cgc acc gtc cac agt gtg ctc aat cgc teg cct cca gag ctg			528
Leu Leu Arg Thr Val His Ser Val Leu Asn Arg Ser Pro Pro Glu Leu			
	165	170	175
gtc gcc gag att gta ctg gtc gac gac ttc agt gat cga gag cac ctg			576
Val Ala Glu Ile Val Leu Val Asp Asp Phe Ser Asp Arg Glu His Leu			
	180	185	190
aag aag cct ctt gaa gac tac atg gcc ctt ttc ccc agt gtg agg att			624
Lys Lys Pro Leu Glu Asp Tyr Met Ala Leu Phe Pro Ser Val Arg Ile			

195	200	205	
ctt cga acc aag aaa cgg gaa ggg ctg ata agg acc cga atg ctg ggg			672
Leu Arg Thr Lys Lys Arg Glu Gly Leu Ile Arg Thr Arg Met Leu Gly			
210	215	220	
gcc tca gtg gca act ggg gat gtc atc aca ttc ttg gat tca cac tgt			720
Ala Ser Val Ala Thr Gly Asp Val Ile Thr Phe Leu Asp Ser His Cys			
225	230	235	240
gaa gcc aat gtc aac tgg ctt ccc ccc ttg ctt gac cgc att gct cgg			768
Glu Ala Asn Val Asn Trp Leu Pro Pro Leu Leu Asp Arg Ile Ala Arg			
	245	250	255
aac cgc aag acc att gtg tgc ccg atg att gat gta att gac cat gac			816
Asn Arg Lys Thr Ile Val Cys Pro Met Ile Asp Val Ile Asp His Asp			
	260	265	270
gac ttt cgg tac gag aca cag gca ggg gat gcc atg cgg gga gcc ttt			864
Asp Phe Arg Tyr Glu Thr Gln Ala Gly Asp Ala Met Arg Gly Ala Phe			
	275	280	285
gac tgg gag atg tac tac aag cgg atc ccg aic cct cca gaa cig cag			912
Asp Trp Glu Met Tyr Tyr Lys Arg Ile Pro Ile Pro Pro Glu Leu Gln			
	290	295	300
aaa gct gac ccc agc gac cca ttt gag tct ccc gtg atg gcc ggt gga			960
Lys Ala Asp Pro Ser Asp Pro Phe Glu Ser Pro Val Met Ala Gly Gly			
305	310	315	320
ctg ttc gcc gtg gat cgg aag tgg ttc tgg gaa ctc ggc ggg tat gac			1008
Leu Phe Ala Val Asp Arg Lys Trp Phe Trp Glu Leu Gly Gly Tyr Asp			
	325	330	335
cca ggc ttg gag atc tgg gga ggg gag cag tat gaa atc tcc ttc aag			1056
Pro Gly Leu Glu Ile Trp Gly Gly Glu Gln Tyr Glu Ile Ser Phe Lys			

340	345	350	
gtg tgg atg tgt ggg ggc cgc atg gag gac atc ccc tgc tcc agg gtg			1104
Val Trp Met Cys Gly Gly Arg Met Glu Asp Ile Pro Cys Ser Arg Val			
355	360	365	
ggc cat atc tac agg aag tat gtg ccc tac aag gtc ccg gcc gga gtc			1152
Gly His Ile Tyr Arg Lys Tyr Val Pro Tyr Lys Val Pro Ala Gly Val			
370	375	380	
agc ctg gcc cgg aac ctt aag cgg gtg gcc gaa gtg tgg atg gai gag			1200
Ser Leu Ala Arg Asn Leu Lys Arg Val Ala Glu Val Trp Met Asp Glu			
385	390	395	400
tac gca gag tac att tac cag cgc cgg cct gaa tac cgc cac ctc tcc			1248
Tyr Ala Glu Tyr Ile Tyr Gln Arg Arg Pro Glu Tyr Arg His Leu Ser			
405	410	415	
gct ggg gat gtc gca gtc cag aaa aag ctc cgc agc tcc ctt aac tgc			1296
Ala Gly Asp Val Ala Val Gln Lys Lys Leu Arg Ser Ser Leu Asn Cys			
420	425	430	
aag agt ttc aag tgg ttt atg acg aag ata gcc tgg gac ctg ccc aaa			1344
Lys Ser Phe Lys Trp Phe Met Thr Lys Ile Ala Trp Asp Leu Pro Lys			
435	440	445	
ttc tac cca ccc gtg gag ccc ccg gct gca gct tgg ggg gag atc cga			1392
Phe Tyr Pro Pro Val Glu Pro Pro Ala Ala Ala Trp Gly Glu Ile Arg			
450	455	460	
aat gtg ggc aca ggg ctg tgt gca gac aca aag cac ggg gcc ttg ggc			1440
Asn Val Gly Thr Gly Leu Cys Ala Asp Thr Lys His Gly Ala Leu Gly			
465	470	475	480
tcc cca cta agg cta gag ggc tgc gtc cga ggc cgt ggg gag gct gcc			1488
Ser Pro Leu Arg Leu Glu Gly Cys Val Arg Gly Arg Gly Glu Ala Ala			

485	490	495	
lgg aac aac atg cag gta ttc acc ttc acc tgg aga gag gac atc cgg			1536
Trp Asn Asn Met Gln Val Phe Thr Phe Thr Trp Arg Glu Asp Ile Arg			
500	505	510	
cct gga gac ccc cag tac acc aag aag ttc tgc ttt gat gcc att tcc			1584
Pro Gly Asp Pro Gln Tyr Thr Lys Lys Phe Cys Phe Asp Ala Ile Ser			
515	520	525	
cac acc agc cct gtc acg ctg tac gac tgc cac agc atg aag ggc aac			1632
His Thr Ser Pro Val Thr Leu Tyr Asp Cys His Ser Met Lys Gly Asn			
530	535	540	
cag ctg tgg aaa tac cgc aaa gac aag acc ctg tac cac cct gtc agt			1680
Gln Leu Trp Lys Tyr Arg Lys Asp Lys Thr Leu Tyr His Pro Val Ser			
545	550	555	560
ggc agc tgc atg gac tgc agt gaa agt gac cat agg atc ttc atg aac			1728
Gly Ser Cys Met Asp Cys Ser Glu Ser Asp His Arg Ile Phe Met Asn			
565	570	575	
acc tgc aac cca tcc tct ctc acc cag caa tgg ctg ttt gaa cac acc			1776
Thr Cys Asn Pro Ser Ser Leu Thr Gln Gln Trp Leu Phe Glu His Thr			
580	585	590	
aac tca aca gtc ttg gaa aaa ttc aat agg aac tga			1812
Asn Ser Thr Val Leu Glu Lys Phe Asn Arg Asn			
595	600		

<210> 11

<211> 1746

<212> DNA

<213> Homo sapiens

<400> 11

tat aat gaa gcc tgg tca act ctc ctt cgg aca gtt tac agt gtc ctt	480
Tyr Asn Glu Ala Trp Ser Thr Leu Leu Arg Thr Val Tyr Ser Val Leu	
145	150
155	160
gag aca tcc ccg gat atc ctg cta gaa gaa gtg atc ctt gta gat gac	528
Glu Thr Ser Pro Asp Ile Leu Leu Glu Glu Val Ile Leu Val Asp Asp	
165	170
175	
tac agt gat aga gag cac ctg aag gag cgc ttg gcc aat gag ctt tcg	576
Tyr Ser Asp Arg Glu His Leu Lys Glu Arg Leu Ala Asn Glu Leu Ser	
180	185
190	
gga ctg ccc aag gtg cgc ctg atc cgc gcc aac aag aga gag ggc ctg	624
Gly Leu Pro Lys Val Arg Leu Ile Arg Ala Asn Lys Arg Glu Gly Leu	
195	200
205	
gtg cga gcc cgg ctg ctg ggg gcg tct gcg gcg agg ggc gat gtt ctg	672
Val Arg Ala Arg Leu Leu Gly Ala Ser Ala Ala Arg Gly Asp Val Leu	
210	215
220	
acc ttc ctg gac tgt cac tgt gag tgc cac gaa ggg tgg ctg gag ccg	720
Thr Phe Leu Asp Cys His Cys Glu Cys His Glu Gly Trp Leu Glu Pro	
225	230
235	240
ctg ctg cag agg atc cat gaa gag gag tcg gca gtg gtg tgc ccg gtg	768
Leu Leu Gln Arg Ile His Glu Glu Glu Ser Ala Val Val Cys Pro Val	
245	250
255	
att gat gtg atc gac tgg aac acc ttc gaa tac ctg ggg aac tcc ggg	816
Ile Asp Val Ile Asp Trp Asn Thr Phe Glu Tyr Leu Gly Asn Ser Gly	
260	265
270	
gag ccc cag atc ggc ggt ttc gac tgg agg ctg gtg ttc acg tgg cac	864
Glu Pro Gln Ile Gly Gly Phe Asp Trp Arg Leu Val Phe Thr Trp His	
275	280
285	

aca gtt cct gag agg gag agg ata cgg atg caa tcc ccc gtc gat gtc 912
 Thr Val Pro Glu Arg Glu Arg Ile Arg Met Gln Ser Pro Val Asp Val
 290 295 300

atc agg tct cca aca atg gct ggt ggg ctg ttt gct gtg agt aag aaa 960
 Ile Arg Ser Pro Thr Met Ala Gly Gly Leu Phe Ala Val Ser Lys Lys
 305 310 315 320

tat ttt gaa tat ctg ggg tct tat gat aca gga atg gaa gtt tgg gga 1008
 Tyr Phe Glu Tyr Leu Gly Ser Tyr Asp Thr Gly Met Glu Val Trp Gly
 325 330 335

gga gaa aac ctc gaa ttt tcc ttt agg atc tgg cag tgt ggt ggg gtt 1056
 Gly Glu Asn Leu Glu Phe Ser Phe Arg Ile Trp Gln Cys Gly Gly Val
 340 345 350

ctg gaa aca cac cca tgt tcc cat gtt ggc cat gtt ttc ccc aag caa 1104
 Leu Glu Thr His Pro Cys Ser His Val Gly His Val Phe Pro Lys Gln
 355 360 365

gct ccc tac tcc cgc aac aag gct ctg gcc aac agt gtt cgt gca gct 1152
 Ala Pro Tyr Ser Arg Asn Lys Ala Leu Ala Asn Ser Val Arg Ala Ala
 370 375 380

gaa gta tgg atg gat gaa ttt aaa gag ctc tac tac cat cgc aac ccc 1200
 Glu Val Trp Met Asp Glu Phe Lys Glu Leu Tyr Tyr His Arg Asn Pro
 385 390 395 400

cgt gcc cgc ttg gaa cct ttt ggg gat gtg aca gag agg aag cag ctc 1248
 Arg Ala Arg Leu Glu Pro Phe Gly Asp Val Thr Glu Arg Lys Gln Leu
 405 410 415

cgg gac aag ctc cag tgt aaa gac ttc aag tgg ttc ttg gag act gtg 1296
 Arg Asp Lys Leu Gln Cys Lys Asp Phe Lys Trp Phe Leu Glu Thr Val
 420 425 430

tal cca gaa ctg cat gtg cct gag gac agg cct ggc ttc ttc ggg atg	1344
Tyr Pro Glu Leu His Val Pro Glu Asp Arg Pro Gly Phe Phe Gly Met	
435	440
ctc cag aac aaa gga cta aca gac tac tgc ttt gac tat aac cct ccc	1392
Leu Gln Asn Lys Gly Leu Thr Asp Tyr Cys Phe Asp Tyr Asn Pro Pro	
450	460
gat gaa aac cag att gtg gga cac cag gtc att ctg tac ctc tgt cat	1440
Asp Glu Asn Gln Ile Val Gly His Gln Val Ile Leu Tyr Leu Cys His	
465	470
ggg atg ggc cag aat cag ttt ttc gag tac acg tcc cag aaa gaa aia	1488
Gly Met Gly Gln Asn Gln Phe Phe Glu Tyr Thr Ser Gln Lys Glu Ile	
485	490
cgc tat aac acc cac cag cct gag ggc tgc att gct gtg gaa gca gga	1536
Arg Tyr Asn Thr His Gln Pro Glu Gly Cys Ile Ala Val Glu Ala Gly	
500	505
atg gat acc ctt atc atg cat ctc tgc gaa gaa act gcc cca gag aat	1584
Met Asp Thr Leu Ile Met His Leu Cys Glu Glu Thr Ala Pro Glu Asn	
515	520
cag aag ttc atc ttg cag gag gat gga tct tta ttt cac gaa cag tcc	1632
Gln Lys Phe Ile Leu Gln Glu Asp Gly Ser Leu Phe His Glu Gln Ser	
530	535
aag aaa tgt gtc cag gct gcg agg aag gag tcg agt gac agt ttc gtt	1680
Lys Lys Cys Val Gln Ala Ala Arg Lys Glu Ser Ser Asp Ser Phe Val	
545	550
cca ctc tta cga gac tgc acc aac tcg gat cat cag aaa tgg ttc ttc	1728
Pro Leu Leu Arg Asp Cys Thr Asn Ser Asp His Gln Lys Trp Phe Phe	
565	570

aaa gag cgc atg tta tga 1746
 Lys Glu Arg Met Leu
 580

<210> 12
 <211> 1824
 <212> DNA
 <213> Homo sapiens

<400> 12
 atg gtg tgc acc agg aag acc aaa act ttg gtg tcc act tgc gtg atc 48
 Met Val Cys Thr Arg Lys Thr Lys Thr Leu Val Ser Thr Cys Val Ile
 1 5 10 15
 ctg agc ggc atg act aac atc atc igc ctg ctc tac gtg ggc tgg gtc 96
 Leu Ser Gly Met Thr Asn Ile Ile Cys Leu Leu Tyr Val Gly Trp Val
 20 25 30
 acc aac tac atc gcc agc gtg tat gtg cgg ggg cag gag ccg gcg ccc 144
 Thr Asn Tyr Ile Ala Ser Val Tyr Val Arg Gly Gln Glu Pro Ala Pro
 35 40 45
 gac aag aag ctg gag gaa gac aaa ggg gac act ctg aag att att gag 192
 Asp Lys Lys Leu Glu Glu Asp Lys Gly Asp Thr Leu Lys Ile Ile Glu
 50 55 60
 cgg ctg gac cac ctg gag aat gtc atc aag cag cac att caa gag gct 240
 Arg Leu Asp His Leu Glu Asn Val Ile Lys Gln His Ile Gln Glu Ala
 65 70 75 80
 cct gcc aag cct gag gag gca gag gcc gag ccc ttc aca gac tcc tct 288
 Pro Ala Lys Pro Glu Glu Ala Glu Ala Glu Pro Phe Thr Asp Ser Ser
 85 90 95
 ctg ttt gca cac tgg ggc cag gag ctc agc ccc gaa ggc cgg cgc gtg 336
 Leu Phe Ala His Trp Gly Gln Glu Leu Ser Pro Glu Gly Arg Arg Val

100	105	110	
gcc ctg aag caa ttc cag tac tac ggc tac aac gcc tac ctc agc gac			384
Ala Leu Lys Gln Phe Gln Tyr Tyr Gly Tyr Asn Ala Tyr Leu Ser Asp			
115	120	125	
cgc ctg ccc ctg gac cgg ccc ctg cct gac ctc aga ccc agt ggg tgc			432
Arg Leu Pro Leu Asp Arg Pro Leu Pro Asp Leu Arg Pro Ser Gly Cys			
130	135	140	
cgt aac ctc tca ttt cct gac agc ctg cca gag gtc agc atc gtg ttc			480
Arg Asn Leu Ser Phe Pro Asp Ser Leu Pro Glu Val Ser Ile Val Phe			
145	150	155	160
atc ttc gtc aat gaa ggc ctt tca gtg ctg ctg cgc tcc atc cac tcg			528
Ile Phe Val Asn Glu Ala Leu Ser Val Leu Leu Arg Ser Ile His Ser			
165	170	175	
gcc atg gaa cgc acg ccc cca cat ctg ctc aag gag atc att ctg gtg			576
Ala Met Glu Arg Thr Pro Pro His Leu Leu Lys Glu Ile Ile Leu Val			
180	185	190	
gat gac aac agc agt aac gag gaa ctg aag gag aag ctg acc gaa tat			624
Asp Asp Asn Ser Ser Asn Glu Glu Leu Lys Glu Lys Leu Thr Glu Tyr			
195	200	205	
gtg gac aag gtg aac agc cag aag cca ggc ttc atc aaa gtc gtg cgt			672
Val Asp Lys Val Asn Ser Gln Lys Pro Gly Phe Ile Lys Val Val Arg			
210	215	220	
cac agc aag cag gaa ggc ctc atc cgc tcc agg gtc agt ggc tgg agg			720
His Ser Lys Gln Glu Gly Leu Ile Arg Ser Arg Val Ser Gly Trp Arg			
225	230	235	240
gcg gcc act gcc cct gtg gtg gca ctc ttt gat gcc cac gtg gag ttc			768
Ala Ala Thr Ala Pro Val Val Ala Leu Phe Asp Ala His Val Glu Phe			

245	250	255	
aat gtg ggc tgg gct gaa cct gta ctc acc cgc atc aag gag aac cgg			816
Asn Val Gly Trp Ala Glu Pro Val Leu Thr Arg Ile Lys Glu Asn Arg			
260	265	270	
aag cgg atc atc tcg cca tcc ttt gat aac atc aaa tat gac aac ttt			864
Lys Arg Ile Ile Ser Pro Ser Phe Asp Asn Ile Lys Tyr Asp Asn Phe			
275	280	285	
gag ata gaa gag tac ccg ctg gct gcc cag ggc ttt gac tgg gag ctg			912
Glu Ile Glu Glu Tyr Pro Leu Ala Ala Gln Gly Phe Asp Trp Glu Leu			
290	295	300	
tgg tgc cgc tac cta aat ccc ccc aag gcc tgg tgg aag ctg gag aac			960
Trp Cys Arg Tyr Leu Asn Pro Pro Lys Ala Trp Trp Lys Leu Glu Asn			
305	310	315	320
icc aca gcg cca atc agg agc cct gcc ctc att ggc tgc ttc att gtg			1008
Ser Thr Ala Pro Ile Arg Ser Pro Ala Leu Ile Gly Cys Phe Ile Val			
325	330	335	
gac cgg cag tac ttc cag gag atc ggc ctg ctg gac gaa ggc atg gaa			1056
Asp Arg Gln Tyr Phe Gln Glu Ile Gly Leu Leu Asp Glu Gly Met Glu			
340	345	350	
gtc tac ggg ggc gag aat gtg gag ctt ggg atc agg gtg tgg cag tgt			1104
Val Tyr Gly Gly Glu Asn Val Glu Leu Gly Ile Arg Val Trp Gln Cys			
355	360	365	
ggc ggg agt gtg gag gtc ctg ccc tgc tca cgg att gcc cac att gag			1152
Gly Gly Ser Val Glu Val Leu Pro Cys Ser Arg Ile Ala His Ile Glu			
370	375	380	
cga gcc cac aag ccc tac aca gag gac ctc acc gcc cat gtc cgc agg			1200
Arg Ala His Lys Pro Tyr Thr Glu Asp Leu Thr Ala His Val Arg Arg			

385	390	395	400	
aac gct ctc agg gtg gct gaa gtc tgg atg gat gaa ttt aaa agc cac				1248
Asn Ala Leu Arg Val Ala Glu Val Trp Met Asp Glu Phe Lys Ser His				
	405	410	415	
gtc tac atg gca tgg aac ata ccg cag gag gac tca gga att gac att				1296
Val Tyr Met Ala Trp Asn Ile Pro Gln Glu Asp Ser Gly Ile Asp Ile				
	420	425	430	
ggg gac atc act gca agg aag gct ctc agg aaa cag ctg cag tgc aag				1344
Gly Asp Ile Thr Ala Arg Lys Ala Leu Arg Lys Gln Leu Gln Cys Lys				
	435	440	445	
acc ttc cgg tgg tac ctg gtc agc gtg tac cca gag atg agg atg tac				1392
Thr Phe Arg Trp Tyr Leu Val Ser Val Tyr Pro Glu Met Arg Met Tyr				
	450	455	460	
tcc gac atc att gcc tat gga gtg ctg cag aat tct ctg aag act gat				1440
Ser Asp Ile Ile Ala Tyr Gly Val Leu Gln Asn Ser Leu Lys Thr Asp				
	465	470	475	480
ttg tgt ctt gac cag ggg cca gat aca gag aat gtc ccc atc atg tac				1488
Leu Cys Leu Asp Gln Gly Pro Asp Thr Glu Asn Val Pro Ile Met Tyr				
	485	490	495	
atc tgc cat ggg atg acg cct cag aac gtg tac tac acg agc agt cag				1536
Ile Cys His Gly Met Thr Pro Gln Asn Val Tyr Tyr Thr Ser Ser Gln				
	500	505	510	
cag atc cat gtg ggc att ctg agc ccc acc gtg gat gat gat gac aac				1584
Gln Ile His Val Gly Ile Leu Ser Pro Thr Val Asp Asp Asp Asp Asn				
	515	520	525	
cga tgc ctg gtg gac gtc aac agc cgg ecc cgg ctc atc gaa tgc agc				1632
Arg Cys Leu Val Asp Val Asn Ser Arg Pro Arg Leu Ile Glu Cys Ser				

gcc aac ctc agc gcg cac agc gcc agc ccc atc cag gat gcg gtc ctg	192
Ala Asn Leu Ser Ala His Ser Ala Ser Pro Ile Gln Asp Ala Val Leu	
50 55 60	
aag cgc ctg tcg ctg ctg gag gac atc gtg tac cgg cag ctg aat ggc	240
Lys Arg Leu Ser Leu Leu Glu Asp Ile Val Tyr Arg Gln Leu Asn Gly	
65 70 75 80	
tta tcc aaa tcc ctt ggg ctc att gaa ggt tat ggt ggg cgg ggt aaa	288
Leu Ser Lys Ser Leu Gly Leu Ile Glu Gly Tyr Gly Gly Arg Gly Lys	
85 90 95	
ggg ggc ctt ccg gct act ctt tcc ccg gct gaa gaa gaa aag gct aag	336
Gly Gly Leu Pro Ala Thr Leu Ser Pro Ala Glu Glu Glu Lys Ala Lys	
100 105 110	
gga ccc cat gag aag tat ggc tac aat tca tac ctc agt gaa aaa att	384
Gly Pro His Glu Lys Tyr Gly Tyr Asn Ser Tyr Leu Ser Glu Lys Ile	
115 120 125	
tca ctg gac cgt tcc att ccg gat tat cgt ccc acc aag tgt aag gag	432
Ser Leu Asp Arg Ser Ile Pro Asp Tyr Arg Pro Thr Lys Cys Lys Glu	
130 135 140	
ctc aag tac tcc aag gac ctg ccc cag ata tcc atc ata ttc atc ttc	480
Leu Lys Tyr Ser Lys Asp Leu Pro Gln Ile Ser Ile Ile Phe Ile Phe	
145 150 155 160	
gtg aac gag gcc ctg tcg gtg atc ctg cgg tcc gtg cac agt gcc gtc	528
Val Asn Glu Ala Leu Ser Val Ile Leu Arg Ser Val His Ser Ala Val	
165 170 175	
aat cac acg ccc aca cac gtg atg aag gaa atc att ttg gtg gat gac	576
Asn His Thr Pro Thr His Val Met Lys Glu Ile Ile Leu Val Asp Asp	
180 185 190	

aac agc gac gaa gag gag ctg aag gtc ccc cta gag gag tat gtc cac	624
Asn Ser Asp Glu Glu Glu Leu Lys Val Pro Leu Glu Glu Tyr Val His	
195	200 205
aaa cgc tac ccc ggg ctg gtg aag gtg gta aga aat cag aag agg gaa	672
Lys Arg Tyr Pro Gly Leu Val Lys Val Val Arg Asn Gln Lys Arg Glu	
210	215 220
ggc ctg atc cgc gct cgc att gag ggc tgg aag gtg gct acc ggg cag	720
Gly Leu Ile Arg Ala Arg Ile Glu Gly Trp Lys Val Ala Thr Gly Gln	
225	230 235 240
gtc act ggc ttc ttt gat gcc cac gtg gaa ttc acc gct ggc tgg gct	768
Val Thr Gly Phe Phe Asp Ala His Val Glu Phe Thr Ala Gly Trp Ala	
245	250 255
gag ccg gtt cta tcc cgc atc cag gaa aac cgg aag cgt gtg atc ctc	816
Glu Pro Val Leu Ser Arg Ile Gln Glu Asn Arg Lys Arg Val Ile Leu	
260	265 270
ccc tcc att gac aac atc aaa cag gac aac ttt gag gtg cag cgg tac	864
Pro Ser Ile Asp Asn Ile Lys Gln Asp Asn Phe Glu Val Gln Arg Tyr	
275	280 285
gag aac tcg gcc cac ggg tac agc tgg gag ctg tgg tgc atg tac atc	912
Glu Asn Ser Ala His Gly Tyr Ser Trp Glu Leu Trp Cys Met Tyr Ile	
290	295 300
agc ccc cca aaa gac tgg tgg gac gcc gga gac cct tct ctc ccc atc	960
Ser Pro Pro Lys Asp Trp Trp Asp Ala Gly Asp Pro Ser Leu Pro Ile	
305	310 315 320
agg acc cca gcc atg ata ggc tgc tcg ttc gtg gtc aac agg aag ttc	1008
Arg Thr Pro Ala Met Ile Gly Cys Ser Phe Val Val Asn Arg Lys Phe	
325	330 335

ttc ggt gaa att ggt ctt ctg gat cct ggc atg gat gta tac gga gga	1056
Phe Gly Glu Ile Gly Leu Leu Asp Pro Gly Met Asp Val Tyr Gly Gly	
340	345
350	
gaa aat att gaa ctg gga atc aag gta tgg ctc tgt ggg ggc agc atg	1104
Glu Asn Ile Glu Leu Gly Ile Lys Val Trp Leu Cys Gly Gly Ser Met	
355	360
365	
gag gtc ctt cct tgc tca cgg gtg gcc cac att gag cgg aag aag aag	1152
Glu Val Leu Pro Cys Ser Arg Val Ala His Ile Glu Arg Lys Lys Lys	
370	375
380	
cca tat aat agc aac att ggc ttc tac acc aag agg aat gct ctt cgc	1200
Pro Tyr Asn Ser Asn Ile Gly Phe Tyr Thr Lys Arg Asn Ala Leu Arg	
385	390
395	400
gtt gct gag gtc tgg atg gac gat tac aag tct cat gtg tac ata gcg	1248
Val Ala Glu Val Trp Met Asp Asp Tyr Lys Ser His Val Tyr Ile Ala	
405	410
415	
tgg aac ctg ccg ctg gag aat ccg gga att gac atc ggt gat gtc tcc	1296
Trp Asn Leu Pro Leu Glu Asn Pro Gly Ile Asp Ile Gly Asp Val Ser	
420	425
430	
gaa aga aga gca tta agg aaa agt tta aag tgt aag aat ttc cag tgg	1344
Glu Arg Arg Ala Leu Arg Lys Ser Leu Lys Cys Lys Asn Phe Gln Trp	
435	440
445	
tac ctg gac cat gtt tac cca gaa atg aga aga tac aat aat acc gtt	1392
Tyr Leu Asp His Val Tyr Pro Glu Met Arg Arg Tyr Asn Asn Thr Val	
450	455
460	
gct tac ggg gag ctt cgc aac aac aag gca aaa gac gtc tgc ttg gac	1440
Ala Tyr Gly Glu Leu Arg Asn Asn Lys Ala Lys Asp Val Cys Leu Asp	
465	470
475	480

cag ggg ccg ctg gag aac cac aca gca ata ttg tat ccg tgc cat ggc 1488
 Gln Gly Pro Leu Glu Asn His Thr Ala Ile Leu Tyr Pro Cys His Gly
 485 490 495

tgg gga cca cag ctt gcc cgc tac acc aag gaa ggc ttc ctg cac ttg 1536
 Trp Gly Pro Gln Leu Ala Arg Tyr Thr Lys Glu Gly Phe Leu His Leu
 500 505 510

ggt gcc ctg ggg acc acc aca ctc ctc cct gac acc cgc tgc ctg gtg 1584
 Gly Ala Leu Gly Thr Thr Thr Leu Leu Pro Asp Thr Arg Cys Leu Val
 515 520 525

gac aac tcc aag agt cgg ctg ccc cag ctc ctg gac tgc gac aag gtc 1632
 Asp Asn Ser Lys Ser Arg Leu Pro Gln Leu Leu Asp Cys Asp Lys Val
 530 535 540

aag agc agc ctg tac aag cgc tgg aac ttc atc cag aat gga gcc atc 1680
 Lys Ser Ser Leu Tyr Lys Arg Trp Asn Phe Ile Gln Asn Gly Ala Ile
 545 550 555 560

atg aac aag ggc acg gga cgc tgc ctg gag gtg gag aac cgg ggc ctg 1728
 Met Asn Lys Gly Thr Gly Arg Cys Leu Glu Val Glu Asn Arg Gly Leu
 565 570 575

gct ggc atc gac ctc atc ctc cgc agc tgc aca ggt cag agg tgg acc 1776
 Ala Gly Ile Asp Leu Ile Leu Arg Ser Cys Thr Gly Gln Arg Trp Thr
 580 585 590

att aag aac tcc atc aag tag 1797
 Ile Lys Asn Ser Ile Lys
 595

<210> 14

<211> 1920

<212> DNA

<213> Homo sapiens

<400> 14

atg ctc cta agg aag cga tac agg cac aga cca tgc aga ctc cag ttc 48
 Met Leu Leu Arg Lys Arg Tyr Arg His Arg Pro Cys Arg Leu Gln Phe
 1 5 10 15

ctc ctg ctg ctc ctg atg ctg gga tgc gtc ctg atg atg gtg gcg atg 96
 Leu Leu Leu Leu Leu Met Leu Gly Cys Val Leu Met Met Val Ala Met
 20 25 30

TTG CAC CCT CCC CAC CAC ACC CTG CAC CAG ACT GTC ACA GCC CAA GCC 144
 Leu His Pro Pro His His Thr Leu His Gln Thr Val Thr Ala Gln Ala
 35 40 45

agc aag cac agc cct gaa gcc agg tac cgc ctg gac ttt ggg gaa tcc 192
 Ser Lys His Ser Pro Glu Ala Arg Tyr Arg Leu Asp Phe Gly Glu Ser
 50 55 60

cag gat igg gta ctg gaa gct gag gat gag ggt gaa gag tac agc cct 240
 Gln Asp Trp Val Leu Glu Ala Glu Asp Glu Gly Glu Glu Tyr Ser Pro
 65 70 75 80

ctg gag ggc ctg cca ccc ttt atc tca ctg egg gag gat cag ctg ctg 288
 Leu Glu Gly Leu Pro Pro Phe Ile Ser Leu Arg Glu Asp Gln Leu Leu
 85 90 95

gtg gcc gtg gcc tta ccc cag gcc aga agg aac cag agc cag ggc agg 336
 Val Ala Val Ala Leu Pro Gln Ala Arg Arg Asn Gln Ser Gln Gly Arg
 100 105 110

aga ggt ggg agc tac cgc ctc atc aag cag cca agg agg cag gat aag 384
 Arg Gly Gly Ser Tyr Arg Leu Ile Lys Gln Pro Arg Arg Gln Asp Lys
 115 120 125

gaa gcc cca aag agg gac tgg ggg gct gat gag gac ggg gag gtg tct 432
 Glu Ala Pro Lys Arg Asp Trp Gly Ala Asp Glu Asp Gly Glu Val Ser

130	135	140	
gaa gaa gag gag ttg acc ccg ttc agc ctg gac cca cgt ggc ctc cag			480
Glu Glu Glu Glu Leu Thr Pro Phe Ser Leu Asp Pro Arg Gly Leu Gln			
145	150	155	160
gag gca ctc agt gcc cgc atc ccc ctc cag agg gct ctg ccc gag gtc			528
Glu Ala Leu Ser Ala Arg Ile Pro Leu Gln Arg Ala Leu Pro Glu Val			
	165	170	175
cgg cac cca ctg tgt ctg cag cag cac cct cag gac agc ctg ccc aca			576
Arg His Pro Leu Cys Leu Gln Gln His Pro Gln Asp Ser Leu Pro Thr			
	180	185	190
gcc agc gtc atc ctc tgt ttc cat gat gag gcc tgg tcc act ctc ctg			624
Ala Ser Val Ile Leu Cys Phe His Asp Glu Ala Trp Ser Thr Leu Leu			
	195	200	205
cgg act gta cac agc atc ctc gac aca gtg ccc agg gcc ttc ctg aag			672
Arg Thr Val His Ser Ile Leu Asp Thr Val Pro Arg Ala Phe Leu Lys			
	210	215	220
gag atc atc ctc gtg gac gac ctc agc cag caa gga caa ctc aag tct			720
Glu Ile Ile Leu Val Asp Asp Leu Ser Gln Gln Gly Gln Leu Lys Ser			
225	230	235	240
gct ctc agc gaa tat gtg gcc agg ctg gag ggg gtg aag tta ctc agg			768
Ala Leu Ser Glu Tyr Val Ala Arg Leu Glu Gly Val Lys Leu Leu Arg			
	245	250	255
agc aac aag agg ctg ggt gcc atc agg gcc cgg atg ctg ggg gcc acc			816
Ser Asn Lys Arg Leu Gly Ala Ile Arg Ala Arg Met Leu Gly Ala Thr			
	260	265	270
aga gcc acc ggg gat gtg ctc gtc ttc atg gat gcc cac tgc gag tgc			864
Arg Ala Thr Gly Asp Val Leu Val Phe Met Asp Ala His Cys Glu Cys			

275	280	285	
cac cca ggc tgg ctg gag ccc ctc ctc agc aga ata gct ggt gac agg			912
His Pro Gly Trp Leu Glu Pro Leu Leu Ser Arg Ile Ala Gly Asp Arg			
290	295	300	
agc cga gtg gta tct ccg gtg ata gat gtg att gac tgg aag act ttc			960
Ser Arg Val Val Ser Pro Val Ile Asp Val Ile Asp Trp Lys Thr Phe			
305	310	315	320
cag iat tac ccc tca aag gac ctg cag cgt ggg gtg ttg gac tgg aag			1008
Gln Tyr Tyr Pro Ser Lys Asp Leu Gln Arg Gly Val Leu Asp Trp Lys			
	325	330	335
ctg gat ttc cac tgg gaa cct ttg cca gag cat gtg agg aag gcc ctc			1056
Leu Asp Phe His Trp Glu Pro Leu Pro Glu His Val Arg Lys Ala Leu			
	340	345	350
cag tcc ccc ata agc ccc atc agg agc cct gtg gtg ccc gga gag gtg			1104
Gln Ser Pro Ile Ser Pro Ile Arg Ser Pro Val Val Pro Gly Glu Val			
	355	360	365
gtg gcc atg gac aga cat tac itc caa aac act gga gcg tat gac tct			1152
Val Ala Met Asp Arg His Tyr Phe Gln Asn Thr Gly Ala Tyr Asp Ser			
	370	375	380
ctt atg tcg ctg cga ggt ggt gaa aac ctc gaa ctg tct ttc aag gcc			1200
Leu Met Ser Leu Arg Gly Gly Glu Asn Leu Glu Leu Ser Phe Lys Ala			
385	390	395	400
tgg ctc tgt ggt ggc tct gtt gaa atc ctt ccc tgc tct cgg gta gga			1248
Trp Leu Cys Gly Gly Ser Val Glu Ile Leu Pro Cys Ser Arg Val Gly			
	405	410	415
cac atc tac caa aat cag gat tcc cat tcc ccc ctc gac cag gag gcc			1296
His Ile Tyr Gln Asn Gln Asp Ser His Ser Pro Leu Asp Gln Glu Ala			

420	425	430	
acc ctg agg aac agg gtt cgc att gct gag acc tgg ctg ggg tca ttc			1344
Thr Leu Arg Asn Arg Val Arg Ile Ala Glu Thr Trp Leu Gly Ser Phe			
435	440	445	
aaa gaa acc ttc tac aag cat agc cca gag gcc ttc tcc ttg agc aag			1392
Lys Glu Thr Phe Tyr Lys His Ser Pro Glu Ala Phe Ser Leu Ser Lys			
450	455	460	
gct gag aag cca gac tgc atg gaa cgc ttg cag ctg caa agg aga ctg			1440
Ala Glu Lys Pro Asp Cys Met Glu Arg Leu Gln Leu Gln Arg Arg Leu			
465	470	475	480
ggt tgt cgg aca ttc cac tgg ttt ctg gct aat gtc tac cct gag ctg			1488
Gly Cys Arg Thr Phe His Trp Phe Leu Ala Asn Val Tyr Pro Glu Leu			
485	490	495	
tac cca tct gaa ccc agg ccc agt ttc tct gga aag ctc cac aac act			1536
Tyr Pro Ser Glu Pro Arg Pro Ser Phe Ser Gly Lys Leu His Asn Thr			
500	505	510	
gga ctt ggg ctc tgt gca gac tgc cag gca gaa ggg gac atc ctg ggc			1584
Gly Leu Gly Leu Cys Ala Asp Cys Gln Ala Glu Gly Asp Ile Leu Gly			
515	520	525	
igt ccc atg gtg ttg gct cct tgc agt gac agc cgg cag caa cag tac			1632
Cys Pro Met Val Leu Ala Pro Cys Ser Asp Ser Arg Gln Gln Gln Tyr			
530	535	540	
ctg cag cac acc agc agg aag gag att cac ttt ggc agc cca cag cac			1680
Leu Gln His Thr Ser Arg Lys Glu Ile His Phe Gly Ser Pro Gln His			
545	550	555	560
ctg tgc ttt gct gtc agg cag gag cag gtg att ctt cag aac tgc acg			1728
Leu Cys Phe Ala Val Arg Gln Glu Gln Val Ile Leu Gln Asn Cys Thr			

565

570

575

gag gaa ggc ctg gcc atc cac cag cag cac tgg gac ttc cag gag aat 1776

Glu Glu Gly Leu Ala Ile His Gln Gln His Trp Asp Phe Gln Glu Asn

580

585

590

ggg atg att gtc cac att ctt tct ggg aaa tgc atg gaa gct gtg gtg 1824

Gly Met Ile Val His Ile Leu Ser Gly Lys Cys Met Glu Ala Val Val

595

600

605

caa gaa aac aat aaa gat ttg tac ctg cgt ccg tgt gat gga aaa gcc 1872

Gln Glu Asn Asn Lys Asp Leu Tyr Leu Arg Pro Cys Asp Gly Lys Ala

610

615

620

cgc cag cag tgg cgt ttt gac cag atc aat gct gtg gat gaa cga tga 1920

Arg Gln Gln Trp Arg Phe Asp Gln Ile Asn Ala Val Asp Glu Arg

625

630

635

<210> 15

<211> 71

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 15

ggggacaagt ttgtacaaaa aagcaggctt cgaaggagat agaaccatga ggaagatccg 60

cgccaatgcc a

71

<210> 16

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 16

ggggacaagt ttgtacaaaa aagcaggctt ccggggcgcg cagagggcag gcagga 56

<210> 17

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 17

ggggaccact ttgtacaaga aagetgggtc ccacagaatcc tcgctctgaa gggc 54

<210> 18

<211> 70

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 18

ggggacaagt ttgtacaaaa aagcaggctt cgaaggagat agaaccatgc ggcgctgac 60

tcgteggctg

70

<210> 19

<211> 59

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 19

ggggacaagt ttgtacaaaa aagcaggctt cgtgcagacc cctaagcctt cggacgctg 59

<210> 20

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 20

ggggaccact ttgtacaaga aagctgggtc ctaagagctc accatgtccc agtgctg 57

<210> 21

<211> 70

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 21

ggggacaagt ttgtacaaaa aagcaggctt cgaaggagat agaacatga ggcggaagga 60

gaagcggctc

70

<210> 22

<211> 58

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 22

ggggacaagt ttgtacaaaa aagcaggcct cccgcagccc gacggcaccc ciggggga 58

<210> 23

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 23

ggggaccact ttgtacaaga aagctgggtc tcagttccta ttgaatttt ccaa 54

<210> 24

<211> 70

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 24

ggggacaagt ttgtacaaaa aagcaggcct cgaaggagat agaaccatgt gggggcgcac 60

ggcgcggcgg

70

<210> 25

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 25

ggggacaagt ttgtacaaaa aagcaggcct cggcgcgag cgggtcatgc cgcggc 56

<210> 26

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 26

ggggaccact ttgtacaaga aagctgggtc tcataacatg cgcctttga agaacca 57

<210> 27

<211> 73

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 27

ggggacaagt ttgtacaaaa aagcaggcct cgaaggagat agaaccatgg tgtgcaccag 60

gaagaccaaa act

<210> 28

<211> 59

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 28

ggggacaagt ttgtacaaaa aagcaggctt ccgggggcag gagccggcgc ccgacaag 58

<210> 29

<211> 55

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 29

ggggaccact ttgtacaaga aagctgggtc icaggacgcg aggctcctca ggacg 55

<210> 30

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 30

ggggacaagt ttgtacaaaa aagcaggctt catggcttca ctgagaagag tcaaag 56

<210> 31

<211> 55

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 31

ggggacaagt ttgtacaaaa aagcaggctt ccgcagcgga gacgccttc acgag 55

<210> 32

<211> 58

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 32

ggggaccact ttgtacaaga aagctgggtc ctacttgatg gaggcttaa tgggccac 58

<210> 33

<211> 58

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 33

ggggacaagt ttgtacaaaa aagcaggctt catgctccta aggaagcgat acaggcac 58

<210> 34

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 34

ggggacaagi ttgtacaaaa aagcaggcct cgccagggtac cgccctggact ttgggg 56

<210> 35

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 35

ggggaccact ttgtacaaga aagctggggtc ctcatcgttc atccacagca ttgatct 57

<210> 36

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 36

ccggctcatc gaatgccgct acgcc 25

<210> 37
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> primer for PCR amplification

<400> 37
agagcattcc tcttgggtga gaagcc 26

<210> 38
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> primer for PCR amplification

<400> 38
ttccgctcaa tgtgggccac ccgtg 25

<210> 39
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer for PCR amplification

<400> 39
ggtctccggc gtcccaccag 20

<210> 40
<211> 22
<212> PRT
<213> Artificial Sequence

<220>
<223> Igk sequence for secreting the expressed protein

<400> 40
Met His Phe Gln Val Gln Ile Phe Ser Phe Leu Leu Ile Ser Ala
1 5 10 15
Ser Val Ile Met Ser Arg Gly
20 22

<210> 41
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> FLAG tag sequence for purifying the expressed protein

<400> 41
Asp Tyr Lys Asp Asp Asp Asp Lys
1 5 8

<210> 42
<211> 94
<212> DNA
<213> Artificial Sequence

<220>
<223> template DNA for PCR amplification

<400> 42

gatcatgcat tttcaagigc agattttcag ctctctgcta atcagtgcc t cagtcataat 60

gtcacgtgga gattacaagg acgacgatga caag

94

<210> 43

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 43

cgggatccat gcattttcaa gtgcag 26

<210> 44

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for PCR amplification

<400> 44

ggaattcttg tcatcgtcgt ccttg 25

<210> 45

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FITC at N-terminus position

for assay of GalNAc-T enzyme assay

<400> 45

Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg
 1 5 10 11

<210> 46

<211> 24

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FITC at C-terminus position
 for assay of GalNAc-T enzyme assay

<400> 46

Pro Thr Thr Thr Pro Ile Thr Thr Thr Thr Thr Val Thr Pro Thr
 1 5 10 15
 Pro Thr Pro Thr Gly Thr Gln Thr Lys
 20 24

<210> 47

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FITC at N-terminus position
 for assay of GalNAc-T enzyme assay

<400> 47

Gly Thr Thr Pro Ser Pro Val Pro Thr Thr Ser Thr Thr Ser Ala
 1 5 10 15

<210> 48

<211> 23

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FITC at C-terminus position
for assay of GalNAc-T enzyme assay

<400> 48

Pro Thr Pro Ser Ala Thr Thr Pro Ala Pro Pro Ser Ser Ser Ala
1 5 10 15
Pro Pro Glu Thr Thr Ala Ala Lys
20 23

<210> 49

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FITC at C-terminus position
for assay of GalNAc-T enzyme assay

<400> 49

Pro Thr Thr Asp Ser Thr Thr Pro Ala Pro Thr Thr Lys
1 5 10 13

<210> 50

<211> 20

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FITC at C-terminus position
for assay of GalNAc-T enzyme assay

<400> 50

Val Pro Ser Thr Pro Pro Thr Pro Ser Pro Ser Thr Pro Pro Thr
1 5 10 15
Pro Ser Pro Ser Lys
20

<210> 51

<211> 19

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FITC at N-terminus position
for assay of GalNAc-T enzyme assay

<400> 51

Gly Pro Thr Glu Val Ala Gln Thr Pro Thr Pro Glu Thr Phe Leu
1 5 10 15
Thr Thr Ile Lys
19

<210> 52

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FAM (5-carboxyfluorescein
succinimidyl ester) at N-terminus position for assay of GalNAc-T
enzyme assay

<400> 52

Ala His Gly Val Thr Ser Ala Pro Asp Thr Arg
1 5 10 11

<210> 53

<211> 24

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FAM (5-carboxyfluorescein succinimidyl ester) at C-terminus position for assay of GalNAc-T enzyme assay

<400> 53

Pro Thr Thr Thr Pro Ile Thr Thr Thr Thr Thr Val Thr Pro Thr
1 5 10 15
Pro Thr Pro Thr Gly Thr Gln Thr Lys
20 24

<210> 54

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FAM (5-carboxyfluorescein succinimidyl ester) at N-terminus position for assay of GalNAc-T enzyme assay

<400> 54

Gly Thr Thr Pro Ser Pro Val Pro Thr Thr Ser Thr Thr Ser Ala
1 5 10 15

<210> 55

<211> 23

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FAM (5-carboxyfluorescein succinimidyl ester) at C-terminus position for assay of GalNAc-T

enzyme assay

<400> 55

Pro Thr Pro Ser Ala Thr Thr Pro Ala Pro Pro Ser Ser Ser Ala

1 5 10 15

Pro Pro Glu Thr Thr Ala Ala Lys

20 23

<210> 56

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FAM (5-carboxyfluorescein succinimidyl ester) at C-terminus position for assay of GalNac-T enzyme assay

<400> 56

Pro Thr Thr Asp Ser Thr Thr Pro Ala Pro Thr Thr Lys

1 5 10 13

<210> 57

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 57

cgtgcccgct tggaac 16

<210> 58

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 58

tgttctggag catcccgaag 20

<210> 59

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> probe for real time PCR

<400> 59

tgggatgtg acagagagga agcagct 27

<210> 60

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 60

cgcgtaggaa atggagaaca a 21

<210> 61

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 61

tcggtaigcc tgaiccactc t 21

<210> 62

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> probe for real time PCR

<400> 62

agaccittacc ccatgaccga tgctga 26

<210> 63

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 63

gtcggctggt tctgccagtc ttcg 24

<210> 64

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 64

tggaaggtcc gtgcaataca ccag 24

<210> 65

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 65

tgagatgcac actgctggtg tattg 25

<210> 66

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 66

gggtgcggtt taatacactg cg 22

<210> 67

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> prove for real time PCR

<400> 67

aacgaggccc gctccacgct gct 23

<210> 68

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 68

agcgtcatca tcaccttcca 20

<210> 69

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 69

ttcggtcag gacactcttc ac 22

<210> 70

<211> 18

<212> DNA
<213> Artificial Sequence

<220>
<223> probe for real time PCR

<400> 70
ccgttcacc ctgcgcgc 18

<210> 71
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> primer for real time PCR

<400> 71
ggatgtgctc gtcttcatgg a 21

<210> 72
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> primer for real time PCR

<400> 72
gataccactc ggctcctgctc a 21

<210> 73
<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> primer for real time PCR

<400> 73

ctggagcccc tectcagcag aatagc 26

<210> 74

<211> 17

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic oligopeptide labeled with FAM (5-carboxyfluorescein succinimidyl ester) at C-terminus position for assay of GalNAc-T enzyme assay

<400> 74

Ala Asn Thr Pro Ser Phe Pro Thr Ala Thr Ser Pro Ala Pro Pro

1 5 10 15

Ile Lys

17

【図面の簡単な説明】

図1は、F I T CラベルしたE A 2サンプルをH P L C分析で保持時間20.27分のピークとして分取し、凍結乾燥後、M A L D I - T O F質量分析装置で分析した結果を示すチャートである。

図2は、E A 2を受容体基質として、既知のU D P - N - アセチル - D - ガラクトサミン：ポリペプチドN - アセチルガラクトサミン転移酵素（G a l N A c - T 1 0）と反応させて得られた反応液をH P L C分析し、受容体基質E A 2ピーク（保持時間20.27分）と0.72分差を示す単一の産物ピーク（保持時間19.55分）を分取し、凍結乾燥後、M A L D I - T O F質量分析装置で分析した結果を示すチャートである。

図3は、F A MラベルされたM u c 1 aペプチドを基質としたG a l N A c - T 1 4の反応の経時変化の追跡を示すH P L Cチャートである。

図4は、F A MラベルされたM u c 5 A Cペプチドを基質としたG a l N A c - T 1 4の反応の経時変化の追跡を示すH P L Cチャートである。

図5は、種々のF A Mラベルされた受容体基質を用いたG a l N A c - T 1 4の一次G a l N A c - T活性の測定における、反応前と反応後のH P L Cチャートである：Sは反応前の基質ピークを表し、Pは生成物ピークを示す。

図6は、一箇所G a l N A c化を受けた種々のF A Mラベルされた受容体基質を用いたG a l N A c - T 1 4の二次G a l N A c - T活性の測定における、反応前と反応後のH P

LCチャートである：Sは反応前の基質ピークを表し、Pは生成物ピークを示す。

図7は、リアルタイムPCRによる種々のヒト組織におけるGalNAc-T14転写物の定量解析の結果を示すグラフである。

図8は、リアルタイムPCRによる種々のヒト細胞系列におけるGalNAc-T14転写物の定量解析の結果を示すグラフである。

図9は、種々のFAMラベルされた受容体基質を用いたGalNAc-T13の反応前と反応後のHPLCチャートである：Sは反応前の基質ピークを表し、Pは生成物ピークを示す。

図10は、FAMラベルされたMuc1aペプチドを基質としたGalNAc-T13（パネルA）及びGalNAc-T6（パネルB）の反応の経時変化の追跡を示すHPLCチャートである。

10

図11は、FAMラベルされたMuc1aペプチドを基質としたGalNAc-T13（パネルA）及びGalNAc-T6（パネルB）の反応の経時変化の追跡を示すHPLCチャートである。

図12は、EA2ペプチド（S）及びGalNAc-T13（P1ないしP5）及び-T6（P1'ないしP2'）によって生じた反応生成物のHPLCチャート（パネルA）、及び単離した各ピークに対するMALDI-TOF質量分析スペクトル（パネルB）を示す：ジ、ジ-糖ペプチド；トリ、トリ-糖ペプチド；テトラ、テトラ-糖ペプチド。

図13は、EA2糖ペプチド基質及びGalNAc-T13によって生じた反応生成物のHPLCチャートである。S"は、GalNAc-T13の二次活性アッセイ用の糖ペプチド基質として使用されるモノGalNAc-EA2に対応する。P1"、P2"及びP3"は、GalNAc-EA2を用いてGalNAc-T13によって生じた生成物に対応する。

20

図14は、リアルタイムPCRによる種々のヒト組織におけるGalNAc-T13転写物の定量解析の結果を示すグラフである。

図15は、リアルタイムPCRによる種々のヒト細胞系列におけるGalNAc-T13転写物の定量解析の結果を示すグラフである。

図16は、GalNAc-T12転写物のAlternatively Spliced Isoformの同定について示す。

図17は、FAMラベルされた受容体基質（Muc5AC（パネルA）、Muc2（パネルB）およびMuc13-58（パネルC））を用いたGalNAc-T12の一次GalNAc-T活性の測定における、反応前と反応後のHPLCチャートである。各基質の残存量が、I：50%、II：20%、およびIII：0%である場合のHPLCチャートを示す。

30

図18は、リアルタイムPCRによる種々のヒト組織（パネルA）および細胞系列（パネルB）におけるGalNAc-T12転写物の定量解析を示すグラフである。

図19は、種々のFAMラベルされた受容体基質を用いたGalNAc-T11の一次GalNAc-T活性の測定における、反応前と反応後のHPLCチャートである。

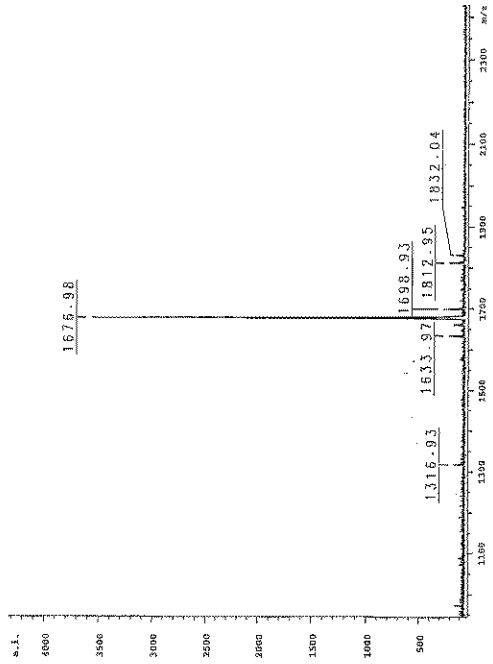
図20は、リアルタイムPCRによる種々のヒト組織におけるGalNAc-T11転写物の定量解析を示すグラフである。

40

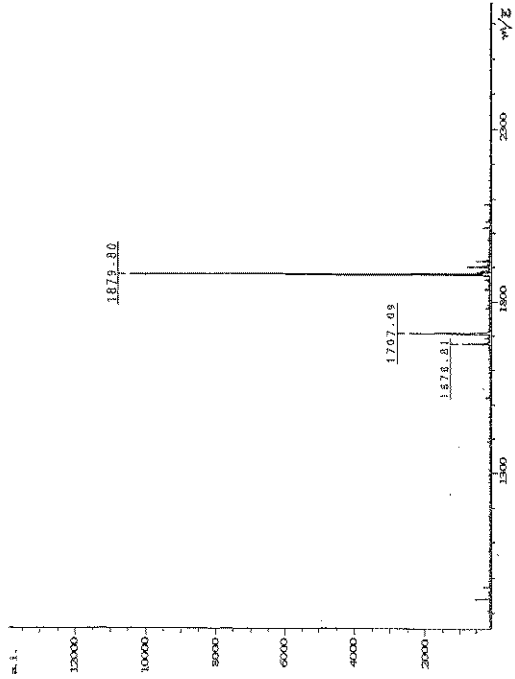
図21は、種々のFAMラベルされた受容体基質を用いたGalNAc-T17の一次GalNAc-T活性の測定における、反応前と反応後のHPLCチャートである。

図22は、リアルタイムPCRによる種々のヒト組織（パネルA）および細胞系列（パネルB）におけるGalNAc-T17転写物の定量解析を示すグラフである。

【 図 1 】
図 1

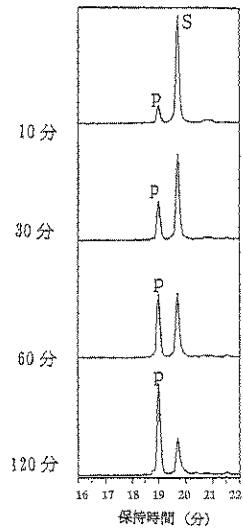


【 図 2 】
図 2



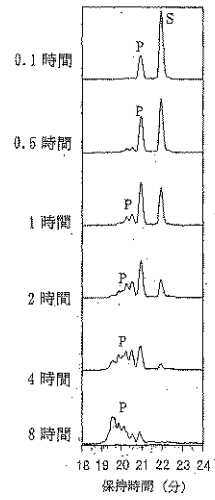
【 図 3 】
図 3

AHGVTSAPDTR (Muc1a)

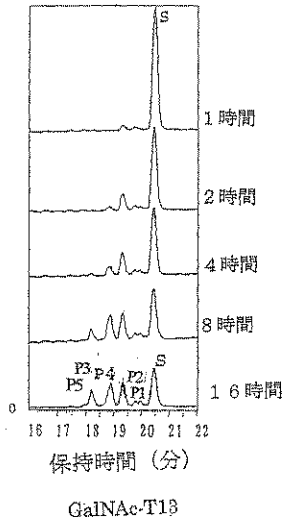


【 図 4 】
図 4

GTTPSPVPTTSTSA
(Muc5AC)



【 図 1 1 】
図 1 1



【 図 1 2 】

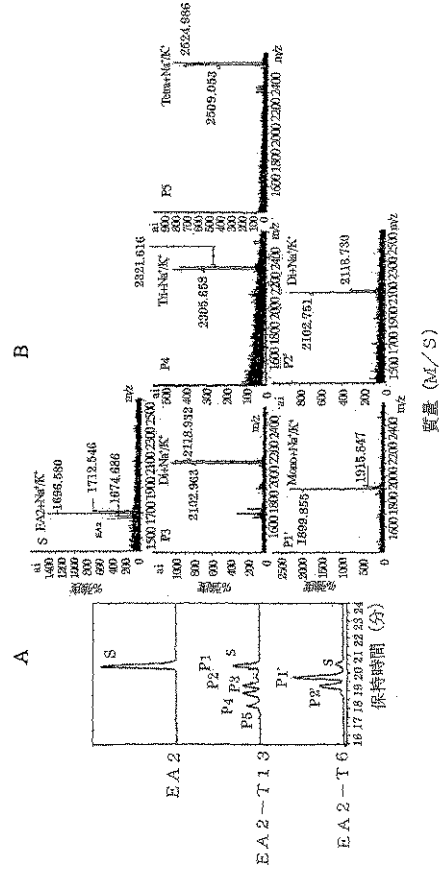
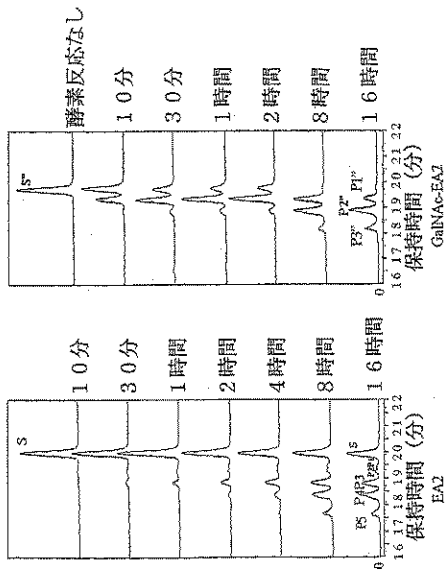


図 1 2

【 図 1 3 】
図 1 3



【 図 1 4 】

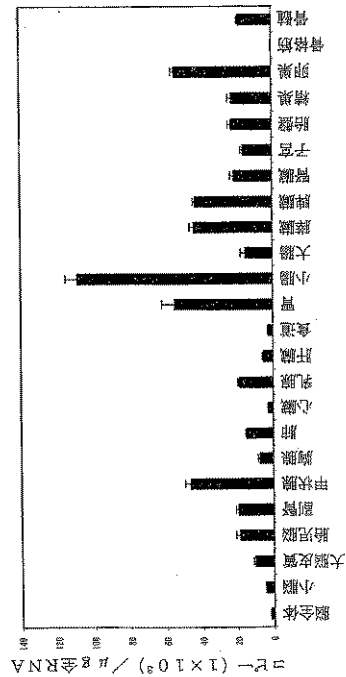


図 1 4

【 図 1 5 】

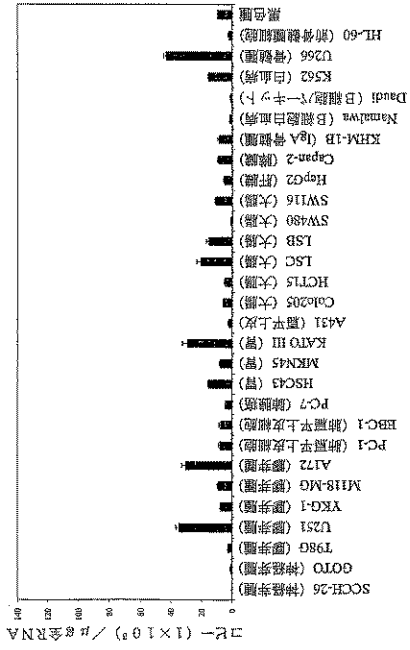


図 1 5

【 図 1 6 】

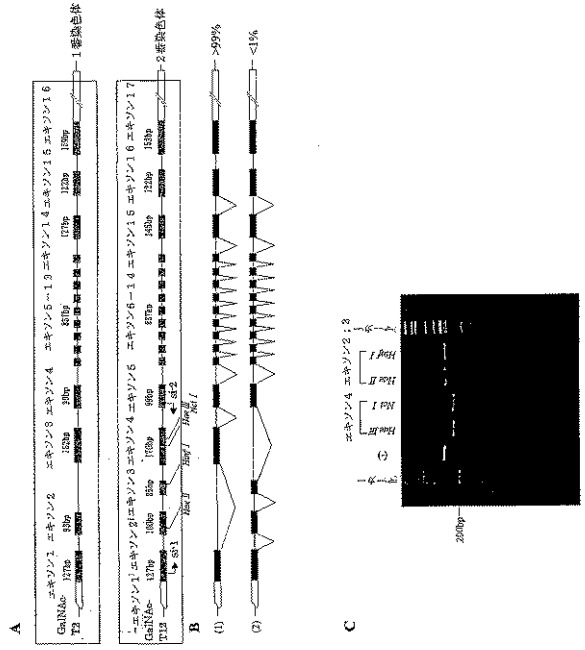
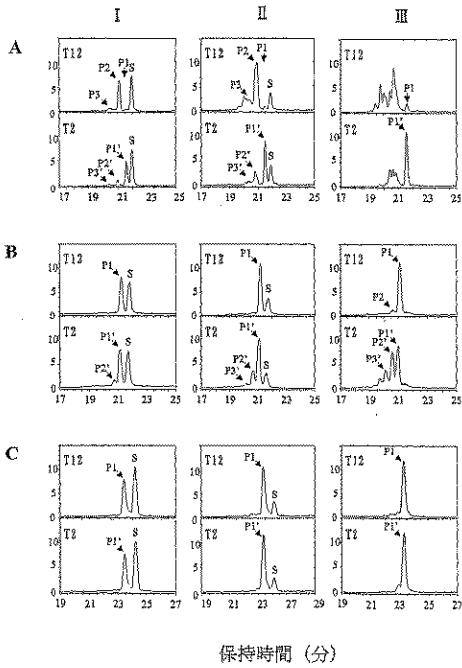


図 1 6

【 図 1 7 】

図 1 7



【 図 1 8 】

図 1 8

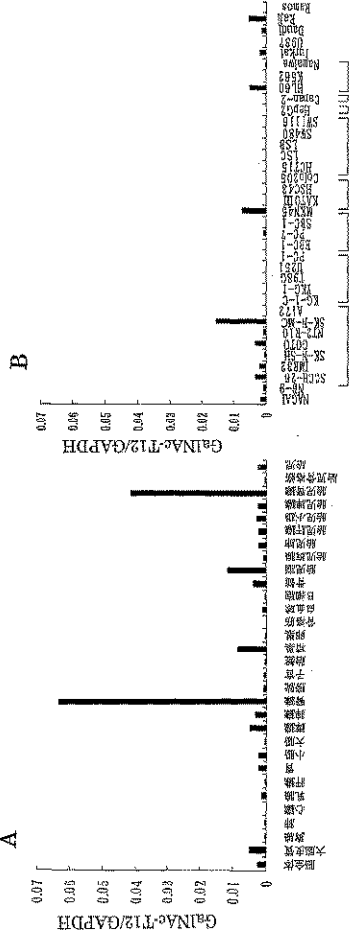


図 1 8

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/JP03/00008
A. CLASSIFICATION OF SUBJECT MATTER Int.Cl. ⁷ C12N15/54, C12N9/10, C12N5/00, C12Q1/68, G01N33/50, A61K38/43, A61K48/00, A61P35/00		
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) Int.Cl. ⁷ C12N15/54, C12N9/10, C12N5/00, C12Q1/68, G01N33/50, A61K38/43, A61K48/00, A61P35/00		
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 practicable, search terms used) SwissProt/PIR/GeneSeq, WPI (DIALOG), BIOSIS (DIALOG), MEDLINE (STN)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	WO 02/46426 A2 (Incyte Genomics Inc.), 13 June, 2002 (13.06.02), Particularly, SEQ ID NO.6 & AU 2002-239570 A	1-4, 6, 7, 9-17, 19
X	GB 2288401 A (Bay Development Corp.), 18 October, 1995 (18.10.95), Particularly, Fig. 4 (Family: none)	1-17, 19
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family 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 04 April, 2003 (04.04.03)		Date of mailing of the international search report 30 April, 2003 (30.04.03)
Name and mailing address of the ISA/ Japanese Patent Office		Authorized officer
Facsimile No.		Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/00008

Box I Observations where certain claims were found unsearchable (Continuation of item 2 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.: 18
because they relate to subject matter not required to be searched by this Authority, namely:
The invention as set forth in this claim pertains to methods for treatment of the human body by therapy.
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 3 of first sheet)This International Searching Authority found multiple inventions in this international application, as follows:
(See extra sheet)

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.:
In the inventions as set forth in claims 1 to 17 and 19, those relating to SEQ ID NOS:1 and 8.
- Remark on Protest The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/00008

Continuation of Box No.II of continuation of first sheet(1)

Although the relationships between the claimed proteins having the amino acid sequences represented by SEQ ID NOS: 1 to 7 and having an activity of transferring N-acetylgalactosamine (GalNAc) via an α 1-bond to a hydroxyl group of serine or threonine and the nucleic acids encoding the same having the base sequences represented by SEQ ID NOS: 8 to 14 respectively fulfill the requirement of unity of invention, there is no common chemical structure among the individual proteins or the individual nucleic acids. Namely, they are common to each other exclusively in "having an activity of transferring N-acetylgalactosamine (GalNAc) via an α 1-bond to a hydroxyl group of serine or threonine".

As stated in the description (p. 1, see, for example, the following documents 1 and 2), however, there had been already known 8 types of human enzymes having an activity of transferring N-acetylgalactosamine (GalNAc) via an α 1-bond to a hydroxyl group of serine or threonine before the application of the present case. Thus, it cannot be considered as a special technical feature in the meaning as described in PCT Rule 13.2 that the claimed proteins and nucleic acids have the GalNAc transferase activity.

According to PCT Rule 13.3, the determination of unity of invention shall be made without regard to whether the inventions are claimed in separate claims or as alternatives within a single claim.

Therefore, the inventions relating to the proteins represented by SEQ ID NOS: 1 to 7, from among the claimed inventions, cannot be considered as relating to a group of inventions so linked as to form a single general inventive concept. That is to say, it is recognized that there are groups having 7 inventions relating to 7 proteins or nucleic acids encoding the same which are different from each other, though the requirement of unity of invention is fulfilled concerning the relationship between a single protein and a nucleic acid encoding the same.

Document 1: J. Biol. Chem., 1996, Vol.271, pp.17006-17012

Document 2: J. Biol. Chem., 1995, Vol.270, pp.24116-24126

国際調査報告		国際出願番号 PCT/JPO3/00008
A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int. Cl ⁷ C12N15/54, C12N9/10, C12N5/10, C12Q1/68, G01N33/50, A61K38/43, A61K48/00, A61P35/00		
B. 調査を行った分野 調査を行った最小限資料 (国際特許分類 (IPC)) Int. Cl ⁷ C12N15/54, C12N9/10, C12N5/10, C12Q1/68, G01N33/50, A61K38/43, A61K48/00, A61P35/00		
最小限資料以外の資料で調査を行った分野に含まれるもの		
国際調査で使用了電子データベース (データベースの名称、調査に使用した用語) SwissProt/PIR/Geneseq, WPI(DIALOG)、BIOSIS(DIALOG)、MEDLINE(STN)		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
P X	WO 02/46426 A2 (INCYTE GENOMICS INC) 2002.06.13 & AU 2002-239570 A 特にSEQ ID NO:6参照	1-4, 6, 7, 9-17, 19
X	GB 2288401 A (Bay Development Corp.) 1995.10.18 (ファミリーなし) 特にFig.4参照	1-17, 19
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの 「L」 優先権主張に疑義を提起する文献又は他の文献の発行人若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」 口頭による開示、使用、展示等に言及する文献 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願日の後に公表された文献 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」 同一パテントファミリー文献		
国際調査を完了した日 04.04.03	国際調査報告の発送日 30.04.03	
国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 田村 明 照	4B 8412
	電話番号 03-3581-1101 内線 3448	

国際調査報告

国際出願番号 PCT/JP03/00008

第I欄 請求の範囲の一部の調査ができないときの意見 (第1ページの2の続き)

法第8条第3項 (PCT17条(2)(a)) の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

1. 請求の範囲 18 は、この国際調査機関が調査をすることを要しない対象に係るものである。つまり、
人体の治療方法に係る発明が記載されている。
2. 請求の範囲 _____ は、有意義な国際調査をすることができる程度まで所定の要件を満たしていない国際出願の部分に係るものである。つまり、
3. 請求の範囲 _____ は、従属請求の範囲であってPCT規則6.4(a)の第2文及び第3文の規定に従って記載されていない。

第II欄 発明の単一性が欠如しているときの意見 (第1ページの3の続き)

次に述べるようにこの国際出願に二以上の発明があるとこの国際調査機関は認めた。

特別ページ参照

1. 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求の範囲について作成した。
2. 追加調査手数料を要求するまでもなく、すべての調査可能な請求の範囲について調査することができたので、追加調査手数料の納付を求めなかった。
3. 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったため、この国際調査報告は、手数料の納付のあった次の請求の範囲のみについて作成した。
4. 出願人が必要な追加調査手数料を期間内に納付しなかったため、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求の範囲について作成した。

請求の範囲 1-17、19に記載された発明のうち、配列番号1及び8に係る発明

追加調査手数料の異議の申立てに関する注意

- 追加調査手数料の納付と共に出願人から異議申立てがあった。
 追加調査手数料の納付と共に出願人から異議申立てがなかった。

国際調査報告

国際出願番号 PCT/JP03/00008

(第II欄の続き)

請求の範囲に記載された配列番号1-7で示されるアミノ酸配列を有するセリン、スレオニン残基の水酸基にN-アセチルガラクトサミン (GalNAc) を α 1結合で転移する活性を有するタンパク質と、それをコードした配列番号8-14に示される塩基配列を有する核酸との関係においてはそれぞれ発明の単一性を満たすものの、個々のタンパク質或いは個々の核酸の間には共通の化学構造を有するものでなく、「セリン、スレオニン残基の水酸基にN-アセチルガラクトサミン (GalNAc) を α 1結合で転移する活性を有する」ことにおいてのみ共通する。

しかしながら、明細書第1頁(例えば下記文献1、2)に記載されているように、セリン、スレオニン残基の水酸基にN-アセチルガラクトサミン (GalNAc) を α 1結合で転移する活性を有するヒトの酵素は出願前に既に8種類が知られている。したがって、請求の範囲に記載されたタンパク質及び核酸がGalNAc転移酵素活性を有することはPCT規則13.2における特別な技術的特徴であるとはいえない。

ここで、PCT規則13.3によると、発明の単一性の判断はこれらの発明が別個の請求の範囲に記載されているか単一の請求の範囲に択一的な形式によって記載されているかを考慮することなく行われるべきものである。

よって、請求の範囲に記載された発明のうち配列番号1-7で表されるタンパク質に関する発明は、単一の一般的発明概念を形成するように連関している一群の発明であるとはいえず、1つのタンパク質とそれをコードした核酸との関係においてはそれぞれ発明の単一性を満たすものの、異なった7種のタンパク質或いはそれをコードした核酸それぞれに関する7個の発明からなる発明群であると認める。

文献1 : J. Biol. Chem. , 1996, Vol. 271, pp. 17006-17012

文献2 : J. Biol. Chem. , 1995, Vol. 270, pp. 24116-24126

フロントページの続き

(51)Int.Cl. ⁷		F I		
C 1 2 N	1/19	C 1 2 N	1/21	
C 1 2 N	1/21	C 1 2 N	9/10	
C 1 2 N	5/10	C 1 2 Q	1/68	A
C 1 2 N	9/10	G 0 1 N	33/53	M
C 1 2 Q	1/68	G 0 1 N	37/00	1 0 2
G 0 1 N	33/53	C 1 2 N	5/00	A
G 0 1 N	37/00	C 1 2 N	15/00	F
		A 6 1 K	37/02	

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW, ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES, FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,N O,NZ,OM,PH,PL,PT,RO,RU,SC,SD,SE,SG,SI,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,YU,ZA,ZM,ZW

(74)代理人 100075270

弁理士 小林 泰

(74)代理人 100080137

弁理士 千葉 昭男

(74)代理人 100096013

弁理士 富田 博行

(74)代理人 100107386

弁理士 泉谷 玲子

(72)発明者 成松 久

茨城県つくば市東 1 - 1 - 1 独立行政法人産業技術総合研究所つくばセンター第二事業所内

(72)発明者 張 延

茨城県つくば市東 1 - 1 - 1 独立行政法人産業技術総合研究所つくばセンター第二事業所内

(72)発明者 後藤 雅式

茨城県つくば市東 1 - 1 - 1 独立行政法人産業技術総合研究所つくばセンター第二事業所内

(注) この公表は、国際事務局(WIPO)により国際公開された公報を基に作成したものである。なおこの公表に係る日本語特許出願(日本語実用新案登録出願)の国際公開の効果は、特許法第184条の10第1項(実用新案法第48条の13第2項)により生ずるものであり、本掲載とは関係ありません。

专利名称(译)	新型UDP-N-乙酰基-D-半乳糖胺：多肽N-乙酰半乳糖胺基转移酶及其编码核酸		
公开(公告)号	JPWO2003057887A1	公开(公告)日	2005-05-19
申请号	JP2003558181	申请日	2003-01-06
[标]申请(专利权)人(译)	Amersham公司生物科学 富士瑞必欧株式会社		
申请(专利权)人(译)	先进工业科学和技术研究院 Amersham公司生物科学公司 FUJIREBIO		
[标]发明人	成松久 張延 後藤雅式		
发明人	成松久 張延 後藤雅式		
IPC分类号	C12N15/09 A61K38/00 A61P35/00 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N9/10 C12N15/54 C12Q1/68 G01N33/53 G01N37/00		
CPC分类号	C12N9/1003		
FI分类号	C12N15/00.ZNA.A A61P35/00 C12M1/00.A C12N1/15 C12N1/19 C12N1/21 C12N9/10 C12Q1/68.A G01N33/53.M G01N37/00.102 C12N5/00.A C12N15/00.F A61K37/02		
代理人(译)	小林 泰 千叶昭夫		
优先权	2001401507 2001-12-28 JP 2002163832 2002-06-05 JP 2002203696 2002-07-12 JP		
外部链接	Espacenet		

摘要(译)

分离具有通过 α 1键将N-乙酰基半乳糖胺转移至核心蛋白或肽序列的丝氨酸和苏氨酸残基的羟基的活性的酶，并且通过阐明基因的结构，可以进行酶等的基因工程。基于该酶的GalNAc修饰的肽或GalNAc修饰的蛋白的生产，生产作为底物，鉴定基于该酶添加了未知的O-连接的糖链的蛋白质或未知的肽的O-连接的糖链修饰位置 并且能够基于基因等诊断疾病。取代或缺失序列表SEQ ID NO：1-7所示氨基酸序列中的一个或多个氨基酸或氨基酸序列，或在氨基酸序列中插入或添加一个或多个氨基酸 使用具有氨基酸序列并且具有通过 α 1键将N-乙酰基半乳糖胺 (GalNAc) 转移至核心蛋白质或肽的丝氨酸或苏氨酸残基的羟基作为底物的活性的蛋白质和编码该蛋白质的核酸。