

(19)日本国特許庁 (J P)

(12)特許公報 (B 1)

(11)特許番号

特許第3012928号

(P 3 0 1 2 9 2 8)

(45)発行日 平成12年 2月28日(2000.2.28)

(24)登録日 平成11年12月17日(1999.12.17)

(51)Int.Cl. ⁷	識別記号	F I
C12N 15/09	ZNA	C12N 15/00
// C12N 9/50		9/50
(C12N 15/09	ZNA	
C12R 1:91)		

請求項の数 2 (全10頁)

(21)出願番号	特願平10 - 327537	(73)特許権者	591031360 農林水産省食品総合研究所長 茨城県つくば市観音台 2 丁目 1 - 2
(22)出願日	平成10年11月 4 日(1998.11.4)	(72)発明者	荒平 正緒美 茨城県つくば市並木 4 - 10 - 1 903棟3 02号
審査請求日	平成10年11月 4 日(1998.11.4)	(72)発明者	深澤 親房 茨城県つくば市小野川 9 - 35
		(74)代理人	100074077 弁理士 久保田 藤郎 (外 1 名)
		審査官	内田 俊生

最終頁に続く

(54)【発明の名称】植物由来アスパラギン残基特異的エンドプロテアーゼ c D N A および遺伝子

1

(57)【特許請求の範囲】

【請求項 1】 以下の (a) 又は (b) の植物由来のアスパラギン残基特異的エンドプロテアーゼをコードする c D N A。

(a) 配列表の配列番号 1 記載のアミノ酸配列からなる植物由来のアスパラギン残基特異的エンドプロテアーゼ

(b) 配列表の配列番号 1 記載のアミノ酸配列に対し 1 もしくは数個のアミノ酸残基の欠失、置換または付加がなされたアミノ酸配列からなり、かつ植物由来のアスパラギン残基特異的エンドプロテアーゼ活性を有するタンパク質

【請求項 2】 以下の (a) 又は (b) の植物由来のアスパラギン残基特異的エンドプロテアーゼをコードする遺伝子。

(a) 配列表の配列番号 2 記載のアミノ酸配列からなる

2

植物由来のアスパラギン残基特異的エンドプロテアーゼ (b) 配列表の配列番号 2 記載のアミノ酸配列に対し 1 もしくは数個のアミノ酸残基の欠失、置換または付加がなされたアミノ酸配列からなり、かつ植物由来のアスパラギン残基特異的エンドプロテアーゼ活性を有するタンパク質

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】本発明は、植物由来のアスパラギン残基特異的エンドプロテアーゼ c D N A および遺伝子に関するものである。

【 0 0 0 2 】

【従来の技術】アスパラギン残基特異的エンドプロテアーゼとは、ペプチドまたはタンパク質のアミノ酸配列のうち、アスパラギン残基の C - 末端でペプチド等を切断

10

する酵素である。この中でも、特に植物由来のアスパラギン残基特異的エンドプロテアーゼをレグマチュレインと称されている。これまでに、レグマチュレインは種々の高等植物において酵素活性が確認されているが、非常に不安定なため、本発明者らの他には高度に精製した例がない。

【 0 0 0 3 】植物の貯蔵タンパク質は、当初プレプロタンパク質として合成され、レグマチュレインの作用によりはじめて成熟タンパク質の形となる。この植物の貯蔵タンパク質は、安全性の点で心配がない上に、優れた蛋白資源であることから、食品工業界では大量生産が要望されている。したがって、この貯蔵タンパク質の合成に必要な純度の高いレグマチュレインの量産も強く望まれている。しかし、前述した理由から、レグマチュレインの精製は極めて困難であり、未だ要望には応えられていない状況である。

【 0 0 0 4 】

【発明が解決しようとする課題】レグマチュレインの cDNA および遺伝子が取得され、これらを利用することができれば、本酵素の大量生産への道が開かれる可能性がある。また、部位特異的変異法を利用して本酵素の安定性の増大や酵素活性の改良等にも利用できる。したがって、本発明の目的は、レグマチュレインの遺伝的情報を得て、本酵素の量産化を可能とすることである。

【 0 0 0 5 】

【課題を解決するための手段】本発明者らは、レグマチュレインを高度に精製し、そのアミノ酸配列を基にして、レグマチュレインの cDNA および遺伝子のクローニングを試みた。その結果、その cDNA および遺伝子の全構造を解明することに成功し、本発明に到達した。

【 0 0 0 6 】すなわち、高度に精製されたレグマチュレインのアミノ酸配列の解析から、縮重のある合成オリゴヌクレオチドを作成した。次に、ダイズ登熟初期の種子より mRNA を抽出し、それに相補的な cDNA を合成した。作成した合成オリゴヌクレオチドと cDNA との間で PCR を行い、レグマチュレインの部分的な cDNA をクローニングした。

【 0 0 0 7 】一方、同じダイズ登熟初期の種子の mRNA を用いて、2本鎖 cDNA を合成し、ファージベクターにアダプターを介して組み込み、cDNA ライブラリーを構築した。この cDNA ライブラリーより、PCR 法により得た部分的なレグマチュレイン cDNA をラベルし、これをプローブとしてスクリーニングを行って、完全長のレグマチュレイン cDNA をクローニングした。この cDNA の一次構造解析により、シグナルペプチドの配列を含むレグマチュレインの全アミノ酸配列が決定された。

【 0 0 0 8 】cDNA の配列を基にしたオリゴヌクレオチドプライマーを合成し、発芽したダイズより各 DNA を抽出・精製し、これを鋳型として PCR 法を行い、レ

グマチュレイン遺伝子をクローニングした。

【 0 0 0 9 】すなわち、請求項 1 記載の本発明は、以下の (a) 又は (b) の植物由来のアスパラギン残基特異的エンドプロテアーゼをコードする cDNA である。

(a) 配列表の配列番号 1 記載のアミノ酸配列からなる植物由来のアスパラギン残基特異的エンドプロテアーゼ (b) 配列表の配列番号 1 記載のアミノ酸配列に対し 1 もしくは数個のアミノ酸残基の欠失、置換または付加がなされたアミノ酸配列からなり、かつ植物由来のアスパラギン残基特異的エンドプロテアーゼ活性を有するタンパク質

10

20

20

(a) 配列表の配列番号 2 記載のアミノ酸配列からなる植物由来のアスパラギン残基特異的エンドプロテアーゼ (b) 配列表の配列番号 2 記載のアミノ酸配列に対し 1 もしくは数個のアミノ酸残基の欠失、置換または付加がなされたアミノ酸配列からなり、かつ植物由来のアスパラギン残基特異的エンドプロテアーゼ活性を有するタンパク質

【 0 0 1 0 】

【発明の実施の形態】本発明について以下に詳しく説明する。請求項 1 記載の本発明は、以下の (a) 又は (b) の植物由来のアスパラギン残基特異的エンドプロテアーゼをコードする cDNA である。

(a) 配列表の配列番号 1 記載のアミノ酸配列からなる植物由来のアスパラギン残基特異的エンドプロテアーゼ (b) 配列表の配列番号 1 記載のアミノ酸配列に対し 1 もしくは数個のアミノ酸残基の欠失、置換または付加がなされたアミノ酸配列からなり、かつ植物由来のアスパラギン残基特異的エンドプロテアーゼ活性を有するタンパク質

30

40

【 0 0 1 1 】本発明の対象とする植物由来のアスパラギン残基特異的エンドプロテアーゼ、すなわちレグマチュレインは、上述したように、ペプチドやタンパク質のアミノ酸配列のうち、アスパラギン残基の C - 末端側でペプチド等を切断する作用を有している酵素である。

【 0 0 1 2 】本発明者らは、以下のようにして請求項 1 記載の cDNA の取得に成功した。

40

① まず、植物中に含まれるレグマチュレインを抽出、精製した。レグマチュレインは各種の植物、特に種子中に含まれる貯蔵タンパク質に含まれている。例えば、ダイズの種子からレグマチュレインを抽出する場合は、登熟初期から完熟期まで、特に登熟初期から登熟後期までのものを用いることが好ましい。

【 0 0 1 3 】ここで、レグマチュレインのみを得るためには、材料である植物を高度に精製する必要がある。1 例を示すと、登熟初期から完熟期までの植物種子を粉碎し、これを適当な緩衝液で抽出し、その可溶性画分を透析後、固液分離する。得られる粗酵素液を疎水クロマト

50

グラフィー、ゲルろ過等にかけることにより、ほぼ純粋なレグマチュレインまで精製することが可能である。

【0014】② 得られた精製レグマチュレインのアミノ酸配列のうち、N - 末端のアミノ酸配列を決定した。アミノ酸配列の決定は、Edman 分解により行うことができ、自動アミノ酸シーケンサーを用いると簡便に行うことができる。

【0015】③ 一方、上記②とは別に、精製レグマチュレインの部分アミノ酸配列を得た。まず、レグマチュレインを短いペプチドに酵素分解し、それぞれのペプチドサンプルを高速液体クロマトグラフィー等により分画した上で、各サンプルのアミノ酸配列、すなわちレグマチュレインの内部のアミノ酸配列を決定した。アミノ酸配列の決定は、上記②と同様の方法により行うことができる。各アミノ酸配列のうち、縮重のあるオリゴヌクレオチド2つを選択し、後述のPCR反応でプライマーとして用いた(配列表の配列番号3および4参照)。

【0016】④ 次に、植物の種子よりRNAを抽出し、これに相補的なcDNAを合成した。植物からのRNA抽出は、Fukazawa C. et al., Journal of Biological Chemistry, 200, 6234-6239 (1985)に従い、SDS - フェノール法を用いて行うことができる。抽出した全RNAからPoly(A) + RNAを調製し、これにオリゴ(dT)をプライマーとした逆転写酵素による逆転写反応を行って、1本鎖cDNAを得た。

【0017】⑤ この1本鎖cDNAを鋳型として、先に作成した合成オリゴヌクレオチド(配列表の配列番号3および4参照)との間でPCRを行い、レグマチュレインの部分cDNA(400bp)をクローニングした。得られたバンドをベクターにサブクローニングし、先に決定したレグマチュレインの内部のアミノ酸配列を含んでいるかどうか解析した。その結果、該バンドは、レグマチュレインの部分cDNAであることが確認された。

【0018】⑥ 次に、同じ植物種子のmRNAを用いて2本鎖cDNAを合成し、ファージベクターにアダプターを介して組み込み、cDNAライブラリーを構築した。cDNAライブラリーの作成には、岡山 - Berg法、ガブラー - ホフマン(Gubler-Hoffman)法を採用できるが、簡略な点から後者の方が好ましい。以下、ガブラー - ホフマン法を採用した場合について述べる。

【0019】まず、先述のpoly(A) + RNAから2本鎖のcDNAを合成する。このcDNAに、EcoRI、NotI、BamHI等の制限酵素部位を含むアダプターをライゲーションし、制限酵素によりcDNAを切断する。ポリアクリルアミドゲル電気泳動法により500bp以上の長さのものだけをゲルから切り出して集めると共に、アダプターを除去する。その後、得られたヌクレオチドの5'位にリン酸基を導入し、制限酵素で切断し、脱リン酸化してあるgt10ファージベク

ターアーム(宝酒造社製)にライゲーションする。これをファージにパッケージングした。こうして、cDNAライブラリーを作成した。

【0020】⑦ 上記⑤のPCR法により得た部分的なレグマチュレインcDNAをラベルし、これをプローブとして、⑥のcDNAライブラリーからスクリーニングを行った。次いで、完全長のレグマチュレインcDNAをクローニングした。スクリーニングは、先に得たレグマチュレイン部分cDNAを放射標識等したものをプローブとして、約10万プラークのcDNAライブラリーにプラークハイブリダイゼーションすることにより行うことができる。その結果、約10万プラークのうち、数個のプラークが陽性を示したので、これを分離した。

【0021】分離した陽性プラークを純化し、大腸菌に感染させファージを増やした後に、ファージ粒子を精製した。ファージDNAの精製は、CsClステップワイズ密度勾配を用いた超遠心法等により行うことができる。精製したファージDNAから制限酵素でインサートを切り出し、アガロースゲル等で精製した。これをプラスミドベクターにサブクローニングした後、DNAシーケンスを行った。

【0022】その結果、配列表の配列番号1記載のアミノ酸配列を有するcDNAが得られた。このcDNAが、請求項1記載の本発明のcDNAである。該配列はレグマチュレインの全長を含んでいる。また、EMBL、GenBank等のデータベースと比較したところ、全く新規な配列であることが明らかとなった。

【0023】請求項1記載の本発明のcDNAは、全長1055bpからなり、うち14bpがpoly Aであった。アミノ酸数は、開始メチオニンからの全体で241残基、シグナルペプチドは21残基であり、成熟型タンパク質のN - 末端配列からクローニングされたcDNAから推定されるアミノ酸のC - 末端まで220残基であった。C - 末端解析から得られた結果およびイオンスプレームススペクトロメトリーで得られた分子量の結果から、レグマチュレインは、C - 末端側でプロセスを受けるプレプロ型のタンパク質であることが明らかとなった。

【0024】この配列中、レグマチュレイン活性に関与する糖鎖の結合部位は、配列表の配列番号1記載のアミノ酸配列のうち、71~78番目の部分であった。この部分の構造は、N - 末端側から、塩基性アミノ酸 - 疎水性アミノ酸 - アスパラギン(糖鎖) - 疎水性アミノ酸 - スレオニン/セリン - 疎水性アミノ酸 - 酸性アミノ酸 - システインとなっている。この糖鎖結合部位は、アスパラギン残基のC - 末端側でアミノ酸を特異的に切断する活性を有するレグマチュレインの安定化に関与すると考えられる。この糖鎖が欠失すると、酵素活性の急速な低下がみられる。

【0025】次に、請求項2記載の本発明について説明

する。請求項 2 記載の本発明は、請求項 1 記載の本発明の cDNA の一次構造解析により得られるレグマチュレインの全遺伝子である。まず、配列表の配列番号 1 記載の cDNA の塩基配列をもとに、25mer と 26mer の 2 つのオリゴヌクレオチドプライマー（配列表の配列番号 5、6 参照）を合成した。

【0026】一方、ダイズの芽生えから合成した DNA を調製した。調製は、CTAB 法 [Fukazawa C. et al., Journal of Biological Chemistry, 200, 6234-6239 (1985)] 等により行うことができる。この DNA を鋳型にして配列表の配列番号 5、6 記載のプライマーを用いて PCR 法を行った。

【0027】その結果、約 3.1bp の単一バンドが得られた。このバンドを TA ベクター [pCR2.1] (Invitrogen 社製) にサブクローニング後、上記の方法で DNA シークエンスを行った。その結果、配列表の配列番号 2 に示す塩基配列が得られた。

【0028】請求項 2 記載の本発明の遺伝子の全長（配列表の配列番号 2 参照）は 3077bp で、4 つのイントロンにより 5 つのエクソンに分断されている。各イントロンの塩基数は、第 1 イントロンよりそれぞれ 67、705、1113、200bp であった。一方、各エクソンの塩基数は、第 1 エクソンよりそれぞれ 321、66、141、93、420bp であった（配列表の配列番号 2 参照）。前記 cDNA に存在した活性に関する糖鎖結合部位は、第 1 エクソン上に存在していた（配列表の配列番号 2 記載のアミノ酸配列の 71～78 番目参照）。

【0029】請求項 2 記載の本発明は、以下の (a) 又は (b) の植物由来のアスパラギン残基特異的エンドプロテアーゼをコードする遺伝子である。

(a) 配列表の配列番号 2 記載のアミノ酸配列からなる植物由来のアスパラギン残基特異的エンドプロテアーゼ

(b) 配列表の配列番号 2 記載のアミノ酸配列に対し 1 もしくは数個のアミノ酸残基の欠失、置換または付加がなされたアミノ酸配列からなり、かつ植物由来のアスパラギン残基特異的エンドプロテアーゼ活性を有するタンパク質

【0030】このようにして得られた請求項 1 記載の本発明の cDNA および請求項 2 記載の本発明の遺伝子を利用することによって、アスパラギン残基特異的エンドプロテアーゼ（レグマチュレイン）を大量に、しかも効率よく生産することが可能となる。

【0031】なお、配列表の配列番号 1 記載のアミノ酸配列に対して 1 もしくは数個のアミノ酸残基の欠失、置換または付加がなされたアミノ酸配列を有するものであっても、同様の作用効果を奏するものは、本発明に含まれる。同じく、配列表の配列番号 2 記載のアミノ酸配列に対して 1 もしくは数個のアミノ酸残基の欠失、置換または付加がなされたアミノ酸配列を有するものであ

ても、同様の作用効果を奏するものは、本発明に含まれる。

【0032】

【実施例】以下、実施例により本発明を詳細に説明するが、本発明はこれらにより何ら制限を受けるものではない。

実施例 1 [レグマチュレインのアミノ酸配列の決定および PCR 法を用いた部分 cDNA のクローニング]

(1) 鋳型の作成

10 開花後 13 日目のダイズの種子より、Fukazawa C. et al., Journal of Biological Chemistry, 200, 6234-6239 (1985) に従い、SDS-フェノール法を用いて全 RNA を抽出した。抽出した全 RNA から Poly A T r a c t キット（プロメガ社製）により、Poly (A) + RNA を調製した。

【0033】調製した Poly (A) + RNA 1 μg につき、オリゴ(dT)をプライマーとし、逆転写酵素 RAV-2（宝酒造社製）100U を用いて 42、60 分間逆転写反応を行い、1本鎖の cDNA を得た。残存する RNA を分解するため、終濃度 0.5N-NaOH で 25、一晩アルカリ分解した。アルカリ分解後、終濃度 0.5M の HEPES 緩衝液 pH 7.2 で中和後、セファデックス G-50（アマシャム-ファルマシア社製）を充填したゲルろ過カラム（0.8cm × 27cm）を通し、分解した RNA を除去した。こうして得られる 1本鎖の cDNA を以下の PCR に用いた。

【0034】(2) プライマーの作成

高度に生成したレグマチュレインの N-末端アミノ酸配列を、気相式アミノ酸シーケンサー（パーキンエルマー社製 477A 型）により決定した。一方、Arahira M. and Fukazawa C., Plant Molecular Biology, 25, 597-605 (1994) の方法に従い、V8 プロテアーゼ（宝酒造社製）およびリジルエンドペプチダーゼ（和光純薬社製）により、レグマチュレインを分解した。

【0035】分解後、それぞれのサンプルを逆相カラム（東ソー社製、Silica ODS 120T 4.6mm × 150mm）を接続した高速液体クロマトグラフィー（島津社製 LC-6AD）を用いて、0.1% TFA - 0.1% TFA / 60% アセトニトリルの直線濃度勾配をかけた溶媒系で分画した。得られたペプチドを、N-末端のアミノ酸配列決定と同様に、気相式アミノ酸シーケンサー（パーキンエルマー社製、477A 型）で決定し、内部のアミノ酸配列とした。このアミノ酸配列のうち、PCR のプライマーとして適する配列を検討し、20mer と 32mer の 2 つの縮重のあるプライマーを合成した（配列表の配列番号 3 および 4 参照）。

【0036】(3) PCR によるレグマチュレインの部分 cDNA クローニング

50 上記(2)で合成したオリゴヌクレオチドプライマーを

用いて、前記(1)で合成した一本鎖cDNAを鋳型としてPCR反応を行った。PCRの条件としては、TaqポリメラーゼはTakara ExTaq(宝酒造社製)を1反応当たり2.5U用いた。また、プライマーは終濃度0.4μM、dNTPmixは0.2mM、鋳型となる一本鎖cDNAは10ng使用した。遺伝子増幅装置は、Takara PCR Thermal Cycler 480(宝酒造社製)を用い、変性温度95、30秒、アニーリング温度56、30秒、伸長反応72、1分30秒、35サイクルの条件で行った。

【0037】その結果、約400bpの単一のバンドが得られたので、これをTAベクター[pCR2.1](Invitrogen社製)にサブクローニングした。クローン化した大腸菌から、プラスミドDNAを常法[Maniatis T. et al., "Molecular cloning" Cold Spring Harbor Labo. (1982)]に従い調製した。

【0038】調製したプラスミドについて、島津社製、DNAシーケンサーDSQ1000を用いた蛍光オートシーケンスを行った。その結果、サブクローニングしたバンドは、前記(2)で決定したレグマチュレインの内部のアミノ酸配列を含んでいることが明らかとなった。このことから、クローニングしたバンドは、レグマチュレインの部分cDNAであることが確認された。

【0039】実施例2〔登熟初期のcDNAライブラリーの作成〕

実施例1の(1)で得られたpoly(A)+RNA 2.5μgを、ガブラー-ホフマン(Gubler-Hoffman)法の原理に基づいたcDNA合成キット(アマシャム-ファルマシア社製)を用いて、 $[^{32}\text{P}]$ -dCTPを加えて合成をモニターしながら2本鎖DNAを合成した。

【0040】1nmolのEcoRI-NotI-BamHIアダプター(宝酒造社製)をニッポンジーン社製、ライゲーションパックでライゲーションした。このcDNAをFukazawa C. et al., Journal of Biological Chemistry, 200, 6234-6239(1985)の方法に従い、ポリアクリルアミドゲル電気泳動法により、500bp以上の長さのものをゲルから切り出して集めると共に、アダプターを除去した。アダプターには、5'位にリン酸がないので、T4ポリヌクレオチドキナーゼ(ニッポンジーン社製)を用いてリン酸基を導入した。

【0041】その後、EcoRIで切断し、脱リン酸化してあるgt10ファージベクターアーム(宝酒造社製)に、上記と同様に、ニッポンジーン社製、ライゲーションパックでライゲーションし、*in vitro* パッケージングキット(アマシャム-ファルマシア社製)でラムダファージにパッケージングを行った。このようにして、開花後13日目のダイズ登熟種子cDNAライブラリーを作成した。

【0042】実施例3〔完全長レグマチュレインcDNAのクローニング〕

実施例1で得たレグマチュレイン部分cDNAを、PCR法を用いて $[^{32}\text{P}]$ -dCTPで、Arahira M. and Fukazawa C., Plant Molecular Biology, 25, 597-605 (1994)の方法に従いラベルした。これをプローブとして、実施例2で作成した開花後13日目のダイズの登熟種子ライブラリーから、Arahira M. and Fukazawa C., Plant Molecular Biology, 25, 597-605 (1994)の方法に従って、ブランクハイブリダイゼーションを実施した。その結果、cDNA約10万ブランクから、数個の陽性ブランクが分離できた。

【0043】陽性ブランクを純化し、大腸菌に感染させてファージを増やした後に、ファージ粒子をCsClステップウイズ密度勾配を用いた超遠心法により精製し、ファージDNAを精製した。精製したファージDNAから、BamHIでインサートを切り出し、アガロースゲルから精製した後に、pUC19プラスミドベクターにサブクローニングした。その後、実施例1と同様の方法で、DNAシーケンスを行った。解析の結果、得られたcDNAは、レグマチュレインの全長を含んでおり、配列表の配列番号1に示す塩基配列およびアミノ酸配列を有していた。このcDNAは、EMBL、GenBank等のデータベースとの比較により、全く新規な配列であることが明らかとなった。

【0044】本cDNAの全長は、1055bp、うち14bpがpoly Aであった。アミノ酸数は、開始メチオニンからの全体で241残基、シグナルペプチドは、21残基であり、成熟型タンパク質のN-末端配列からクローニングされたcDNAから推定されるアミノ酸のC-末端まで220残基であった。C-末端解析から得られた結果およびイオンスプレームススペクトロメトリーで得られた分子量の結果から、レグマチュレインは、C-末端側でプロセスを受けるプレプロ型のタンパク質であることが明らかとなった。

【0045】また、活性に関与する糖鎖の結合部位は1カ所(配列表の配列番号1のアミノ酸配列の71~78番目)であった。結合部位の配列は、塩基性アミノ酸(アルギニン)-疎水性アミノ酸(アラニン)-アスパラギン(糖鎖)-疎水性アミノ酸(アラニン)-スレオニン/セリン-疎水性アミノ酸(フェニルアラニン)-酸性アミノ酸(アスパラギン酸)-システインであった。

【0046】実施例4〔レグマチュレイン遺伝子のクローニング〕

実施例3で得られたcDNAの塩基配列を基に、25merと26merのオリゴヌクレオチドプライマーを合成した(配列表の配列番号5および6参照)。ダイズの芽生えからCTAB法[Fukazawa C. et al., Nucleic Acid Research, 8117-8117 (1987)]により核DNAを調製した。このDNA10ngを鋳型として、実施例1と同じ条件でPCR法を行った。

【 0 0 4 7 】 その結果、約 3 . 1 k b p の単一バンドが得られた。T A ベクター [p C R 2 . 1] (Invitrogen社製) にサブクローニングした後、実施例 3 に記載した方法で DNA シークエンスを行った。解析の結果、クローニングした遺伝子の塩基配列を配列表の配列番号 2 に示す通りであった。

【 0 0 4 8 】 該遺伝子の全長は 3 0 7 7 b p で、4 つのイントロンにより 5 つのエクソンに分断されていた。各イントロンの塩基数は、第 1 イントロンよりそれぞれ、6 7、7 0 5、1 1 1 3、2 0 0 b p であった。一方、各エクソンの塩基数は、第 1 エクソンよりそれぞれ 3 2 1、6 6、1 4 1、9 3、4 2 0 b p であった (配列表

の配列番号 2 参照)。なお、活性に関与する糖鎖結合部位は、第 1 エクソン上に存在していた。

【 0 0 4 9 】

【発明の効果】本発明の c D N A および遺伝子を利用することにより、不安定なために精製が困難と考えられていたアスパラギン残基特異的エンドプロテアーゼ (レグマチュレイン) の大量、かつ効率的な生産が可能となる。また、部位特異的変異法を利用してレグマチュレインの安定性の増大や酵素活性の改良等に寄与することができる。

【配列表】

SEQUENCE LISTING

```

<110> Director of National Food Research Institute, Ministry of Agriculture, Forestry and Fisheries
<120> cDNA and genomic sequence of asparagin-bond specific endoprotease
<130> P100997K
<160> 6
<210> 1
<211> 1055
<212> DNA
<213> Glycine max
<220>
<221> CDS
<222> (85)...(810)
<400> 1
ttgggtgtag ctgttacgag tcagcaagag agaaatatct tgttgagttg atttctgatc 60
ccaaattgag tttgagtttc agcc atg gta tct gtt tct ctg ctc tta ttc 111
Met Val Ser Val Ser Leu Leu Leu Phe
1 5
cta tct ttc ttt gcc ttc ttt gca ccc tct gaa tct cac gac aaa gtc 159
Leu Ser Phe Phe Ala Phe Phe Ala Pro Ser Glu Ser His Asp Lys Val
10 15 20 25
tcg ttg gaa ttg tac tac gaa agc tta tgc ccc tac agc gcc aac ttc 207
Ser Leu Glu Leu Tyr Tyr Glu Ser Leu Cys Pro Tyr Ser Ala Asn Phe
30 35 40
atc gtg aat cac ctc cca aaa atc ttc acc cca gat ctt gcc cca atc 255
Ile Val Asn His Leu Pro Lys Ile Phe Thr Pro Asp Leu Ala Pro Ile
45 50 55
gtt cac ctc aaa ctc gtt cct tgg ggc aat gcc aaa ctc cgt gcc aac 303
Val His Leu Lys Leu Val Pro Trp Gly Asn Ala Lys Leu Arg Ala Asn
60 65 70
gcc acc ttt gac tgc cag cat ggg cca tat gag tgc ttg ctg aac aca 351
Ala Thr Phe Asp Cys Gln His Gly Pro Tyr Glu Cys Leu Leu Asn Thr
75 80 85
gtt gaa gcc tgt gca att cac atc tgg oct caa ctc agc aaa cat ttt 399
Val Glu Ala Cys Ala Ile His Ile Trp Pro Gln Leu Ser Lys His Phe
90 95 100 105
cct ttc atc tac tgt gtt gag gat ctg gtg ttt cag agt aag cgt gag 447
Pro Phe Ile Tyr Cys Val Glu Asp Leu Val Phe Gln Ser Lys Arg Glu

```

13
110 115 120
gaa tgg gaa tct tgt ttt gag aaa ctg gat ctt gat tcg gaa cct att 495
Glu Trp Glu Ser Cys Phe Glu Lys Leu Asp Leu Asp Ser Glu Pro Ile
125 130 135
aat cag tgt tat aat agt gaa cat gga aaa cag ttg gag cta caa tat 543
Asn Gln Cys Tyr Asn Ser Glu His Gly Lys Gln Leu Glu Leu Gln Tyr
140 145 150
gca gct gaa aca agt gct ctg gag cct cct cac aag tat gtt cct tgg 591
Ala Ala Glu Thr Ser Ala Leu Glu Pro Pro His Lys Tyr Val Pro Trp
155 160 165
gta gtt gtg gat gga gaa cca ctc tat gag gat tat gag aac ttc tta 639
Val Val Val Asp Gly Glu Pro Leu Tyr Glu Asp Tyr Glu Asn Phe Leu
170 175 180 185
agc tat ctt tgt aag gct tat aaa ggc act gtt aca ccc caa agt tgc 687
Ser Tyr Leu Cys Lys Ala Tyr Lys Gly Thr Val Thr Pro Gln Ser Cys
190 195 200
acc caa gca tca tac cta aga gaa gtg aat gca aag cct aag cat tcg 735

Thr Gln Ala Ser Tyr Leu Arg Glu Val Asn Ala Lys Pro Lys His Ser
205 210 215
gtt tgc tac aag gac agt ggg att atg cta aca tgg gaa aaa gtg agg 783
Val Cys Tyr Lys Asp Ser Gly Ile Met Leu Thr Trp Glu Lys Val Arg
220 225 230
tca acc att gca tca tgg atg cac tag atgaatcttg gcagtcatt 830
Ser Thr Ile Ala Ser Trp Met His Stop
235 240
ttagatgtgc caatgctgca tttagtagca gtttgttcgt gttatcttgt gtgtagttgt 890
gttgatccct ccaaaagtgc aaatagaact tgtgatgcac ataaaacat atcgactct 950
cataaaaaaa ataaaacat gatgtgatg tgtatgagg acttaagtac aatatata 1010
gagagagaaa aaaaaaagta agtgcaattc taaaaaaaa aaaaa 1055
<210> 2
<211> 1055
<212> DNA
<213> Glycine max
<220>
<221> CDS
<222> (85)-(321), (389)-(454), (1160)-(1300), (2414)-(2506), (2707)-(2895)
<400> 2
ttggtgtag ctgttacgag tcagcaagag agaaatatct tgttgagttg atttctgatc 60
ccaaattgag tttgagtttc agcc atg gta tct gtt tct ctg ctc tta ttc 111
Met Val Ser Val Ser Leu Leu Leu Phe
1 5
cta tct ttc ttt gcc ttc ttt gca ccc tct gaa tct cac gac aaa gtc 159
Leu Ser Phe Phe Ala Phe Phe Ala Pro Ser Glu Ser His Asp Lys Val
10 15 20 25
tcg ttg gaa ttg tac tac gaa agc tta tgc ccc tac agc gcc aac ttc 207
Ser Leu Glu Leu Tyr Tyr Glu Ser Leu Cys Pro Tyr Ser Ala Asn Phe
30 35 40
atc gtg aat cac ctc cca aaa atc ttc acc cca gat ctt gcc cca atc 255

15 16

Ile Val Asn His Leu Pro Lys Ile Phe Thr Pro Asp Leu Ala Pro Ile
45 50 55

gtt cac ctc aaa ctc gtt cct tgg ggc aat gcc aaa ctc cgt gcc aac 303
Val His Leu Lys Leu Val Pro Trp Gly Asn Ala Lys Leu Arg Ala Asn
60 65 70

gcc acc ttt gac tgc cag gtactctctt ctccaatctt atacacactc 351
Ala Thr Phe Asp Cys Gln
75

ttcacctatt tttatftaat tctgatgaca attgcag cat ggg cca tat gag tgc 406
His Gly Pro Tyr Glu Cys
80 85

ttg ctg aac aca gtt gaa gcc tgt gca att cac atc tgg cct caa ctc 454
Leu Leu Asn Thr Val Glu Ala Cys Ala Ile His Ile Trp Pro Gln Leu
90 95 100

gtaagtaaca attgctgtgc ctttgcccct ttctttcttt ctttcttttt tcttttttct 514
ttgatttggg atcagaactt aagaatgctt tttttattat tataatatta tgttattatt 574
acggattatt cagcatcgat tttttcagat gattaatfff tattggattt tgattaggtt 634
atftttctcc aattatfttt tttatccaaa aaggagatga agttgttggg gaacttagga 694
gtgcctttat tattattgtt attatgttat tgtaaggag tattcattga ttttatagat 754
gattaatfff cattggatca cttttgatta ggttatttcc tttcaattat tttatftttt 814
atgaaaaagg agatgaagtc tggtgccacc tggacgaatg agttgttggg gaacttagaa 874
gtactattat aattatgata aatfttgaaa aataaaaaata aaaaaggagt actftttggtc 934
agtgtaacag tgcttttagaa agatgctcta gcttgagcac tgtgcttaag ttgctataaa 994
ttcctgttat ccgagtttga ttccatgga taaagttgtt attataaatt ctggcacct 1054
gagttcagat tcaatfttga cggataaaaa agtttagcag tgctttgcat gtgtacctcg 1114
gtatctctgt tatgagtgtt ttatgtaaat cacactcttt tgcag agc aaa cat ttt 1171
Ser Lys His Phe
105

cct ttc atc tac tgt gtt gag gat ctg gtg ttt cag agt aag cgt gag 1219
Pro Phe Ile Tyr Cys Val Glu Asp Leu Val Phe Gln Ser Lys Arg Glu
110 115 120

gaa tgg gaa tct tgt ttt gag aaa ctg gat ctt gat tcg gaa cct att 1267
Glu Trp Glu Ser Cys Phe Glu Lys Leu Asp Leu Asp Ser Glu Pro Ile
125 130 135

aat cag tgt tat aat agt gaa cat gga aaa cag gttgggtgtt ttctctgatt 1320
Asn Gln Cys Tyr Asn Ser Glu His Gly Lys Gln
140 145

tcgctcttgt ccttgcttga gctaccatgt tgtgattgag atagtattat gtacattagt 1380
ttagttttca tgcctccga gtgtaagtat gacttgtaaa ctagtacta caaagtattc 1440
taagatccta tttggataaa atftgttaca aatacttga ggtgaagaaa ataagaagat 1500
aaaaaaatt gagfttttct tccgtgagtt aaaaatcaat caacttcggc acctcggatt 1560
ttggaagagt taaatgagac tgagaacttt tatggaagtt tagtgtaag ttgattttag 1620
cttgtgagag aagctcaatt tftttactt tcttattttc tcatcttata agtgcttatg 1680
aaaaagtga tctaaacagg actftaaaat tttttaacta acctttctct tctatggttg 1740
ftttttatga acctaaagt tatactaaca ttgctaaatg atacaagcta gtggaagtt 1800
gggaattaga aaatftgaac agttatacag ttacagggtt ggtggtgact tctcaattcc 1860
ftatftaaac taatcttaag tgtcatagat aattgaaatc tcttcgaggt gagatataat 1920
ccatgtcaac tagccgtaac tgggctctgt tatgtgaaac ttggatggtg aggtctggtg 1980
fttgttatgt gccggtctac tactttgctt ftattcttca agacaacaaa atgaaatagc 2040
taatgagaag taatftaac ftattccagc aattacttca aaaaatggtc atgttcttca 2100

17
 tttttacaaa actctaaagc agtgtaatat gatctacaaa ttaccttatt gtcatacatg 2160
 attgctcttt gctacagata ctgccacacg atctaataga actgccaatt tgtagcttct 2220
 gaacattcat ctttcttcaa gctaatttct caaaatcctt tttcaatctt gctgaatgtg 2280
 aagactatct cttgatgaaa aacaaaacat atgacagatg tgctgaatac atcattgttt 2340
 gactggaatt ttgtctcaag tagcatttac tatcttgttt tgatcattag taacatttac 2400
 tatctttatg tag ttg gag cta caa tat gca gct gaa aca agt gct ctg 2449
 Leu Glu Leu Gln Tyr Ala Ala Glu Thr Ser Ala Leu
 150 155 160
 gag cct cct cac aag tat gtt cct tgg gta gtt gtg gat gga gaa cca 2497
 Glu Pro Pro His Lys Tyr Val Pro Trp Val Val Val Asp Gly Glu Pro
 165 170 175
 ctc tat gag gtca gttttgcaat tttatatcgg ctgaaccaag ccccatcccc 2550
 Leu Tyr Glu
 agtgtttga attttcacta aaaaatactt ctaattaaca aataacttat aacactgaaa 2610
 tttgtgtttt acatatgtaa gttgtgttat attaatgaat cgaattagaa gtgtacagga 2670
 aattcaaaaa gaacttgcaa atttcttttg ctgcag gat tat gag aac ttc tta 2724
 Asp Tyr Glu Asn Phe Leu
 180 185
 agc tat ctt tgt aag gct tat aaa ggc act gtt aca ccc caa agt tgc 2772
 Ser Tyr Leu Cys Lys Ala Tyr Lys Gly Thr Val Thr Pro Gln Ser Cys
 190 195 200
 acc caa gca tca tac cta aga gaa gtg aat gca aag cct aag cat tgc 2820
 Thr Gln Ala Ser Tyr Leu Arg Glu Val Asn Ala Lys Pro Lys His Ser
 205 210 215
 gtt tgc tac aag gac agt ggg att atg cta aca tgg gaa aaa gtg agg 2868
 Val Cys Tyr Lys Asp Ser Gly Ile Met Leu Thr Trp Glu Lys Val Arg
 220 225 230
 tca acc att gca tca tgg atg cac tag atgaa tcttggcagt gcattttaga 2920
 Ser Thr Ile Ala Ser Trp Met His Stop
 235 240
 tgtgccaatg ctgcatttag tagcagtttg ttcgtgttat cttgtgtgta gttgtgttga 2980
 tccctccaaa agtgcaaata gaacttgta tgacataaaa accatatcgt actctcataa 3040
 aaaaaataaa accatgatgt gtatgtgtat gaggtac 3077
 <210> 3
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Designed nucleotide based on legumaturain amino acid sequence.
 <400> 3
 gcnaayttya thgtnaayca 20
 <210> 4
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Designed nucleotide based on legumaturain amino acid sequence.
 <400> 4
 tcncrctna cnacnaccca nggnacrta tt 32
 <210> 5

19
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Designed nucleotide based on legumaturain cDNA sequence.
 <400> 5
 ttggtggtag ctgttacgag tcagc 25
 <210> 6
 <211> 26
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Designed nucleotide based on legumaturain cDNA sequence.
 <400> 6
 gtacctcata cacatacaca tcatgg 26

20

【要約】

【課題】 アスパラギン残基特異的エンドプロテアーゼの遺伝的情報を得て、本酵素の量産化を可能にすること。

【解決手段】 特定のアミノ酸配列を有する植物由来の

アスパラギン残基特異的エンドプロテアーゼをコードする cDNA および配列表の配列番号 2 記載のアミノ酸配列を有する植物由来のアスパラギン残基特異的エンドプロテアーゼをコードする遺伝子。

フロントページの続き

(58)調査した分野(Int.Cl.⁷, DB名)

C12N 15/00 - 15/90

BIOSIS (DIALOG)

WPI (DIALOG)

GenBank / EMBL / DDBJ / G

eneSeq